Proceedings of the LSA

Is my bilingual your bilingual? Researchers' definitions and operationalizations of multilingual terminology

Kelly Kendro & Scott Jarvis*

Abstract. Comparisons of results from studies of "bilinguals" often assume that the participants come from a comparable "bilingual" population. This study examines consistency in researchers' use and definition of seven terms describing an individual's language experience: monolingual, bilingual, trilingual, multilingual, heritage language user, language learner, and language attriter. We report preliminary data from 19 language researchers who completed a four-part survey. This paper explores responses to two of the four parts: the vignette and Likert-scale items. The vignette portion asked respondents to evaluate how well each of the seven target terms fits vignette participant descriptions. The Likert-scale portion collected responses to statements about beliefs and practices. The results showed wide variation in classification of "participant" language experience in the vignettes, as well as research practices such as control group comparisons and researcher beliefs about concepts such as the critical period and trustworthiness of self-report proficiency data. Altogether, the results highlight inconsistency in use of language experience descriptors and the implications for conducting and interpreting empirical research.

Keywords. Bilingualism, language experience, terminological variation, research methods

1. Introduction. Though multilingualism exists on a spectrum, researchers often employ labels such as "bilingual," "monolingual," or "heritage speaker" to describe their participants. Jessner (2008:20) suggests that "finding a definition of multilingualism can be described as one of the most daunting research questions of current linguistics." Indeed, despite the widespread use of categorical labels for language experience (used here as an umbrella term to refer to an individual's combination of language knowledge, proficiency, and use), these terms lack agreement among linguists and laypeople about the boundaries of their definitions (e.g., Kremin & Byers-Heinlein, 2021; Surrain & Luk, 2019; Wagner et al., 2022). Uncertainty in the scope of defining these terms poses challenges across stages of the research process for both researchers and laypeople.

When language experience labels are not explicitly defined in research, they invite inconsistency in how they are applied, which can complicate recruitment, interpretation, replicability, and science communication. In a systematic review of monolingual/bilingual comparisons in research, Surrain and Luk (2019) found that researchers' criteria for "bilingual" varied widely, with language use (79%), proficiency (77%), age of acquisition (67%), and language of schooling (60%) most commonly used for "bilingual" classification; however, "bilingual" operationalizations varied within these criteria. In a scoping review of developmental bilingualism research, Rocha-Hidalgo and Barr (2023) found that 86% of papers in their sample provided some definition of bilingual, with the most used criteria as exposure to a second language (85%) and bilingual status of the primary caregiver (13%). However, minimum exposure thresholds

* Authors: Kelly Kendro, Northern Arizona University (Kelly.Kendro@nau.edu) & Scott Jarvis, Northern Arizona University (Scott.Jarvis@nau.edu).

varied widely across studies, and 20% of studies did not report their threshold. Castro et al. (2022) raise similar concerns, asking whether the term "bilingual" should be restricted to a certain frequency or context of language use, or whether it should describe anyone who knows multiple languages. In short, the use of underspecified or underdefined terms can lead to confusion and methodological uncertainty when, for instance, a "heritage speaker" in one study would be "bilingual" in another.

Lack of terminological consensus harms research theory and practice, a growing concern as indicated by recent scholarship. De Houwer (2023) advocates for responsible bilingualism research in which extensive information about all language knowledge and exposure is collected. Rothman et al. (2023) point out that language knowledge should be considered continuous rather than dichotomous and question the empirical soundness of including "monolingual" control groups in multilingualism research. Namboodiripad et al.'s (2026) consensus paper, authored by linguists and other language researchers from allied fields, notes that essentialist characterization of labels like "native speaker," when used without clarification of the criteria for membership in such a group, impose artificial homogeneity on language experience. In sum, use of categorical descriptors of language experience may lead to false equivalence between participant groups, particularly when these labels are used without explicit definition.

The problem of inconsistent use of language experience labels also extends to interactions with the general public. When recruiting "bilingual" participants, for instance, self-identification may vary widely and does not directly reflect language proficiency (Fishman, 1965). Even "monolingual" participants may have additional, undisclosed language background beyond a single language (Castro et al., 2022). In a largescale survey, Sia and Dewaele (2006) found that some individuals did not self-identify as "bilingual" despite average self-rated proficiency as 5.8/10. Beyond participant recruitment, when research about "bilingualism" is disseminated to the wider public, layperson readers may not know whether the research applies to them or if they "count" as "bilingual."

While previous work has extensively problematized the use of underspecified categorical descriptions of language experience in language research, there is sparse direct evidence regarding research practices in relation to language experience characterization from the people who conduct this work. Thus, the current study has three major aims:

1. Quantifying variation in use of multilingual terms/labels.
2. Identifying researcher beliefs and practices in the study of bilingualism.
3. Exploring (sub)disciplinary trends in responses.

2. Method.

2.1 RESPONDENTS. We recruited language researchers through postings on the "Applied Linguistics Research Methods—Discussion" Facebook group, listservs related to the study of bilingualism (i.e., ISB_list, ROLE Collective), and through word-of-mouth. Here, we present preliminary data from 19 researchers who have been active in linguistics and related fields for 4 to 47 years (M = 17.8, SD = 11.3). They self-reported expertise in the subfields of psycholinguistics, neurolinguistics, education, speech-language pathology, and sociolinguistics.

2.2 MATERIALS. The survey was built in Qualtrics and consisted of four parts. The first part asked about the beliefs, frameworks, and practices present in their research related to language experience. The second part consisted of evaluating how well each of the seven terms fit vignette participant descriptions. The third part collected Likert-scale responses to statements about

beliefs and practices. The fourth part asked about the researcher's demographics and training. We detail parts 2 (vignettes) and 3 (Likert-scale items) in this paper.

We chose to use vignettes in this study because they offer realistic and plausible situations that require respondents to interpret information beyond simply self-reporting (Goetze, 2023). This study included a total of 18 vignettes which were modeled on either descriptions of monolingual, heritage speaker, bilingual, trilingual, language learner, or language attriter participant samples (12 vignettes) from published research or from self-descriptions (6 vignettes) of participants recruited by Kendro and Jarvis (2025). After reading the vignette, respondents indicated how well each of the seven target labels described the model participant (see Figure 1). The presentation order of vignettes was randomized by Qualtrics.

Figure 1. Response matrix for vignette items.

There were 22 Likert-scale items in total, each containing a statement about research practices (e.g., "I compare my bilingual participants to a control group of monolinguals.") or beliefs (e.g., "I believe that participants overestimate their language knowledge."). These items were displayed via multiple choice radio buttons to allow a single response to each question. The Likert scale options were either "strongly agree," "agree," "disagree," or "strongly disagree" (10 statements) or "always," "frequently," "sometimes," "rarely," "never," or "not ap-plicable" (12 statements).

3. Results. Due to length constraints, we present here four vignettes demonstrating wide variability in language experience classifications and ten Likert scale items. We report general trends in results as our sample is not yet large enough for the planned inferential statistical analyses.

3.1 VIGNETTES. The first vignette was adapted from the description of trilingual participants in Lemhöfer et al. (2004):

> Johanna (aged 25) lives in the Netherlands. She learned Dutch at home as a child, and starting at age 12 she began taking English and German language courses. She is now a university student with a concentration in German, and she is comfortable communicating in any of these languages.

As shown in Figure 2, most respondents agreed that "language attriter" did not fit Johanna at all and that "multilingual" fit her well. However, respondents were divided on whether she should be considered "bilingual," and only 73.7% of respondents believed the label "trilingual" fit her "very" or "somewhat" well despite the vignette being an adapted description of trilingual participants.

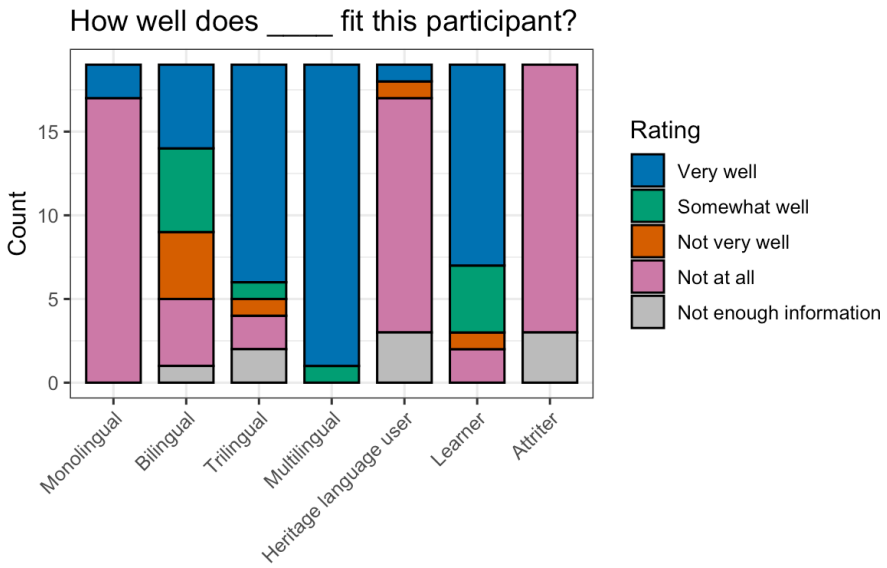

Figure 2. Classification of "Johanna."

The second vignette was adapted from a self-reported bilingual participant's self-description of language experience from Kendro and Jarvis (2025):

> Anika (aged 38) lives in South Africa. Her father's family is Afrikaner, and her paternal grandfather was a Dutch emigrant. Her mother's family is English. Anika's schooling was in English, and she took Afrikaans language courses. She knows limited Tswana and Zulu from the community where she grew up, and her grandfather sometimes used a combination of Dutch and Afrikaans, but she cannot hold a conversation in a language other than English and Afrikaans.

As shown in Figure 3, most respondents believed the label "monolingual" did not fit Anika at all and that "bilingual" and "multilingual" fit her "very" or "somewhat" well. However, respondents were divided on whether she should be considered "trilingual," a "heritage language user," and/or a "learner."

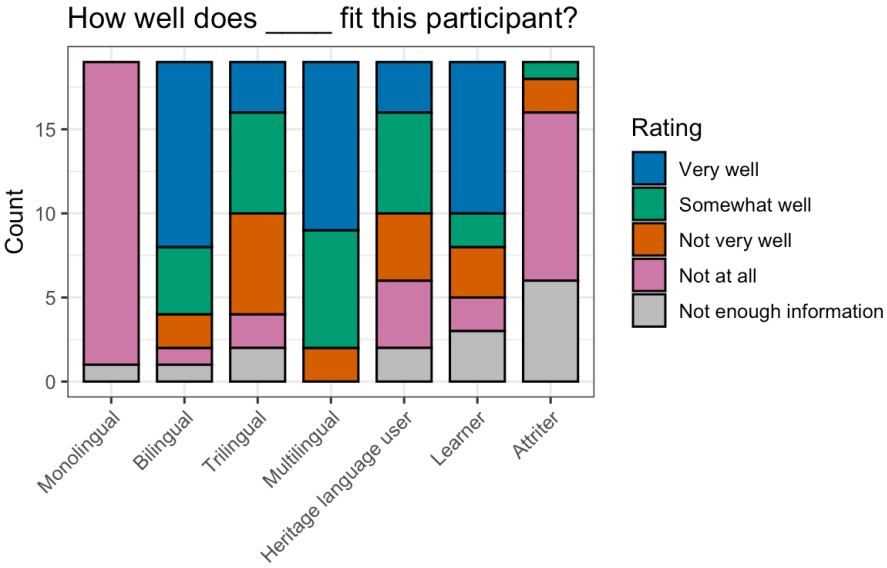

Figure 3. Classification of "Anika."

The third vignette was adapted from the description of heritage speaker participants in Polinsky (2008):

> Olga (aged 27) is a graduate student at a Canadian university. She was born in the Czech Republic, and her family moved to Canada when she was 4 years old. She has not spoken Czech since about the age of 5, but she can understand the language when others speak it. She has never been literate in Czech, and she only uses English to communicate.

As shown in Figure 4, most respondents believed that "trilingual" did not describe Olga at all. Respondents did show wide variation in whether Olga should be considered an "attriter," learner, and/or "multilingual." Similar to the first vignette, while most respondents (68.4%) said that "heritage language user" fit Olga "very" or "somewhat" well, this was not a unanimous designation despite the vignette's origin as a description of "heritage speaker" from the literature.

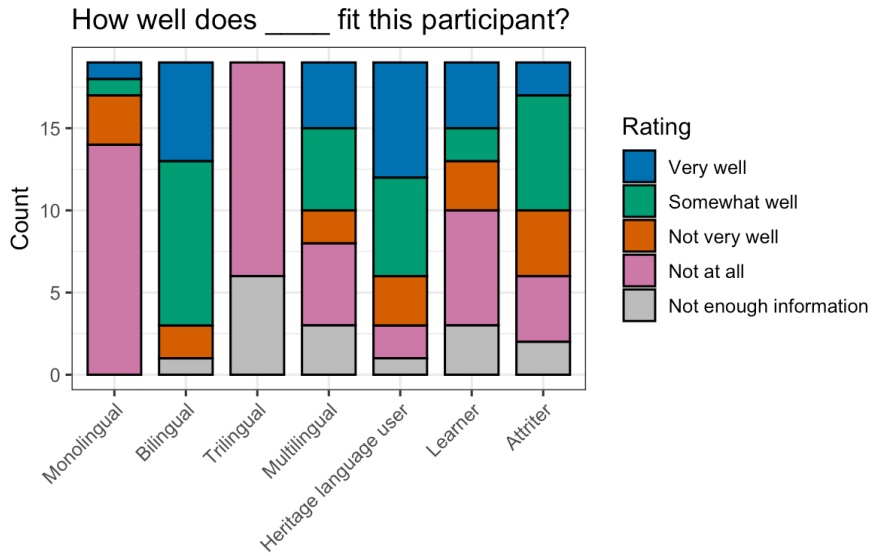

Figure 4. Classification of "Olga."

The fourth vignette was adapted from a self-reported bilingual participant's self-description of language experience from Kendro and Jarvis (2025):

> Claudia (aged 22) grew up in America, and her grandmother is Japanese. Though her grandmother did not teach her children or grandchildren the language, she spoke to her family in Japanese often enough that the children were able to understand (but not speak) the language. Claudia attended a Spanish immersion school as a child, but she did not continue with Spanish in university and no longer feels comfortable speaking the language.

As shown in Figure 5, respondents showed a lack of consensus for most terms, broadly disagreeing on whether terms such as "bilingual," "trilingual," "multilingual," "heritage language user," and/or "attriter," were appropriate characterizations of Claudia's language experience.

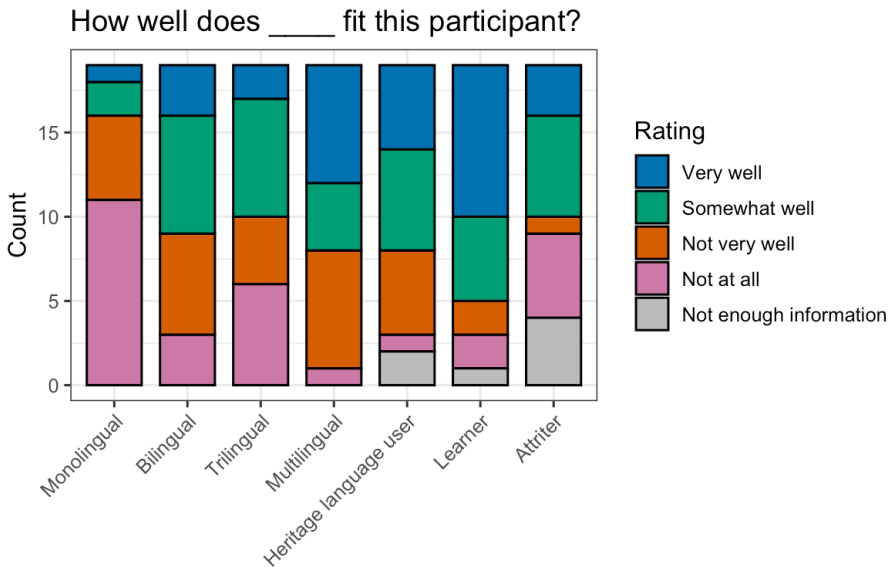

Figure 5. Classification of "Claudia."

3.2 LIKERT SCALE ITEMS. Our Likert scale items focused on researchers' beliefs and practices, and we report responses to 10 of the 22 total items.

Figure 6 shows responses regarding a bilingual advantage and a critical or sensitive period. Generally, respondents believed in some sort of bilingual advantage (84.2% agree or strongly agree), while belief in a critical period in second language acquisition was much lower (15.8% agree or strongly agree) than belief in a *sensitive* period (78.9% agree or strongly agree).

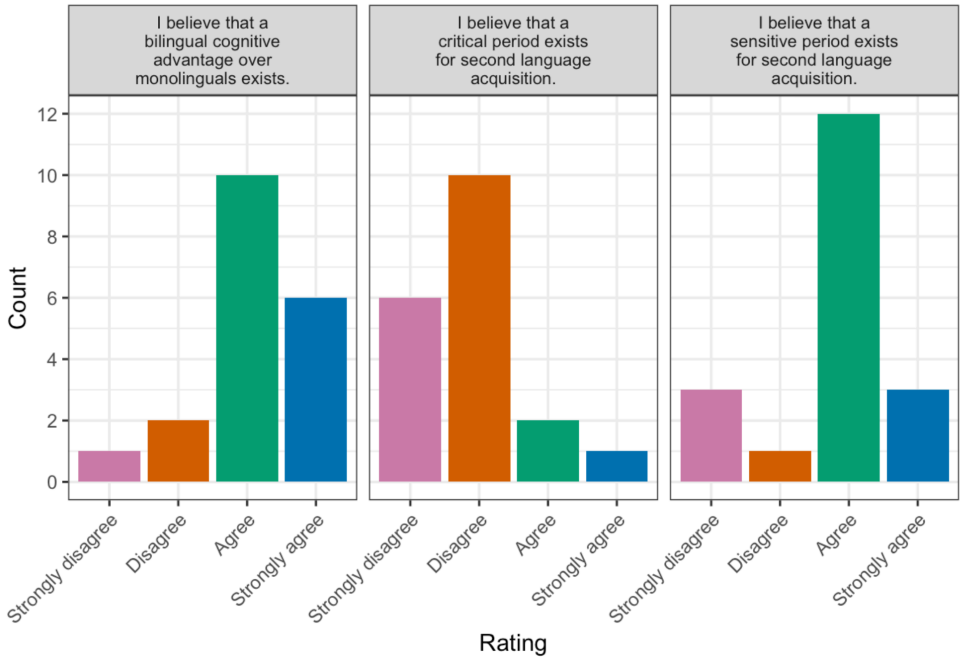

Figure 6. Beliefs about the bilingual advantage and a sensitive or critical period.

Figure 7 shows information about researchers' practices related to language experience. While all respondents said they either "frequently" or "always" collected comprehensive

language data, only 35.3% of respondents reported considering attrition "frequently" or "always," potentially missing important language background information. In other words, comprehensive language experience information does not appear to always include attrition. Many of the respondents collected qualitative information about language experience (78.9% frequently or always), and all respondents reported information regarding a participant's "native language" (cf. methodological concerns about "nativeness" raised by Cheng et al., 2021).

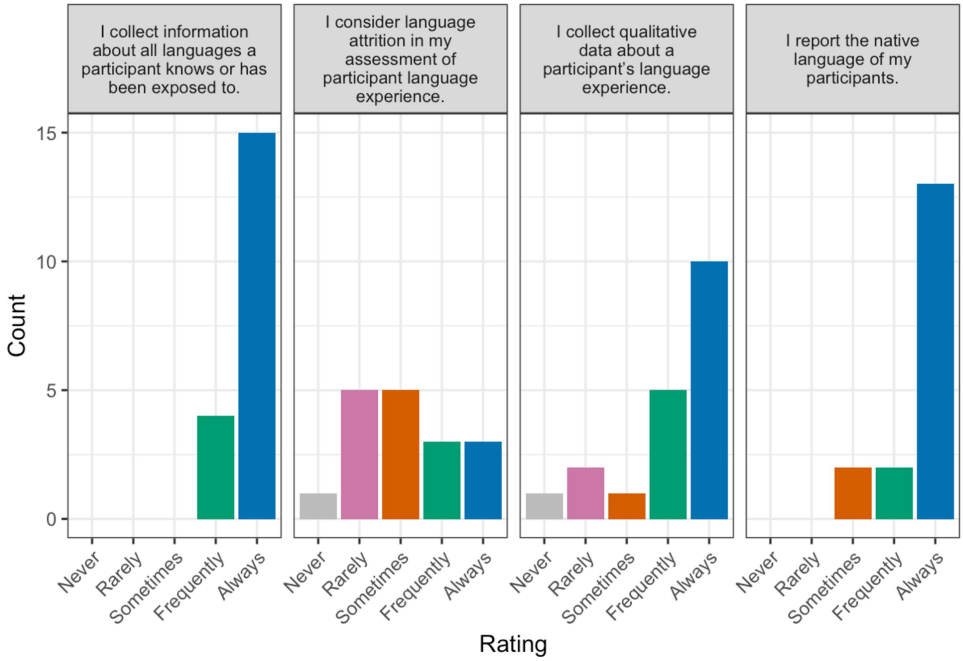

Figure 7. Practices regarding participant language experience information.

Figure 8 shows information about researchers' beliefs related to the term "bilingual" and multilingualism research more broadly. Responses show major concerns with the comparability of "bilingual" across researchers and studies, as 78.9% of respondents said they disagreed or strongly disagreed that the term is used consistently. Responses to whether self-identified bilinguals should be considered "bilingual" were mixed, with 42.1% disagreeing and 57.9% agreeing. Most respondents (73.7%) believe that, to some extent, there is a reproducibility problem in multilingualism research.

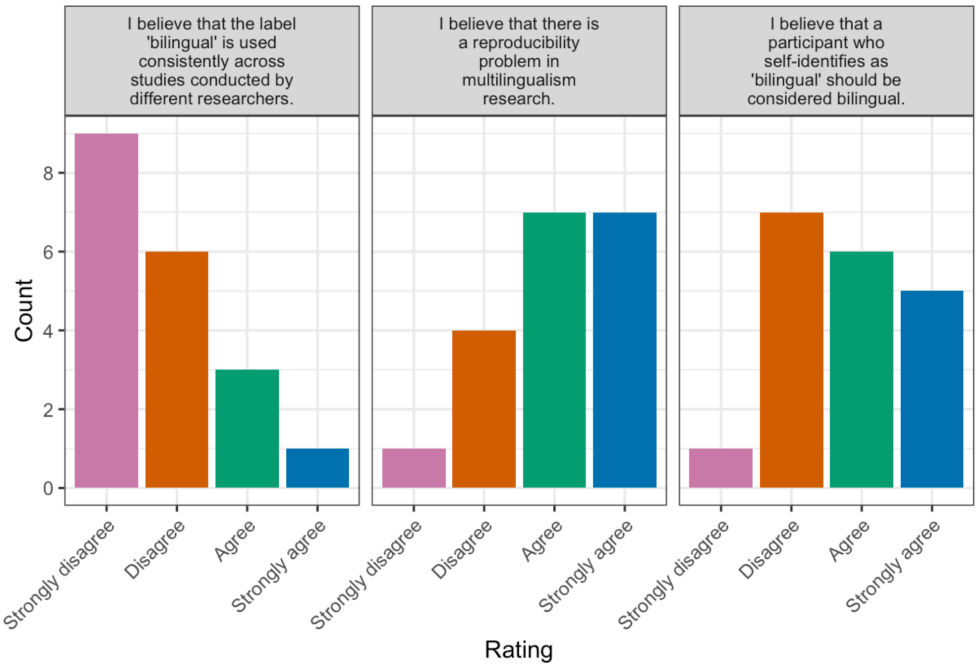

Figure 8. Beliefs about the term "bilingual" and reproducibility in multilingualism research.

4. Discussion. Our empirical results from the vignettes illustrate how widely researchers vary in attributing a population-level label such as "bilingual" to an individual's language experience. Across all responses, there was no vignette for which a respondent believed only one label fit "very" well. Instead, the results demonstrate the overlapping, non-discrete boundaries of language experience descriptors, even when the vignettes are directly adapted from published descriptions of specific types of participant language experience as in the first and third vignettes.

When there is ambiguity in terminology at the conceptual level, decisions about participant classification can unintentionally complicate or hinder the interpretation of results. These data underscore that broad language experience labels like "bilingual" are context-dependent and lack consistent or discrete boundaries, which in turn compromise cross-study comparison in the absence of explicit and detailed operational definitions. To mitigate these issues, researchers can collect more comprehensive language experience data from all participants (see De Houwer, 2023; Kendro & Jarvis, 2025) and report detailed criteria/thresholds for group membership when including categorical comparisons.

Our empirical results from the Likert scale items shed light on the research beliefs and practices underpinning the respondents' vignette responses. Even though respondents largely do not believe that "bilingual" is used consistently across studies, there is still widespread support for the notion of a "bilingual advantage." Stronger belief in the existence of a "sensitive period" over a "critical period," as well as split beliefs about the replicability of multilingualism research, is consistent with a shift from categorical concepts of language knowledge to more gradient or continuous representations of language experience. Though researchers say they collect comprehensive language experience information, potential attrition is overlooked; as attrition may be systematically underreported by participants, even when using major surveys of language experience (Kendro & Jarvis, 2025), it is important for researchers to explicitly ask about attrition to build a more complete representation of a participant's language experience. Thus, these

results highlight a disconnect between researchers' awareness of terminological variation and their continued use of these constructs and labels.

Altogether, the vignette and Likert scale items point to a major underlying issue: language experience labels like "bilingual" do not have stable conceptualizations or operationalizations across language research. Inconsistent operationalizations of imprecise, underspecified, or conflated terminology can lead to methodological instability and ambiguous interpretation of results. Methodological instability may manifest as studies identifying different participant populations but using the same underspecified language experience label, applying different inclusion criteria, or employing different measurement techniques. In turn, the interpretation of such results may reflect the differences in operationalization rather than the predicted population, construct, or task effects. These issues can be mitigated by further consensus regarding the use of terms like "bilingual," additional data collection and reporting of qualitative information about participants' language experience (as many of our respondents report already doing), and a shift away from categorical representations of language experience in favor of discussions that acknowledge the individual complexity and multidimensionality of multilingualism. Above all, these discussions need to occur during both theory building and methodological implementation, with discussion of each iteratively informing the other. Without addressing variation in language experience terminology, researchers risk conflating separate populations under a single, underspecified label or dividing a single population into multiple discrete categories.

5. Future steps. As of writing, data collection is ongoing. Eventually, we plan to use inferential statistics to explore possible patterns in responses related to subdisciplines, training, and/or other individual variables. Following recent approval from our institutional review board, we will be expanding our recruitment efforts to include additional venues such as Bluesky and Discord, as well as continuing recruitment efforts via listservs.

The ultimate goal of this research is not necessarily a unified definition of any of these terms (cf. approaches to definition consensus from Bishop, 2017; Hanks, 2025; Kendro & Cabrera Pérez, 2025): terms like "monolingual" and "bilingual" have transcended linguistics to join the layperson lexicon as descriptors of language experience, and other terms (e.g., heritage language user) may be used differently across different (sub)disciplines of language research. Instead, this research seeks to empirically document variation in how language experience labels are used and advocate for including clear working definitions to help others understand how a given term is operationalized. Even if no consensus can be reached, we seek to promote higher consistency in use and operationalization of language experience labels across language research.

References

Bishop, Dorothy V. M. 2017. Why is it so hard to reach agreement on terminology? The case of developmental language disorder (DLD). *International Journal of Language & Communication Disorders*, 52(6), 671–680. https://doi.org/10.1111/1460-6984.12335.

Castro, Sofía, Zofia Wodniecka, & Kalinka Timmer. 2022. Am I truly monolingual? Exploring foreign language experiences in monolinguals. *PloS one*, 17(3). e0265563. https://doi.org/10.1371/journal.pone.0265563.

Cheng, Lauretta S. P., Danielle Burgess, Natasha Vernooij, Cecilia Solís-Barroso, Ashley McDermott, & Savithry Namboodiripad. 2021. The problematic concept of native speaker in psycholinguistics: Replacing vague and harmful terminology with inclusive and accurate measures. *Frontiers in Psychology,* 12. 715843. https://doi.org/10.3389/fpsyg.2021.715843.

De Houwer, Annick. 2023. The danger of bilingual–monolingual comparisons in applied psycholinguistic research. *Applied Psycholinguistics*, 44(3). 343–357. https://doi.org/10.1017/S014271642200042X.

Fishman, Joshua A. 1965. Who Speaks What Language to Whom and When? *La Linguistique*, 1(2). 67–88. https://www.jstor.org/stable/30248773.

Goetze, Julia. 2023. Vignette methodology in applied linguistics. *Research Methods in Applied Linguistics,* 2(3). 100078. https://doi.org/10.1016/j.rmal.2023.100078.

Hanks, Elizabeth. 2025. *Mapping Out American English Conversation: Central and Peripheral Features of Intra-Register Variation*. Flagstaff, AZ: NAU Dissertation. https://www.proquest.com/openview/439569f758d87321304675f116158102/1?pq-origsite=gscholar&cbl=18750&diss=y.

Jessner, Ulrike. 2008. Teaching third languages: Findings, trends and challenges. *Language Teaching*, 41(1), 15–56. https://doi.org/10.1017/S0261444807004739.

Kendro, Kelly, & Reinaldo Cabrera Pérez. 2025. Precision matters: Synthesis through terminological review in applied linguistics. Manuscript. SSRN. https://dx.doi.org/10.2139/ssrn.5704625.

Kendro, Kelly, & Scott Jarvis. 2025. *Assessing potential language attrition via the Supplemental Language Attrition Questionnaire*. Manuscript. PsyArXiv. https://doi.org/10.31234/osf.io/wrgq4_v1.

Kremin, Lena V., & Krista Byers-Heinlein. 2021. Why not both? Rethinking categorical and continuous approaches to bilingualism. *International Journal of Bilingualism*, 25(6). 1560–1575. https://doi.org/10.1177/13670069211031986.

Lemhöfer, Kristin, Ton Dijkstra, & Marije Michel. 2004. Three languages, one ECHO: Cognate effects in trilingual word recognition. *Language and Cognitive Processes*, 19(5), 585–611. https://doi.org/10.1080/01690960444000007.

Namboodiripad, Savithry, Ethan Kutlu, Anna Babel, Molly Babel, Melissa Baese-Berk, Paras B. Bassuk, Adeli Block, Reinaldo Cabrera Pérez, Matthew T. Carlson, Sita Carraturo, Andrew Cheng, Lauretta S. P. Cheng, Philip Combiths, Ruthe Foushee, Anne Therese Frederiksen, Devin Grammon, Rachel Hayes-Harb, Eve Higby, Kelly Kendro, Elena Koulidobrova, Chien-Jer Charles Lin, Zhicheng Lin, Alicia Luque, Kevin B. McGowan, John Muegge, Christine Shea, Alayo Tripp, Sarah J. Woods, & Kelly E. Wright. 2026. *Cognition*, 271, 106444. Finding our ROLE: How and why to reframe essentialist approaches to language. https://doi.org/10.1016/j.cognition.2026.106444.

Polinsky, Maria. 2008. Gender under Incomplete Acquisition: Heritage Speakers' Knowledge of Noun Categorization. *Heritage Language Journal*, 6(1), 40–71. https://doi.org/10.46538/hlj6.1.3.

Rocha-Hidalgo, Joscelin, & Rachel Barr. 2023. Defining bilingualism in infancy and toddlerhood: A scoping review. *International Journal of Bilingualism,* 27(3), 253–274. https://doi.org/10.1177/13670069211069067.

Rothman, Jason, Faith Bayram, Vincent DeLuca, Grazia Di Pisa, Jon Andoni Duñabeitia, Khadij Gharibi, Jiuzhou Hao, Nadine Kolb, Maki Kubota, Tanja Kupisch, Tim Laméris, Alicia Luque, Brechje van Osch, Sergio Miguel Pereira Soares, Yanina Prystauka, Deniz Tat, Aleksandra Tomić, Toms Voits, & Stefanie Wulff. 2023. Monolingual comparative normativity in bilingualism research is out of "control": Arguments and alternatives. *Applied Psycholinguistics*, 44(3). 316–329. https://doi.org/10.1017/S0142716422000315.

Sia, Jennifer, & Jean-Marc Dewaele. 2006. Are you bilingual? *Birkbeck Studies in Applied Linguistics,* 1, 1–19. https://eprints.bbk.ac.uk/id/eprint/15551/

Surrain, Sarah, & Gigi Luk. 2019. Describing bilinguals: A systematic review of labels and descriptions used in the literature between 2005–2015. *Bilingualism: Language and Cognition*, 22(2). 401–415. https://doi.org/10.1017/S1366728917000682.

Wagner, Danika, Ellen Bialystok, & John G. Grundy. 2022. What is a language? Who is bilingual? Perceptions underlying self-assessment in studies of bilingualism. *Frontiers in Psychology*, 13. 863991. https://doi.org/10.3389/fpsyg.2022.863991.