

ORIGINAL ARTICLE

Examining linguistic and experimenter biases through “non-native” versus “native” speech

Rachel Elizabeth Weissler , Shiloh Drake , Ksenia Kampf, Carissa Diantoro ,
Kurtis Foster, Audrey Kirkpatrick, Isabel Preligera, Orion Wesson,
Anna Wood and Melissa M. Baese-Berk* 

Department of Linguistics, University of Oregon, Eugene, OR, USA

*Corresponding author. Email: mbaesebe@uoregon.edu

(Received 1 June 2022; revised 21 January 2023; accepted 2 February 2023; first published online 13 March 2023)

Abstract

There is a consensus in psycholinguistic research that listening to unfamiliar speech constitutes a challenging listening situation. In this commentary, we explore the problems with the construct of *non-native* and ask whether using this construct in research is useful, specifically to shift the communicative burden from the language learner to the perceiver, who often occupies a position of power. We examine what factors affect perception of non-native talkers. We frame this question by addressing the observation that not all “difficult” listening conditions provide equal challenges. Given this, we ask how cognitive and social factors impact perception of unfamiliar accents and ask what our psycholinguistic measurements are capturing. We close by making recommendations for future work. We propose that the issue is less with the terminology of *native* versus *non-native*, but rather how our unexamined biases affect the methodological assumptions that we make. We propose that we can use the existing dichotomy to create research programs that focus on teaching perceivers to better understand talkers more generally. Finally, we call on perceivers and researchers alike to question the idea of speech being “native,” “non-native,” “unfamiliar,” and “accented” to better align with reality as opposed to our inherently biased views.

Keywords: speech perception; speech production

The way researchers frame *non-native* speech in experimental research is crucial to understanding the structure, methods, and findings of studies. Often, the framing introduces bias and affects how research is being done. Framing often has its roots in belief systems that are held by researchers from what they have read in previous research, pop science beliefs, or socialization. A common belief historically held by researchers and laypeople alike is that *non-native* speech is harder to understand than *native* speech. This is supported by studies that have shown that listeners exert more effort when they listen to speech with unfamiliar accents (Adank et al., 2009;

© The Author(s), 2023. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Porretta & Tucker, 2019; Van Engen & Peelle, 2014). However, there are multiple factors that affect speech perception, including talker or accent familiarity. In this commentary, we discuss the concept of *native* versus *non-native* speech and argue that, as researchers, we are better served by conceptualizing it as a continuum of familiarity, and adjusting our experiments and protocols to therefore be just and equitable, rather than situating research under this umbrella within a dichotomy that more often than not ends up privileging certain speakers over others.

In the context of language perception and production, the term “(non-)native speaker” is used often. However, the term “non-native speaker” encompasses a wide range of individuals who do not necessarily form a homogenous group (Cheng et al., 2021). This term is used in different ways across different studies, and this makes comparison across individual studies difficult because the terms may be referring to quite disparate populations. For the purposes of this paper, we are using the terms “native” and “non-native” as they have been used in the papers we discuss; we are not putting any value judgments on the terms, thus not treating the terms as analytical objects. We find that regardless of the terminology we use, we do not mitigate the bigger issues of interest—those being the way people who speak minoritized languages and dialects are treated and written about in the field. Instead, we propose that the issues with terminology are symptoms of a broader issue. We suggest that our experimental design decisions can be one of the loci to effect change within the field.

While some recent work (e.g., Cheng et al., 2021) has suggested that terms like “native” and “non-native” are problematic and may introduce bias into research, we instead consider the imbalance regarding the listener versus the speaker in speech perception studies. While it is undeniable that everyone is inherently biased in their own beliefs (e.g., Kutlu, 2020; Kutlu et al., 2022; McGowan, 2015), equitable inquiry is paramount in science. This includes considering both grammatical and communicative competencies, and social factors being taken into account during processing. The framing of research questions regarding *native* and *non-native* talkers importantly reflects methodological assumptions that in turn affect the kinds of experimental tasks we choose. Often, researchers frame their questions focusing on a binary distinction between *native* versus *non-native*. They treat native speakers as a group of speakers with particular qualities, and any speaker who does not meet those criteria as a non-native speaker. This implies fluency as something one has or lacks, for example, rather than a gradient scale of linguistic competence. Although it is natural for humans to categorically discriminate, the way we have been unidirectionally framing our research questions concerning *native* versus *non-native* limits our understanding of language users in these populations. We must consider the purposes of different tasks and how they serve greater understandings of all language users.

In this paper, we first lay out both the problems that the term “non-native” entails and the potential usefulness of using this term. We then focus on factors that affect the perception of non-native talkers and how listening conditions and social expectations play a significant role in accent perception. This is followed by a breakdown of the kinds of measurements being used to examine perception of non-native talkers and whether these psycholinguistic measurements capture what researchers intended. Finally, we outline recommendations for future work that moves toward more equitable psycholinguistics research in this area.

Positionality statement

The authors of this paper are members of the Speech Perception and Production Laboratory at the University of Oregon, including faculty, postdoctoral fellows, graduate students, and undergraduate students. The authors come from diverse linguistic, socioeconomic, and racial backgrounds and have had diverse experiences, which impact their perspectives, biases, and research interests. We refer the reader to this webpage <https://www.speechperceptionproductionlab.com/positionality-statments>, where each of our positionality statements is available.

Problems with “non-native”

“Non-native,” a term used frequently in linguistics and psychological research, more generally refers to something that is not endemic to a region or place. This is ironic within a linguistic context as most of the dominant languages in the world have their dominant status due to dissemination and colonization of places and peoples (Cooper, 1982). The term has been used in linguistics for decades; Leonard Bloomfield used it to describe the first language that someone acquires (1933). The term remained largely uninterrogated and broadly accepted as a useful delineator between speakers. The definition resulted in the dichotomized nature linguists use to describe the differences between variable ways of speaking (Chomsky, 1965), and while this view evolved to include not only grammatical competence but also communicative competence (Chomsky, 1980), the prevailing notion became that a native speaker was the only reliable source for linguistic data. The notion of “non-nativeness” is exacerbated by nomenclature, but, importantly, the issues of discrimination and mistreatment of “non-native” speakers persist no matter what label is used, particularly within the frame of experimental design.

Experimental research in psycholinguistics is a lucrative area for the language learning industry, which has become a space that further perpetuates the native/non-native binary. The funding that goes into these sorts of programs could be guiding the kinds of questions researchers are asked to investigate (Kilman *et al.*, 2014; Tamminen *et al.*, 2015). The mass interest in language learning covertly reinforces the need to teach speakers rather than gaining a better understanding of how listeners perceive speech (Flores & Rosa, 2015; Ramjattan, 2019). Thus, the communicative burden is typically placed on those considered non-native, insinuating the need to be very proficient in the language of wider communication to be successful. Baese-Berk *et al.*'s (2020) review on perception of non-native speech discusses the many factors that impact its perception: The relationship one has between their native language and people who speak it, their relation with people who speak a non-native variety, as well as perception of non-native speech drives the kinds of research questions people ask.

To reframe non-native speech as *unfamiliar speech* would benefit the biggest goal of communication: to be mutually understood in terms of intelligibility, no matter what one's L1 is. However, this reframing also has more specific benefits. It could help researchers clarify precisely what questions they are addressing and incorporate these questions in a broader literature. It is clear from a wide body of psycholinguistic literature that listeners process familiar talkers (*i.e.*, those talkers they have heard before either through life experience or through experimental exposure)

differently from unfamiliar talkers (i.e., talkers they have not heard previously; e.g., Levi et al., 2011; Nygaard et al., 1994; Nygaard & Pisoni, 1998), and this benefit has been shown to occur for both talkers with accents that the listener has prior exposure to (e.g., Nygaard & Pisoni, 1998) and accents that a listener has little to no experience with (e.g., Levi et al., 2019). Given this understanding that unfamiliar speech of various types is challenging for listeners, we are able to acknowledge communicative difficulties that affect *all* people, regardless of language background. We propose harnessing the power of the dichotomies that currently exist (i.e., *native vs non-native*) to shift how we as experimentalists are framing *non-native* speakers.

Framing *non-native* speech as *unfamiliar speech* shifts the communicative burden from the speaker to the listener, who often occupies a position of power. This framing, however, maintains a binary, rather than recognizing that there is likely a continuum of familiarity. For example, one may be much more familiar with the voices and speech of their close family and friends than with acquaintances; however, this same individual is likely much more familiar with the voices of these acquaintances than with a speaker from a language background they have not previously encountered at all. Therefore, we would expect *familiarity* to be a gradient construct, likely co-varying with experience, that may have gradient results on psycholinguistic work. Further, this familiarity is likely to vary as a function of context. For example, an individual living in the American Southwest may not speak Spanish themselves but is more likely to have encountered Spanish, varieties of English influenced by Spanish, or terms that have come out of Spanish/English contact than varieties of English influenced by, say, Khmer¹. There is a growing body of psycholinguistic work that considers factors that impact listeners' perception of and adaptation to unfamiliar speech; this work does shift the communication burden in a practical way (e.g., Kutlu et al., 2021).

Another key issue is *what* precisely is unfamiliar to a listener in these scenarios. As an anonymous reviewer noted, a listener is likely to be unfamiliar with a wide array of cues ranging from phonetic to lexical to cultural. Which of these cues are most relevant for a listener is likely to depend on the precise research question being investigated. As we think about research design going forward, definitions of familiarity should be explicitly stated according to the questions (and answers) each study is hoping to address. It is also important to consider the power of social information when constituting unfamiliarity, as social information can dictate to a listener that they are hearing something in the speech signal that is not acoustically there, from something as simple as a fabrication regarding where the speaker is from or a picture of a face (Niedzielski, 1999; Rubin, 1992).

Using *familiar* and *unfamiliar* as framing devices provides an opportunity to consider perception of *non-native* speech not as a specific problem to be solved that is distinct from other issues in speech perception but rather as a specific instance of a broader problem which is addressed throughout related literature—how a listeners' familiarity with a particular voice, accent, or dialect impacts their perception. Instead of focusing exclusively on properties of the talker that make speech perception challenging, we can instead point to the fact that listeners face challenges across a range of listening conditions, and this particular condition may just be one instance of this challenge. This framing shifts the communicative burden from speaker to listener, benefitting all speakers. Still, reframing does not solve the basic

problem we see in this work. The issue at hand is not fundamentally about terminology but more centrally about the biases brought to bear in the research setting, and how we can use the existing dichotomies to our advantage in psycholinguistic research specifically. We know that the discerning nature of humans can lead to discrimination and bias (Weissler, 2022), and therefore, there needs to be dedicated consideration by the psycholinguistics community, and by people more broadly, about the powerful influence of ideology in every facet of our lives, including in how we design our scientific experiments.

Factors affecting the perception of non-native talkers

We see in many studies that listening to speech with unfamiliar accents requires more effort than speech with familiar accents, even if the speakers are fully intelligible (e.g., McLaughlin & Van Engen, 2020). One of the claims we see in the literature addressing perception of native and non-native talkers is that speech from non-native talkers is inherently more difficult to perceive (Bent & Frush Holt, 2013; Leikin *et al.*, 2009; Munro & Derwing, 1995). However, knowing what makes speech more difficult to perceive is complex because not all difficult listening conditions are created equal (see Mattys *et al.*, 2012 for a review). For example, speech presented in noise is processed differently from accented speech (Adank *et al.*, 2015), and even different types of noise (such as white noise vs. cocktail party babble) seem to affect speech recognition in different ways, at different times in processing (Mattys *et al.*, 2012; McLaughlin *et al.*, 2018). Furthermore, other types of unfamiliar speech, like regional accents and dialects, or speech from people with dysarthria and other speech disorders are also difficult to perceive but in ways that are distinct from non-native speech (Bent *et al.*, 2016). Researchers have also pointed out the wide range of variation in individuals' ability to understand speech in noise, accented speech, or disordered speech (Bent *et al.*, 2016).

In addition to how speech in general is perceived under different conditions, there are both cognitive and social factors that affect how we perceive unfamiliar accents. Cognitive factors include the size of our receptive vocabulary (i.e., the number of words we understand but not necessarily the number of words that we use in our own speech) and our working memory capacity (e.g., Adank *et al.*, 2009; Banks *et al.*, 2015, 2016; McLaughlin *et al.*, 2018). Social factors, on the other hand, include how we associate accents with faces, voices, and identity markers, often races: when our assumptions about what type of person has what kind of accent are violated, we find the speech more difficult to perceive (Hanulíková, 2021; Hanulíková *et al.*, 2012; Kutlu, 2020; Kutlu *et al.*, 2022; McGowan, 2015; Rubin, 1992; Vaughn, 2019). We discuss these factors in more detail below.

In speech perception research, we use the term “intelligibility” to refer to the extent to which listeners can correctly recognize the speech stimuli they hear, often as measured by the number of correctly transcribed words from an auditory stimulus. This measure represents how accurately a speech signal is perceived, which is a beneficial framework to use when considering unfamiliar accent perception. A number of studies have relied on intelligibility measures to discover the effects of adverse conditions on speech processing due to its ease of interpretation and

possibility of cross-study comparison. Multiple studies have found that the size of a participant's receptive vocabulary is a good predictor of their intelligibility score. Listeners with larger receptive vocabularies tend to provide higher intelligibility scores, while the opposite is true for participants with smaller receptive vocabularies.

In experiments where information about where a speaker is from is manipulated, listeners familiar with the region and accent perceive the same recorded vowel productions differently depending on the social or dialect group they believe the speaker to be a part of (Hay et al., 2006; Hay & Drager, 2010; Niedzielski, 1999). Perceived social group identity as indexed by the physical appearance of a speaker has also been shown to affect listener perception through the use of matched-guise tasks or pairing images of speakers' faces with audio. Rubin's (1992) classic finding that the same recording would be rated as more accented when paired with images of Chinese faces than when paired with images of white faces has been investigated further in recent studies. For example, speech paired with South Asian faces in a matched-guise task was rated as more accented than speech paired with white faces in a mostly monolingual setting (Kutlu, 2020) but not when participants are located in an area with more exposure to multilingual speakers (i.e., Montréal; Kutlu et al., 2022). Babel and Russell (2015) similarly found that speech paired with photos of Chinese faces led to lower intelligibility scores and higher accentedness ratings than speech paired with white faces. They found these results despite another group of listeners being unable to reliably distinguish between the Chinese and white talkers when the speech was presented without photos, even though listeners resided in a Canadian neighborhood with a large multilingual and multicultural population.

Similar evidence of a complex interaction between listener experience with and expectations regarding speech varieties, speakers, and contexts and the way speech is evaluated has also been found in matched-guise tasks. Speech is transcribed less accurately when an image of the purported talker did not match the listener's expectations (e.g., a white face paired with Mandarin-accented utterances; McGowan, 2015). Listeners have also been shown to transcribe speech more accurately when they are given information about the speaker's accent than when they receive no information about the accent (Vaughn, 2019).

Neuroimaging studies have also shown that listeners' processing is modulated by how the speaker is perceived. When listening to speech, electrophysiological responses to morphosyntactic errors (Hanulíková et al., 2012) and semantic anomalies (Hanulíková et al., 2012; Romero-Rivas et al., 2015) are elicited when the speech comes from a native speaker but not when it comes from a non-native speaker. Listeners also show differences in the synchronization of their neural oscillations when they listen to speech in their native language, a foreign language that they have learned, and a language completely unknown to them (Jin et al., 2014; Pérez et al., 2015). Studies have also found a neural response bias in favor of the listener's own accent, with decreased neural response to accents of social groups the listener is not a part of (Bestmeyer et al., 2015).

The perceptual effects that we see in behavioral and neuroimaging studies occur with unfamiliar accents and listeners' perception is shaped by co-occurring social information, like photos and descriptions of racioethnic traits. This points to the factors affecting perception arising from both properties of the listener and the speech signal itself. In addition, facets of listener experience that may also impact

perception of unfamiliar accent include factors as disparate as musical experience (Kraus & Chandrasekaran, 2010; Kraus *et al.*, 2014; Moreno *et al.*, 2009; Musacchia *et al.*, 2007; Qin *et al.*, 2021; Zhao & Kuhl, 2016; Zhao *et al.*, 2022) and motivation (Gardner & Lambert, 1959; Gardner *et al.*, 1997; Saito *et al.*, 2018; Tsang, 2022). The impacts of such diverse factors complicate approaches to measuring the perception of unfamiliar speech; researchers must take into account the contribution of the listener to speech perception tasks and assess whether the questions we ask and empirical methods we employ accurately address all of the factors at play in perception of different speech varieties. With this point in mind, we must ask whether the tasks we use to measure perception of non-native speech are, in fact, accurately measuring what we believe they are. We discuss this point in the following section.

What are our psycholinguistic measurements capturing?

The selection of which task to use in a study is a complex decision, as the type of task a participant is asked to do affects how they interact with both the task and the materials used in the experiment. For example, listeners demonstrate different performance if they are asked to do a simple discrimination task compared to a more complex task requiring listeners to compare stimuli across trials (e.g., Pisoni & Lazarus, 1974). The assumptions about a task also affect how results are interpreted. All of these factors together mean that we must know whether our tasks are measuring what we think they are. We are choosing to focus on transcription tasks, but similar questions can (and should) be asked of other tasks.

Transcription tasks provide an intelligibility score based on the number of words a listener correctly transcribes from each utterance in a stimuli set. However, apart from demonstrating the accuracy of response, the metric does not provide an insight into underlying causes of listeners' performance. Moreover, even when participants are able to accurately transcribe all of a talker's speech, there is evidence that they exert additional effort to understand non-native talkers than with native talkers (Brown *et al.*, 2020; McLaughlin & Van Engen, 2020). This discrepancy between the required listening effort and intelligibility suggests that there are additional factors that an intelligibility score is not capturing. "Listening effort" itself is another complex construct made up of multiple factors, which are difficult to tease apart (McGarrigle *et al.*, 2014; Van Engen & Peelle, 2014). The exact nature of the reasons behind increased effort when listening to non-native speech remains unclear. It is likely an indication of deviations between L1 and L2 language patterns that require more cognitive resources to correctly map those representations (Brown *et al.*, 2020). Regardless of the source of difficulty, it is important to highlight the need to better understand how non-native speech affects perception and the inability of the intelligibility measure in isolation to provide the full picture of all listening challenges (Baese-Berk *et al.*, 2020, *under review*).

Because intelligibility tasks are not very sensitive, alternative or additional measures should be considered as important means to capture the nuance of what exactly is "difficult" about perceiving accented speech. Both additional research methods and additional measures available to researchers using intelligibility tasks can provide insights on the complex nature of this area of perception. For example,

instead of using the percentage of correctly transcribed words alone to investigate perception of various accents, researchers can classify the types of errors that listeners made during transcription (Winn & Teece, 2021). As Winn and Teece point out, participants may have similar intelligibility scores despite committing different errors—and different types of errors that suggest different amounts of cognitive effort are scored the same way in a standard intelligibility measure. Errorful responses that are more plausible than the intended sentence may, in fact, require less cognitive effort than providing the correct response. Even when intelligibility is fully achieved, correct responses may demand various degrees of cognitive effort. Pupillometry data reflect processing of non-native speech in a more nuanced way by capturing an increase in listening effort when intelligibility is at ceiling, which would otherwise remain undetected in standard transcription tasks (McLaughlin & Van Engen, 2020).

In addition to qualitative examinations of error type, it may be useful to incorporate passive physiological measures that are associated with processing difficulty, such as pupillometry, skin conductance, and various types of neural responses. Due to the level of sensitivity to online processing they afford, these passive measures collect additional information about listener effort when processing speech. Gathering fine-grained information about listener processing is vital. Studies using these passive measures have found that variation in listener-related factors such as level of experience with the accent being perceived (Brown et al., 2020; Porretta & Tucker, 2019) and listener personality traits (Francis et al., 2021) affect processing effort during perception of non-native speech.

Passive measures have the additional advantage of demonstrating the nature of online processing, before the transcription is provided and analyzed in addition to the accuracy of the transcription or other offline measures. Utilizing multiple offline measures may be just as informative as online measures. For example, researchers also use subjective (e.g., self-report questionnaires; Koelewijn et al., 2012) and behavioral measures (e.g., response times; Munro & Derwing, 1995; Floccia et al., 2009) to provide additional insight into listening effort exerted during the perception of non-native speech. These approaches used in combination with transcription tasks can make our analysis more informative: while relying on binary responses (“correct” or “incorrect”) limits us to interpreting only the result of the perception, adding other measures helps us better understand the cognitive processes determining certain behavioral response and get a better sense of the factors that combine to form the construct of intelligibility.

By using measures that describe what drives listeners’ behavior and reflect underlying cognition during speech perception, we focus on the listener and the effort they do or do not exert. Looking at intelligibility from this perspective has the potential to bring us closer to understanding how to minimize the cognitive load of the listener and improve their perceptive abilities without placing the communicative burden entirely on the speaker and the characteristics of their speech. Furthermore, while the listening subject is often perceived as white (Flores & Rosa, 2015), we believe the issues of communicative burden also harm white speakers, particularly through the intersection of socioeconomic background and disability. We thus do not want to imply that “nativeness” is inherently white, rather that those who are considered native are typically those who are in power in a given society.

Recommendations for future work

Thus far, we have described the notion of *non-native* from various vantage points, particularly relating this term to both linguistic and experimenter biases. Here, we provide recommendations for future work surrounding how we as scientists can reframe the notion of “non-native” to help shift the communicative burden from speaker to listeners of all language backgrounds.

The issue of *non-native* within psycholinguistics has begun to be approached by finding more accurate ways to characterize the people whose language we are studying and working with (Cheng *et al.*, 2021), such as removing the dichotomy of *native* and *non-native* to more accurately represent speakers’ history, identity, and other continuous measures of speaker performance. This is a useful step forward, though as we described, the issue of terminology is reflective of a bigger issue: the focus of much work in this area focuses on making the non-native speaker “better,” in terms of their speech being more intelligible, more aligned with a prestigious variety, or easier to understand. However, it is crucial to understand that these terms were not created by minoritized speakers themselves and are not typically focused on minoritized speaker equality or liberation. In reality, most labels will inherently be exclusionary in one way or another, and even the most inclusive terminology does not solve the overarching cause of the problem—“native” listeners often have challenges in understanding unfamiliar accents that sometimes have little to do with the accents themselves and more to do with cognitive and social factors of the listener themselves. We suggest that using the terms that already exist can be helpful to frame experiments as how listeners in general interact with unfamiliar speech on a continuum, rather than as a dichotomy.

We also need to be cognizant of the globalized society we live in and the ideological backgrounds and socialization that our participants come from. This entails addressing real-world marginalization. One approach to address these issues is to investigate ways to encourage listeners to improve and challenge their listening abilities, which can be difficult when there is a seemingly baseline notion of validity of talkers (e.g., Chomsky’s (1965) definition of the ideal speaker being a native speaker of a language). Further, examining our understanding of the motivations for and results of learning additional languages is crucial. As discussed above, many extralinguistic factors affect how a listener perceives speech, including the listener’s own biases, expectations, and background. To more accurately represent these facets of the listener, it would be useful to collect and include more data in the analysis of the listener’s responses—for example, language background/experience questionnaires or attitudinal surveys in addition to accuracy or intelligibility scores, reaction times, or neuroimaging data.

Some steps in the direction toward equitable psycholinguistics research practice are already taking effect. There are examples of psycholinguistic work that embrace the dichotomy and shift the communicative burden to the listener. For example, Kutlu *et al.* (2021, 2022) further the field’s understanding of perception of non-native speech, finding that listeners with less racially diverse social networks give higher accentedness judgments to non-native speakers, further suggesting the impact of social perception (or lack of exemplars) and speech perception.

Linguistic anthropology and education are related fields that also explore these problems through the lens of the white listening subject, making a concrete link

between familiarity of languages and intercultural competence (Hannerz, 1973; Jung, 2010; Flores & Rosa, 2015). The fundamental strategy of framing linguistic knowledge as cultural competence would be beneficial for psycholinguistics to pull from when conceptualizing experiments. The framework of recognizing that all people are socialized to perceive *white as right* requires an interrogation of what biases we have when listening to speech. Gerald (2020) suggests teaching the white perceiver to raise their communicative awareness, potentially by meeting the listeners where they are. By emphasizing that the ways we perceive can be augmented and shifted, listeners can gain a clearer sense of what they are tuning into when listening to all speech. Being clear that the perceptions we have are *learned* knowledge and not *inherent* knowledge can combat the notions that conceive of “native” speech as better than non-native. In addition to using established theories from related fields to inform our research, we can revise existing theories and frameworks within linguistics to account for the social factors that affect language processing: consider Chomsky’s shift from his 1965 work to his 1980 work considering only grammatical competence to including communicative competence as well. An ideal theory would expand further to incorporate theory from the aforementioned fields as well as speech familiarity to reflect how our linguistic knowledge changes over time.

We want to acknowledge that some tools we have mentioned in this paper might not be accessible to every researcher. As such, we suggest behavioral data collection tools that are free or inexpensive to license and flexible to use with many methodologies, such as Gorilla, PCIbex, PsychoPy/Pavlovia, and others. We would also like to suggest incorporating multiple types of data that are collected, using both passive and active measures, online and offline measures, or multiple measurements of all types. Finally, we would also like to highlight the various online recruitment platforms that exist as of this writing, which can help expand the reach of our experiments to more than the usual college student participant population and ideally make our results more broadly applicable and relevant to multiple populations.

Conclusion

Based on examples in an individual’s immediate vicinity, many people think their way of speaking is *normal*, *neutral*, or *unmarked*, and that everyone else around them sounds *different*. This is understandable, as, in general, people are more inclined to be used to one’s own way of speaking and the ways of those we spend much time with than those we spend less time with. However, we also see this issue potentially arising in the way we frame our research along a dichotomy rather than along a continuum, which can feed back into how some speech is perceived by the general public as either *correct* or *incorrect*.

In this paper, we propose that the root of this issue is not with the terminology used but rather how unexamined biases affect methodological assumptions that are exemplified and perpetuated from conclusions that come from research on non-native speech. Considering the environments that linguists and laypeople come from, there are plenty of preconceived notions that there are right and wrong ways to speak. This occurs when we center some speech as *baseline* and other speech is

framed as *non-standard*. We will not be able to change that perspective in one or two experimental studies, but what we can do is harness the notion of familiarity and ask *all* listeners how they perceive speech that may be unfamiliar to them. Psycholinguists often have research goals of figuring out what results in ease of processing, but the questions asked and results concluded can often result in unidirectional outcomes that further perpetuate bias (e.g., only being concerned how L1 Spanish speakers perceive English, rather than also considering how L1 English speakers perceive Spanish).

We have mentioned some interventions at different levels of experiment building. The discrepancy between required listening effort and intelligibility suggests that researchers have work to do to more adequately capture intelligibility. Some efforts in this domain include subjective questionnaires and additional behavioral measures, focusing our effort on exploring what leads to the reduction of cognitive load being different for all kinds of speech, even if they are rated and scored as equally intelligible. This type of work in turn rules out confounds such that we can focus on how to make listeners better with unfamiliar speech perception.

We end with a call to action, specifically that researchers broaden their notions of what they want the outcomes of their research to be. We want to encourage more openness to question notions of how listeners broadly interact with unfamiliar speech. When we do ask questions of particular minoritized subgroups, we frame research questions through how listeners can adapt to the unique features of a group's speech. By reframing our research, we hope to also help people shift their internal narratives that are rife with standard language ideology. For example, a person may realize it is not really the doctor's fault if they feel like they cannot understand the doctor, or they may become curious about how to interact with someone in a language which might not be the language of wider communication in a given society. A solution here would be to look inward, reflect on bias, and consider the *facts* at play, not the *assumptions* based in bias. The burden shifts when there is room for understanding. Ultimately, the questions we ask as scientists matter, so it is critical to be thoughtful about what we are interrogating, how we ask the questions, and how we achieve these goals within the experimental frame. By more mindfully considering how we frame questions around non-native speech at every level of conceptualization in the research process, we can make our science more equitable and accessible to a broader understanding of human language. Attention toward bias in research and centering equity has the potential to make listeners in general better at perceiving unfamiliar speech.

Conflict of interest. We have no conflicts of interest to disclose.

Note

1 We thank an anonymous reviewer for this example.

References

- Adank, P., Evans, B. G., Stuart-Smith, J., & Scott, S. K. (2009). Comprehension of familiar and unfamiliar native accents under adverse listening conditions. *Journal of Experimental Psychology: Human Perception and Performance*, 35(2), 520–529. <https://doi.org/10.1037/a0013552>

- Adank, P., Nuttall, H. E., Banks, B., & Kennedy-Higgins, D. (2015). Neural bases of accented speech perception. *Frontiers in Human Neuroscience*, *9*. <https://doi.org/10.3389/fnhum.2015.00558>
- Babel, M., & Russell, J. (2015). Expectations and speech intelligibility. *The Journal of the Acoustical Society of America*, *137*(5), 2823–2833. <https://doi.org/10.1121/1.4919317>
- Baese-Berk, M. M., Levi, S. V., & Van Engen, K. J. (under review). *Intelligibility as a measure of speech perception: Current approaches, challenges, and recommendations*.
- Baese-Berk, M. M., McLaughlin, D. J., & McGowan, K. B. (2020). Perception of non-native speech. *Language and Linguistics Compass*, *14*(7), e12375. <https://doi.org/10.1111/lnc3.12375>
- Banks, B., Gowen, E., Munro, K. J., & Adank, P. (2015). Cognitive predictors of perceptual adaptation to accented speech. *The Journal of the Acoustical Society of America*, *137*(4), 2015–2024. <https://doi.org/10.1121/1.4916265>
- Bent, T., Baese-Berk, M., Borrie, S. A., & McKee, M. (2016). Individual differences in the perception of regional, nonnative, and disordered speech varieties. *The Journal of the Acoustical Society of America*, *140*(5), 3775–3786. <https://doi.org/10.1121/1.4966677>
- Bent, T., & Frush Holt, R. (2013). The influence of talker and foreign-accent variability on spoken word identification. *The Journal of the Acoustical Society of America*, *133*(3), 1677–1686. <https://doi.org/10.1121/1.4776212>
- Bestmeyer, P. E. G., Belin, P., & Ladd, D. R. (2015). A neural marker for social bias toward in-group accents. *Cerebral Cortex (New York, NY)*, *25*(10), 3953–3961. <https://doi.org/10.1093/cercor/bhu282>
- Bloomfield, L. (1933). *Language and Linguistics*. George Allen & LTD.
- Brown, V. A., McLaughlin, D. J., Strand, J. F., & Van Engen, K. J. (2020). Rapid adaptation to fully intelligible nonnative-accented speech reduces listening effort. *Quarterly Journal of Experimental Psychology*, *73*(9), 1431–1443. <https://doi.org/10.1177/1747021820916726>
- Cheng, L. S. P., Burgess, D., Vernooij, N., Solís-Barroso, C., McDermott, A., & Namboodiripad, S. (2021). The problematic concept of native speaker in psycholinguistics: replacing vague and harmful terminology with inclusive and accurate measures. *Frontiers in Psychology*, *12*. <https://doi.org/10.3389/fpsyg.2021.715843>
- Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.
- Chomsky, N. (1980). *Rules and representations*. Columbia University Press.
- Cooper, R. L. (Ed.). (1982). *Language spread: Studies in diffusion and social change*. Indiana University Press.
- Floccia, C., Butler, J., Goslin, J., & Ellis, L. (2009). Regional and foreign accent processing in English: Can listeners adapt? *Journal of Psycholinguistic Research*, *38*(4), 379–412. <https://doi.org/10.1007/s10936-008-9097-8>
- Flores, N., & Rosa, J. (2015). Undoing appropriateness: Raciolinguistic ideologies and language diversity in education. *Harvard Educational Review*, *85*(2), 149–171. <https://doi.org/10.17763/0017-8055.85.2.149>
- Francis, A. L., Bent, T., Schumaker, J., Love, J., & Silbert, N. (2021). Listener characteristics differentially affect self-reported and physiological measures of effort associated with two challenging listening conditions. *Attention, Perception, & Psychophysics*, *83*(4), 1818–1841. <https://doi.org/10.3758/s13414-020-02195-9>
- Gardner, R. C., & Lambert, W. E. (1959). Motivational variables in second-language acquisition. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *13*(4), 266. <https://doi.org/10.1037/h0083787>
- Gardner, R. C., Tremblay, P. F., & Masgoret, A.-M. (1997). Towards a full model of second language learning: An empirical investigation. *The Modern Language Journal*, *81*(3), 344–362. <https://doi.org/10.1111/j.1540-4781.1997.tb05495.x>
- Gerald, J. P. B. (2020). Worth the risk: Towards decentring whiteness in English language teaching. *BC TEAL Journal*, *5*(1), 44–54. <https://doi.org/10.14288/bctj.v5i1.345>
- Hannerz, U. (1973). The second language: An anthropological view. *TESOL Quarterly*, *7*(3), 235–248. <https://doi.org/10.2307/3585672>
- Hanulíková, A. (2021). Do faces speak volumes? Social expectations in speech comprehension and evaluation across three age groups. *PLOS ONE*, *16*(10), e0259230. <https://doi.org/10.1371/journal.pone.0259230>

- Hanulíková, A., van Alphen, P. M., van Goch, M. M., & Weber, A. (2012). When one person's mistake is another's standard usage: The effect of foreign accent on syntactic processing. *Journal of Cognitive Neuroscience*, *24*(4), 878–887. https://doi.org/10.1162/jocn_a_00103
- Hay, J., & Drager, K. (2010). Stuffed toys and speech perception. *Linguistics*, *48*(4), 865–892. <https://doi.org/10.1515/LING.2010.027>
- Hay, J., Warren, P., & Drager, K. (2006). Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics*, *34*(4), 458–484. <https://doi.org/10.1016/j.wocn.2005.10.001>
- Jin, Y., Díaz, B., Colomer, M., & Sebastián-Gallés, N. (2014). Oscillation encoding of individual differences in speech perception. *PLOS ONE*, *9*(7), e100901. <https://doi.org/10.1371/journal.pone.0100901>
- Jung, M.-Y. (2010). The intelligibility and comprehensibility of World Englishes to non-native speakers. *Journal of Pan-Pacific Association of Applied Linguistics*, *14*(2), 141–163.
- Kilman, L., Zekveld, A., Hällgren, M., & Rönnerberg, J. (2014). The influence of non-native language proficiency on speech perception performance. *Frontiers in Psychology*, *5*. <https://doi.org/10.3389/fpsyg.2014.00651>
- Koelewijn, T., Zekveld, A. A., Festen, J. M., Rönnerberg, J., & Kramer, S. E. (2012). Processing load induced by informational masking is related to linguistic abilities. *International Journal of Otolaryngology*, *2012*, e865731. <https://doi.org/10.1155/2012/865731>
- Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience*, *11*(8), 599–605. <https://doi.org/10.1038/nrn2882>
- Kraus, N., Slater, J., Thompson, E. C., Hornickel, J., Strait, D. L., Nicol, T., & White-Schwoch, T. (2014). Music enrichment programs improve the neural encoding of speech in at-risk children. *Journal of Neuroscience*, *34*(36), 11913–11918. <https://doi.org/10.1523/JNEUROSCI.1881-14.2014>
- Kutlu, E. (2020). Now you see me, now you mishear me: Raciolinguistic accounts of speech perception in different English varieties. *Journal of Multilingual and Multicultural Development*, *0*(0), 1–15. <https://doi.org/10.1080/01434632.2020.1835929>
- Kutlu, E., Tiv, M., Wulff, S., & Titone, D. (2021). The impact of race on speech perception and accentedness judgements in racially diverse and non-diverse groups. *Applied Linguistics*, amab072. <https://doi.org/10.1093/applin/amab072>
- Kutlu, E., Tiv, M., Wulff, S., & Titone, D. (2022). Does race impact speech perception? An account of accented speech in two different multilingual locales. *Cognitive Research: Principles and Implications*, *7*(1), 7. <https://doi.org/10.1017/gn94jg>
- Leikin, M., Ibrahim, R., Eviatar, Z., & Sapir, S. (2009). Listening with an accent: Speech perception in a second language by late bilinguals. *Journal of Psycholinguistic Research*, *38*(5), 447. <https://doi.org/10.1007/s10936-009-9099-1>
- Levi, S. V., Harel, D., & Schwartz, R. G. (2019). Language ability and the familiar talker advantage: Generalizing to unfamiliar talkers is what matters. *Journal of Speech, Language, and Hearing Research*, *62*(5), 1427–1436. https://doi.org/10.1044/2019_JSLHR-L-18-0160
- Levi, S. V., Winters, S. J., & Pisoni, D. B. (2011). Effects of cross-language voice training on speech perception: Whose familiar voices are more intelligible? *The Journal of the Acoustical Society of America*, *130*(6), 4053–4062. <https://doi.org/10.1121/1.3651816>
- Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*, *27*(7–8), 953–978. <https://doi.org/10.1080/01690965.2012.705006>
- McGarrigle, R., Munro, K. J., Dawes, P., Stewart, A. J., Moore, D. R., Barry, J. G., & Amitay, S. (2014). Listening effort and fatigue: What exactly are we measuring? A British society of audiology cognition in hearing special interest group 'white paper.' *International Journal of Audiology*, *53*(7), 433–445. <https://doi.org/10.3109/14992027.2014.890296>
- McGowan, K. B. (2015). Social expectation improves speech perception in noise. *Language and Speech*, *58*(4), 502–521. <https://doi.org/10.1177/0023830914565191>
- McLaughlin, D. J., Baese-Berk, M. M., Bent, T., Borrie, S. A., & Van Engen, K. J. (2018). Coping with adversity: Individual differences in the perception of noisy and accented speech. *Attention, Perception, & Psychophysics*, *80*(6), 1559–1570. <https://doi.org/10.3758/s13414-018-1537-4>

- McLaughlin, D. J., & Van Engen, K. J. (2020). Task-evoked pupil response for accurately recognized accented speech. *The Journal of the Acoustical Society of America*, *147*(2), EL151–EL156. <https://doi.org/10.1121/10.0000718>
- Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S. L., & Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cerebral Cortex*, *19*(3), 712–723. <https://doi.org/10.1093/cercor/bhn120>
- Munro, M. J., & Derwing, T. M. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech*, *38*(3), 289–306. <https://doi.org/10.1177/002383099503800305>
- Muscaccia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences*, *104*(40), 15894–15898. <https://doi.org/10.1073/pnas.0701498104>
- Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology*, *18*(1), 62–85. <https://doi.org/10.1177/0261927X99018001005>
- Nygaard, L. C., Sommers, M. S., & Pisoni, D. B. (1994). Speech perception as a talker-contingent process. *Psychological Science*, *5*(1), 42–46. <https://doi.org/10.1111/j.1467-9280.1994.tb00612.x>
- Nygaard, L. C., & Pisoni, D. B. (1998). Talker-specific learning in speech perception. *Perception & Psychophysics*, *60*(3), 355–376. <https://doi.org/10.3758/BF03206860>
- Pérez, A., Carreiras, M., Gillon Dowens, M., & Duñabeitia, J. A. (2015). Differential oscillatory encoding of foreign speech. *Brain and Language*, *147*, 51–57. <https://doi.org/10.1016/j.bandl.2015.05.008>
- Pisoni, D. B., & Lazarus, J. H. (1974). Categorical and noncategorical modes of speech perception along the voicing continuum. *The Journal of the Acoustical Society of America*, *55*(2), 328–333. <https://doi.org/10.1121/1.1914506>
- Porretta, V., & Tucker, B. V. (2019). Eyes wide open: Pupillary response to a foreign accent varying in intelligibility. *Frontiers in Communication*, *4*. <https://doi.org/10/gn94qv>
- Qin, Z., Zhang, C., & Wang, W. S. (2021). The effect of Mandarin listeners' musical and pitch aptitude on perceptual learning of Cantonese level-tones. *The Journal of the Acoustical Society of America*, *149*(1), 435–446. <https://doi.org/10.1121/10.0003330>
- Ramjattan, V. A. (2019). The white native speaker and inequality regimes in the private English language school. *Intercultural Education*, *30*(2), 126–140. <https://doi.org/10.1080/14675986.2018.1538043>
- Romero-Rivas, C., Martin, C. D., & Costa, A. (2015). Processing changes when listening to foreign-accented speech. *Frontiers in Human Neuroscience*, *9*. <https://doi.org/10.3389/fnhum.2015.00167>
- Rubin, D. L. (1992). Nonlanguage factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants. *Research in Higher Education*, *33*(4), 511–531. <https://doi.org/10.1007/bf00973770>
- Saito, K., Dewaele, J.-M., Abe, M., & In'nami, Y. (2018). Motivation, emotion, learning experience, and second language comprehensibility development in classroom settings: A cross-sectional and longitudinal study. *Language Learning*, *68*(3), 709–743. <https://doi.org/10.1111/lang.12297>
- Tamminen, H., Peltola, M. S., Kujala, T., & Näätänen, R. (2015). Phonetic training and non-native speech perception—New memory traces evolve in just three days as indexed by the mismatch negativity (MMN) and behavioural measures. *International Journal of Psychophysiology*, *97*(1), 23–29. <https://doi.org/10.1016/j.ijpsycho.2015.04.020>
- Tsang, A. (2022). FL listening motivation, interest, linguistic self-confidence, and overall proficiency: A complete double mediation model. *System*, *108*, 102833. <https://doi.org/10.1016/j.system.2022.102833>
- Van Engen, K. J., & Peelle, J. E. (2014). Listening effort and accented speech. *Frontiers in Human Neuroscience*, *8*. <https://doi.org/10.3389/fnhum.2014.00577>
- Vaughn, C. R. (2019). Expectations about the source of a speaker's accent affect accent adaptation. *The Journal of the Acoustical Society of America*, *145*(5), 3218–3232. <https://doi.org/10.1121/1.5108831>
- Weissler, R. E. (2022). A meeting of the minds: Broadening horizons in the study of linguistic discrimination and social justice through sociolinguistic and psycholinguistic approaches. *Annual Review of Applied Linguistics*, *42*, 137–143. <https://doi.org/10.1017/S0267190521000131>
- Winn, M. B., & Teece, K. H. (2021). Listening effort is not the same as speech intelligibility score. *Trends in Hearing*, *25*, 23312165211027690. <https://doi.org/10.1177/23312165211027688>

- Zhao, T. C., & Kuhl, P. K.** (2016). Musical intervention enhances infants' neural processing of temporal structure in music and speech. *Proceedings of the National Academy of Sciences*, 113(19), 5212–5217. <https://doi.org/10.1073/pnas.1603984113>
- Zhao, T. C., Llanos, F., Chandrasekaran, B., & Kuhl, P. K.** (2022). Language experience during the sensitive period narrows infants' sensory encoding of lexical tones—Music intervention reverses it. *Frontiers in Human Neuroscience*, 16. <https://doi.org/10.3389/fnhum.2022.941853>

Cite this article: Weissler, RE., Drake, S., Kampf, K., Diantoro, C., Foster, K., Kirkpatrick, A., Prelinger, I., Wesson, O., Wood, A., and Baese-Berk, MM. (2023). Examining linguistic and experimenter biases through “non-native” versus “native” speech. *Applied Psycholinguistics* 44, 460–474. <https://doi.org/10.1017/S0142716423000115>