Sound communities:

A quantitative proposal for studying bilingualism in context

Charlie Nagle[1], Melissa Baese-Berk[2], Mark Amengual[3], and Joseph V. Casillas[4]

[1]College of Liberal Arts, Department of Spanish & Portuguese, University of Texas at Austin

[2]Division of the Humanities, Department of Linguistics, University of Chicago

[3]Humanities Division, Department of Languages and Applied Linguistics, University of California, Santa Cruz

[4]College of Arts and Sciences, Department of Spanish & Portuguese, Rutgers University

Author Note

Charlie Nagle http://orcid.org/0000-0003-2712-2705

Melissa Baese-Berk http://orcid.org/0000-0002-4855-9319

Mark Amengual http://orcid.org/0000-0003-2007-9687

Joseph V. Casillas http://orcid.org/0000-0001-8735-9910

Author roles were classified using the Contributor Role Taxonomy (CRediT; https://credit.niso.org/) as follows: *Charlie Nagle***:** conceptualization and writing – original draft. *Melissa Baese-Berk***:** conceptualization and writing – original draft. *Mark Amengual***:** conceptualization and writing – original draft. *Joseph V. Casillas***:** conceptualization and writing – original draft. Correspondence concerning this article should be addressed to Charlie Nagle, College of Liberal Arts, Department of Spanish & Portuguese, University of Texas at Austin, 150 W 21st St., Austin, TX 78712, Email: cnagle@austin.utexas.edu

Abstract

Bilingualism researchers have intensively studied how learning and using multiple languages affects all levels of linguistic structure. In this strand, examining diversity in the bilingual experience and the extent to which variables like language dominance regulate crosslinguistic interaction has been of special interest. However, most studies sample small groups of bilinguals from a single research site, creating a twofold generalizability problem. First, with small samples it is unlikely that researchers will be able to fully capture and quantify the range of variables known to affect findings. Second, when bilinguals are recruited from a single site, it is impossible to determine if findings are site-specific or apply to bilinguals more broadly. To address these issues, we propose a large(r)-scale, multisite approach to bilingualism research. We believe that such an approach, when informed by open science practices, has the potential to significantly advance the state of the art.

Keywords: bilingualism, multisite research, big team science, open science

Sound communities: A quantitative proposal for studying bilingualism in context

Introduction

In recent years, interest in bilingualism and second language acquisition has rapidly expanded. Researchers have recognized that referencing a monolingual norm ignores most individuals' linguistic experience because bi/multilingualism is more common globally than monolingualism (Grosjean, 1997; Kroll & Bialystok, 2013). One major challenge in this line of research is terminological because the terms "bilingual" and "multilingual" often refer to individuals from radically different linguistic and sociocultural backgrounds who show non-trivial differences in the experiential and attitudinal variables that are considered fundamental to the bi/multilingual experience. Indeed, this is precisely why Bialystok (2021) characterized bilingualism as a "package of Swiss cheese, with different manifestations of bilingual experience placing the holes in different places that together define the experience" (p. 2). In this sense, one can see how characterizing someone as "bilingual" without a rich understanding of their background can lead to a false impression that one "bilingual" should somehow resemble another when those individuals might actually be drawn from distinct bilingual populations. Historically, bilingualism research has focused on the individual but asked questions about bilingualism as a whole. This approach assumes that bilingualism is a single, coherent characteristic, when, in reality, it is a highly multidimensional, variable phenomenon whose manifestation depends on myriad factors. We contend that, to date, individual studies of bilingualism reveal more about the experiences and behaviors of specific bilingual samples than about bilingualism itself. Studying bilingualism in all of its complexity therefore requires thinking about both bilingual individuals and the (bilingual) communities, geographies, and social strata they inhabit.

In what follows, we build the case that bilingualism research would benefit from a multisite approach supported by the principles of open science. We also contend that bilingualism is inherently dynamic, and, as such, necessitates large-scale, longitudinal assessment. To support this view, we provide illustrative examples from specific variables of interest in bilingualism research, such as language dominance, and focus on how our proposed approach would benefit a concrete area, speech research.

As Bialystok's "Swiss cheese" metaphor alludes to, grouping units are important because individuals are embedded into layers of structure that shape their behaviors. Bilingualism can certainly be conceived of as an individual characteristic because there is no doubt that each bilingual holds attitudes towards and makes choices about their bilingual identity, and those attitudes and choices affect their behaviors, linguistic and otherwise. At the same time, those attitudes, choices, and behaviors are also constrained, or at least influenced, by community-level language ideologies, patterns of language dominance and use, and so on. We have no doubt that bilingualism research must continue to focus on individuals and their characteristics, but it will be difficult to understand the full scope of bilingual behavior if the geographies where bilinguals live and work are not considered. To date, most studies have sampled bilinguals from a single community, which means that it is currently not possible to understand the extent to which individual behavior is shaped by community-level patterns. This might explain why, for instance, studies sometimes report contradictory findings. In fact, those findings may not be contradictory at all. Instead, they may simply reflect variation in the bilingual communities from which participants are drawn, underscoring the fact that bilingualism in one community or region may not look anything like bilingualism in another. As a result, effects observed in one area may not

be observed in others, and when observed, the magnitude of the effects is likely to vary considerably.

In bilingualism research, sampling can be a serious challenge. If research is carried out in a single bilingual community, researchers might expect to recruit a relatively homogenous sample of bilinguals. However, even in this context bilinguals are likely to show substantial individual differences in language use, social networks, language attitudes, and proficiency, to name but a few key variables. Certainly, some variation is expected and, from a quantitative perspective, necessary and desirable, but we do not yet have a firm understanding of how individuals vary within and across communities and how this variation leads to variation in the types of behaviors observed for diverse bilingual populations. As a concrete example, consider Spanish-English bilinguals in Chicago. Opportunities to use Spanish and English might vary neighborhood-by-neighborhood, and that variation might condition attitudes toward the use of one language or the other, as well as attitudes toward language mixing and code-switching. Individuals also come with their own personal and intergenerational histories. Perhaps some are recent immigrants to the US, whereas others are first or second generation immigrants. Compare this sample to Spanish-English bilinguals in major urban centers in Texas, where community attitudes toward bilingualism are likely to be quite different, given proximity to Mexico and prevailing attitudes toward Spanish in the region. Additionally, in Chicago there are more opportunities for dialectal leveling and accommodation, due to extensive contact between Puerto Rican and Mexican Spanish speakers (see Potowski & Torres, 2023), whereas in cities like Houston, El Paso, and San Antonio, despite substantial sociodemographic diversity, there may be fewer features that arise from dialect contact due to the presence of a larger Mexican(-American) Spanish-speaking community (but see Bayley et al., 2012). To put a point on it, what is possible

and acceptable language use for a bilingual in Chicago may not be possible and acceptable for a bilingual in El Paso and vice versa. Likewise, what is possible and acceptable for one bilingual in El Paso may not be possible and acceptable for another, depending on the social groups they interact with, the beliefs they hold, and the identities they construct for themselves (cf. Achugar & Pessoa, 2009; Velázquez, 2013). This means that we should perhaps not expect to observe consistent effects across geographies, communities, and individuals, but rather highly nuanced, complex, and even contradictory (i.e., context-specific) ones.

There are myriad forms of bilingualism that could be investigated, but Spanish-English bilingualism presents a unique and significant opportunity because Spanish is the most frequently spoken additional language in the United States (Bureau, 2022). Its diverse speakers include individuals raised in a Spanish-speaking country who have immigrated to the US, individuals raised in the US who have grown up speaking Spanish, and English speakers who have acquired Spanish through a combination of immersion and/or formal instruction. States like California, Arizona, Texas, and Florida have large and well-documented bilingual populations, but there are also populations of bilingual speakers around prominent US urban centers throughout the country and large populations of bilingual speakers in agricultural regions. These communities have their own characteristics in terms of language contact and use, as do their speakers. As such, they offer an unprecedented opportunity to study how contact between Spanish and English is reshaping linguistic structure in both languages. For instance, Spanish-English contact may be creating new phonetic norms or even new phonological targets in US varieties of Spanish and English. And, to our point about community-level sampling, in communities with balanced language use profiles, or communities where code-switching and language mixing are the norm, there may be more phonetic blending. The only way to address

these types of questions, questions that can shed light on the very nature of language change, is by revisiting methodological practices in bilingualism research.

In this paper, we examine current modes of inquiry in bilingualism. It would be impossible to address all aspects of research methodology, so we focus on considerations that are often not directly acknowledged or addressed in empirical research but nonetheless have a hidden impact on research findings. We conclude the paper with a series of recommendations for engaging in the type of multisite research that is likely to drive the field forward, anchoring practices in our immediate area of interest: bi/multilingual speech research.

Bilingualism Across Space and Time

A bilingual is typically defined as an individual who has command of more than one language (Grosjean, 1997). One challenge is that the division between monolingual and bilingual is not clear cut because the two categories represent a continuum rather than a binary classification. Most researchers would agree that an individual who knows one language well (i.e., their native language) and only knows how to count in a second language is likely not bilingual. However, researchers also understand that a "fully-balanced" bilingual, who demonstrates equal proficiency in both their languages in all domains and for all contexts and situations, is theoretically possible but not common (see Grosjean, 1989). Indeed, the continuum from monolingual to bilingual encompasses a wide range of speakers, including simultaneous and sequential childhood bilinguals, heritage language learners, and second language learners, among many others (and these terms reflect our collective effort to derive discrete categories that capture important experiential and acquisitional characteristics about bilinguals). Furthermore, looking within instead of across individuals, the language dominance and proficiency of a given

person is likely to change over the lifespan, which means that bilingualism is best understood as a dynamic phenomenon and, as a result, perhaps best measured longitudinally.

The same logic applies to communities, insofar as synchronic variation between communities is also accompanied by diachronic variation within them. In an era of globalization and mass migration, the bilingual communities that have historically been the object of study may look and therefore behave very differently over time. On the one hand, this variation presents an incredible research opportunity: If we can systematically tag studies by decade and recover some of the sociodemographic data associated with those communities, then we could end up with a powerful lens into diachronic patterns of stability and change in demographics, attitudes, and behaviors (see e.g., Mejías et al., 2003). On the other hand, this variation presents challenges because it cannot be assumed that data collected from a bilingual community 30 or 40 years ago is representative of that community today. By extension, we should not assume that bilingualism findings published decades ago necessarily hold now. In short, the very nature of what it means to be bilingual has changed and will continue to do so. Thus, there is an important temporal dimension to bilingualism research that, in our view, we do not adequately account for with current research practices. That is not to say that findings from older studies are somehow invalid, but rather that those findings are at least partially a property of when the data were collected. One need only consider the rapid evolution of technology to realize that bilinguals today have opportunities for language learning and interaction that were inconceivable a few decades ago. Scholarly inquiry has also changed dramatically over the lifetime of contemporary bilingualism research, leading to updates in how bilingualism is conceptualized and measured. To further illustrate the dynamic nature of bilingualism, we briefly consider one especially influential way of quantifying the bilingual experience: language dominance.

In the context of bilingualism, language dominance "refers to observed asymmetries of skill in, or use of, one language over the other" (Birdsong, 2014, p. 374). At the level of the bilingual individual, this asymmetry may play out, for example, in speaking or reading aloud more words per minute in one language than in the other, in more frequent everyday use of one language over the other, or in greater ease processing natural speech in one language over the other. Importantly, dominance is best understood as a relative and gradient construct, both within and across individuals. For instance, a second generation English-Spanish bilingual speaker in California may be more dominant in English than another English-Spanish bilingual in the same community.

Language dominance is domain-specific and contextually-sensitive. Bilinguals are not monolithic when it comes to the contexts and frequencies of use of each language. An appropriate understanding of dominance is predicated on careful consideration of the situations and contexts in which each language is used at work and at school, with friends or family, for the purpose of caregiving, inner speech, and so on. As Grosjean (2010) observed when explaining the complementarity principle, "bilinguals usually acquire and use their languages for different purposes, in different domains of life, with different people. Different aspects of life often require different languages" (p. 29). In other words, language dominance is inherently dynamic. Over the course of a bilingual's lifetime the relationship between their languages may show dominance shifts (relative dominance changes as a matter of degree, but not reversal) or even dominance switches in cases of extensive language exposure to and use of the non-dominant language leading to attrition in the dominant language (one language replaces the other as the dominant). These shifts and switches can be brought about by diverse circumstances, including temporary relocation or permanent immigration to a different linguistic community, change of

employment, change of educational setting, caregiving arrangements, or even engagement with a new social network. Changing circumstances ultimately modulate the relative frequency of contact with, and use of, the two languages, leading to shifts and switches in domain-based dominance. Here too, sampling is critical. Many studies on bilingualism rely on college-aged undergraduate student participants, and for work on heritage speakers or speakers of minority languages in the US, these participants may commonly find themselves in a predominantly English-speaking environment for most of their lives. This observation underscores the value of longitudinal studies (Grosjean, 2010) that situate the effects of dominance shifts and switches within the general context of bilingualism as a lifelong phenomenon in which dominance is impacted by situational and transitory language-relevant events. Work on the speech production of Steffi Graff (Leeuw, 2019), Arnold Schwarzenegger (Kornder & Mennen, 2021), and Ole Gunnar Solskjaer (Kelly, 2022) are compelling individual examples, but longitudinal work is also needed on bilingual communities at large.

Since its introduction as a measurable trait in bilingualism by Lambert (1955), and its subsequent rise in visibility (e.g., Bahrick et al., 1994; Grosjean, 2010; Hamann et al., 2019; Silva-Corvalán & Treffers-Daller, 2016), dominance has emerged as a notable predictive factor in studies on bilingualism. As a result, in recent years a number of instruments have been widely used to assess language dominance, such as the Bilingual Language Profile (Birdsong et al., 2012), Bilingual Dominance Scale (Dunn & Fox Tree, 2009), and Self-Report Classification Tool (Lim et al., 2008). Other questionnaires and tests that address bilingualism more generally may contain items about dominance; these include Language Experience and Proficiency Questionnaire (Marian et al., 2007), Language and Social Background Questionnaire (Luk & Bialystok, 2013), the Bilingual Input-Output Survey subset of Bilingual English-Spanish

Assessment (Peña et al., 2014), and L2 Language History Questionnaire (Li et al., 2006). In addition to these instruments, language dominance has also been assessed with home-grown questionnaires that include self-ratings of relative language use along with self-ratings of proficiency (Kartushina & Martin, 2019), questionnaires on language background and exposure, language use, as well as other social variables (Tomé Lourido & Evans, 2019), self-assessments of speaking, understanding, writing, and reading in each language (Flege et al., 2002) or self-ratings of fluency and daily language use of each language (Antoniou et al., 2012).

Diversity in instrumentation in the measurement of language dominance is matched by diversity in how dominance is conceptualized and how its impact on bilingual behavior is analyzed. To illustrate this point, consider Amengual and Chamorro (2016) and Tomé Lourido and Evans (2019), two studies examining the Galician mid-vowel contrasts in the production and perception in Spanish-Galician bilinguals. Tomé Lourido and Evans (2019) used a language background questionnaire to obtain information about language dominance and partition the bilinguals into groups (Spanish dominant, Galician dominant). Amengual and Chamorro (2016), for their part, also separated the bilinguals into groups based on language dominance, in this case via the BLP (Birdsong et al., 2012). In addition, Amengual and Chamorro (2016) included the language dominance scores derived from the BLP as predictors in subsequent bivariate models, which allowed them to uncover relevant patterns of between-speaker variation. Both studies found that language dominance played an important role in the production and perception of the Galician mid-vowel contrast. In a similar vein, Melnik-Leroy et al. (2022) examined vowel production and perception in bilingual speech, though in this case focusing on English-speaking late learners of French. Whereas Amengual and Chamorro (2016) and Tomé Lourido and Evans (2019) found robust differences in the production and perception of the early bilinguals based on

language dominance, Melnik-Leroy et al. (2022) used the BDS (Dunn & Fox Tree, 2009) to ensure heterogeneity amongst their late learners, focusing on language dominance and self-report to contextualize late learner behaviour in a qualitative manner. All three studies operationalized language dominance in some way and included it as part of their analytic strategy. They differ in how and the extent to which the construct was implemented. Importantly, these studies highlight how researchers have made use of the same construct, dominance, in different, yet effective, ways to explore bilingualism.

If a current goal in the field is to create a comprehensive picture of bilingualism and its evaluation in distinct speech communities, our current data collection and reporting practices make it difficult to do so. As a consequence, broad generalizations about bilingual populations are likely to be inaccurate, if not altogether misguided. Without considering the specific characteristics and contexts of each speech community, such generalizations risk oversimplifying the complex nature of bilingualism. Comprehensive and context-sensitive research is essential for accurately capturing the dynamics of bilingualism and understanding how it evolves within different communities over time. In other words, we contend that in our current research environment, we must begin to address what makes various communities of study distinct and begin to address the ways in which our studies can be made more comparable to one another so that generalizations about these communities may be more robust. One way to achieve this goal is via open science practices. If we conduct large-scale, multi-site research, then the results of the same study can be compared across bilingual populations, allowing us to determine which factors are unique to a particular community and which are representative of broader bilingual populations (i.e., part and parcel of bilingualism itself). Returning to language dominance, a variable we have singled out as an illustrative example, we note that our field will benefit from a

consistent conceptualization and operationalization of language dominance that enables direct comparisons within and across the bilingual communities we investigate. In the next section, we give an overview of the open science movement before making a case for embracing its principles in bilingual speech research.

Open Science and Multisite Research

Open science is not new but has gained increasing traction over the last decade. In part, this traction can be traced back to the replication crisis in psychology, where concerns related to the non-replication of sensational findings led to the creation of the Open Science Collaboration (2015), a large-scale endeavor to probe reproducibility in the psychological sciences. One of the basic principles of open science is to ensure research quality through transparency and reproducibility. Around the time psychology was grappling with these issues, language researchers began to do the same. For instance, in 2016, Heidi Byrnes, then editor of *The Modern Language Journal*, described an uptick in the collective methodological consciousness of applied linguistics researchers, stating that "methodological issues… demand a kind of professional scrutiny that goes directly to the core of what we do and what we know and what we can tell our publics that we know" (p. 825), presaging a surge in methods-oriented research. In (applied) linguistics, initiatives such as the Instruments and Data for Research in Language Studies (IRIS, 2011–, www.iris-database.org) and Open Accessible Summaries in Language Studies (OASIS, 2018–, www.oasis-database.org) are signs that researchers are deeply concerned with the transparency, reproducibility, and quality of language research, which has been systematically evaluated in several publications (e.g., Plonsky, 2013). Furthermore, journal-based initiatives, such as replication studies (McManus, 2024) and Registered Reports (Marsden et al., 2018), further reinforce that now more than ever open-research practices are a mainstay of language

science (for an overview, see Marsden & Morgan-Short, 2023). One especially compelling feature of this new research mode is an emphasis on a multisite approach (Morgan-Short et al., 2018).

One of the main benefits of a multisite approach is increased sample diversity, which we view as critical for bilingualism research. It is well documented that most research participants are drawn from predominantly Western, educated, industrialized, rich, and democratic (WEIRD) populations (see, e.g., Henrich et al., 2010; Nielsen et al., 2017). Linguistic samples are certainly WEIRD (see Bochynska et al., 2023). They also tend to be heavily skewed toward academic convenience samples consisting predominantly of university undergraduates (Andringa & Godfroid, 2020). An overrepresentation of university undergraduates brings with it a whole host of hidden sampling variables related to socioeconomic status and educational access. And, as previously mentioned, for bilingualism, a focus on young adults typically entails capturing their bilingual profile at a very specific moment in their language trajectories.

When sampling practices focus on a narrow slice of the population, researchers run the risk of constructing a normative or modal response against which all others should be compared. To give but one example, monolingual native speakers were traditionally positioned as the benchmark against which bilinguals should be compared, a methodological and analytical choice that can be traced back to early research on age, attainment, and critical periods for language learning. Yet even purportedly monolingual native speakers (and here, it begs the question: In our globalized world, is anyone truly monolingual anymore and if so, how is monolingualism best operationalized?) are not a homogenous group when representative samples are constructed based on sociodemographic population-level characteristics (Andringa, 2014). Recent efforts in the field of psycholinguistics have begun to push back on these and other related issues (see

Kutlu & Hayes-Harb, 2023), which likely represents a larger trend in linguistics at large. At the same time, researchers face serious practical challenges in creating such samples.

First, language researchers cannot engage in truly random sampling, where each member of the population has an equal likelihood of inclusion in the sample. Researchers are bound by the research contexts, geographic, bureaucratic, and institutional, in which they work. Yet, where a single researcher or research team might struggle, multisite research can succeed. Multisite studies are all but guaranteed to generate a larger, more diverse, and therefore more representative, sample (Moranski & Ziegler, 2021), which means that they have a higher chance of achieving generalizability and robust external validity compared to their single-site counterparts. Another advantage to multisite research is that at the analytical stage, researchers can examine the extent to which findings vary across clustering variables such as site. In this way, both the aggregate effect, pooling across sites, and site-specific variation can be modeled. Such an approach has clear interpretational advantages given that it affords the research group the opportunity to compare findings at multiple levels of data structure (Morgan-Short et al., 2018).

Another clear advantage to a multisite approach is the fact that it demands clarity with respect to research tasks and procedures, insofar as they must be aligned across sites (Morgan-Short et al., 2018). Transparent, reproducible practices ensure that observed differences are due to true differences rather than to non-trivial deviations in study implementation. Such deviations are not the product of researcher apathy or negligence. Indeed, most, if not all, researchers want to carry out high quality studies that have the greatest chance of lending real insight into the research topic with real implications for the target domain(s). Yet, even under the best circumstances, “the lifecycle of any research study is beset by a series of decisions, many of

which are essentially arbitrary, whose consequences are usually unknown" (Bolibaugh et al., 2021, p. 804). Making the research process transparent and reproducible pays dividends during the project and afterwards: during the project, taking the time to create reproducible workflows ensures fidelity of implementation and afterwards, future researchers can replicate the project while critically reassessing the methodological decisions that were made at the time. In an ideal scenario, if research tasks and procedures are well documented and publicly available, then researchers outside of the core research network may be able to participate and contribute data to the study (Morgan-Short et al., 2018). This has the added benefit of creating an equitable research landscape where scholars from around the globe can collaborate and interact, which in turn should lead to even more representative samples and research findings.

As data sets grow larger and more complex, so too do analytical choices. Even if tasks and procedures are rigorously implemented and followed, researchers face a garden of forking paths at the analysis stage, where researcher degrees of freedom can lead experts to engage in radically different analytical strategies. These strategies, like their methodological and procedural correlates, can be completely justified (for examples in phonetic science, see Roettger et al., 2019) yet still have a serious effect on the conclusions that researchers reach (Coretta et al., 2023). The issue is further complicated in bilingualism research, as it is an inherently complex and heterogeneous phenomenon. It is no surprise, then, that analyzing bilingual data involves additional steps that often present the researcher with even more decision points, or researcher degrees of freedom, all of which can have downstream consequences. Preregistration, including preregistered analysis plans (Roettger, 2019), can help mitigate researcher degrees of freedom, but even when analyses are preregistered, the realities of the data often necessitate changes (e.g., planned models do not converge). Models are objective statistical tools, but the

people who implement them are not, and any analysis requires substantial statistical decision-making. Where bilingualism is concerned, there tends to be a larger set of potential between- and within-subjects (and between- and within-language) comparisons that can be made, and, indeed, diverse ways of approaching those comparisons, making analytical transparency critical.

Opportunities for Pronunciation Research

We believe that the larger-scale, multisite approach we advocate for will be useful for all types of bilingual research. However, this approach also holds special promise for two specific areas of research, both of which are related to phonetic and phonological aspects of language. These two areas have consequences for the intersection of bilingualism and linguistic theory and for individuals who are bilingual. The first is sound change induced by language contact. The second is accentedness and the role of accentedness in the perception of both speech and speakers.

Bilingual contexts, and pronunciation research specifically, can provide insight into contact-induced language change because bilingual communities provide opportunities for both cross-sectional data (i.e., older and younger speakers within the same community) and longitudinal investigations of how contact between languages may shift pronunciation in each language over time. Bilingualism has often been credited as the source of sound change (Sankoff, 2002; Weinreich, 1968), and recent research has focused on integrating community- and individual-level sound change in theories of bilingualism and contact-induced sound change (e.g., Filipović & Hawkins, 2019; Hawkins & Filipović, 2024; Yao & Chang, 2016).

However, even studies that explicitly acknowledge the role of bilingualism and language contact in sound change typically focus on a single community. If we leverage the heterogeneity of linguistic communities where bilingualism occurs, we can better understand the typicality of

particular types of sound change and the frequency of such changes. Investigations of communities across multiple sites and contexts will provide us with powerful tools to inform our theories of language change and phonological theory more broadly.

Another area of special opportunity is related to accent. A speaker's accent is used by listeners to determine facts about the speaker, including, perhaps most obviously, their likely place of origin (e.g., Carmichael, 2016). However, a speaker's accent is also used by listeners for a variety of other aspects of perception including properties of the speaker that are not explicit in the speech, including intelligence, trustworthiness, and competence (e.g., Jackson & Denis, 2024; Lev-Ari & Keysar, 2010). These features then impact how listeners perceive the speech produced by the speaker (Jiang et al., 2020). That is, a speaker's accent is very likely to significantly impact how both people perceive the speaker and how they perceive their speech.

In the context of bilingualism, this is especially interesting because bilingual speech is often classified as "other" in both languages. For example, Hispanic English is a specific dialect and has a specific accent associated with it, even for speakers who do not speak Spanish. This variety is distinct from what one might consider "Spanish-accented English" (e.g., Mendoza-Denton, 2014). Similarly, Spanish spoken by Heritage Spanish Speakers is sometimes recognized as being distinct from "native" varieties of Spanish (e.g., Potowski, 2012). In bilingual contexts, one can imagine that the accents in each language variety may differ as a function of both issues of dominance, as described above, and a variety of social and identity factors. However, crucially, the varieties of each language spoken will also likely differ as a function of the specific context. For example, Wolfram et al. (2004) described an emergent variety of Hispanic English in the American South, which shares some phonological features with other varieties of Hispanic English but is also distinct from those varieties in important

ways. Studying production and perception of accents in bilingual communities can provide insight into both the linguistic and social features of those communities. The multisite approach championed here would allow researchers to investigate what properties of perception and production of accents are specific to a given context and which are broader features of a bilingual community more generally.

Experimental Design and Analysis in Bilingual (Speech) Research

Most studies on bilingual speech use some variation on a common design. Typically, a (convenience) sample of 20–30 bilinguals is recruited and tested in both of their languages. There may be a language mode manipulation, where participants are exposed to and given instructions in one of the languages to encourage language-specific processing (Lozano-Argüelles et al., 2021), or the target stimuli themselves may be designed with the goal of influencing processing strategies (Casillas & Simonet, 2018). The participants may differ from one another across a range of key variables such as age of onset of L2 learning, quantity of L2 input and exposure, or composite variables such as language dominance (e.g., Birdsong et al., 2012). Then, their performance on perception and production tasks is related to their individual differences profiles, processing mode, or another independent variable that is the focus of the study.

This research paradigm has generated a large and important body of work on the (socio)linguistic and contextual variables that affect the extent to which bilinguals maintain and deploy language-specific representations in speech perception and production. Researchers have now begun to use the tools of meta-analysis to sort through this research, demonstrating, for instance, that the “compromise” phonetic categories that bilinguals appear to show in their languages may be more an artifact of study design than a generalizable characteristic of bilingual

behavior (Casillas, 2021).

Meta-analyses are powerful, but they are limited by the study samples on which they are based. Estimates can be adjusted based on study characteristics, including potential biases that become evident as studies are aggregated and analyzed, but the original gaps remain (i.e., what does not exist cannot be synthesized). Crucially, although we focus intensively on the bilingual speakers that provide data, it bears mentioning that the same argumentation applies to studies involving listener-based evaluations of speech. After all, listeners are at least as complex as speakers, insofar as they bring with them expectations and biases that shape their perception of the speech and the speaker (Lindemann & Subtirelu, 2013).

In our view, current practice has created two critical gaps. First, the modest sample size of most studies makes it difficult, if not impossible, to accurately and comprehensively characterize the individuals included in the research. Thus, whether those individuals represent a normally-distributed continuum of language dominance, use, and other key quantitative variables remains an open question. Another gap that we have outlined throughout this paper is that studies have been conducted at a single site, which makes it impossible to separate the characteristics of the individuals from the characteristics of the communities they live in. Site-specific studies were a sensible and necessary starting point for bilingualism research. However, technological advances in data storage, processing, and analysis have made larger-scale, multisite projects more feasible than ever. If individual bilingual behaviors coalesce into community standards and practices, then large-sample, multi-community studies are an important frontier for future research. With respect to speech, such studies are uniquely positioned to shed light on how bilingual characteristics like dominance and bilingual behaviors such as codeswitching and

language mixing create new phonetic and phonological targets in each language, catalyzing sound changes that over time form the basis of new accents and dialect regions.

Given the vast potential of open, multisite designs for bilingualism research, it is somewhat surprising that, to our knowledge, no such studies have been undertaken. A preliminary search of *Bilingualism: Language and Cognition* and *Journal of Phonetics*, prominent journals in bilingualism and speech research, revealed no studies that self-identified as "multisite." Clear examples of what such research could look like abound in psychology (Open Science Collaboration, 2015), and important, albeit smaller-scale, examples are also evident in L2 research (Moranski & Zalbidea, 2022; Moranski & Ziegler, 2021; Morgan-Short et al., 2018). Certainly an absence of bilingualism studies self-identifying as multisite does not mean that such studies do not exist. Instead, it could be that researchers have chosen to emphasize other aspects of the research concept and design. Nevertheless, the fact that we could find no evidence of multisite tagging in our search is telling, insofar as it suggests that this methodological feature, which is increasingly common in other research areas, has yet to gain equal prominence in the domain of bilingualism.

Recommendations

As is clear, we recommend researchers conduct multisite research using best practices from open science. However, aside from these broad recommendations, we also offer a series of recommendations that bilingualism researchers could consider as they plan, conduct, and report results from studies.

Collect and report extensive language background information

It is crucially important for us to understand who the participants in a given study are and what bilingual populations they best represent. Given this, it is likely that more extensive

language background information should be collected and reported. In addition to questions around dominance and proficiency, it would be very helpful for the research community to collect information about social networks and their language use, geographic factors including the participant's neighborhood, as well as language attitude information. One could imagine that a participant who speaks Spanish approximately 50% of the time but only to members of their immediate family may differ from an individual with the same overall amount of use but to a wider variety of people in their community. Collecting and reporting this information is crucial to better understand the landscape of bilingualism speech research. Basic sociodemographic data would also be useful because as Sugden and Moulson (2015) observed, "even if not interested in the impact of race, culture, and ethnicity on the phenomenon under study, including these variables can add depth to researchers' understanding of a phenomenon" (p. 2). We believe this to be true of bilingualism because it likely interacts with several other social constructs and categories, which may covertly shape how bilingualism manifests in a particular community or individual.

Consider implications given samples

Pressures of academic publishing often encourage researchers to make broad generalizations about their work in order to demonstrate "impact." However, as we have discussed above, this can result in sweeping generalizations that do not actually represent the bilingual community that is being studied. We recommend that researchers closely consider how their results may be specific to the population they are examining. Often, in our own work, this consideration takes the form of a sentence or two stating that "different populations may demonstrate different results." We challenge researchers (including ourselves!) to consider engaging with these statements in more detail. Could we make predictions for how these results

might change as a function of proficiency, dominance, geography or myriad other factors? We are not suggesting that researchers must engage with all possible options, but rather that we begin to more deeply examine how our samples impact our results. Of course, recruiting larger and more representative samples (and characterizing those samples precisely, as mentioned in our first point) would be advantageous, but we recognize that there are practical barriers that may make doing so impossible. To that end, we believe that multi-site approaches, which necessarily collect data from multiple samples, will be especially valuable.

Interpretation of conflicting results

One major challenge for a researcher is when their results conflict with previously published results. Often, our instinct is to search for which set of results is correct. This may mean developing arguments that support our own results compared to previous work or exploring a large number of competing perspectives, all with the goal of justifying one set of results compared to others. We recommend that rather than adopting such an approach, we (as a field) acknowledge that in most cases all results reflect some aspect of the true phenomenon or effect we wish to quantify. This perspective implies a more meta-analytic mindset, where researchers attempt to understand how differences in sampling, methodology, and analytical choices might drive differences in outcomes and conclusions.

Forge multi-site collaborations

Our core recommendation, to forge multi-site collaborations, may initially seem daunting, especially to junior scientists or anyone accustomed to a more independent style of research. How to get started in this area may differ depending on a researcher's preference and personality. It is possible to contact other researchers in your area at conferences, via email, or through personal connections; however, we encourage all researchers to consider expanding their

own willingness to collaborate on questions that could benefit from such multisite designs. One way to get started in multi-site research is to sign up and participate in a large-scale collaboration spearheaded by ManyLanguages (https://many-languages.com/), whose purpose is to "[…] facilitate the connection between language science researchers to diversify the languages, participants, researchers, and projects represented in the language sciences." As a field, we will also need to envision and invent new ways to promote and facilitate participation in large-scale multi-site collaborations. Thus, we feel that individual researchers should have the goal of participating in these initiatives, but the field should also invest in making participation feasible and worthwhile.

Conclusion

In order to understand bilingualism, we must understand that a single site and a single sample are unlikely to be representative of the bilingual experience broadly speaking, just as a single individual is not necessarily representative of an entire community of language users. At first glance, this may seem a truism, but the reality is that we may tacitly adopt this perspective when designing, implementing, and interpreting the results of our studies. To put a point on it, one site is incapable of capturing all of the complexity associated with bilingual language use, and each study is likely to inadvertently capture a heterogeneous sample. As siloed researchers, we are likely to continue to discover findings that are, on their face, divergent, when divergence was in fact an inevitable outcome due to non-trivial differences in sampling and study design.

In addition to the methodological concerns raised above it is important to recognize the opportunities that multisite research would create in terms of theoretical advancement in our field. For example, examining what is ostensibly the same population in multiple locations would allow us to investigate synchronic variation that may be driven by un(der)investigated

factors that fly under the radar in many linguistic studies (e.g., geography, cf: work in sociolinguistics, Reed, 2020). This work will also pave the way for relatively novel areas of exploration. Language contact via bilingualism is creating new language varieties, with new lexical items, phonetic and phonological targets, and even syntactic properties. The ability to investigate language change in progress and to ask complex questions about diachronic variation is a strength of bilingualism research as a whole, a strength that in our view can best be leveraged through larger-scale, collaborative, open research projects.

It is important to have sound methodology and work that advances theory in our field, but it is also important to remember that bilingualism research can have a real impact on language and education policy and therefore a real impact on people. Findings that are based on narrowly constructed samples, if taken to be representative of all bilinguals, could have negative consequences on the lives of many individuals, including in the availability of translators in critical settings such as healthcare and legal environments or in services for and assessment of bilingual children. Thus, the challenge facing the bilingual research community is bigger than a single phonetic variable or syntactic structure. Instead, it is a challenge that strikes at the very heart of science and research ethics itself. As a community, if we fail to ensure that our methods are sound and that our findings are widely and appropriately applicable to the communities we research, we are not appropriately serving those communities. Rethinking our methodology means reappraising and reinvesting in the ethical side of our work.

References

Achugar, M., & Pessoa, S. (2009). Power and place: Language attitudes towards Spanish in a bilingual academic community in Southwest Texas. *Spanish in Context*, *6*(2), 199–223. https://doi.org/10.1075/sic.6.2.03ach

Amengual, M., & Chamorro, P. (2016). The effects of language dominance in the perception and production of the Galician mid vowel contrasts. *Phonetica*, *72*(4), 207–236. https://doi.org/10.1159/000439406

Andringa, S. (2014). The use of native speaker norms in critical period hypothesis research. *Studies in Second Language Acquisition*, *36*(3), 565–596. https://doi.org/10.1017/S0272263113000600

Andringa, S., & Godfroid, A. (2020). Sampling bias and the problem of generalizability in applied linguistics. *Annual Review of Applied Linguistics*, *40*, 134–142. https://doi.org/10.1017/S0267190520000033

Antoniou, M., Tyler, M. D., & Best, C. T. (2012). Two ways to listen: Do L2-dominant bilinguals perceive stop voicing according to language mode? *Journal of Phonetics*, *40*(4), 582–594. https://doi.org/10.1016/j.wocn.2012.05.005

Bahrick, H. P., Hall, L. K., Goggin, J. P., Bahrick, L. E., & Berger, S. A. (1994). Fifty years of language maintenance and language dominance in bilingual hispanic immigrants. *Journal of Experimental Psychology: General*, *123*(3), 264–283. https://doi.org/10.1037/0096-3445.123.3.264

Bayley, R., Cárdenas, N. L., Schouten, B. T., & Salas, C. M. V. (2012). Spanish dialect contact in San Antonio, Texas: An exploratory study. *Selected Proceedings of the 14th Hispanic Linguistics Symposium*, 48–60.

Bialystok, E. (2021). Bilingualism as a slice of Swiss cheese. *Frontiers in Psychology*, *12*, 1–6. https://doi.org/10.3389/fpsyg.2021.769323

Birdsong, D. (2014). Dominance and age in bilingualism. *Applied Linguistics*, *35*(4), 374–392. https://doi.org/10.1093/applin/amu031

Birdsong, D., Gertken, L. M., & Amengual, M. (2012). *Bilingual Language Profile: An easy-to-use instrument to assess bilingualism*. COERLL, University of Texas at Austin. https://sites.la.utexas.edu/bilingual/

Bochynska, A., Keeble, L., Halfacre, C., Casillas, J. V., Champagne, I.-A., Chen, K., Röthlisberger, M., Buchanan, E. M., & Roettger, T. (2023). Reproducible research practices and transparency across linguistics. *Glossa Psycholinguistics*, *2*(1), 1–36. https://doi.org/10.5070/G6011239

Bolibaugh, C., Vanek, N., & Marsden, E. J. (2021). Towards a credibility revolution in bilingualism research: Open data and materials as stepping stones to more reproducible and replicable research. *Bilingualism: Language and Cognition*, *24*(5), 801–806. https://doi.org/10.1017/S1366728921000535

Bureau, U. S. C. (2022). *The american community survey: 50 years of personal stories*. U.S. Department of Commerce. https://www.census.gov/content/dam/Census/library/publications/2022/acs/acs-50.pdf

Byrnes, H. (2016). Notes from the editor. *The Modern Language Journal*, 3–18. https://doi.org/10.1111/j.1540-4781.2013.12051.x

Carmichael, K. (2016). Place-linked expectations and listener awareness of regional accents. In A. M. Babel (Ed.), *Awareness and control in sociolinguistic research* (pp. 123–151). Cambridge University Press. https://doi.org/10.1017/CBO9781139680448.009

Casillas, J. V. (2021). Interlingual interactions elicit performance mismatches not “compromise” categories in early bilinguals: Evidence from meta-analysis and coronal stops. *Languages*, *6*(9), 1–20. https://doi.org/10.3390/languages6010009

Casillas, J. V., & Simonet, M. (2018). Perceptual categorization and bilingual language modes: Assessing the double phonemic boundary in early and late bilinguals. *Journal of Phonetics*, *71*, 51–64. https://doi.org/10.1016/j.wocn.2018.07.002

Coretta, S., Casillas, J. V., Roessig, S., Franke, M., Ahn, B., Al-Hoorie, A. H., Al-Tamimi, J., Alotaibi, N. E., AlShakhori, M. K., Altmiller, R. M., Arantes, P., Athanasopoulou, A., Baese-Berk, M. M., Bailey, G., Sangma, C. B. A., Beier, E. J., Benavides, G. M., Benker, N., BensonMeyer, E. P., … Roettger, T. B. (2023). Multidimensional signals and analytic flexibility: Estimating degrees of freedom in human-speech analyses. *Advances in Methods and Practices in Psychological Science*, *6*(3), 1–29. https://doi.org/10.1177/25152459231162567

Dunn, A. L., & Fox Tree, J. E. (2009). A quick, gradient bilingual dominance scale. *Bilingualism: Language and Cognition*, *12*(3), 273–289. https://doi.org/10.1017/S1366728909990113

Filipović, L., & Hawkins, J. A. (2019). The Complex Adaptive System Principles model for bilingualism: Language interactions within and across bilingual minds. *International Journal of Bilingualism*, *23*(6), 1223–1248. https://doi.org/10.1177/1367006918781076

Flege, J. E., MacKay, I. R. A., & Piske, T. (2002). Assessing bilingual dominance. *Applied Psycholinguistics*, *23*(4), 567–598. https://doi.org/10.1017/S0142716402004046

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, *36*(1), 3–15. https://doi.org/10.1016/0093-934X(89)90048-5

Grosjean, F. (1997). The bilingual individual. *Interpreting*, *2*(1-2), 163–187. https://doi.org/10.1075/intp.2.1-2.07gro

Grosjean, F. (2010). *Bilingual: Life and reality*. Harvard university press. https://doi.org/10.4159/9780674056459

Hamann, C., Rinke, E., & Genevska-Hanke, D. (2019). Editorial: Bilingual Language Development: The role of dominance. *Frontiers in Psychology*, *10*, 1064. https://doi.org/10.3389/fpsyg.2019.01064

Hawkins, J., & Filipović, L. (2024). Bilingualism-induced language change: What can change, when, and why? *Linguistics Vanguard*, *10*(s2), 115–124. https://doi.org/10.1515/lingvan-2023-0089

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Beyond WEIRD: Towards a broad-based behavioral science. *Behavioral and Brain Sciences*, *33*(2-3), 111. https://doi.org/10.1017/S0140525X10000725

Jackson, S., & Denis, D. (2024). What I say, or how I say it? Ethnic accents and hiring evaluations in the Greater Toronto Area. *Language*, *100*(2), e27–e62. https://doi.org/10.1353/lan.2024.a929753

Jiang, X., Gossack-Keenan, K., & Pell, M. D. (2020). To believe or not to believe? How voice and accent information in speech alter listener impressions of trust. *Quarterly Journal of Experimental Psychology*, *73*(1), 55–79. https://doi.org/10.1177/1747021819865833

Kartushina, N., & Martin, C. D. (2019). Third-language learning affects bilinguals' production in both their native languages: A longitudinal study of dynamic changes in L1, L2 and L3 vowel production. *Journal of Phonetics*, *77*, 1–21. https://doi.org/10.1016/j.wocn.2019.100920

Kelly, N. (2022). Change across time in L2 intonation vs. Segments: A longitudinal study of the English of Ole Gunnar Solskjaer. *Languages*, *7*(3), 1–17. https://doi.org/10.3390/languages7030210

Kornder, L., & Mennen, I. (2021). Longitudinal developments in bilingual second language acquisition and first language attrition of speech: The case of Arnold Schwarzenegger. *Languages*, *6*(2), 1–25. https://doi.org/10.3390/languages6020061

Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology*, *25*(5), 497–514. https://doi.org/10.1080/20445911.2013.799170

Kutlu, E., & Hayes-Harb, R. (2023). Towards a just and equitable applied psycholinguistics. *Applied Psycholinguistics*, *44*(3), 293–300. https://doi.org/10.1017/S0142716423000280

Lambert, W. E. (1955). Measurement of the linguistic dominance of bilinguals. *The Journal of Abnormal and Social Psychology*, *50*(2), 197–200. https://doi.org/10.1037/h0042120

Leeuw, E. de. (2019). Native speech plasticity in the German-English late bilingual Stefanie Graf: A longitudinal study over four decades. *Journal of Phonetics*, *73*, 24–39. https://doi.org/10.1016/j.wocn.2018.12.002

Lev-Ari, S., & Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology*, *46*(6), 1093–1096. https://doi.org/10.1016/j.jesp.2010.05.025

Li, P., Sepanski, S., & Zhao, X. (2006). Language history questionnaire: A web-based interface for bilingual research. *Behavior Research Methods*, *38*(2), 202–210. https://doi.org/10.3758/BF03192770

Lim, V. P. C., Liow, S. J. R., Lincoln, M., Chan, Y. H., & Onslow, M. (2008). Determining language dominance in English-Mandarin bilinguals: Development of a self-report classification tool for clinical use. *Applied Psycholinguistics*, *29*(3), 389–412. https://doi.org/10.1017/S0142716408080181

Lindemann, S., & Subtirelu, N. C. (2013). Reliably biased: The role of listener expectation in the perception of second language speech. *Language Learning*, *63*(3), 567–594. https://doi.org/10.1111/lang.12014

Lozano-Argüelles, C., Arroyo, L. F., Rodríguez, N., López, E. M. D., Pozú, J. J. G., Markovits, J., Varela, J. P., Rocafiguera, N. de, & Casillas, J. V. (2021). Conceptually cued perceptual categorization in adult L2 learners. *Studies in Second Language Acquisition*, *43*(1), 204–219. https://doi.org/10.1017/S0272263120000273

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology*, *25*(5), 605–621. https://doi.org/10.1080/20445911.2013.795574

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940–967. https://doi.org/10.1044/1092-4388(2007/067)

Marsden, E., & Morgan-Short, K. (2023). (Why) are open research practices the future for the study of language learning? *Language Learning*, *73*, 344–387. https://doi.org/10.1111/lang.12568

Marsden, E., Morgan-Short, K., Trofimovich, P., & Ellis, N. C. (2018). Introducing registered reports at language learning: Promoting transparency, replication, and a synthetic ethic in the language sciences. *Language Learning*, *68*(2), 309–320. https://doi.org/10.1111/lang.12284

McManus, K. (2024). Replication studies in second language acquisition research: Definitions, issues, resources, and future directions: Introduction to the special issue. *Studies in Second Language Acquisition*, 1–21. https://doi.org/10.1017/S0272263124000652

Mejías, H. A., Anderson-Mejías, P. L., & Carlson, R. (2003). Attitude update: Spanish on the South Texas border. *Hispania*, 138–150. https://doi.org/10.2307/20062822

Melnik-Leroy, G. A., Turnbull, R., & Peperkamp, S. (2022). On the relationship between perception and production of L2 sounds: Evidence from anglophones' processing of the French /u/–/y/ contrast. *Second Language Research*, *38*(3), 581–605. https://doi.org/10.1177/026765832098806

Mendoza-Denton, N. (2014). *Homegirls: Language and cultural practice among latina youth gangs*. John Wiley & Sons. https://doi.org/10.1002/9780470693728

Moranski, K., & Zalbidea, J. (2022). Context and generalizability in multisite L2 classroom research: The impact of deductive versus guided inductive instruction. *Language Learning*, *72*(S1), 41–82. https://doi.org/10.1111/lang.12487

Moranski, K., & Ziegler, N. (2021). A case for multisite second language acquisition research: Challenges, risks, and rewards. *Language Learning*, *71*(1), 204–242. https://doi.org/10.1111/lang.12434

Morgan-Short, K., Marsden, E., Heil, J., Issa Ii, B. I., Leow, R. P., Mikhaylova, A., Mikołajczak, S., Moreno, N., Slabakova, R., & Szudarski, P. (2018). Multisite replication in second language acquisition research: Attention to form during listening and reading comprehension. *Language Learning*, *68*(2), 392–437. https://doi.org/10.1111/lang.12292

Nielsen, M., Haun, D., Kärtner, J., & Legare, C. H. (2017). The persistent sampling bias in developmental psychology: A call to action. *Journal of Experimental Child Psychology*, *162*, 31–38. https://doi.org/10.1016/j.jecp.2017.04.017

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716. https://doi.org/10.1126/science.aac4716

Peña, E. D., Gillam, R. B., & Bedore, L. M. (2014). Dynamic assessment of narrative ability in English accurately identifies language impairment in English language learners. *Journal of Speech, Language, and Hearing Research*, *57*(6), 2208–2220. https://doi.org/10.1044/2014_JSLHR-L-13-0151

Plonsky, L. (2013). Study quality in SLA: An assessment of designs, analyses, and reporting practices in quantitative L2 research. *Studies in Second Language Acquisition*, *35*(4), 655–687. https://doi.org/10.1017/S0272263113000399

Potowski, K. (2012). Identity and heritage learners: Moving beyond essentializations. In S. M. Beaudrie & M. Fairclough (Eds.), *Spanish as a heritage language in the United States: The state of the field* (pp. 179–202). Georgetown University Press.

Potowski, K., & Torres, L. (2023). *Spanish in Chicago*. Oxford University Press.

Reed, P. E. (2020). Place and language: Links between speech, region, and connection to place. *Wiley Interdisciplinary Reviews: Cognitive Science*, *11*(3), 1–11. https://doi.org/10.1002/wcs.1524

Roettger, T. B. (2019). Researcher degrees of freedom in phonetic research. *Laboratory Phonology*, *10*(1), 1–27. https://doi.org/10.5334/labphon.147

Roettger, T. B., Winter, B., & Baayen, H. (2019). Emergent data analysis in phonetic sciences: Towards pluralism and reproducibility. *Journal of Phonetics*, *73*, 1–7. https://doi.org/10.1016/j.wocn.2018.12.001

Sankoff, G. (2002). Linguistic outcomes of language contact. In J. K. Chambers, P. Trudgill, & N. Schilling-Estes (Eds.), *The handbook of language variation and change* (Vol. 1, pp. 638–668). Blackwell. https://doi.org/10.1002/9781118335598

Silva-Corvalán, C., & Treffers-Daller, J. (2016). *Language dominance in bilinguals: Issues of measurement and operationalization*. Cambridge University Press. https://doi.org/10.1017/CBO9781107375345

Sugden, N. A., & Moulson, M. C. (2015). Recruitment strategies should not be randomly selected: Empirically improving recruitment success and diversity in developmental psychology research. *Frontiers in Psychology*, *6*, 523. https://doi.org/10.3389/fpsyg.2015.00523

Tomé Lourido, G., & Evans, B. (2019). The effects of language dominance switch in bilinguals: Galician new speakers' speech production and perception. *Bilingualism: Language and Cognition*, *22*(3), 637–654. https://doi.org/10.1017/S1366728918000603

Velázquez, I. (2013). Individual discourse, language ideology and Spanish transmission in El Paso, Texas. *Critical Discourse Studies*, *10*(3), 245–262. https://doi.org/10.1080/17405904.2013.789975

Weinreich, U. (1968). *Languages in contact: Findings and problems*. Mouton de Gruyter. https://doi.org/https://doi.org/10.1515/9783110802177

Wolfram, W., Carter, P., & Moriello, B. (2004). Emerging Hispanic English: New dialect formation in the American South. *Journal of Sociolinguistics*, *8*(3), 339–358. https://doi.org/10.1111/j.1467-9841.2004.00264.x

Yao, Y., & Chang, C. B. (2016). On the cognitive basis of contact-induced sound change: Vowel merger reversal in Shanghainese. *Language*, *92*(2), 433–467. https://doi.org/10.1353/lan.2016.0031