

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Assessing the Language and Cognitive Profiles of Deaf ASL-English Bilinguals Across the Lifespan

Permalink

<https://escholarship.org/uc/item/1pt9d72z>

ISBN

9798190927771

Author

Camilliere, Sadie

Publication Date

2026-09-11

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

SAN DIEGO STATE UNIVERSITY

Assessing the Language and Cognitive Profiles of Deaf ASL-English Bilinguals
Across the Lifespan

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Language and Communicative Disorders

by

Sadie Elizabeth Camilliere

Committee in Charge:

University of California, San Diego

Professor Tamar H. Gollan, Co-Chair

Professor Rachel Mayberry

Professor David Salmon

San Diego State University

Professor Karen Emmorey, Chair

Professor Henrike K. Blumenfeld

©

Sadie Elizabeth Camilliere, 2026

All rights reserved.

DISSERTATION APPROVAL PAGE

The Dissertation of Sadie Elizabeth Camilliere is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

San Diego State University
University of California San Diego

2026

TABLE OF CONTENTS

DISSERTATION APPROVAL PAGE.....	iii
TABLE OF CONTENTS	iv
LIST OF FIGURES	v
LIST OF TABLES	vi
ACKNOWLEDGEMENTS.....	vii
VITA	x
ABSTRACT OF THE DISSERTATION	xi
CHAPTER 1: General Introduction.....	2
CHAPTER 2: Influence of Testing Language and Aging on Verbal List Memory in Deaf American Sign Language-English Bilinguals.....	7
CHAPTER 3: Bigger between-language than within-language Stroop effects in young and older deaf American Sign Language-English bilinguals.....	53
CHAPTER 4: Semantic and Letter Fluency in Young and Older Deaf ASL-English Bilinguals: Effects of Aging and Testing Language	86
CHAPTER 5: General Discussion	114

LIST OF FIGURES

Figure 2.1. Mean number of items recalled in ASL and English on Trial 3 and after a Delay	28
Figure 2.2. Learning-adjusted forgetting scores in each language	29
Figure 2.3. Serial position curves showing the proportion of ASL and English items recalled for younger and older participants.....	30
Figure 2.4. Correlations between language proficiency and list memory recall	33
Figure 3.1. ASL and English Stimuli	67
Figure 3.2. Scatterplots illustrating the effects of age on baseline adjusted Stroop effects in RTs with ASL vs. with English stimuli (n=108)	72
Figure 3.3. Scatterplots illustrating the null relationship between English proficiency and English adjusted baseline Stroop score (left) and the null relationship between ASL proficiency and ASL adjusted baseline Stroop score (right).....	72
Figure 3.4. Baseline adjusted Stroop effect in RTs in four age groups (see Table 3.1) with ASL and English stimuli. Error bars denote standard error.....	73
Figure 4.1. Scatterplots illustrating the relationships between ASL category fluency and age (a) and between English category and English letter fluency and age (b-c).	101
Figure 4.2. Performance on fluency tasks by age group.....	104

LIST OF TABLES

Table 2.1. Averages and standard deviations for participant characteristics and demographic information.	17
Table 2.2. Lexical characteristics of ASL and English stimuli in the list memory tests.	19
Table 2.3. Order of items in ASL and English lists on Trials 1-3.	21
Table 2.4. Means (and SDs) of younger, middle-aged, and older deaf signers on the list memory tests.	25
Table 2.5. Correlations between language dominance scores and number of items recalled on Trials 1-3 and the delayed recall trial.....	32
Table 2.6. Means and standard deviations for recall scores on Trial 3 and the delayed trials by language and presentation order	35
Table 3.1. Participant characteristics divided into four age-groups.....	67
Table 3.2. Correlations between demographics, proficiency tests and Stroop measures (n=108) 69	
Table 3.3. Means (and SDs) for raw RTs and percent errors by condition in each age group.....	71
Table 4.1. Participant Characteristics.....	98
Table 4.2. Correlation matrix of fluency performance and demographics	100
Table 4.3. Means (and SDs) for fluency tasks by age group	102

ACKNOWLEDGEMENTS

This dissertation is the product of the work I have done over the last five years, and an endeavor that would not have been possible without the support of so many people. First, to Dr. Karen Emmorey, thank you for your consistent advice, feedback, and encouragement throughout my time here. Despite juggling a million responsibilities and jet-setting all over the world, you've always made time to sit down in person to check in whenever I needed it, and I am so grateful for all you have taught me, both explicitly and by example. It truly has been an honor and a dream to work with you, I couldn't have asked for a better mentor. To Dr. Tamar Gollan, thank you for your patience with me as I navigated the realm of neuropsychological assessment, and for sharing your expertise and enthusiasm in every part of the research process, I am certainly a better scientist for having worked with you. To the other members of my committee, Drs. Henrike Blumenfeld, Rachel Mayberry, and David Salmon, thank you for your time and feedback on this dissertation, your input has been so valuable.

I'm so grateful for the wonderful team of researchers here at the LLCN, in particular Cindy Farnady and Julie Fisher for their tremendous help running these studies both in San Diego and across state lines, and Allison Bassett for always looking out for me, in every way I could need. Through the last year or so of this program, I have been so privileged to have Caitlin Pearson, a rockstar of a research assistant, to help me with so many aspects of this project—your help has allowed me to breathe.

This research project spans multiple institutions, and I have been lucky to experience the well-oiled machine that is Dr. Peter Hauser's lab, and it's been such a pleasure to work with everyone involved on the RIT side including Dr. Jessica Contreras, Emma Kane, Hayden Orr, Lucy Herbold, Chang-May Tan, Gigi Zheng, and Chehally Soto.

To my current lab mates, Emily Akers, Ayden Kpormegbey, and Sofia Ortega, thank you for the bad book club picks, the afternoons working at Majorette, and for making the experience of getting a Ph.D. not just bearable but enjoyable. And to the ones who came before, Drs. Meghan McGarry, Emily Saunders, Sarah Kimbley, and Brennan Terhune-Cotter, thank you for welcoming me into such a collaborative, supportive lab environment, I'm so grateful for your guidance and friendship through our time at the LLCN and beyond.

To the professors and mentors I was lucky enough to work with at Swarthmore, particularly Drs. Daniel Grodner, Peet Klecha, and Donna Jo Napoli, thank you for helping me figure out who I am as a researcher, a linguist, an advocate. I learned so much from you, and I absolutely found the path I am on thanks to you. To Melanie Drolsbaugh, my first ASL professor, thank you for giving me a window into this world, and for letting me choose my own topic for the first research paper in your class instead of picking from the approved list—you sparked a curiosity in me that's followed me ever since.

It would have been impossible for me to have gotten to this point without the many support systems that have kept me motivated, laughing, and well-fed throughout my life. My time in San Diego would have been much less bright without Jackie Jardel and Averil Lassiter, and I don't think I would have survived college without Evan Orticia, Lilia McGee-Harris, Bella Fiorante, and Andrew Huynh. To Spencer Fitzgerald, Savanna Robbins, Camryn Bittner, and Morgan McGuffin, I know you all are watching and on my side no matter how far we are, and I will never take that for granted. To my brothers, Alex and Peter, who may not understand what I'm doing over here in California, but are supportive nonetheless, thank you for always making me laugh exactly when I need it. And finally, to my parents, Jim and Betsie, thank you for the relentless support and love you've always shown me in every aspect of my life. Everything I

have accomplished has been buoyed by the opportunities you have given me. So, thank you—for everything.

Chapter 2, in full is a reprint of the material as it appears in *Neuropsychology*, 2026.

Camilliere, S., Emmorey, K., Hauser, P. C., Contreras, J., McKee, M. M., & Gollan, T. H. (2026). Influence of testing language and aging on verbal list memory in deaf American Sign Language–English bilinguals. The dissertation author was the primary investigator and author of this paper. This research was supported by NIH grant R01-AG074928.

Chapter 3 has been submitted for publication of the material and is currently under revised peer review. Camilliere, S., Emmorey, K., Hauser, P. C., Contreras, J., McKee, M. M., & Gollan, T. H. (Under review). Bigger between-language than within-language Stroop effects in young and older deaf American Sign Language-English bilinguals. *Journal of International Neuropsychological Society*. The dissertation author was the primary investigator and author of this paper. This research was supported by NIH grant R01-AG074928.

Chapter 4 is being prepared for submission for publication of the material. Camilliere, S., Emmorey, K., Hauser, P. C., Contreras, J., McKee, M. M., Salmon, D.P., & Gollan, T. H. Semantic and Letter Fluency in Young and Older Deaf ASL-English Bilinguals: Effects of Aging and Testing Language (Manuscript in preparation). The dissertation author was the primary investigator and author of this paper. This research was supported by NIH grant R01-AG074928.

VITA

- 2020 Bachelor of Arts in Psychology and Linguistics, Swarthmore College
- 2026 Doctor of Philosophy in Language and Communicative Disorders, Joint Doctoral Program, San Diego State University/ University of California San Diego

PUBLICATIONS

Camilliere, S., Emmorey, K., Hauser, P. C., Contreras, J., McKee, M. M., & Gollan, T. H.
(2026). Influence of testing language and aging on verbal list memory in deaf American Sign Language–English bilinguals. *Neuropsychology*.

Under Review

Camilliere, S., Emmorey, K., Hauser, P. C., Contreras, J., McKee, M. M., & Gollan, T. H.
(Under revision). Bigger between-language than within-language Stroop effects in young and older deaf American Sign Language-English bilinguals. *Journal of the International Neuropsychological Society*.

FIELD OF STUDY

Major Field: Language and Communicative Disorders
Neuropsychological evaluation of deaf signers
Professor Karen Emmorey

ABSTRACT OF THE DISSERTATION

Assessing the Language and Cognitive Profiles of Deaf ASL-English Bilinguals
Across the Lifespan

by

Sadie Elizabeth Camilliere

Doctor of Philosophy in Language and Communicative Disorders

University of California San Diego, 2026
San Diego State University, 2026

Professor Karen Emmorey, Chair
Professor Tamar H. Gollan, Co-Chair

While much is known about the cognitive and linguistic changes associated with aging in hearing monolinguals and bilinguals, less is known about aging effects in deaf bimodal bilinguals who use American Sign Language (ASL) and English. This population presents unique challenges for neuropsychological testing due to a lack of appropriate assessment material, wide variability in exposure to language, and differences in educational backgrounds. This dissertation investigated aging effects in deaf ASL-English bilinguals by adapting three

widely used neuropsychological tests that have shown utility for identifying Alzheimer's disease in hearing monolinguals.

Study 1 examined the effects of aging and testing language on verbal list learning, a common component of cognitive evaluations, in a group of young and older deaf ASL-English bilinguals. Younger participants showed no effect of language on recall and recalled more items than older participants in both languages. Older participants showed better immediate learning for English stimuli, but higher rates of forgetting in English compared to when they were tested in ASL, suggesting that in older signers, the two languages maximize different aspects of memory performance, both of which may be important for clinical evaluation of memory. Study 2 investigated Stroop performance in deaf bilinguals across a wide age range using computerized ASL and English Stroop tasks. Stroop effects were larger with English compared to ASL stimuli for both young and older deaf bilinguals. Stroop interference effects were larger for older than for young participants in both languages, but the aging-related increase in Stroop effects was larger in English than in ASL. Study 3 explored deaf bilinguals' verbal fluency performance using ASL and English semantic category fluency tasks and an English letter fluency task. Of the three verbal fluency tasks administered, ASL category fluency was the most sensitive to aging effects, whereas English category and letter fluency performance was better predicted by education and proficiency level (as measured by independent objective measures of ASL and English proficiency).

Taken together, these studies provide insight into how three commonly used neuropsychological tests must be adapted for use with deaf, signing populations, and provide preliminary normative data for older deaf signers across a wide age range. While some results patterned similarly to what is observed in hearing, unimodal bilinguals, several critical

differences emerged which may be attributed to modality driven effects, effects of language acquisition, or schooling experiences that are specific to deaf individuals. These findings help to characterize some aspects of the unique linguistic and cognitive profiles of older deaf signers, a critical first step in evaluating whether these tasks can be used for clinical assessment of cognitive decline or Alzheimer's disease.

CHAPTER 1: General Introduction

There is a vast body of research investigating the cognitive and linguistic changes associated with aging in hearing, spoken language users, but research investigating typical aging effects in deaf sign language users is scarce. While initial explorations of aging in deaf signers (Corina et al., 2020; Luna et al., 2021) have found some effects parallel to those found in older hearing non-signers, they also reveal differences between populations of elder deaf and hearing individuals, potentially caused by modality differences as well as differential circumstances of language acquisition and schooling experiences. Understanding the aging profiles of older deaf adults is critical for designing sensitive cognitive and language evaluations.

Neuropsychological evaluation of deaf individuals presents several challenges unique to this population. Unique aspects of their language and educational backgrounds, as well as linguistic and modality differences between signed and spoken languages, contribute to these complications. Many deaf children, between 90-95%, are born to hearing parents who have no knowledge of sign language or Deaf culture (Mitchell & Karchmer, 2004), and these children may not be fully exposed to an accessible language (e.g., a sign language) until several years after they are born. This delayed access to a first language can lead to language deprivation which can impact an individual's later linguistic and cognitive abilities (Humphries et al., 2016; Mayberry et al., 2002). Understanding the contribution of delayed Age of Acquisition (AoA) on cognitive and linguistic performance may be a critical part of neuropsychological testing with deaf individuals. For example, Zorzi et al. (2023) compared the performance of native, early, and late signers on standardized lexical and syntactic tests that are used to assess language impairments in adult users of French Sign Language and Italian Sign Language. The researchers found categorical differences in performance between native and non-native (i.e., early and late)

signers, with native signers demonstrating higher accuracy on syntactic tasks than non-native signers. The authors argue that although language based tasks are critical for clinical assessment, there is a need for baseline data that captures AoA effects when assessing language competence in the deaf population. Importantly, Corina et al. (2020) showed that AoA effects were dissociable from effects of chronological aging in an ASL sentence repetition task. The impacts of language deprivation could easily be confounded with aging effects, disordered or typical, if care is not taken to investigate deaf signers with a wide range of ages of acquisition of ASL.

Educational differences between older and younger generations of deaf signers, particularly in America, may also influence language proficiency in both ASL and English. Many older deaf adults were taught using an oral approach which emphasized spoken language and often discouraged or punished sign language use (Moore, 2010). In the 1970s, there was a shift towards a bilingual-bicultural teaching model that encouraged the use of both ASL and English in the classroom (Padden, 1996). Deaf seniors report a wide range of experiences with schooling, ranging from having minimal access to the curriculum as it was presented in spoken English with no interpretation, to having full access at residential schools that supported the use of ASL (Roberson & Shaw, 2015). These generational differences in schooling experience must also be taken into account when examining performance on neuropsychological assessments, particularly those involving English.

Another important consideration is that almost all deaf individuals are bilingual to some extent in a sign language and a spoken/written language (e.g., ASL and written English). In the assessment of hearing bilinguals, it is critical to be able to identify which language is dominant, as testing in the non-dominant language can lead to higher rates of false positives (Gasquoine & Gonzalez, 2012). Testing in the dominant language may also maximize differences between

patients and controls, such as in bilingual picture naming performance (Gollan et al., 2010; Gollan et al., 2023). For many deaf individuals in America, ASL is their primary language of communication (Grosjean, 2010). Many deaf individuals only use written English, and as a result the majority of their daily communication and social interactions is likely in ASL which does not have a written form. Written English provides deaf individuals access to information, sources of entertainment, and a way to communicate with the dominant culture. These contextual differences may contribute to a unique profile of language dominance that is distinct from that of hearing bilinguals. Understanding how language dominance and language modality impact performance on neuropsychological assessments is critical for clinical application of these tests, particularly when making decisions about the language of test administration.

In addition to demographic and linguistic characteristics of deaf populations that can complicate neuropsychological testing, there are also features of the test design and administration that must be accounted for as well. Morere (2013) describes some of the methodological issues that can occur when designing and administering signed-based neuropsychological assessments, including difficulties with the direct translation of English stimuli (see also Peña, 2007) and instructions into ASL. Issues can arise when attempting to translate phrase-level stimuli from English to ASL due to differences in grammatical structures and the visuospatial affordances that sign languages take advantage of. As a result, sentences that are simpler in English may become more complex when expressed in ASL, making interpretation of performance on those items unclear (Paludneviciene et al., 2012).

One of the primary barriers of access to diagnoses and care for deaf individuals is the lack of appropriate screening instruments and norms designed specifically for use with signing bilinguals (McKee et al., 2015; McKee & Paasche-Orlow, 2012). Currently it is common

practice to translate an existing screening tool into ASL (or another sign language), but this leads to the potential for methodological concerns that often come along with translating assessments from English (Peña, 2007). In fact, Dean et al. (2009) evaluated the reliability and validity of a common cognitive screener to demonstrate how the unique cultural and linguistic features of the older deaf population can influence results on a translated screener. Dean et al. (2009) recruited 117 deaf senior citizens without a history of head injury, depression, or dementia who use ASL as one of their primary languages. Participants completed a version of the Mini-Mental State Exam (MMSE; Folstein, Folstein, & McHugh, 1975) that was translated from English to ASL, with several items changed to make them more culturally and linguistically appropriate for the test population. Despite the fact that this group of deaf senior citizens had relatively high levels of education, the mean performance on the MMSE fell below the typical cutoff point for “normal” functioning. The authors present several explanations that may have led to artificially deflated scores in this population, particularly differences between ASL and English grammar that made direct translation of task instructions difficult, and the use of English specific idioms or vocabulary words that do not have ASL equivalents. The authors cautioned against using this screener with deaf populations because it did not account for cultural and linguistic differences which could ultimately lead to a higher rate of false positives for cognitive impairment or dementia (Dean et al., 2009).

Currently there is a gap in the understanding of cognitive and linguistic changes that occur with typical aging in populations of deaf bilinguals who use ASL and written English as their primary modes of communication. This situation is due in part to a lack of appropriate neuropsychological assessment tools suited for the unique cultural and linguistic circumstances of deaf signers. This dissertation aimed to develop and test ASL and written English versions of

three important measures that have been shown to be sensitive to Alzheimer's disease in hearing bilinguals: list memory, Stroop performance, and verbal fluency. Results from these experiments provide a preliminary overview of potential differences in cognitive and linguistic performance in older compared to younger ASL-English bilinguals. Additionally, an in-depth language background survey allowed us to explore the influence of age of acquisition, educational background, and language proficiency on performance on these measures. These are critical considerations for future clinical application of neuropsychological assessments in identifying cognitive decline, in particular Alzheimer's disease, in this population.

CHAPTER 2: Influence of Testing Language and Aging on Verbal List Memory in Deaf
American Sign Language-English Bilinguals

Sadie Camilliere¹, Karen Emmorey², Peter C. Hauser³, Jessica Contreras³, Michael M. McKee⁴,
and Tamar H. Gollan⁵

¹ Joint Doctoral Program in Language and Communicative Disorders, San Diego State
University & University of California, San Diego, CA USA

² School of Speech, Language, and Hearing Science, San Diego State University, San Diego, CA
USA

³ National Technical Institute for the Deaf, Rochester Institute of Technology, Rochester, NY
USA

⁴ Department of Family Medicine, University of Michigan, Ann Arbor, MI USA

⁵ Department of Psychiatry, University of California, San Diego, La Jolla, CA USA

Correspondence concerning this article should be addressed to Sadie Camilliere, Laboratory for
Language and Cognitive Neuroscience, 6495 Alvarado Rd Suite 200, San Diego, CA 92182

Email: sadiecamilliere@gmail.com

Abstract

Objective: The present study examined aging and testing-language effects on verbal list learning in young adult and older deaf bilinguals of American Sign Language (ASL) and written English. It is not known which language maximizes free recall, and no list learning task has been widely adopted for testing this population.

Method: Thirty-two younger (aged 20-45) and 32 older (aged 64-84) deaf ASL-English bilinguals completed list memory tests in each language. Participants were shown videos of 10 ASL signs and 10 written English words one at a time for immediate recall across three learning trials, and a delayed recall trial, with language of testing counterbalanced.

Results: Younger participants showed no effect of language on recall, recalled more items than older participants in both languages, and had higher primacy scores on Trial 1. Older participants showed better learning when tested in English but higher rates of forgetting compared to when they were tested in ASL (a robust interaction between language and group with a medium-to-large effect size, $\eta_p^2=0.11$). Both young and older participants forgot more items in whichever language was tested second. Though most deaf participants reported being more proficient in ASL, an equal number of deaf participants recalled more English than ASL list items as vice versa (i.e., more ASL items than English items).

Conclusions: Use of both ASL and English lists maximizes different aspects of memory performance in older deaf participants, with English benefitting learning and ASL minimizing forgetting. However, a complete characterization of memory requires testing in both languages.

Keywords: *Bimodal bilingualism, Deaf, List Memory, Serial Position, Aging, American Sign Language*

Key Points

Question: Tests of verbal learning and memory form a critical part of neuropsychological evaluations, but it is not known if ASL or English should be used for evaluation of deaf bilinguals, or how aging affects learning and free recall in older deaf signers.

Findings: Using ASL (rather than speech) to recall items in both English and ASL lists, older deaf signers achieved higher initial list learning of written English words than ASL signs but exhibited more rapid forgetting of English words compared to ASL signs, a pattern not seen in younger deaf signers.

Importance: List memory tests with deaf ASL-English bilinguals should be conducted in both languages, and signed recall of list items in both languages should be allowed to fully characterize performance on different aspects of memory functioning.

Next Steps: Additional research is needed to understand why older deaf signers exhibited better learning but more rapid forgetting of English than ASL lists, perhaps by testing learning and recall in young deaf participants under different learning conditions.

Introduction

Neuropsychological assessment of deaf individuals, for clinical and research purposes, presents several challenges that are unique to this population (Morere, 2013a). Most deaf individuals who became deaf early in life communicate primarily using a visual-manual language (e.g., a sign language like American Sign Language, ASL), and almost all deaf individuals are also *bimodal bilinguals*, which means they also know a spoken language. Some deaf signers may be able to lipread or use speech, but more typically they access their second language through the written modality (e.g., written English). Thus, although deaf adults use spoken language to varying degrees, they all access the dominant surrounding language through reading and writing. ASL, like other sign languages, does not have a conventional written form, and therefore deaf bilinguals only read and write in one of their languages – in contrast to most spoken language (unimodal) bilinguals. Currently, there are very few neuropsychological tests that were specifically designed for use with the deaf population, and it is not known how ASL-English bilingualism may affect performance on these tests (McKee et al., 2015; McKee & Paasche-Orlow, 2012). As a result, testing is either administered in English, which is not the dominant language for most deaf individuals (Grosjean, 2010), or testing is adapted on the fly into ASL (McKee & Paasche-Orlow, 2012). Translation of tests developed for other languages, especially on the fly, hampers test validity and interpretation (Peña, 2007).

There are several reasons to expect that deaf ASL-English bilinguals may perform differently on neuropsychological tasks created for hearing individuals. For example, experience with a sign language has been associated with subtle but significant changes in visuo-spatial processing, such as mental imagery (Emmorey et al., 1993, 1998; Kubicek & Quandt, 2021) and visual-spatial memory (Cattani et al., 2007; Hirshorn et al., 2013), which could impact task

performance. Further, modality-specific response options, such as fingerspelling (handshapes are used to represent letters in a word), need to be considered for tasks that require language production, such as verbal fluency (see Sehyr et al., 2018). It is not known how sign language experience or the unique linguistic options available to signers might affect many of the language and cognitive tasks that are commonly used in neuropsychological assessment.

Even less is known about how typical aging influences cognition and language in older deaf signers. Older signers may differ from younger signers due to shifts in deaf education in the USA over the last several decades, which has led to substantial differences in the quality of education for younger compared to older signers. Most older deaf adults were taught using an oral approach which emphasized spoken language and often discouraged the use of sign language, despite the fact that sign language is much more accessible for deaf individuals. Starting in the 1970's, recognition of this serious problem led to a shift towards a bilingual-bicultural teaching model (Moores, 2010). The importance of this shift to include sign language in deaf education is particularly critical for deaf individuals because 90% of deaf children are born to hearing parents who have no knowledge of sign language (Mitchell & Karchmer, 2004). As a result, deaf children often experience language deprivation due to delayed access to a first language (Humphries et al., 2016). Further, age of first language acquisition is quite variable across the deaf population and may also affect performance on cognitive and linguistic tests (e.g., Mayberry et al., 2002). Hearing individuals all acquire a first language from birth, except under very unusual circumstances (e.g., due to child abuse, as with the famous case of "Genie"; Curtiss, 2014). These differences in language modality, educational background, and age of acquisition of a first language position deaf signers as a unique group of bilinguals whose

performance on cognitive and linguistic tasks across the lifespan is likely to be different than hearing bilinguals. This unique population is in great need of investigation.

A prominent question concerning how to adjust cognitive assessments for deaf bilinguals is whether testing should be administered in ASL, in English, or in both languages. Spoken language bilinguals perform worse on language tests when they are tested in their less proficient or nondominant language, and bilinguals are not always able to accurately report which language is dominant (Gollan et al., 2023). These factors can impact the sensitivity of neurocognitive assessments as diagnostic instruments for identifying cognitive decline (Gollan et al., 2010). In particular, testing in the nondominant language leads to higher rates of false positive diagnosis (Gasquoine & Gonzalez, 2012). Crucially, almost nothing is known about the possible impact of language of testing on test performance for deaf bimodal bilinguals.

Tests of verbal list learning and memory are one of the most commonly administered tests in cognitive evaluations. Two previous studies attempted to develop a verbal list learning test for sign language users, and participants were only tested in sign language. The Signed Verbal Learning Test (SVLT; Morere, 2013b) was administered to young deaf signers who reported using ASL and English in their daily lives and were recruited from Gallaudet university ($M\text{ age} = 25.5$, $SD = 6.51$). The SVLT had two 16 item lists of ASL signs, with five learning trials for one list, and the second list served as an interference list (similar to the CVLT; Delis et al., 1987). Overall, deaf participants showed typical learning across the five immediate recall trials, and recall after a delay was consistent with scores from the normative sample of spoken language CVLT scores (as reported in Morere, 2012).

A second study by Denmark et al. (2016) created a verbal list learning and memory test for British Sign Language (BSL¹), which is the only current test developed for use with older deaf signers. The BSL Verbal List Memory Test (BSL-VLMT) included 12 BSL signs from three semantic categories, three immediate recall trials, a delayed recall trial, and a recognition task. Participants included 223 deaf signers aged 50-89, and 12 deaf individuals, aged 54-88, who received clinical diagnoses of dementia from a multidisciplinary panel based on clinical diagnostic criteria for dementia (American Psychiatric Association, 2000). In this study, age was the single factor that most influenced immediate recall, and delayed recall decreased with increasing age (Denmark et al., 2016). Participants in the dementia group scored significantly lower than the control group on both immediate and delayed recall. However, 34% of cognitively healthy participants performed very poorly on the delayed recall trial (i.e., recalled zero items on the delayed trial). Denmark et al. (2016) suggested that some controls may have been confused by the instructions or may have experienced interference from other linguistic tasks that were administered prior to the delayed recall trial. Thus, data from this study may not have accurately reflected memory abilities in this population.

To date no study has examined how language of testing and degree of bilingualism might affect learning and recall in deaf signers. Previous studies showed significant effects of language dominance on aspects of verbal list learning tests that are most sensitive to Alzheimer's disease (AD; Salmon & Bondi, 2009; Twamley et al., 2006), including list learning, delayed recall (Belleville et al., 2017) and primacy effects (Bayley et al., 2000; Bruno et al., 2015; Cunha et al., 2012; Kloth et al., 2020). Primacy effects refer to the tendency for better immediate free recall of

¹ ASL and BSL are mutually unintelligible languages, and thus this test could not be used with ASL-English bilinguals.

items presented at the beginning of a list (Murdock, 1962; Sumbly, 1963). Two studies of young bilinguals reported lower learning scores when bilinguals were tested in their nondominant language when compared to their dominant language, especially in the primacy region (Francis et al., 2020; Yoo & Kaushanskaya, 2016) (Francis et al., 2020; Yoo & Kaushanskaya, 2016). Another study (Gavino et al., submitted) investigated verbal list learning in 40 older Spanish-English bilinguals (mean age = 72.6, $SD = 7.4$) on 10-item word lists, with language of testing order counterbalanced between participants (as in the present study). Bilinguals recalled fewer words in their nondominant language than in their dominant language, but only in the first recall trial of each list. The disadvantage of being tested in the nondominant language was also especially strong in bilinguals for whom one language was much more proficient than the other. These results highlight the importance of understanding how language dominance effects may influence assessments of list memory in older participants.

The current study developed list memory tests in ASL and English and tested 79 deaf ASL-English bilingual adults (aged 20-84) on both tests. We aimed to assess the effects of age and testing language on verbal list learning and free recall. To make it feasible to assess memory in both languages, we modeled our verbal list learning tests on the CERAD (Consortium to Establish a Registry for Alzheimer's Disease; Welsh et al., 1994) list learning task, which is a relatively short test that recently demonstrated sensitivity to preclinical AD in older bilinguals (Mizrahi et al., 2024). For each language, we created a list of ten unrelated items (words or signs) which was presented and recalled over three learning trials with an additional recall trial after a short delay. Use of a relatively short list learning test (with just 10 items in each language), with semantically unrelated words both within and between lists may be especially

useful for testing participants in both languages and for examining primacy effects, which are stronger with short, semantically unrelated word lists (Ward et al., 2010; Bair et al., 2023).

Based on previous studies with spoken language bilinguals, we predicted that young participants would recall more items than older participants, and that primacy scores might be lower in older than in younger participants, but only on Trial 1 (which exhibits the stronger primacy effects than subsequent learning trials; Griffin et al., 2017). Additionally, we predicted participants might recall more list items in ASL than in English, perhaps especially for younger deaf participants who are more likely to have had consistent access to ASL from an early age, although for both young and older deaf individuals ASL is the preferred and more naturally accessible language for communication.

Methods

Participants

Eighty-two deaf ASL-English bilinguals participated in the current study and were tested in person in San Diego or Riverside, CA, or in Rochester, NY. Participants were recruited through “word of mouth” in the Deaf community, presentations at Deaf community events and venues (including Deaf senior citizen centers), and through flyers distributed at these events/venues. Three participants were excluded from analyses: one due to technical errors resulting in loss of video data and one due to self-reported misunderstanding of the task instructions. The third participant was excluded due to self-report of considerable difficulties with memory in activities of daily living, lack of sleep the night before testing, and some possible disorientation with respect to time of appointment. Participants self-reported their race and ethnicity in a background questionnaire². The race breakdown for the included participants

² Participants were asked two questions “*What is your race?*” and “*What is your ethnicity?*” with the following disclaimer: *While we recognize there are other labels for race and ethnic groups, these categories are determined*

was 86% white (68/79), 4% black (3/79), 3% Asian (2/79), 6% reported more than one race (5/79), and 1% declined to answer (1/79). The ethnicity breakdown was 11% Hispanic or Latino (9/79), 84% Not Hispanic or Latino (66/79), and 5% of participants reported that their ethnicity was unknown or declined to answer (4/79). All but 7 participants (91%) were prelingually deaf (i.e., became deaf at or before age two). The analyses reported below focused on the 32 oldest participants aged 64-84 ($n=21/32$ women) and the 32 youngest deaf participants that were tested, aged 20-45 ($n= 14/32$ women). Participant characteristics and performance of 15 middle-aged participants (aged 46-63, $n=7/15$ women) are shown in Tables 2.1 and 2.4, but these participants were excluded from the primary analyses (see Analysis Plan).

Table 2.1 summarizes participant characteristics. All participants reported daily use of ASL as a primary language. The older and younger groups had similar education levels and self-ratings of proficiency in ASL and English.

This research protocol was approved by the University of California, San Diego (UCSD) Institutional Review Board in accordance with the Helsinki Declaration. All participants were given informed consent and were given monetary compensation for participation.

by funding agencies and are used for reporting purposes.” The response options for race were: *Black, American Indian/Alaskan Native, Asian, Hawaiian/Pacific Islander, White, More than one race, Unknown/Decline to answer.* The response options for ethnicity were: *Hispanic or Latino, Not Hispanic or Latino, Unknown/Decline to answer.* Participants selected one response per question.

Table 2.1. Averages and standard deviations for participant characteristics and demographic information

Variable	Younger (<i>n</i> = 32)				Middle-aged (<i>n</i> = 15)				Older (<i>n</i> = 32)			
	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
Age***	34.2	6.7	20	45	55.6	5.6	46	51	70.5	4.7	64	84
Education (by highest degree)	16.2	2.8	12	21	15.1	2.4	12	21	15.8	2.9	12	21
Age of Acquisition of ASL**	3.5	5.7	0	23	3.8	5.3	0	23	9.1	10.5	0	51
Self-ratings of proficiency^a												
ASL production	6.5	0.8	5	7	6.7	0.6	5	7	6.5	0.8	4	7
ASL comprehension	6.7	0.5	5	7	6.8	0.4	6	7	6.4	1.1	2	7
ASL average	6.6	0.6	5	7	6.7	0.5	6	7	6.4	0.9	4	7
English writing†	6.1	0.8	4	7	5.7	1.0	4	7	5.7	1.1	2	7
English reading	5.8	0.9	4	7	5.3	1.2	4	7	5.5	1.1	3	7
Written English average	5.9	0.8	4	7	5.5	1.1	4	7	5.6	1.1	3	7
Spoken English production	3.0	2.2	1	7	2.5	2.0	1	7	3.0	1.9	1	7
Spoken English comprehension	2.8	2.0	1	7	2.5	1.7	1	7	3.1	1.9	1	7
Extent of oral communication^b												
Elementary School	4.6	2.0	1	7	4.7	2.0	1	7	4.0	2.2	1	7
Junior/High School	5.0	1.8	1	7	5.5	1.7	1	7	5.3	2.1	1	7

Note. *** $p \leq .001$, ** $p \leq .01$, * $p < .05$, † $p \leq .10$; p -values comparing younger signers to older signers (all other scores in these columns did not differ significantly, all $ps \geq .11$). ^aThe scale for self-ratings of proficiency ranged from 1-7 with the following descriptions of proficiency: 1=Almost none, 2=Very poor, 3=Fair, 4=Functional, 5=Good, 6=Very good, and 7=Like native. ^bThe scale for extent of oral communication ranged from 1-all oral to 4-equal use of signed/oral language to 7-all manual – participants rated how much of their communication was manual/sign or oral/spoken language for Elementary school and for Junior/High school

Procedure and Stimuli

This study was completed as part of a larger battery of cognitive and language tasks for deaf ASL-English participants. After completing a comprehensive language background questionnaire, participants completed three testing blocks. The first consisted of three tasks completed in English or ASL: a) list memory learning (trials 1-3), b) a computerized Stroop task (not reported here), and c) the list memory delayed recall task. Language of testing order (first, second) was counterbalanced across participants, such that they completed all three tasks first in English then all three tasks in ASL, or vice versa.

Participants completed list learning tasks in each language that were modeled after the CERAD Word List Learning and Recall tasks (Welsh et al., 1994). For each language, participants completed three learning trials, in which they were presented with 10 signs in ASL or 10 written words in English. The CERAD Word List Learning Test requires participants to

read aloud each printed word. However, this is not appropriate for deaf signers who do not use speech, and therefore participants were not asked to say each English word aloud. Both ASL and English lists were presented using timed videos. The 10 ASL signs were produced by a deaf native signer, with her hands returning to her side in between each sign in the list. The video presented one sign every two seconds. Printed English words were presented in 120 point Arial font on a screen at the same rate to match the duration of the sign productions (each word was presented for two seconds in a timed PowerPoint presentation). Following each presentation of the 10 item list, participants were prompted to recall as many of the items as possible (i.e., immediate recall). After the third learning trial, participants completed a computer administered Stroop task that lasted between 8-10 minutes and had no items that overlapped with the list memory stimuli. Following the Stroop task, they were asked again to recall as many of the 10 items from the immediately preceding learning trials as possible (i.e., they completed a delayed recall trial). Task instructions were presented in ASL without specifying which language should be used to recall items – participants were free to sign, speak, or fingerspell the recalled items. Participants were asked to recall as many items as they could in any order. Participants were tested in person, and all responses were videotaped for later scoring.

The English stimuli were taken from the CERAD Word List Learning Test with some adaptations for deaf ASL-English bilinguals. To ensure that all English items could be easily produced in ASL during recall (thus supporting the choice of signed rather than spoken responses for English trials), items from the original CERAD list that did not have ASL translation equivalents were replaced. The words *shore*, *cabin*, and *pole* are not represented by a single ASL sign and would either need to be fingerspelled or expressed using a classifier construction, which takes more time than producing a single word or sign. These three items were replaced with

words matched for frequency and roughly matched in length relative to the original CERAD items according to the Subtitle-Based Word Frequency Estimates for English (SUBTLEX-US; Brysbaert and New, 2009).

The ASL stimuli were selected to match the English CERAD word list on English word frequency, while controlling for several lexical characteristics of the ASL signs, including sign frequency, iconicity, and number of two-handed signs. ASL lexical characteristics were taken from ASL-Lex 2.0 (Sehyr et al., 2021). English word frequency was taken from SUBTLEX-US (Brysbaert and New, 2009). Lexical characteristics for ASL and English stimuli are shown in Table 2.2.

Table 2.2. Lexical characteristics of ASL and English stimuli in the list memory tests

	English Frequency ^a <i>M (SD)</i>	ASL Frequency ^b <i>M (SD)</i>	Iconicity ^c <i>M (SD)</i>	# of two-handed signs
ASL	3.0 (0.4)	4.2 (0.6)	3.6 (2.0)	5/10
English	3.3 (0.3)	4.3 (0.8)	2.9 (1.8)	5/10

Note. ^a English frequency ratings are taken from Brysbaert and New (2009) and reflect log frequency per million.

^b ASL frequency ratings taken from ASL-Lex 2.0 (Sehyr et al., 2021) and reflect deaf signers' judgments of how frequently a sign appears in everyday conversation (1= very infrequently, 7= very frequently).

^c Iconicity ratings taken from ASL-Lex 2.0 (Sehyr et al., 2021) reflect hearing non-signers' judgments of how iconic an ASL sign appears given its English translation (1=not iconic at all, 7=very iconic). Iconicity ratings for English reflect ratings for the ASL translations of the English words.

Items were presented in a different pseudo-randomized order on each learning-recall trial in each language (see Table 2.3). The order of items in the three English lists was determined by the order of the original CERAD test, which presents the words in a different order on each learning trial. To obtain the ASL order, the ASL items were arranged in a pseudo-random order and then each ASL item was paired with an English item to keep the order of appearance of different items in the three recall trials in each language consistent across languages. The pseudo-random order of ASL items was needed to avoid phonological similarity between adjacent signs

in any recall trial (e.g., the signs UNCLE and MIRROR share similar locations, while MOTORCYCLE and TOWEL share the same handshake).³ ASL and English list item orders on each recall trial are shown in Table 2.3, and recordings of each list with the presentation timings can be found at the following link:

https://osf.io/bj8yz/overview?view_only=9f441c0c41404b10a414b3797ffceb29.

Transparency and Openness

We follow the Journal Article Reporting Standards (Kazak, 2018) in reporting all data exclusions, manipulations, and measures in this study. All data, analysis code, and research materials are available at:

https://osf.io/bj8yz/overview?view_only=9f441c0c41404b10a414b3797ffceb29. Data were analyzed using R (R Core Team, 2023) and the R packages *dplyr* (Wickham, François, et al., 2023), *ggplot2* (Wickham, 2016), *ggpubr* (Kassambara, 2023a), *rstatix* (Kassambara, 2023b), and *tidyr* (Wickham, Vaughan, et al, 2023). This study’s design and analysis were not pre-registered.

³ Hyperlinks are to videos from the ASL-LEX database.

Table 2.3. Order of items in ASL and English lists on Trials 1-3

Position in List	ASL Trial 1	ASL Trial 2	ASL Trial 3
1	<u>hearing_aid</u>	certificate	prison
2	<u>farm</u>	flower	mirror
3	<u>umbrella</u>	hearing_aid	farm
4	<u>uncle</u>	umbrella	flower
5	<u>prison</u>	motorcycle	towel
6	<u>flower</u>	farm	umbrella
7	<u>towel</u>	prison	hearing_aid
8	<u>certificate</u>	uncle	motorcycle
9	<u>mirror</u>	towel	certificate
10	<u>motorcycle</u>	mirror	uncle
	English Trial 1	English Trial 2	English Trial 3
1	butter	ticket	queen
2	arm	hotel	grass
3	star	butter	arm
4	letter	star	hotel
5	queen	engine	flag
6	hotel	arm	star
7	flag	queen	butter
8	ticket	letter	engine
9	grass	flag	ticket
10	engine	grass	letter

Results

Participant characteristics

Participants completed a questionnaire about their language background in which they were asked questions about language use and language use history, and they were asked to rate their proficiency level for signing ASL, understanding ASL, writing in English, reading in English, speaking English, and understanding spoken English. Average self-rated proficiency scores for ASL, written English, and spoken English production and comprehension are given in Table 2.1. In addition, overall average self-rating scores for each language were calculated by averaging self-rated proficiency for production and comprehension of ASL, and for writing and reading of English. An averaged measure provides a stronger predictor of objectively measured

proficiency than ratings of each mode of communication on its own (Neveu et al., 2025).

Overall, both young and older participants had similar self-rated proficiency levels in ASL and written English. Based on the average self-rated proficiency in each language, we found that 18/32 young (56%) and 21/32 older (66%) participants rated ASL as more proficient than English, 10/32 young (31%) and 7/32 older (22%) participants had identical self-rating scores in ASL and English, and a small number of participants – 4/32 young (13%) and 4/32 older (13%) – rated English as more proficient than ASL. We note, however, that all participants preferred to use ASL to communicate with experimenters.

Older and younger groups differed significantly in their self-reported age of acquisition (AoA) of ASL, with older participants reporting later and more variable AoA than younger participants (see Table 2.1). Nonetheless, a similar proportion of young and older participants grew up with at least one deaf parent: in the younger group, 12/32 (38%) and in the older group, 10/32 (31%). Concomitantly, in both the younger group [20/32 (62%)] and in the older group [22/32 (69%)] most participants grew up with two hearing parents. The majority of participants (48/62; 75%; 24 young, 24 older) reported having profound hearing loss (i.e., > 90 dB HL); 12/64 (7 young, 5 older) reported severe hearing loss (i.e., 71 to 90 dB HL); 3/64 (1 young, 2 older) reported moderate hearing loss (i.e., 41-55 dB HL); and one older participant reported mild hearing loss (26 to 40 dB HL).

Response mode

For both English and ASL lists, all participants chose to recall items in ASL. Half of the participants (37/64; 20 older, 17 younger) used only ASL signs and half (27/64; 53%; 12 older, 15 younger) used a combination of signs and fingerspelling. One younger participant spoke English words aloud while producing ASL signs. All participants produced at least some English

mouthings while signing, i.e., they produced (unvoiced) mouth movements that resembled the articulation of the English translations of the ASL signs. On average, participants produced mouthings for 94% of items recalled. The use of ASL during recall was not related to self-rated language dominance, i.e., the eight participants (four young; four older signers) who indicated stronger English than ASL proficiency also chose to respond with signs and/or fingerspelling rather than spoken words.

In addition, during recall of both language lists, 17/32 (53%) of young signers and 23/32 (72%) of the older signers spontaneously used the ASL list buoy construction (Liddell, 2003). List buoys are number signs that range from “first” (only the thumb extended) to “five” (all fingers extended) that are produced by the non-dominant hand and can represent ordered entities (e.g., the birth order of siblings). During recall, the list buoy on the non-dominant hand was held in space while the index finger of the dominant hand touched a digit to indicate the number of the list item (e.g., first – thumb, second – index finger, third – middle finger, etc.), either before or after signing the item. To examine whether the use of the list buoy affected recall, we correlated buoy use with performance on each recall trial, on total immediate recall scores (trials 1-3), and on the delayed recall trial. None of these correlations were significant (all $r_s \leq 0.19$, all $p \geq 0.12$).

Memory Performance Analysis Plan

To examine the effects of age group and language on free recall of ASL and English lists we conducted a series of 2x2 ANOVAs crossing age group and language as categorical factors with age of acquisition (AoA) as a continuous covariate, since young and older groups differed significantly in self-reported AoA ($t(62) = -2.68, p = .01$). For all analyses, AoA was not significant as a main effect, and did not interact significantly with any other factor (all $p_s > 0.14$), and

therefore is not reported in detail below. Age groups did not differ significantly in education level ($t < 1$), or in distribution across five categories of education level including 12, 14, 16, 18, or 21 years, ($\chi^2(4,64) = 1.96, p = .743$) and also did not differ significantly in distribution across female/male sex ($\chi^2(1,64) = 3.09, p = .079$). We focused primarily on analyses with age group as a categorical factor to simplify presentation, and to allow for the possibility of nonlinear age effects (without making assumptions about the nature of the relationship between age and performance). However, we repeated each of our primary analyses (Overall Learning, Items recalled in Trial 3 vs. Delay, Adjusted Forgetting, and Primacy Effects) with age as a continuous factor instead of as a group contrast (including all participants whether young, middle-aged, or older). These analyses produced the same effects as the analyses with age as a categorical factor (see Appendix).

First, we examined overall learning (sum of all items recalled across the three learning trials and after a delay). Second, we examined forgetting first by comparing the number of items recalled in each language on Trial 3 versus on the delay trial, and then by calculating ‘proportion forgotten’ scores which controls for initial learning (Petersen et al., 1992). Third, we examined primacy effects by assessing memory for the first three items on Trial 1 for each language. Fourth, we asked if participants’ self-ratings of relative level of proficiency in each language was correlated with relative ability to remember items in each language. Finally, we added language of testing order as a factor to all three analyses to determine if any of the observed effects were modulated by testing order. Means and standard deviations of younger, middle, and older deaf signers on the ASL and English List Learning Tests are shown in Table 2.4.

Table 2.4. Means (and SDs) of younger, middle-aged, and older deaf signers on the list memory tests

	Younger (<i>n</i> = 32)		Middle (<i>n</i> = 15)		Older (<i>n</i> = 32)	
Variable	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
ASL						
Trial 1 correct	4.5	1.4	3.8	1.4	3.4	1.4
Trial 2 correct	6.4	1.5	6.6	1.1	5.6	1.5
Trial 3 correct	7.4	1.4	6.5	1.3	6.1	1.9
Delay correct	6.4	1.6	5.5	1.4	5.4	2.0
Trial 1-3 Correct	18.3	3.5	16.9	3.2	15.1	3.9
Proportion forgotten ^a	0.1	0.2	0.1	0.2	0.1	0.3
Trial 1 primacy ^b	1.7	0.7	1.6	0.6	1.3	0.9
Trial 1 middle	1.5	1.2	1.1	0.9	0.9	1.0
Trial 1 recency	1.3	0.8	1.1	0.8	1.2	0.8
English						
Trial 1 correct	4.6	1.5	3.8	1.3	3.6	1.7
Trial 2 correct	6.7	1.6	6.2	1.6	5.7	1.8
Trial 3 correct	7.5	1.4	7.4	1.5	7.0	1.9
Delay correct	6.5	1.6	5.3	2.3	4.8	2.8
Trial 1-3 Correct	18.8	3.8	17.4	3.5	16.3	4.7
Proportion forgotten ^a	0.1	0.2	0.3	0.3	0.3	0.3
Trial 1 primacy	2.1	0.8	1.3	0.8	1.5	1.2
Trial 1 middle	1.4	1.0	1.3	0.9	0.9	0.9
Trial 1 recency	1.1	0.8	1.2	1.1	1.2	0.8

Note. ^aProportion forgotten scores were calculated to examine delayed recall while controlling for initial learning; scores were calculated by subtracting the number of items recalled on the Delay trial from the number of items recalled on Trial 3, and then dividing by items recalled on Trial 3. ^bPrimacy scores refer to the number of items recalled from the initial three positions on Trial 1 (items 1-3); middle scores refer to the number of items recalled from the middle four positions of Trial 1 (items 4-7); recency scores refer to number of items recalled from the final three positions of Trial 1 (items 8-10). Recency scores did not differ significantly between age groups, all $ps \geq .11$.

Learning

To examine the effects of testing language (English, ASL) and age group (older, younger) on overall learning, we ran a 2 x 2 ANOVA with total items recalled when combining all three learning trials in each language as the dependent variable (i.e. the sum of correct responses from Trials 1-3), and AoA as a continuous covariate. This analysis revealed that younger participants recalled more items than older participants, $F(1, 61) = 10.97$; $\eta_p^2 = 0.15$; $p < .01$. Additionally, participants recalled the same number of items in both languages, though with a tendency

towards better learning in English than ASL overall, $F(1, 61) = 3.88$; $\eta_p^2 = 0.06$; $p = 0.05$. Age effects were equivalent in the two languages, there was no interaction between age group and language $F(1, 61) < 1$; $\eta_p^2 = 0.001$; $p = 0.78$.

Delayed Recall and Forgetting

Trial 3 vs. Delayed Recall. Figure 2.1 shows recall scores on Trial 3 versus after a delay (see also Table 2.4). To investigate the effects of testing language and age group on forgetting, we ran a 2 x 2 x 2 ANOVA with language (English, ASL) and trial (Trial 3, Delay Trial) as within subjects factors, and age group (younger, older) as a between subjects factor, and number of correct responses recalled as the dependent variable, and AoA as a continuous covariate. This analysis revealed that younger participants recalled more items than older participants, a main effect of age group $F(1, 61) = 12.50$; $\eta_p^2 = 0.17$; $p < .001$, and participants recalled fewer items after the delay compared to on Trial 3, a main effect of trial $F(1, 61) = 64.00$; $\eta_p^2 = 0.51$; $p < .001$. However, in older participants the effect of language varied with trial such that they recalled more items in English than in ASL on Trial 3 but more items in ASL than in English after the delay (see Figure 2.1). Overall the 2-way interaction between language and trial was significant, $F(1, 61) = 6.12$; $\eta_p^2 = 0.09$; $p = 0.02$, but interpretation of the 2-way interaction must be considered only in light of a more complex 3-way interaction between age group, language, and trial $F(1, 61) = 7.31$; $\eta_p^2 = 0.11$ $p = 0.01$.

To investigate the nature of the three-way interaction, follow up 2 x 2 ANOVAS were conducted on the younger and older participants separately. Younger participants recalled fewer items after the delay relative to Trial 3, a main effect of trial $F(1, 30) = 26.35$ $\eta_p^2 = 0.47$; $p < .001$, recalled equivalent numbers of items in both languages, no main effect of language $F(1,$

30) < 1 ; $\eta_p^2 = 0.01$; $p = 0.63$, and the effect of delay on recall was equivalent in both languages, there was no interaction between language and trial $F(1, 30) < 1$; $\eta_p^2 < 0.001$ $p = 0.93$.

Older participants also recalled fewer items after the delay, a main effect of trial $F(1, 30) = 36.69$ $\eta_p^2 = 0.55$; $p < .001$, and recalled equivalent numbers of items in both languages overall, there was no main effect of language $F(1, 30) < 1$ $\eta_p^2 = 0.01$; $p = 0.64$. However, older participants recalled significantly more items in English than ASL on Trial 3 ($t(31) = -2.64$, $p = .01$), but *fewer* items in English than in ASL on the delayed recall trial. That is, the effect of language reversed on the delayed recall trial. Although this reversal effect was not significant, $t(31) = 1.09$, $p = .28$, the cross-over interaction between language and trial was robust $F(1, 30) = 9.13$ $\eta_p^2 = 0.23$; $p = 0.005$.

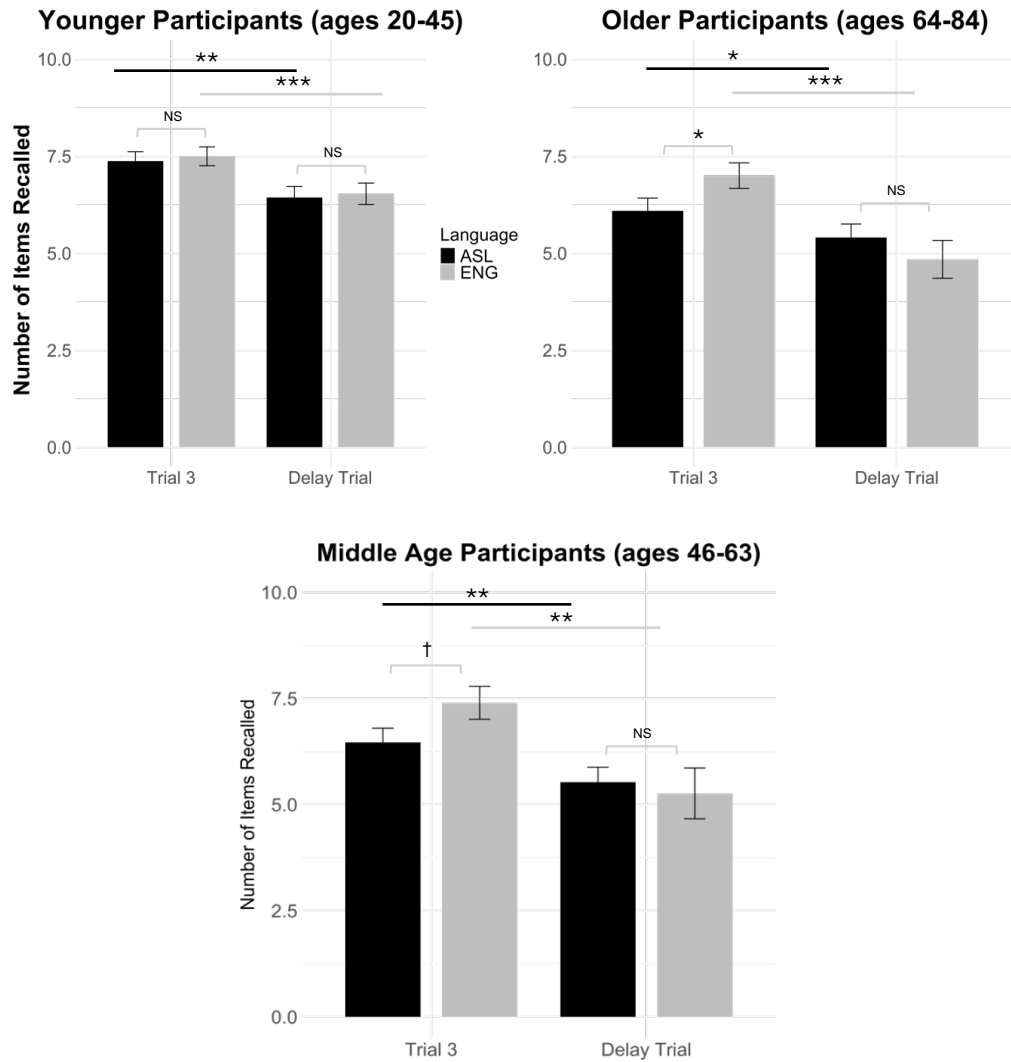


Figure 2.1. Mean number of items recalled in ASL and English on Trial 3 and after a Delay

Note. Younger (left panel), middle-aged (bottom panel), and older (right panel) participants are shown. Error bars denote standard error. *** $p \leq .001$, ** $p \leq .01$, * $p < .05$, † $p \leq .10$

Baseline Adjusted Forgetting. Additionally, to investigate the rate of forgetting after controlling for initial learning we calculated proportion forgotten scores for each participant by subtracting the number of items recalled on the Delay trial from the number of items recalled on Trial 3, and then dividing by items recalled on Trial 3. Adjusted scores for older and younger participants are shown in Table 2.4 and illustrated in Figure 2.2. To investigate the effects of testing language and age group on adjusted forgetting scores, we ran a 2 x 2 ANOVA with

language (English, ASL) as a within subjects factor, and age group (older, younger) as a between subjects factor, and adjusted forgetting scores as the dependent variable. This analysis revealed that older participants forgot more items than younger participants, a main effect of age group $F(1, 61) = 5.17$; $\eta_p^2 = 0.08$; $p = .03$, and participants forgot more in English than in ASL, a main effect of language $F(1, 61) = 6.66$; $\eta_p^2 = 0.10$; $p = 0.01$. However, these main effects were driven by the performance of the older participants who forgot English items at a much higher rate compared to younger participants ($t(62) = -3.19$, $p = 0.002$), but the rate of forgetting ASL items did not differ significantly between younger and older participants ($t(62) = 0.32$, $p = 0.75$), a significant 2-way interaction between language and age group $F(1, 61) = 7.65$ $\eta_p^2 = 0.11$; $p = 0.01$.

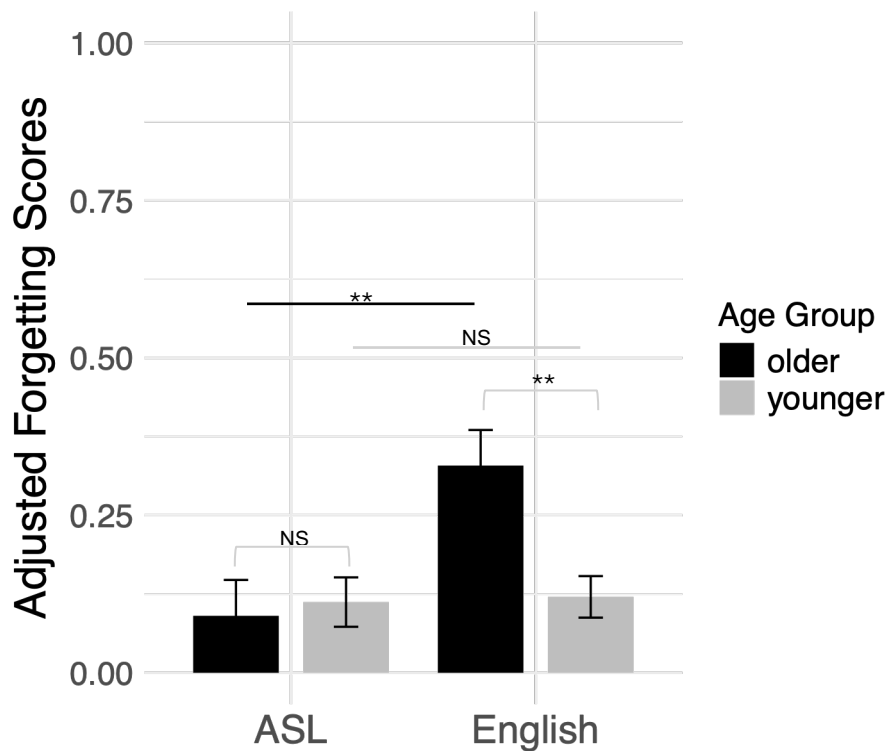


Figure 2.2. Learning-adjusted forgetting scores in each language

Note. Learning adjusted scores were calculated as [number correct on (trial 3 – delay)/Trial 3]. Error bars denote standard error. *** $p \leq .001$, ** $p \leq .01$, * $p < .05$, † $p \leq .10$.

Primacy Effects

Figure 2.3 shows serial position curves in each language for older and younger participants. Primacy scores are shown in Table 2.4 and were calculated as the total number of items recalled from the first three items on the first recall trial (Trial 1; Cunha et al., 2012; Mizrahi et al., 2023). To investigate aging effects on primacy scores, we conducted a 2 x 2 ANOVA with language (English, ASL) as a within subjects factor, and age group (older, younger) as a between subjects factor with primacy scores as the dependent variable. These analyses revealed that older adults recalled fewer primacy items than younger adults, a main effect of age group $F(1, 61) = 7.63$; $\eta_p^2 = 0.11$; $p = 0.01$, primacy scores were equivalent in ASL and English, no main effect of language $F(1, 61) = 3.22$; $\eta_p^2 = 0.05$; $p = 0.08$, and primacy scores were suppressed in older adults to a similar extent in both languages, there was no interaction between age group and language $F(1, 61) < 1$; $\eta_p^2 = 0.01$; $p = 0.55$.

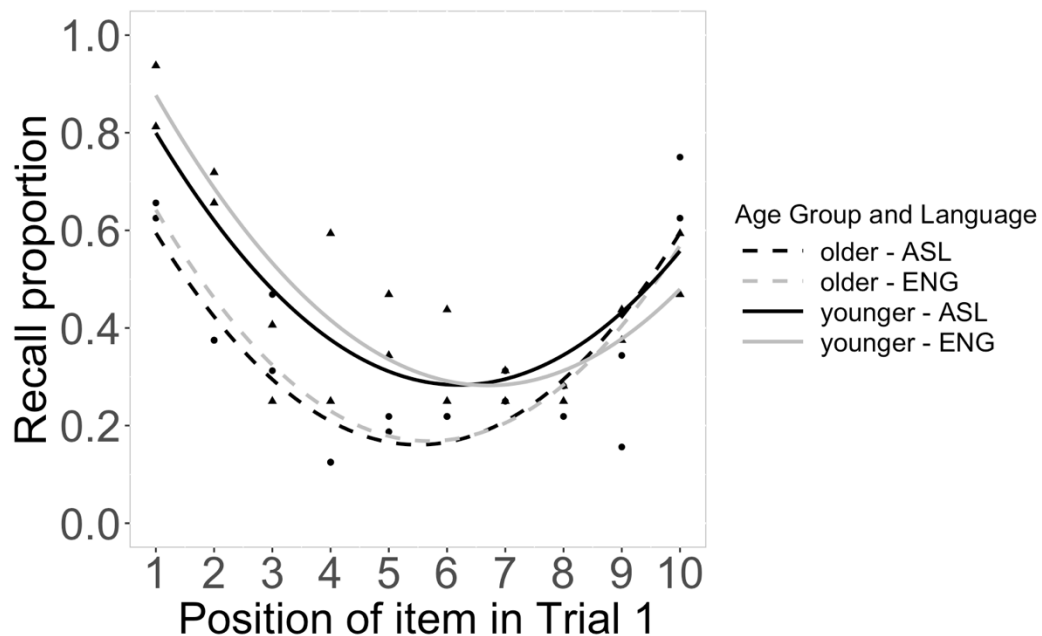


Figure 2.3. Serial position curves showing the proportion of ASL and English items recalled for younger and older participants

Note. Triangles represent younger participants, and dots represent older participants.

Language Proficiency

A common assumption is that bilinguals will score higher in whichever language they report is more proficient. To consider how self-reported language proficiency affected recall scores in deaf participants, we calculated language dominance scores (ASL minus English) for self-rated language proficiency. We used the ASL average and written English average for each person (see Table 2.1; Neveu et al., 2025). We then correlated language proficiency difference scores in each age group with recall accuracy difference scores for each trial (see Table 2.5). In young participants self-rated dominance in proficiency scores were not correlated with differential ability to recall items in ASL versus in English. In older participants, Trials 1 and 2 exhibited moderate to small correlations such that participants who rated their ASL proficiency as higher than their English proficiency tended to recall more items in ASL than in English. However, by Trial 3 and in the Delay recall trial, these associations were no longer significant.

Figure 2.4 illustrates the relationship between self-rated difference in language proficiency and difference in ability to recall ASL versus English on Trial 1 and the Delayed recall trial. Participants who indicated they were more proficient in ASL than in English and recalled more items in ASL than in English are shown in the upper right quadrant, while participants who indicated they were more proficient in English and recalled more items in English than in ASL are shown in the bottom left quadrant. Notably, many participants, both younger and older, recalled more items in the language they rated as *less proficient* (all dots in the upper left, and bottom right quadrants). This pattern may have occurred because participants in this sample had relatively balanced language proficiency, though self-ratings did range from ~40% more English than ASL proficient to ~40% more ASL than English proficient (on the x-axis). On Trial 1 memory scores exhibited a similar range of recall dominance scores (about 40%

better in ASL or English), but by the delayed recall trial older participants (the red dots) exhibited a broader range of recall difference scores (with several at or approaching 60% better in one language than the other). Finally, note that most participants rated themselves as more proficient in ASL than in English (most dots are to the right of the 0 point on the vertical y-axis), but equal numbers of participants learned and recalled more English words than ASL signs as vice versa (many dots are below the 0 point on the horizontal x-axis).

Table 2.5. Pearson bivariate correlations between language dominance scores (ASL minus English) for self-rated proficiency and number of items recalled on Trials 1-3 and the delayed recall trial

	Trial 1	Trial 2	Trial 3	Delay
Older	0.40*	0.35 [†]	-0.10	0.21
Younger	0.05	-0.25	0.01	-0.23

*** $p \leq .001$, ** $p \leq .01$, * $p < .05$, [†] $p \leq .10$.

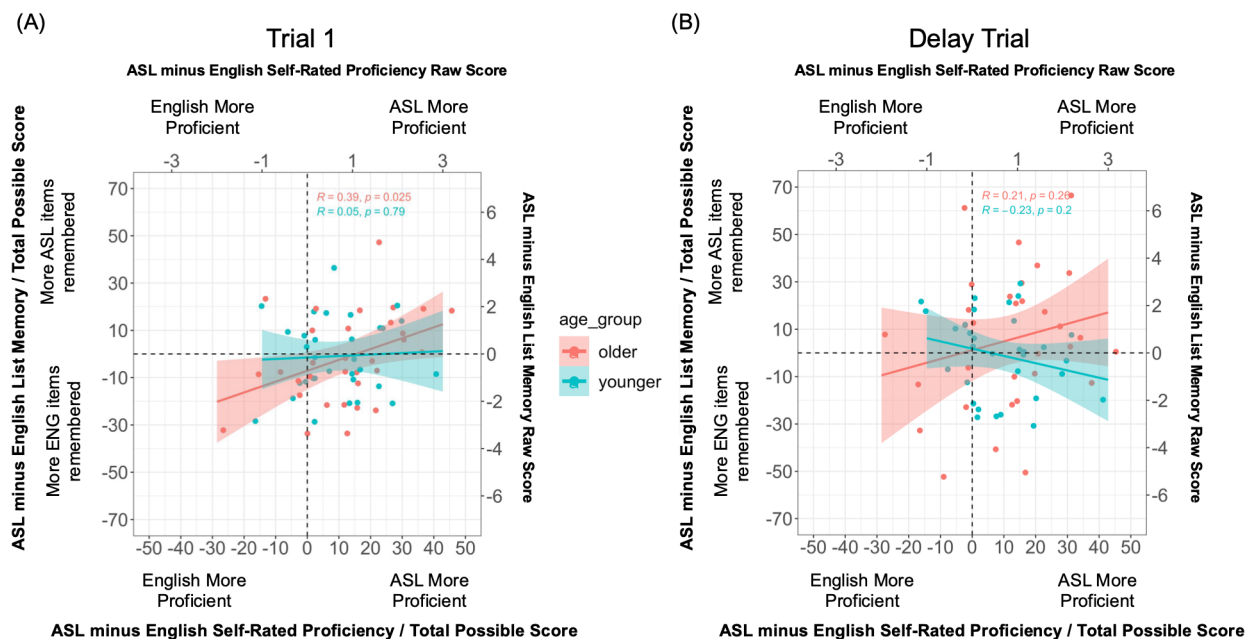


Figure 2.4. Correlations between language proficiency and list memory recall

Note. The figure shows the correlations between the difference in self-rated proficiency level (ASL minus English, maximum difference = 7; shown on x-axis) and ASL minus English list recall scores (shown on y-axis, maximum difference = 10). Raw difference scores between languages are shown on the top and right axes, while on the bottom and left axes, scores were adjusted proportionally relative to maximum possible difference scores to facilitate comparison of differences in self-ratings versus differences in recall scores on the same scale. The left panel shows recall scores from Trial 1 recalled in each language, and the right panel shows recall scores from the delay trial in each language.

Language of Testing Order effects

Language of testing order was counterbalanced across participants. Of the 64 participants included in these analyses, 30 participants (16 young, 14 older) were tested on English lists first and 34 (16 young, 18 older) were tested on the ASL lists first. Descriptive statistics for participants who responded in each order are shown in Table 2.6. To investigate whether results varied with language of testing order, we added language order as a between subjects factor to each of the ANOVA analyses reported above.

Language of testing order did not affect overall learning or primacy scores. When testing order was included in the analysis of overall learning scores, there was no main effect of order

$F(1, 59) < 1$; $\eta_p^2 = 0.004$; $p = 0.63$, and no interactions between order and age group $F(1, 59) < 1$; $\eta_p^2 < 0.001$; $p = 0.95$, order and language $F(1, 59) = 1.59$; $\eta_p^2 = 0.03$; $p = 0.21$, or age group, order, and language $F(1, 59) < 1$; $\eta_p^2 < 0.001$; $p = 0.82$.

When order was added to the analysis of primacy effects, there was no main effect of testing order $F(1, 59) = 1.87$; $\eta_p^2 = 0.03$; $p = 0.18$, and no interactions between order and age group $F(1, 59) < 1$; $\eta_p^2 = 0.01$; $p = 0.40$, order and language $F(1, 59) = 2.65$; $\eta_p^2 = 0.04$; $p = 0.11$, or age group, order, and language $F(1, 59) < 1$; $\eta_p^2 < 0.001$; $p = 0.94$.

With respect to forgetting, there was no main effect of order $F < 1$, no 2-way interactions of age group and order, language and order, trial and order, all $F_s < 1$. The 3-way interaction between age, order, and language, was also not significant $F < 1$, and there was no significant interaction between age, order and trial, $F(1, 59) = 1.47$; $\eta_p^2 = 0.02$; $p = 0.23$. Of greater interest, the 3-way interaction that was already reported above between age, language, and trial remained significant with testing order added to the model, $F(1, 59) = 5.90$; $\eta_p^2 = 0.09$; $p = 0.02$. Additionally, participants forgot more items in whichever language was tested second, a significant three-way interaction between testing order, language, and trial $F(1, 59) = 11.14$; $\eta_p^2 = 0.16$; $p = 0.001$.

To follow-up on this interaction, we conducted two 2 x 2 ANOVAs with order as a between subjects factor, trial as a within subject factor, and items recalled as the dependent variable, for ASL and English items separately. In ASL, participants remembered a similar number of signs overall regardless of testing order, no main effect of order $F(1, 61) < 1$; $\eta_p^2 < 0.001$; $p = 0.77$, and they remembered more signs on Trial 3 than after a delay, a main effect of trial $F(1, 61) = 15.35$; $\eta_p^2 = 0.20$; $p < 0.001$. However, participants who were tested on the ASL list first, forgot fewer ASL items after a delay than participants who were tested on the English

list first, a significant interaction between order and trial $F(1, 61) = 6.55$; $\eta_p^2 = 0.10$; $p = 0.01$. A similar pattern was seen on recall of English words. Participants remembered a similar number of English words overall regardless of testing order, no main effect of order $F(1, 61) < 1$; $\eta_p^2 = 0.01$; $p = 0.51$, and they remembered more English words on Trial 3 than after a delay, a main effect of trial $F(1, 61) = 31.97$; $\eta_p^2 = 0.34$; $p < 0.001$. However, participants who were tested on the English list first, forgot fewer English items after a delay than participants who were tested on the ASL list first, a significant interaction between order and trial $F(1, 61) = 6.55$; $\eta_p^2 = 0.10$; $p = 0.013$. Thus, in both ASL and English, the number of items forgotten was larger in whichever language was tested second.

Table 2.6. Means and standard deviations for recall scores on Trial 3 and the delayed trials by language and presentation order

<i>Language</i>	Participants tested in ASL 1st (<i>n</i> = 35)			Participants tested in English 1st (<i>n</i> = 29)		
	Trial 3 M (<i>SD</i>)	Delay Trial M (<i>SD</i>)	Items Forgotten M (<i>SD</i>)	Trial 3 M (<i>SD</i>)	Delay Trial M (<i>SD</i>)	Items Forgotten M (<i>SD</i>)
ASL	6.5 (1.8)	6.1 (1.8)	0.4 (1.4)	7.1 (1.6)	5.7 (2.0)	1.3 (1.8)
English	7.4 (1.5)	5.3 (2.4)	2.1 (2.0)	7.1 (1.9)	6.2 (2.3)	0.9 (1.4)

Note. The Items Forgotten column shows difference scores (Trial 3 minus Delay Trial within each language).

Discussion

Summary of Results

The current study investigated how testing language and age impacted recall performance on a word list learning task in 32 young and 32 older deaf ASL-English bilinguals. All

participants preferred to respond in ASL for recall of both ASL and English lists, and all produced at least some English mouthing while signing. Some participants used list buoys during recall, but use of this linguistic device did not impact recall performance. Younger participants recalled more items than older participants across all three learning trials, and on the delayed recall trial in both ASL and English. Young participants showed equivalent learning (Trial 3 scores) and equivalent remembering/forgetting (Delayed recall scores) in both languages. Older participants exhibited a more complex pattern in which English benefitted initial learning (Trial 3 scores were higher in English than ASL), but English exhibited more rapid forgetting than ASL so that delayed recall scores were equivalent across languages (and were numerically higher in ASL than in English; see Figures 2.1-2.2). Older participants also showed reduced primacy scores relative to younger participants on Trial 1 in both languages (see Figure 2.3). In young participants there was no significant relationship between self-reported degree of difference in proficiency level between languages and relative ability to recall items in each language. By contrast, in older participants, the difference in self-rated proficiency between languages was moderately correlated with extent of ability to recall more items in the language that was rated as more proficient on Trial 1 and weakly correlated on Trial 2. However, by Trial 3 and on the delayed recall trial recall performance was no longer significantly associated with self-rated proficiency level in each language (see Figure 2.4). Lastly, testing order (i.e., which language was tested first) did not impact total learning or primacy scores, but both young and older participants forgot more items in whichever language was tested second.

How does aging affect free recall in deaf ASL-English bilinguals?

The finding that younger participants recalled more items than older participants is in line with previous studies which demonstrated aging-related decline in overall free recall of items

from word lists (Craik, 1991; Petersen et al., 1992). Interestingly, when participants were tested in ASL, aging effects matched the results of Petersen et al. (1992) who found that older participants (monolingual English speakers) recalled fewer items than young participants after a delay - but critically – *there was no aging-related decline in delayed recall* after adjusting recall scores for amount of material that was learned initially. This is precisely the pattern that we observed for older and younger deaf bilinguals when testing was in ASL. By contrast, testing in English increased learning scores for older deaf participants initially (by Trial 3), but this benefit did not persist over time, so that forgetting in English was proportionally larger in older than in young participants on the delayed recall trial. Since rapid forgetting is the hallmark of Alzheimer’s disease (Karrasch et al., 2005), the most commonly diagnosed form of dementia (Jorm & Jolley, 1998; Morris, 1996), the results of the present study could suggest that testing in English increases the chance of false positive diagnosis in older deaf ASL-English participants. An alternative possibility is that learning is stronger in English, while recall is similar across languages, for older participants. It is possible that high rates of English item learning by older participants leads to exaggerated effects of forgetting in this sample. Additional research will be needed to distinguish between these alternatives.

Primacy scores on Trial 1, which provide another potential marker of dementia (Bayley et al., 2000; Kloth et al., 2020), were also lower in older than in young participants. This aging effect did not significantly interact with language of testing. The relatively weak effects of language on primacy scores also seems consistent with a number of recent studies. For example, a study of free recall on 16 item lists in young hearing native ASL-English bilinguals (i.e., Codas: Children of Deaf Adults) found equivalent primacy and recency effects in ASL and English (Bavelier et al., 2008). In another more recent study of older Spanish-English bilinguals,

primacy scores were only marginally lower in the nondominant than the dominant language (Gavino et al., submitted). Aging may affect primacy scores even less when participants recall items from lists with semantically related subgroups of items (Griffin et al., 2017). Additional research is needed to determine how primacy scores are affected by healthy aging in both hearing and deaf bimodal bilinguals.

A question that arises is why language of testing did not influence recall in younger participants, but in older participants English testing seemed to benefit learning. It is possible that older participants benefited from speech-based encoding for immediate learning. In hearing native ASL-English bilinguals, Hall & Bavelier (2011) examined the effects of language on encoding and immediate serial recall in digit span tests and found a higher span when encoding was speech-based as opposed to sign-based, i.e., when digits were presented in spoken (audiovisual) English versus in ASL. Similarly, Sehyr et al. (2017) found evidence that deaf signers recode both printed words and fingerspelled words into a speech-based code when completing immediate serial recall tasks. Specifically, participants exhibited a speech-based similarity effect for both stimulus types, i.e., word lists with phonologically similar words (e.g., blue, two, who) were remembered less well than lists of dissimilar words (e.g., king, tax, bug) (cf. Baddeley, 1986). Though these two studies investigated serial as opposed to free recall, it is possible that the older deaf participants in the current study benefitted from speech-based encoding immediately following presentation of printed English words. However, it is not clear why older participants would benefit more than younger participants from initial speech-based encoding, particularly since the two groups reported similar amounts of oral (speech-based) education (see Table 2.1), nor is it clear why this benefit would occur with immediate but not delayed recall.

Another reason why English might have benefited learning is that participants translated each word into ASL and preferred to recall items in ASL even when lists were presented in written English. That is, they used ASL to recall items from both the ASL and English lists. This response choice for English lists may not be surprising given that many deaf people prefer not to speak, and fingerspelling the English words would be much slower than signing and would likely hinder recall. The process of translating written English words to ASL signs during encoding and/or recall may have benefitted older more than younger participants, but again it is not clear why this should be the case nor why the benefit would diminish on delayed recall.

Another possible consideration that might have magnified the effects of language on learning during immediate recall trials in older participants is between group differences in educational background. Historically, older deaf individuals are more likely to be limited to oral education during their early schooling years as a result of widespread educational policies in the USA, which restricted access to ASL (Moore, 2010). Table 2.1 shows the means for self-reported extent of oral communication during elementary and junior/high school years. Although older and younger participants did not differ significantly in these measures, it is nonetheless likely that teaching practices implemented for the older participants represented a stronger bias for English education relative to those implemented for the younger participants. This educational background for the older participants could be linked to their better performance in English during the learning trials, in which recall for words was greater than for ASL signs. Additional research will be needed to investigate this possibility, and to determine why English facilitated learning but not delayed recall for older (but not for younger) deaf ASL-English participants.

Language of Testing Order & Self-rated Proficiency Level

An important consideration when conducting cognitive testing with bilingual participants is what language(s) should be used for testing, i.e., whether to test in one or both languages, and if testing in both languages, which language to test first. In the present study, we counterbalanced order of language presentation and found no interactions with age group, and minimal order effects. The only significant order effect was that participants forgot more items in whichever language was presented second. It is possible that this order effect was due to proactive interference, in which learning from a prior list interferes with learning of a new list. Proactive interference can lead to reduced recall of initial learning of a new list compared to initial learning of a previous list. However, proactive interference effects tend to be larger in older than in young participants (Ebert & Anderson, 2009), a pattern not seen in the present study. Furthermore, it is not clear why testing order effects were more evident in the delayed trial compared to immediate learning trials in the current study. A recent study of older Spanish-English bilinguals found no language of testing order effects on recall scores (Gavino et al., submitted), but that study did not include delayed recall. One study of young Spanish-English bilinguals presented four word lists in each language for immediate recall using brief exposure times (each word was presented for just 400 ms) (Gutierrez et al., 2013). In this study, bilinguals recalled fewer English words if they had previously completed the same task in Spanish, but the reverse pattern did not hold, as Spanish recall was equivalent regardless of testing order. Importantly, these participants were strongly English dominant, suggesting that activation of the non-dominant language impeded later ability to learn and recall information in the dominant language, a form of between-language interference that may resemble order effects sometimes found in picture naming tasks (e.g., Branzi et al., 2014; Wodniecka et al., 2020). In the current study, we found effects in both directions (prior recall of English affected later ability to recall items after a delay in ASL, and

prior recall of ASL affected later ability to recall items after a delay in English). The bidirectional effect of interference in the present study could reflect the relatively balanced degree of proficiency in English and ASL in our participants (see Table 2.1).

Our finding of minimal effects of self-reported differences in language proficiency on performance is also consistent with the assumption that ASL and English proficiency was relatively balanced in our participants. Figure 2.4 shows that older participants who self-reported one language as more proficient also tended to recall more items in the more proficient language on Trials 1 and 2, but language dominance did not affect recall scores on Trials 3 and delayed recall. Similar results were reported by Gavino et al. (submitted) who observed a significant correlation between language dominance scores in older bilinguals only on Trial 1 (no young bilinguals were tested in that study). The absence of significant effects in younger participants in the current study could suggest that the young participants were even more balanced in their proficiency levels than the older participants, or that young ASL-English participants were less accurate in their ability to self-rate proficiency level than older participants. Caution is also generally warranted when interpreting self-ratings of proficiency level, which does not necessarily accurately reflect objectively measured proficiency level (Tomoschuk et al., 2019; Garcia & Gollan, 2022; Gollan et al., 2012; Neveu et al., 2025).

However, the most striking aspect of the results shown in Figure 2.4 is the extent of variability between participants. That is, many participants rated one language as more proficient but then recalled many more items in the other language, and vice versa. For example, one participant rated themselves as having better proficiency in ASL than in English (i.e., their self-rated dominance score was positive), yet they consistently remembered more English items than ASL items (i.e., a list memory dominance score of -3 on Trial 1 and -6 on the delay trial).

Another participant who rated themselves as more proficient in English than ASL (i.e., negative self-rated dominance score), recalled more ASL items on both Trial 1 and the delay trial. In Figure 2.4, any participants shown in upper left or bottom right quadrants, exhibited one of these patterns. This pattern may not be fully specific to deaf ASL-English bilinguals as Gavino et al. (submitted) found that some older Spanish-English bilinguals also recalled more items in a language they indicated was less proficient. However, the pattern may be more common for deaf ASL-English bilinguals because many more participants in the present study said they are more proficient in ASL but then recalled more items in English than in ASL on Trial 1. In this respect, testing in just one language might not accurately represent recall performance in many deaf bilinguals.

Limitations

Additional studies are needed to determine if these findings replicate and generalize to older deaf participants with more varied levels of proficiency in ASL and English, as well as those with different educational backgrounds (most participants in our sample were high school graduates with some college education). We suspect that relatively balanced bilingualism may be common in deaf individuals for whom sign language is more accessible and for whom English proficiency is needed for communication within the broader hearing community. However, more work is needed to determine how to assess relative language proficiency in older deaf signers, and how this may affect memory performance. Longitudinal studies will be needed to determine if recall in ASL is more preserved with increased age compared to recall in English in deaf bilinguals, and to determine which memory measures are most sensitive to dementia in older deaf signers. Finally, we focused on recall of relatively short (10 item lists), and additional research is needed to determine if similar results are found in memory for longer lists, and after a

longer delay. Many commonly used list memory tests evaluate delayed recall after 15-20 minutes, whereas in the present study we tested only after a short delay of about 8-10 minutes.

Conclusions

The results of the present study provide critical preliminary information on list memory performance of young and older deaf ASL-English bilinguals. Based on participants' strong preference to recall items in ASL in the present study – regardless of which language was used to present list items – we recommend following this practice for memory assessments of deaf bilinguals. We also recommend testing in both languages, which will be minimally affected by testing order, and may maximize the performance of deaf ASL-English bilinguals on different metrics of memory functioning (i.e., initial learning vs. delayed recall/forgetting). For older participants, testing in ASL may be especially critical to determine if forgetting falls within normal limits. In addition, given the lack of a clear relationship between self-reported language proficiency and memory performance in both young and older participants, testing in both languages will provide a more complete characterization of learning and memory ability for deaf bilinguals.

Appendix

Model Outputs for Age as Continuous Factor

Model outputs with age entered as a continuous factor (instead of as young, old contrast) including all participants tested (middle aged participants aged 46-63 were excluded from the analyses above but were included here to examine the continuous effects of age). AoA refers to age of acquisition of ASL.

Table A1. *Model outputs for primary outcome measures with age as a continuous factor*

Effect	F	p	η_p^2
Overall learning (Trials 1-3)			
Age	18.21	0.00	0.19
AoA	0.98	0.33	0.01
LANGUAGE	3.77	0.06	0.05
Age:LANGUAGE	0.00	0.98	0.00
AoA:LANGUAGE	0.93	0.34	0.01
Trial 3 versus delay trial			
Age	23.02	0.00	0.23
AoA	2.53	0.12	0.03
LANGUAGE	0.86	0.36	0.01
TRIAL	80.95	0.00	0.52
Age:LANGUAGE	0.01	0.91	0.00
AoA:LANGUAGE	0.08	0.78	0.00
Age:TRIAL	4.67	0.03	0.06
AoA:TRIAL	2.33	0.13	0.03
LANGUAGE:TRIAL	10.10	0.00	0.12
Age:LANGUAGE:TRIAL	6.91	0.01	0.08
AoA:LANGUAGE:TRIAL	2.64	0.11	0.03
Adjusted forgetting			
Age	7.52	0.01	0.09
AoA	1.54	0.22	0.02
LANGUAGE	9.95	0.00	0.12
Age:LANGUAGE	7.19	0.01	0.09
AoA:LANGUAGE	2.20	0.14	0.03
Primacy scores			
Age	11.67	0.00	0.13
AoA	1.26	0.27	0.02
LANGUAGE	1.62	0.21	0.02
Age:LANGUAGE	0.51	0.48	0.01
AoA:LANGUAGE	0.47	0.50	0.01

Note.

Note. Significant values appear in bold. AoA = Age of acquisition; Degrees of freedom numerator = 1; Degrees of freedom denominator = 76.

(Appendix continues)

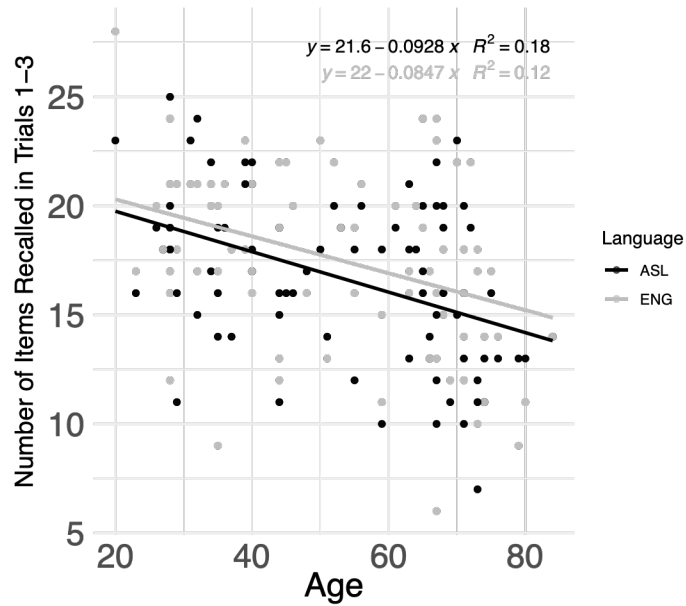


Figure A1. Scatterplot demonstrating the continuous effect of age on learning (sum of Trials 1-3) in all participants

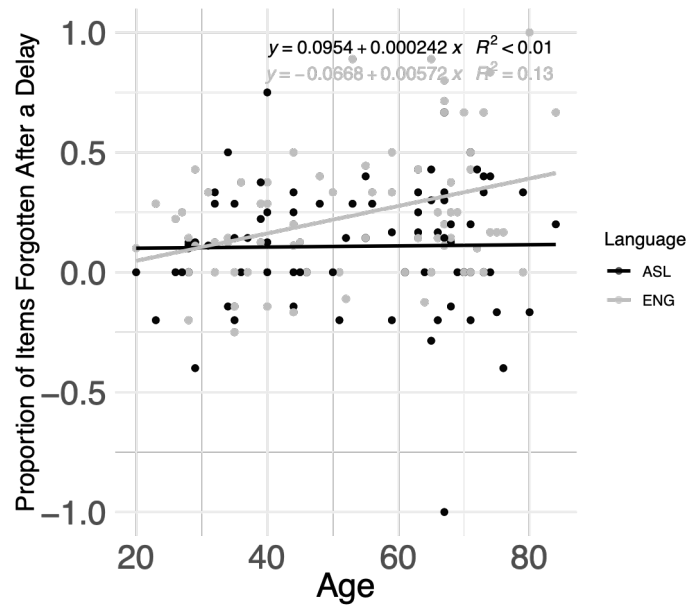


Figure A2. Scatterplot demonstrating the continuous effect of age on adjusted forgetting in all participants

Acknowledgements

Chapter 2, in full is a reprint of the material as it appears in *Neuropsychology*, 2026.

Camilliere, S., Emmorey, K., Hauser, P. C., Contreras, J., McKee, M. M., & Gollan, T. H. (2026). Influence of testing language and aging on verbal list memory in deaf American Sign Language–English bilinguals. The dissertation author was the primary investigator and author of this paper. This research was supported by NIH grant R01-AG074928. The authors have no conflicts of interest to disclose. We would like to thank Lucinda Farnady and Julie Fisher for their help with stimuli development, participant recruitment, and data collection; Chehally Soto, Emma Kane, Hayden Orr, and Gigi Zheng for their help with data collection; and Caitlin Pearson for her help with data coding. We would also like to thank all of our deaf participants.

References

- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC.
- Baddeley, A. (1986). *Working memory*. Oxford: Oxford University Press.
- Bair, J. L., Patrick, Sarah D., Noyes, Emily T., Hale, Andrew C., Campbell, Elizabeth B., Wilson, Addie M., Ransom, Michael T., & Spencer, R. J. (2023). Semantic clustering on common list-learning tasks: A systematic review of the state of the literature and recommendations for future directions. *Journal of Clinical and Experimental Neuropsychology*, 45(7), 652–692. <https://doi.org/10.1080/13803395.2023.2270204>
- Bavelier, D., Newport, E. L., Hall, M., Supalla, T., & Boutla, M. (2008). Ordered short-term memory differs in signers and speakers: Implications for models of short-term memory. *Cognition*, 107(2), 433–459. <https://doi.org/10.1016/j.cognition.2007.10.012>
- Bayley, P. J., Salmon, D. P., Bondi, M. W., Bui, B. K., Olichney, J., Delis, D. C., Thomas, R. G., & Thal, L. J. (2000). Comparison of the serial position effect in very mild Alzheimer's disease, mild Alzheimer's disease, and amnesia associated with electroconvulsive therapy. *Journal of the International Neuropsychological Society*, 6(3), 290–298. <https://doi.org/10.1017/S1355617700633040>
- Belleville, S., Fouquet, C., Hudon, C., Zomahoun, H. T. V., Croteau, J., & Consortium for the Early Identification of Alzheimer's disease-Quebec. (2017). Neuropsychological measures that predict progression from mild cognitive impairment to Alzheimer's type dementia in older adults: A systematic review and meta-analysis. *Neuropsychology Review*, 27(4), 328–353. <https://doi.org/10.1007/s11065-017-9361-5>
- Branzi, F. M., Martin, C. D., Abutalebi, J., & Costa, A. (2014). The after-effects of bilingual language production. *Neuropsychologia*, 52, 102–116. <https://doi.org/10.1016/j.neuropsychologia.2013.09.022>
- Bruno, D., Grothe, M. J., Nierenberg, J., Zetterberg, H., Blennow, K., Teipel, S. J., & Pomara, N. (2015). A study on the specificity of the association between hippocampal volume and delayed primacy performance in cognitively intact elderly individuals. *Neuropsychologia*, 69, 1–8. <https://doi.org/10.1016/j.neuropsychologia.2015.01.025>
- Brysbaert, M. & New, B. (2009). Moving beyond Kučera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, 41, 977–990. doi: 10.3758/BRM.41.4.977
- Camilliere, S. (2025) *ASL Word List Learning*. [Data set]. OSF. Retrieved from https://osf.io/bj8yz/overview?view_only=9f441c0c41404b10a414b3797ffceb29
- Cattani, A., Clibbens, J., & Perfect, T. J. (2007). Visual memory for shapes in deaf signers and nonsigners and in hearing signers and nonsigners: Atypical lateralization and enhancement. *Neuropsychology*, 21(1), 114–121. <https://doi.org/10.1037/0894-4105.21.1.114>

- Craik, F. I. (1991). Memory functions in normal aging. *Memory disorders: Research and clinical practice*, 347-367.
- Cunha, C., Guerreiro, Manuela, de Mendonça, Alexandre, Oliveira, Paulo Eduardo, & Santana, I. (2012). Serial position effects in Alzheimer's disease, mild cognitive impairment, and normal aging: Predictive value for conversion to dementia. *Journal of Clinical and Experimental Neuropsychology*, 34(8), 841–852.
<https://doi.org/10.1080/13803395.2012.689814>
- Curtiss, S. (2014). *Genie: A psycholinguistic study of a modern-day wild child*. Academic Press.
- Delis, D. C., J. H. Kramer, E. Kaplan, and B. A. Ober. 1987. The California Verbal Learning Test. *San Antonio: Psychological Corporation*.
- Denmark, T., Marshall, J., Mummery, C., Roy, P., Woll, B., & Atkinson, J. (2016). Detecting memory impairment in deaf people: A new test of verbal learning and memory in British Sign Language. *Archives of Clinical Neuropsychology*, 31(8), 855–867.
<https://doi.org/10.1093/arclin/acw032>
- Ebert, P. L., & Anderson, N. D. (2009). Proactive and retroactive interference in young adults, healthy older adults, and older adults with amnesic mild cognitive impairment. *Journal of the International Neuropsychological Society*, 15(1), 83–93.
<https://doi.org/10.1017/S1355617708090115>
- Emmorey, K., Klima, E., & Hickok, G. (1998). Mental rotation within linguistic and non-linguistic domains in users of American sign language. *Cognition*, 68(3), 221–246.
[https://doi.org/10.1016/S0010-0277\(98\)00054-7](https://doi.org/10.1016/S0010-0277(98)00054-7)
- Emmorey, K., Kosslyn, S. M., & Bellugi, U. (1993). Visual imagery and visual-spatial language: Enhanced imagery abilities in deaf and hearing ASL signers. *Cognition*, 46(2), 139–181.
[https://doi.org/10.1016/0010-0277\(93\)90017-P](https://doi.org/10.1016/0010-0277(93)90017-P)
- Francis, W. S., Arteaga, M. M., Liaño, M. K., & Taylor, R. S. (2020). Temporal dynamics of free recall: The role of rehearsal efficiency in word frequency and bilingual language proficiency effects. *Journal of Experimental Psychology: General*, 149(8), 1477–1508.
<https://doi.org/10.1037/xge0000732>
- Gasquoine, P. G., & Gonzalez, C. D. (2012). Using monolingual neuropsychological test norms with bilingual Hispanic Americans: Application of an individual comparison standard. *Archives of Clinical Neuropsychology*, 27(3), 268–276.
<https://doi.org/10.1093/arclin/acs004>
- Gavino, M.F., Salmon, D.P., Mizrahi, R. & Gollan, T.H. (in press). *Language dominance effects on verbal list memory in older Spanish-English bilinguals*
- Garcia, D. L., & Gollan, T. H. (2022). The MINT Sprint: Exploring a Fast Administration Procedure with an Expanded Multilingual Naming Test. *Journal of the International Neuropsychological Society*, 28(8), 845–861. doi:10.1017/S1355617721001004
- Gollan, T. H., Salmon, D. P., Montoya, R. I., & da Pena, E. (2010). Accessibility of the nondominant language in picture naming: A counterintuitive effect of dementia on

- bilingual language production. *Neuropsychologia*, 48(5), 1356–1366.
<https://doi.org/10.1016/j.neuropsychologia.2009.12.038>
- Gollan, T. H., Stasenko, A., & Salmon, D. P. (2023). Which language is more affected in bilinguals with Alzheimer's disease? Diagnostic sensitivity of the Multilingual Naming Test. *Neuropsychology*, 37(5), 595–606. <https://doi.org/10.1037/neu0000893>
- Griffin, J. W., John, Samantha E., Adams, Jason W., Bussell, Cara A., Saurman, Jessica L., & Gavett, B. E. (2017). The effects of age on the learning and forgetting of primacy, middle, and recency components of a multi-trial word list. *Journal of Clinical and Experimental Neuropsychology*, 39(9), 900–912. <https://doi.org/10.1080/13803395.2017.1278746>
- Grosjean, F. (2010). Bilingualism, biculturalism, and deafness. *International Journal of Bilingual Education and Bilingualism*, 13(2), 133–145.
<https://doi.org/10.1080/13670050903474051>
- Gutierrez, A., Pilotti, M., Romero, E., Mahamane, S., & Broderick, T. (2013). Proactive interference between languages: Do task demands matter? *International Journal of Bilingualism*, 17(4), 505–524. <https://doi.org/10.1177/1367006912439940>
- Hall, M. L., & Bavelier, D. (2011). Short-term memory stages in sign vs. speech: The source of the serial span discrepancy. *Cognition*, 120(1), 54–66.
<https://doi.org/10.1016/j.cognition.2011.02.014>
- Hirshorn, E. A., Fernandez, N. M., & Bavelier, D. (2013). Routes to short-term memory indexing: Lessons from deaf native users of American Sign Language. In B. Landau (Ed.), *Understanding Cognitive Development*, pp. 85-103, Psychology Press.
- Humphries, T., Kushalnagar, P., Mathur, G., Napoli, D. J., Padden, C., Rathmann, C., & Smith, S. (2016). Avoiding linguistic neglect of deaf children. *Social Service Review*, 90(4), 589–619. <https://doi.org/10.1086/689543>
- Jorm, A. F., & Jolley, D. (1998). The incidence of dementia: A meta-analysis. *Neurology*, 51(3), 728–733. <https://doi.org/10.1212/WNL.51.3.728>
- Karrasch, M., Sinervä, E., Grönholm, P., Rinne, J., & Laine, M. (2005). CERAD test performances in amnesic mild cognitive impairment and Alzheimer's disease. *Acta Neurologica Scandinavica*, 111(3), 172–179. <https://doi.org/10.1111/j.1600-0404.2005.00380.x>
- Kassambara, A. (2023a). `_ggpubr: 'ggplot2' Based Publication Ready Plots_`. R package version 0.6.0, <<https://CRAN.R-project.org/package=ggpubr>>.
- Kassambara A (2023b). `rstatix: Pipe-Friendly Framework for Basic Statistical Tests`. R package version 0.7.2, <https://rpkgs.datanovia.com/rstatix/>.
- Kazak, A. E. (2018). Editorial: Journal article reporting standards. *American Psychologist*, 73(1), 1–2. <https://doi.org/10.1037/amp0000263>
- Kloth, N., Lemke, J., Wiendl, H., Meuth, S. G., Duning, T., & Johnen, A. (2020). Serial position effects rapidly distinguish Alzheimer's from frontotemporal dementia. *Journal of Neurology*, 267(4), 975–983. <https://doi.org/10.1007/s00415-019-09662-w>

- Kubicek, E., & Quandt, L. C. (2021). A Positive Relationship Between Sign Language Comprehension and Mental Rotation Abilities. *The Journal of Deaf Studies and Deaf Education*, 26(1), 1–12. <https://doi.org/10.1093/deafed/enaa030>
- Liddell, S. K. (2003). *Grammar, gesture, and meaning in American Sign Language*. Cambridge University Press.
- Mayberry, R. I., Lock, E., & Kazmi, H. (2002). Linguistic ability and early language exposure. *Nature*, 417(6884), 38–38. <https://doi.org/10.1038/417038a>
- McKee, M. M., & Paasche-Orlow, M. K. (2012). Health literacy and the disenfranchised: The importance of collaboration between limited English proficiency and health literacy researchers. *Journal of Health Communication*, 17(sup3), 7–12. <https://doi.org/10.1080/10810730.2012.712627>
- McKee, M. M., Paasche-Orlow, Michael K., Winters, Paul C., Fiscella, Kevin, Zazove, Philip, Sen, Ananda, & Pearson, T. (2015). Assessing health literacy in deaf American Sign Language users. *Journal of Health Communication*, 20(sup2), 92–100. <https://doi.org/10.1080/10810730.2015.1066468>
- Mitchell, R. E., & Karchmer, M. A. (2004). Chasing the mythical ten percent: Parental hearing status of deaf and hard of hearing students in the United States. *Sign Language Studies*, 4(2), 138–163.
- Mizrahi, R., Cromheecke, O., Salmon, D. P., & Gollan, T. H. (2024). Disruption of the serial position effect as an early marker of Alzheimer’s disease in Spanish–English bilinguals. *Journal of the International Neuropsychological Society*, 30(2), 162–171. <https://doi.org/10.1017/S1355617723000310>
- Moore, D. F. (2010). The history of language and communication issues in deaf education. *The Oxford Handbook of Deaf Studies, Language, and Education*, 2, 17–30.
- Morere, D. A. (2012). Measures of memory and learning. In D. Morere & T. Allen (Eds.), *Assessing Literacy in Deaf Individuals: Neurocognitive Measurement and Predictors* (pp. 75–104). Springer. https://doi.org/10.1007/978-1-4614-5269-0_5
- Morere, D. A. (2013a). Methodological issues associated with sign-based neuropsychological assessment. *Sign Language Studies*, 14(1), 8–20.
- Morere, D. A. (2013b). The Signed Verbal Learning Test: Assessing verbal memory of deaf signers. *Sign Language Studies*, 14(1), 39–57.
- Morris, J. C. (1996). Classification of dementia and Alzheimer’s disease. *Acta Neurologica Scandinavica*, 94(S165), 41–50. <https://doi.org/10.1111/j.1600-0404.1996.tb05871.x>
- Murdock, B. B. (1962). The serial position effect of free recall. *Journal of experimental psychology*, 64(5), 48. doi: doi.org/10.1037/h0045106
- Neveu, A., Garcia, D. L., Escobedo, B., Vazquez, P. E., Mejia, M., Hoversten, L. J., & Gollan, T. H. (2025). Predicting proficiency. *Bilingualism: Language and Cognition*, 1–18. <https://doi.org/10.1017/S1366728925000367>

- Peña, E. D. (2007). Lost in translation: Methodological considerations in cross-cultural research. *Child Development*, 78(4), 1255–1264. <https://doi.org/10.1111/j.1467-8624.2007.01064.x>
- Petersen, R. C., Smith, G., Kokmen, E., Ivnik, R. J., & Tangalos, E. G. (n.d.). *Memory function in normal aging*.
- Petersen, R. C., Smith, G., Kokmen, E., Ivnik, R. J., & Tangalos, E. G. (1992). Memory function in normal aging. *Neurology*, 42(2), 396–396. <https://doi.org/10.1212/WNL.42.2.396>
- R Core Team (2023). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <<https://www.R-project.org/>>.
- Salmon, D. P., & Bondi, M. W. (2009). Neuropsychological assessment of dementia*. *Annual Review of Psychology*, 60(Volume 60, 2009), 257–282. <https://doi.org/10.1146/annurev.psych.57.102904.190024>
- Sehyr, Z. S., Caselli, N., Cohen-Goldberg, A. M., & Emmorey, K. (2021). The ASL-LEX 2.0 project: A database of lexical and phonological properties for 2,723 signs in American Sign Language. *The Journal of Deaf Studies and Deaf Education*, 26(2), 263–277. <https://doi.org/10.1093/deafed/enaa038>
- Sehyr, Z. S., Giezen, M. R., & Emmorey, K. (2018). Comparing semantic fluency in American Sign Language and English. *The Journal of Deaf Studies and Deaf Education*, 23(4), 399–407. <https://doi.org/10.1093/deafed/eny013>
- Sehyr, Z. S., Petrich, J., & Emmorey, K. (2017). Fingerspelled and printed words are recoded into a speech-based code in short-term memory. *The Journal of Deaf Studies and Deaf Education*, 22(1), 72–87. <https://doi.org/10.1093/deafed/enw068>
- Sumby, W. H. (1963). Word frequency and serial position effects. *Journal of Memory and Language*, 1(6), 443.
- Tomoschuk, B., Ferreira, V. S., & Gollan T. H. (2019). When a seven is not a seven: Self-ratings of bilingual language proficiency differ between and within language populations. *Bilingualism: Language and Cognition*, 22(3), 516–536. [doi:10.1017/S1366728918000421](https://doi.org/10.1017/S1366728918000421)
- Ward, G., Tan, L., & Grenfell-Essam, R. (2010). Examining the relationship between free recall and immediate serial recall: The effects of list length and output order. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(5), 1207–1241. doi.org/10.1037/a0020122
- Twamley, E. W., Ropacki, S. a. L., & Bondi, M. W. (2006). Neuropsychological and neuroimaging changes in preclinical Alzheimer’s disease. *Journal of the International Neuropsychological Society*, 12(5), 707–735. <https://doi.org/10.1017/S1355617706060863>
- Welsh, K. A., Butters, N., Mohs, R. C., Beekly, D., Edland, S., Fillenbaum, G., & Heyman, A. (1994). The Consortium to Establish a Registry for Alzheimer’s Disease (CERAD). Part V. A normative study of the neuropsychological battery. *Neurology*, 44(4), 609–609. <https://doi.org/10.1212/WNL.44.4.609>

- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <<https://ggplot2.tidyverse.org>>
- Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). *_dplyr: A Grammar of Data Manipulation_*. R package version 1.1.4, <<https://CRAN.Rproject.org/package=dplyr>>.
- Wickham, H., Vaughan, D., & Girlich, M. (2023). *_tidyr: Tidy Messy Data_*. R package version 1.3.0, <<https://CRAN.R-project.org/package=tidyr>>.
- Wodniecka, Z., Szewczyk, J., Kałamała, P., Mandera, P., & Durlík, J. (2020). When a second language hits a native language. What ERPs (do and do not) tell us about language retrieval difficulty in bilingual language production. *Neuropsychologia*, *141*, 107390. <https://doi.org/10.1016/j.neuropsychologia.2020.107390>
- Yoo, J., & and Kaushanskaya, M. (2016). Serial-position effects on a free-recall task in bilinguals. *Memory*, *24*(3), 409–422. <https://doi.org/10.1080/09658211.2015.1013557>

CHAPTER 3: Bigger between-language than within-language Stroop effects in young and older deaf American Sign Language-English bilinguals

Sadie Camilliere¹, Karen Emmorey², Peter C. Hauser³, Jessica Contreras³, Michael M. McKee⁴,
and Tamar H. Gollan⁵

¹ Joint Doctoral Program in Language and Communicative Disorders, San Diego State

University & University of California, San Diego, CA USA

² School of Speech, Language, and Hearing Science, San Diego State University, San Diego, CA

USA

³ National Technical Institute for the Deaf, Rochester Institute of Technology, Rochester, NY

USA

⁴ Department of Family Medicine, University of Michigan, Ann Arbor, MI USA

⁵ Department of Psychiatry, University of California, San Diego, La Jolla, CA USA

Correspondence concerning this article should be addressed to Sadie Camilliere, Laboratory for Language and Cognitive Neuroscience, 6495 Alvarado Rd Suite 200, San Diego, CA 92182

Email: sadiecamilliere@gmail.com

Abstract

Objective: The present study investigated Stroop performance in young adult and older deaf American Sign Language (ASL) and English bilinguals with both ASL and English stimuli.

Method: One hundred nine deaf ASL-English bilinguals (aged 21-88) signed their responses in two computerized Stroop tasks, one with ASL videos of color signs and one with printed English color words. Participants saw 120 trials in each language with 84 congruent and 36 incongruent trials.

Results: Even though participants responded in ASL in both tasks, Stroop effects were larger with English than with ASL stimuli in both response times (RTs) and errors, a pattern of results not typically seen in spoken-language bilinguals (in which Stroop effects are often smaller between-languages than within-languages). Older participants showed larger Stroop effects than younger participants in RTs, but not in errors, and the Stroop effect increased continuously with age.

Conclusions: The finding of larger Stroop effects in deaf signers with English rather than ASL stimuli suggests a possibly critical role for modality of presentation in eliciting Stroop effects. Additional work is needed to determine which version of the Stroop task is most sensitive to detect cognitive impairment in deaf signing bilinguals.

Keywords: *Bimodal bilingualism; Bilingual Stroop; American Sign Language; Aging; Deaf; Between-language Stroop*

Statement of Research Significance

Research Question(s) or Topic(s): Previous work has demonstrated that young deaf American Sign Language-English bilinguals show Stroop effects when presented with American Sign Language (ASL) stimuli, but no study has examined between-language Stroop effects, nor aging effects on the Stroop task in deaf signers.

Main Findings: Presenting deaf signers with an English Stroop task elicited larger Stroop effects in response times and errors compared to a within-ASL Stroop task, and older deaf participants showed larger Stroop effects than younger deaf participants.

Study Contributions: The current study is the first to examine both between-language and within-language Stroop effects using a naturalistic response mode for deaf participants (i.e., signing), and across a wide age range of deaf ASL-English bilinguals. This work constitutes a critical first step in determining potential diagnostic sensitivity of Stroop tests to cognitive decline in this population by demonstrating how testing language and aging modulate Stroop effects in deaf bilinguals.

Introduction

The Stroop task (Stroop, 1935) is one of the most widely studied experimental paradigms, giving insight into processes such as attentional control (MacLeod, 1991) and automaticity in lexical access (Brown et al., 2002). *Color Stroop interference* refers to the robust phenomenon in which participants have difficulty ignoring the automaticity of reading instead of naming the ink color in which the word is written. For example, it is difficult to say *red* when the word *blue* is written in red ink. Stroop tasks have been used to probe interference effects across languages in bilinguals, e.g., prompting individuals to look at a word written in Spanish, but to name the color of the ink in English; the results of these cross-language studies have informed our understanding of how two languages are represented and interact within a single cognitive system (Fang et al., 1981; Rosselli et al., 2002; Tzelgov et al., 1990). The extent to which Stroop effects transfer across languages is modulated by language proficiency and language dominance (Naylor et al., 2012; Rosselli et al., 2002; Suarez et al., 2014). Stroop effects have also been useful for understanding aging-related changes in cognitive processing. Stroop effects are larger in older than in young participants (Cohn et al., 1984; Comalli Jr. et al., 1962; Nicosia et al., 2021), indexing aging-related declines in processing speed and attentional control. Production of errors in Stroop tasks is especially magnified in early stage Alzheimer's dementia (AD; Balota et al., 2010; Belleville et al., 2006, 2017), making the Stroop task clinically useful.

Despite the vast body of work investigating Stroop effects in hearing individuals across the lifespan (see MacLeod, 1991 and Nicosia et al., 2021 for reviews), little research has examined Stroop effects in deaf bilinguals who know American Sign Language (ASL) and English, and no studies have evaluated Stroop effects in older deaf signers. We can address several questions of interest by investigating Stroop performance in this unique group of deaf

bimodal bilinguals, i.e., deaf individuals who know a sign language and a spoken language and who may differ from *unimodal bilinguals* who know two spoken languages. These include theoretical questions about the automaticity of cognitive processing of sign languages, how testing language affects Stroop interference in different types of bilinguals, and the nature of possible interactions between aging and bilingualism on Stroop interference effects. These are critical considerations for evaluating if Stroop tasks are clinically useful in the diagnosis of cognitive decline in deaf signers, and for understanding this unique form of bilingualism.

Stroop Effects in Hearing Unimodal Bilinguals

In unimodal bilinguals who know two spoken languages, Stroop interference effects are often larger when the stimulus and response languages match (e.g., participants respond in Hebrew and the words are written in Hebrew) compared to when they mismatch (e.g., participants respond in Arabic and the words are written in Hebrew; Tzelgov et al., 1990). The finding of stronger within-language than between-language Stroop effects is also more pronounced when the two languages do not share orthography. For example, Spanish-English bilinguals show more similarly sized within-language and between-language Stroop effects compared to Chinese-English bilinguals, who show much larger within-orthography Stroop effects than between-orthography Stroop effects (Fang et al., 1981). These results suggest that there is greater competition between languages in bilinguals who use languages that rely on overlapping orthographies because the two languages engage the same script processing mechanisms. In contrast, bilinguals reading a logographic script (Chinese) do not experience as much competition from an alphabetic script (English), which leads to reduced Stroop effects between the two languages.

Language proficiency and dominance can also impact Stroop performance; unimodal bilinguals sometimes show greater Stroop interference in their more proficient or dominant language (Naylor et al., 2012), and Stroop interference effects increase with increasing proficiency level (Rosselli et al., 2002; Suarez et al., 2014). In addition, Stroop tasks involve automatic processing (i.e., automatic lexical retrieval of the printed word) as well as cognitive control (i.e., ability to inhibit the automatic reading of the printed word), both of which can vary with proficiency in the testing language. Tzelgov et al. (1990) suggested that the relationship between language proficiency and Stroop interference may be U-shaped; some proficiency is needed to show any interference between languages, but at the highest level of proficiency bilinguals may be better able to control Stroop interference. They found larger within-language than between-language Stroop interference in balanced bilinguals, whereas in unbalanced bilinguals, Stroop effects were larger whenever stimuli were presented in the dominant language. Thus, Stroop effects could both under- or overestimate cognitive control abilities in bilinguals, depending on proficiency level and the language of administration.

Stroop Effects in Deaf Signers

Deaf bilinguals differ from hearing bilinguals in many aspects of language use and exposure. Some deaf individuals use speech or lip reading to access English, but many use English primarily in the written modality, and many learn English through reading and writing. By contrast, hearing bilinguals typically learn a second language by hearing and speaking it (in addition to reading and writing it). Additional considerations are important for predicting how these factors may influence Stroop performance in deaf ASL-English bilinguals. ASL and other sign languages do not have a widely adopted written form, and consequently deaf bilinguals read

and write in just one of their two languages. All of these factors potentially impact Stroop performance, yet few studies have investigated Stroop effects in this population.

Initial attempts to elicit Stroop effects in a sign language used static drawings of ASL color signs and asked deaf participants to sign the name of the drawing's ink color, but these studies yielded inconclusive results with nonsignificant Stroop effects (Marschark, 1988; Marschark & Shroyer, 1993). We suggest that the null results found in these early studies were likely due to the use of static images. Movement is an essential feature of ASL color signs and without movement, the color sign drawings resemble ASL letter signs because most color signs have a handshape that represents the first letter of the English translation. For example, the sign YELLOW involves moving the Y handshape with a repetitive twisting movement, but a static Y handshape is the sign for the letter "Y". Static drawings of letter names may simply not be sufficient to elicit activation of ASL color signs and thus are not appropriate stimuli for an ASL Stroop task.

To date only two studies elicited Stroop effects using dynamic videos of ASL color signs. Dupuis & Berent (2015) used videos of ASL signers from the torso up producing color signs, overlaid with a color that was either congruent or incongruent with the sign being produced. A small group of 10 deaf ASL signers (age not reported) completed two ASL Stroop tasks in which they were instructed to name the color of the hand while ignoring the sign; one task used button-pressed responses and one used signed responses. Regardless of response mode, participants were slower and less accurate when responding to incongruent trials (the hand color and the color sign were different) compared to congruent trials (the hand color and the color sign were the same), thus demonstrating significant Stroop effects for ASL. This result was taken as evidence for the automatic activation of the semantics of ASL signs, since participants

experienced interference from the signed lexical items, even though they were instructed to ignore the meaning of the sign and respond only to the color of the hand (Dupuis & Berent, 2015).

More recently, Bosworth et al. (2021) tested 15 deaf ASL-English bilinguals, ages 19-39, using both ASL and English Stroop tasks with button-pressed responses. The color responses (*blue, green, yellow, purple*) were labeled on a keyboard with color patches that participants used to respond to either ASL signs or printed English words. For ASL, participants were shown dynamic videos of hands producing the ASL color signs, which were all produced in a neutral location with the same movement and differed only in handshape. For both languages, participants completed both a traditional Stroop interference task (i.e., they identified the color of the hand for signs, and the font color for English words), and a control task in which they indicated the color expressed by the sign or word. Both the ASL and English Stroop tasks exhibited robust Stroop effects. Additionally, within the ASL task, participants with higher ASL proficiency (measured by the ASL-Comprehension Test; Hauser et al., 2016), exhibited larger ASL Stroop effects (English proficiency was not assessed). This result replicates the association between proficiency level and the magnitude of Stroop interference found in unimodal bilinguals (Rosselli et al., 2002; Tse & Altarriba, 2012). Thus, this initial investigation into deaf ASL-English bilinguals' Stroop performance reveals some similarities to the literature on Stroop effects in hearing unimodal bilinguals. However, additional work is needed with larger sample sizes to better understand Stroop performance in deaf ASL-English bilinguals (the previous studies included $n=10$, Dupuis & Berent, 2015; and $n=15$ deaf participants, Bosworth et al., 2021). In addition, Bosworth et al. (2021) did not investigate the impact of English proficiency on English Stroop performance in deaf bimodal bilinguals, nor did they investigate between

language Stroop effects (i.e., English stimuli with ASL responses). Finally, neither study tested participants above the age of 40, leaving a critical gap in understanding of Stroop performance in older deaf bilinguals.

Current Study

Investigating deaf bimodal bilinguals' Stroop effects in both ASL and English presents an opportunity to explore a variety of theoretical questions regarding the impact of bilingualism, language proficiency, and modality on Stroop performance. The absence of a writing system for ASL makes it more difficult to predict the expected magnitude of within-language vs. between-language Stroop effects in deaf bilinguals. The lack of a shared orthographic system could suggest similar within- and between-language Stroop effects for ASL-English bilinguals, given the lack of competition at the perceptual form level. However, it is unknown how the cross-modal nature of the between-language task (i.e., signing the color of the printed English word) might influence Stroop effects.

Language proficiency may also be critical in assessing Stroop effects. Although many deaf individuals consider ASL to be their dominant language, they also typically use written English on a daily basis and may become skilled readers (e.g., Sehyr & Emmorey, 2022). English also represents the main form of connection to the hearing-dominant society in which deaf people are immersed. Thus, if language proficiency is an important predictor of Stroop performance in this population, ASL-English bilinguals with high levels of proficiency in both of their languages might exhibit similar within-language and between-language Stroop effects, as has been reported for balanced bilinguals (Rosselli et al., 2002).

Another factor that may be important to consider is the age of acquisition (AoA) of ASL which is variable across deaf individuals and has been shown to affect performance on some cognitive and linguistic tasks (see Mayberry, 2002, for review). In contrast to hearing bilinguals, the AoA of ASL can reflect the timing of acquisition a *first* language because deafness prevents early acquisition of spoken English, and most deaf people (90-95%) are born to hearing parents who do not know a sign language (Mitchell & Karchmer, 2004). Deaf adults with a later ASL AoA have been shown to have poorer language skills in both English and ASL and are at risk for early language deprivation which impacts cognition and mental health (Hall et al., 2017). Thus, it is possible that individuals who acquired ASL later in life will show smaller Stroop effects as a result of lower levels of language proficiency. However, another possibility is that those with a later ASL AoA could show larger Stroop effects due to the detrimental effects of language deprivation on executive control abilities (Botting et al., 2017; Merchán et al., 2022).

Finally, given the above-mentioned meta-analysis by Nicosia et al. (2021) which concluded that Stroop effects are significantly larger in older than in younger hearing adults, even after controlling for general aging-related slowing in response times we expected that older deaf bilinguals would show larger Stroop interference effects compared to younger deaf bilinguals. However, it is unknown whether aging effects on Stroop performance will be equal for both languages or larger in one language than the other.

To characterize Stroop performance in deaf ASL-English bilinguals across a wide age range we administered Stroop tasks using both ASL and English stimuli while allowing participants to respond with ASL (the preferred and dominant language of communication for many deaf individuals). We aimed to understand how testing language and modality influences Stroop performance, potentially revealing differences between how unimodal and bimodal

bilinguals engage in cognitive control when asked to complete a between-language Stroop task. Results of this study may also inform clinical applications of the Stroop task with this population, particularly which language may be better for presentation of the task. Finally, understanding how aging impacts Stroop performance in deaf older adults is critical for interpreting Stroop performance as a possible diagnostic tool to identify the earliest signs of cognitive decline, and to differentiate changes associated with typical aging from those of disordered aging (e.g., AD) in deaf signers.

Methods

Participants

One hundred twelve deaf ASL-English bilinguals ranging in age from 21-88 completed two computerized Stroop tasks, one with videos of dynamic ASL color signs and one with printed English color words. Three participants were excluded from analysis due to self-reported difficulties with memory and struggles to complete many tasks presented on the day of examination. Sixty-four participants were tested in San Diego or Riverside, CA, and 45 were tested in Rochester, NY. Participants were recruited through presentations at Deaf community events, communication with Deaf senior citizen centers, and through flyers distributed at community events. Participants self-reported their race and ethnicity in a background questionnaire. The race breakdown for the participants was 87% white (95/109), 3.7 % black (4/109), 1.8 % Asian (2/109), 5.5 % reported more than one race (6/109), and 1.8 % declined to answer (2/109). The ethnicity breakdown was 11 % Hispanic or Latino (112/109), 85 % not Hispanic or Latino (93/109), and 4% of participants reported that their ethnicity was unknown or declined to answer (4/109). The reported gender breakdown was 53% female (58/109), 45% male (49/109), and 1.8% preferred not to say (2/109). Participants were divided into four age

groups: Young, ranging from age 21-45 ($n = 34$); Middle, from 46-63 ($n = 17$); Young-Old, from 64-70 ($n = 29$); and Older, from 71-88 ($n = 29$). The cutoffs for the young and middle age groups followed those reported in our previous study on list memory performance in this population (Camilliere et al., 2026). The cutoffs for the young-old and older groups were determined by conducting a median split on the group of participants above the age of 64. All participants reported that they use ASL in their daily life and as a primary language⁴.

This research protocol was approved by the University of California, San Diego (UCSD) Institutional Review Board in accordance with the Helsinki Declaration. All participants were given informed consent and were given monetary compensation for participation.

Procedure and Stimuli

Participants completed the Stroop tasks as part of a larger battery of cognitive and language tasks for deaf ASL-English participants. Following a comprehensive language background and demographics questionnaire, participants completed three testing blocks. Stroop tasks were administered following a ten-item list memory task (conducted in the same language; Camilliere et al., 2026). Language of testing order (first, second) was counterbalanced across participants, such that they completed the memory and Stroop tasks in English then both tasks in ASL, or vice versa.

Testing was conducted in ASL by a signing researcher. The ASL and English Stroop tasks were presented via PsychoPy v2.2 (Peirce et al., 2019) in a counterbalanced order between

⁴ A reviewer asked which other languages our participants used in addition to ASL and English. Of our 109 participants, 23 reported knowing another signed/written/spoken language, including Spanish, Chinese, French, British Sign Language, Mexican Sign Language, and others. Of those 23 only one reported using their other language for 50-75% of the time during their everyday life, all others reported using the other language 0-XX% of the time. Our primary inclusion criteria for language was ASL or English as the first acquired language, and daily use ASL and English.

participants. Task instructions were given in both ASL and written English for each task, and participants were instructed to sign the color of the hand/font while ignoring the sign/word. Initially, participants were instructed to respond in ASL for the ASL task and were given a choice between speaking and signing their responses for the English task. This choice was necessary as many deaf individuals do not use speech in their everyday life, and speaking would be an unnatural response mode for those participants. After the first 25 participants (who all preferred to use ASL to respond), we decided it was not necessary to give the choice to respond in English and instructed participants to respond in ASL. Participants completed a 12-item practice block in each language to familiarize themselves with the procedure, during which they were given feedback on their performance. For each trial, participants were instructed to depress the spacebar until the stimuli appeared on the screen. As soon as they identified the color of the hand/font, participants were instructed to lift their hand and sign their response. The stimuli remained on the screen for 2 s or until the spacebar was released. Trials were self-paced, as trial initiation was triggered by the participant depressing the space bar. Reaction time was recorded from the stimulus onset to the response onset. Response onset was indicated by the release of the spacebar, which was taken as a proxy for the onset of sign production (see Emmorey et al., 2012). Response accuracy was scored after the experiment by two ASL-fluent coders with reference to the video recording. Ten percent of data was independently scored by each coder to establish interrater reliability, and scores were in agreement for 98.9 percent of trials. Trials were coded as errors if the participant produced an incorrect handshape or sign following the release of the spacebar. Hesitations, in which a participant released their hand from the space bar and produced a neutral handshape with a marked pause, were not scored as errors but were removed from reaction time analyses. Hesitations accounted for 2% of total trials.

ASL stimuli for this experiment were adapted from Bosworth et al. (2021) and included three ASL color signs (BLUE, GREEN, and YELLOW) and their corresponding English words (see Figure 3.1). These ASL color signs are articulated in the same neutral space and differ only in handshape. English stimuli were presented in Impact font in the center of the screen. ASL videos were filmed and cropped to match the size of the English print stimuli (52 point font). Nine stimuli were created for each language, with three congruent color-word/sign stimuli and six incongruent color-word/sign stimuli. Each participant saw 120 trials in each language, with 84 congruent and 36 incongruent trials. This ratio was selected following a previous study that found larger Stroop effects in participants with MCI (Mild cognitive impairment) and AD (Alzheimer's Disease) compared to healthy controls when fewer incongruent than congruent trials were presented (Bélanger et al., 2010)

As part of the larger testing battery, participants also completed several tests of ASL and English proficiency, including the American Sign Language Comprehension Test (ASL-CT; Hauser et al., 2016) and the English version of the Multilingual Naming Test (MINT Sprint 2.0 version; Gollan et al., 2024). The ASL-CT consists of 30 multiple choice questions probing understanding of ASL complex signs and sentences. The English MINT Sprint consists of 80 pictures that participants were asked to name in English with the option to speak, fingerspell, or write their responses.⁵ Of the 107 participants with English MINT scores available, 87 chose to fingerspell (81%), 13 chose to speak (12%), and 7 chose to write (7%).

⁵ ASL-CT scores were unavailable for one participant, and English MINT Sprint scores were unavailable for two participants due to loss of video data during administration.



Figure 3.1. ASL and English Stimuli

Note. The ASL signs show still images taken from the video stimuli of color signs used. The English stimuli matching the ASL signs are shown above each image. The top row (a) shows examples of congruent stimuli in ASL and English, and the bottom row (b) shows examples of incongruent stimuli in ASL and English.

Table 3.1. Participant characteristics divided into four age-groups

	Younger (<i>n</i> = 34)		Middle (<i>n</i> = 17)		Younger-Old (<i>n</i> = 28)		Older (<i>n</i> = 29)	
Characteristic	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Age	33.6	6.8	56.2	5.6	67.4	2.0	75.8	4.6
Years Education	16.1	2.7	14.9	2.4	16.5	2.7	15.6	2.4
ASL AoA ^A	3.3	5.5	4.6	5.5	8.4	10.8	9.8	7.2
ASL CT score ^B	25.2	2.3	24.2	3.9	22.6	3.1	20.8	4.0
English MINT Sprint 2.0 ^C	67.9	9.0	69.2	5.5	68.4	9.2	69.1	5.9

Note. ^AAoA = reported age of acquisition of ASL. ^BASL CT = ASL Comprehension Test. ^CTotal number of pictures named correctly in English (out of 80) on the Sprint 2.0 version of the Multilingual Naming Test (MINT).

Results

Analysis Plan

Numbers shown in Table 3.1 reflect means and SDs after removing one participant who was excluded from analysis for producing errors on 55% of incongruent trials in the ASL Stroop task (this was more than 5 *SDs* greater than the mean). Preliminary exploratory correlations including the remaining $n=108$ participants and are shown in Table 3.2. The outcome variables used to define Stroop performance were baseline adjusted Stroop response times (RTs) [Baseline adjusted Stroop RT = ((Incongruent RT – Congruent RT)/Incongruent RT)] and Stroop error production (i.e., proportion errors on incongruent trials – proportion errors on congruent trials). These analyses revealed positive correlations between the size of English Stroop effects in baseline adjusted RTs and age (see Figure 3.2), but no significant correlations with ASL Stroop effects in baseline adjusted RTs⁶, Stroop effects in error rates in either language, age of acquisition of ASL, ASL proficiency (measured by the ASL Comprehension Test; Hauser et al., 2016) or English proficiency (measured by the English MINT Sprint 2.0; Gollan et al., 2024; Neveu et al., 2025). See Figure 3.3 for an illustration of the null correlation between proficiency and Stroop performance.

To further illustrate the effects of age and stimulus language we conducted 2 x 2 ANOVAs with RTs and errors as dependent variables, age as a continuous covariate, and stimulus language (ASL, English) and trial type (incongruent, congruent) as within-subjects factors. Finally, to ensure that the effects we observed were not driven by baseline differences in

⁶ When including the one participant with very high error rates this correlation was significant, but still relatively smaller in size than the analogous correlation with baseline adjusted English RTs, $r=.2$ $p=.03$.

response speed, with baseline adjusted Stroop effects as the dependent variable instead of raw RTs.

Table 3.2. Correlations between demographics, proficiency tests and Stroop measures (n=108)

	Age	Education	ASL AoA ^A	ASL CT ^B	English MINT ^C	ASL Stroop RT ^D	English Stroop RT	ASL Stroop Errors
Age								
Education	-0.02							
ASL AoA	0.34***	-0.05						
ASL CT	-0.45***	0.17	-0.36***					
English MINT	0.04	0.42***	0.06	0.21*				
ASL Stroop RT	0.19	0.04	-0.01	0.002	0.09			
English Stroop RT	0.33***	0.01	0.15	-0.01	0.01	0.44***		
ASL Stroop Errors	0.02	-0.04	0.02	-0.14	-0.17	0.10	-0.28**	
English Stroop Errors	0.09	0.05	-0.04	-0.13	0.01	0.01	-0.27**	0.57***

Note. ^A AoA = reported age of acquisition of ASL. ^B ASL CT = ASL Comprehension Test. ^C Total number of pictures named correctly in English (out of 80) on the Sprint 2.0 version of the Multilingual Naming Test (MINT). ^D Stroop RTs reported here are baseline adjusted Stroop RT Effects ((Incongruent RT – Congruent RT)/Incongruent RT)). *** $p \leq .001$, ** $p \leq .01$, * $p < .05$.

Response Times (RTs)

The analysis of the raw RTs revealed robust Stroop effects; incongruent trials elicited slower RTs than congruent trials, $F(1,104) = 168.7$; $\eta_p^2 = 0.62$; $p < 0.001$, a main effect of trial type.

Additionally, participants responded more slowly to ASL than to English stimuli, $F(1, 104) = 29.5$; $\eta_p^2 = 0.22$; $p < 0.001$, a main effect of language. Stimulus language also modulated the size

of the Stroop effect, which was larger with English than with ASL stimuli, indicated by a significant interaction between trial type and language, $F(1, 104) = 16.1$; $\eta_p^2 = 0.13$; $p < 0.001$. Older participants tended to respond more slowly overall, a significant main effect of age, $F(1, 104) = 7$; $\eta_p^2 = 0.06$; $p = 0.009$. Additionally, Stroop effects were significantly larger in older than in younger participants, a significant interaction between trial type and age, $F(1, 104) = 12.2$; $\eta_p^2 = 0.11$; $p < 0.001$, and the magnitude of this increase was larger for English compared to ASL stimuli, a significant three way interaction between trial type, age group, and language, $F(1, 104) = 6.3$; $\eta_p^2 = 0.06$; $p = 0.01$. Both young and older participants responded slower with ASL than with English stimuli, i.e., there was no interaction between age and language, $F < 1$.

Analysis of baseline adjusted Stroop RT effects showed that the main effects survived the control for baseline response speed, including the effect of age, $F(1, 104) = 12.3$; $\eta_p^2 = 0.11$; $p < 0.001$, and stimulus language, $F(1, 104) = 20.7$; $\eta_p^2 = 0.17$; $p < 0.001$. In fact, Stroop effects were especially larger with English than with ASL stimuli, after controlling for baseline speed (given that ASL elicited slower response times, Stroop effects were proportionally even smaller, see Figure 3.4). As reported for raw RTs, age effects were larger for English than ASL stimuli, i.e., there was a significant interaction between age and language in baseline adjusted Stroop effects, $F(1, 104) = 5.8$; $\eta_p^2 = 0.05$; $p = 0.018$.

Error Rates

The analysis of error rates showed robust Stroop interference effects, evidenced by more errors on incongruent trials compared to congruent trials, $F(1, 104) = 109.8$; $\eta_p^2 = 0.51$; $p < 0.001$, a main effect of trial type. Additionally, participants produced more errors when stimuli were presented in English versus in ASL, $F(1, 104) = 3.9$; $\eta_p^2 = 0.04$; $p = 0.05$, a main effect of language. Young and older participants had similar error rates, $F < 1$, but participants showed

larger Stroop effects in English compared to in ASL, producing more errors on incongruent compared to congruent trials in English than in ASL, $F(1, 104) = 4.6$; $\eta_p^2 = 0.04$; $p = 0.03$, a significant interaction between trial type and language. None of the other interactions were significant (all F s < 1 for interactions between age and language, age and trial type, and the 3-way interaction).

Table 3.3. Means (and SDs) for raw RTs and percent errors by condition in each age group

	Younger ($n = 34$)		Middle ($n = 17$)		Younger-Old ($n = 29$)		Older ($n = 29$)	
Variable	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
ASL								
Congruent RT	1655	204	1805	334	1746	323	1807	335
Incongruent RT	1707	237	1873	311	1836	368	1892	362
Stroop effect in ms	53	56	68	73	90	98	85	59
Congruent percent errors	0.4	0.9	0.4	0.8	0.2	0.6	0.5	1.1
Incongruent percent errors	8	7.1	7.1	7.2	10.8	11.8	6.8	7.3
Stroop effect in errors	7.7	6.8	6.7	6.9	10.6	11.7	6.4	6.6
English								
Congruent RT	1541	153	1641	190	1599	150	1644	144
Incongruent RT	1608	195	1728	250	1696	241	1803	225
Stroop effect in ms	67	55	87	86	97	119	158	110
Congruent percent errors	0.4	0.9	0.4	0.7	0.1	0.4	0.4	1.4
Incongruent percent errors	9	10.1	7.9	9.9	14.1	11.5	9.4	13
Stroop effect in errors	8.6	10.2	7.5	10	14	11.5	9	12.1

Note. RT = response time; ms = milliseconds; Stroop effect = Incongruent trials minus Congruent trials in raw RTs (for baseline adjusted Stroop effects see Figure 3.2).

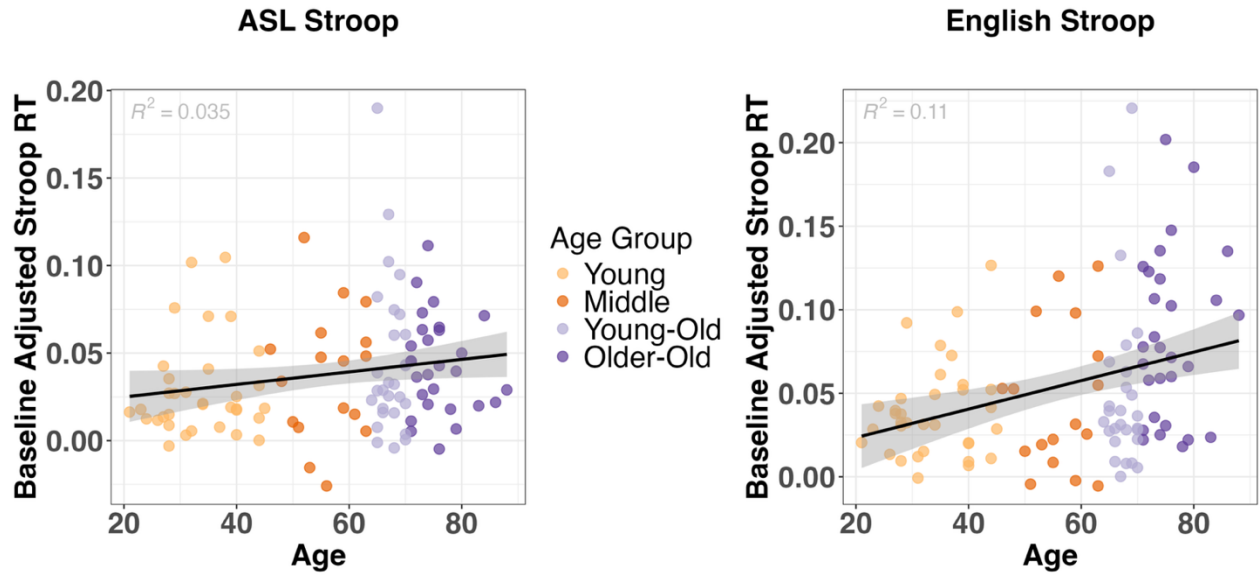


Figure 3.2. Scatterplots illustrating the effects of age on baseline adjusted Stroop effects in RTs with ASL vs. with English stimuli (n=108)

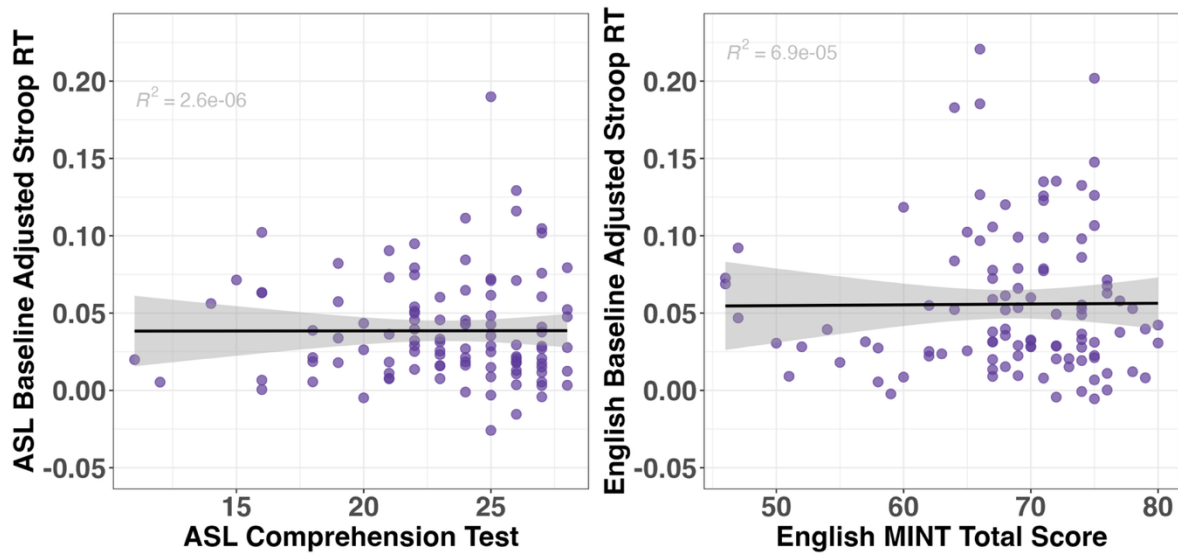


Figure 3.3. Scatterplots illustrating the null relationship between English proficiency and English adjusted baseline Stroop score (left) and the null relationship between ASL proficiency and ASL adjusted baseline Stroop score (right)

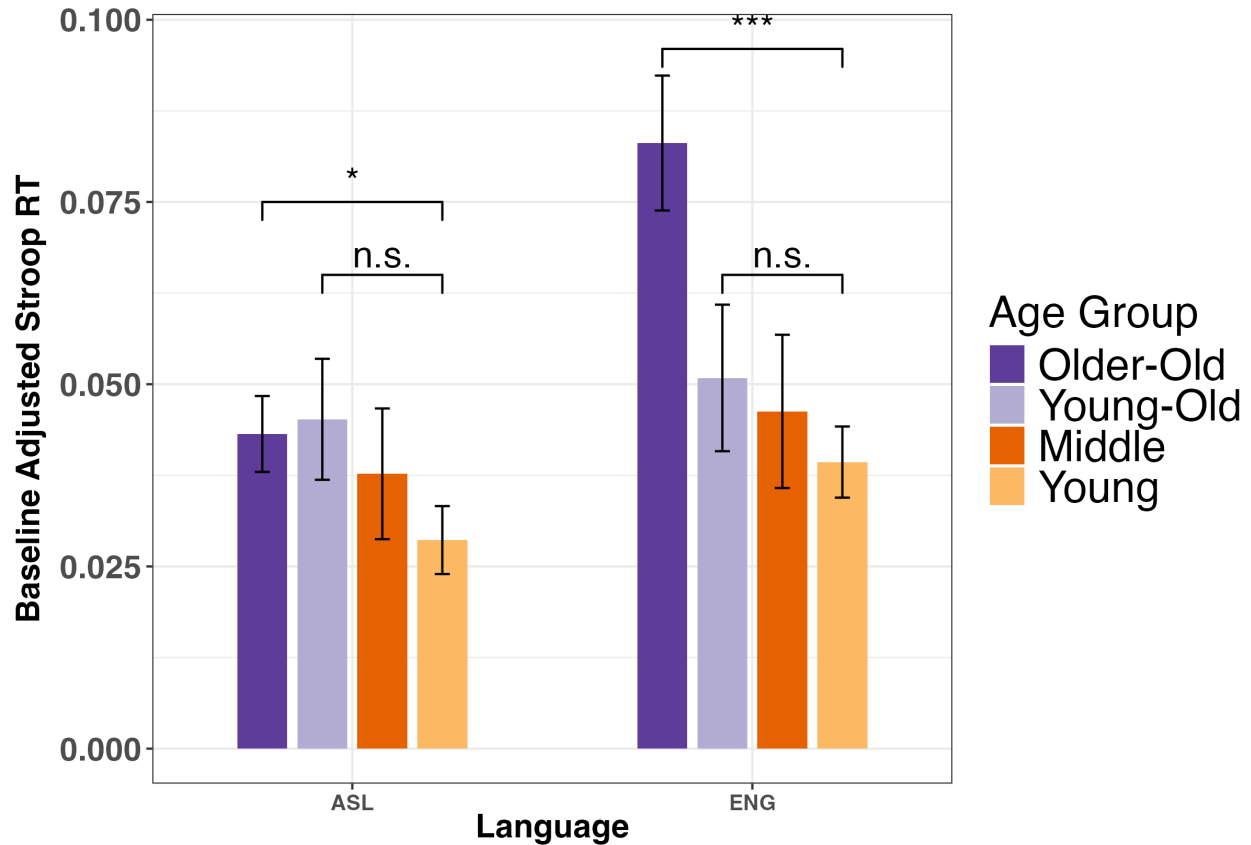


Figure 3.4. Baseline adjusted Stroop effect in RTs in four age groups (see Table 3.1) with ASL and English stimuli

Note. *** $p \leq .001$, ** $p \leq .01$, * $p < .05$. Error bars denote standard error.

Discussion

Summary of Key Results

The present study revealed robust Stroop effects both within ASL (i.e., ASL stimuli with ASL responses) and between languages (i.e., English stimuli with ASL responses). Of greatest interest, Stroop effects were larger with English than with ASL stimuli for both older and younger participants, an effect of stimulus language that was significant in both RTs and errors. Additionally, participants responded more slowly when stimuli were presented in ASL compared to written English words. The language difference was even larger after adjusting for baseline response speed. Older participants showed significantly larger Stroop effects than younger

participants in RTs (see Tables 3.2-3.3; Figures 3.2 and 3.4), but not in error rates (see Table 3.2-3.3). The age-related increase in Stroop effect in the RTs survived control for baseline speed (which was marginally slower overall in older than in younger participants). There was a significant interaction between language and age with older participants showing larger baseline adjusted Stroop effects in English than ASL, also illustrated by a significant correlation between age and baseline adjusted Stroop RT effects in English that was smaller and not significant in ASL (see Table 3.2). Finally, ASL and English proficiency did not appear to be related to Stroop performance, as there were no significant correlations between Stroop RTs or errors and scores on independent measures of ASL and English proficiency.

Testing Language Effects on Stroop Performance

Our observation of robust Stroop effects with ASL and with English stimuli (i.e., longer RTs and more errors on incongruent than congruent trials) replicate those of Dupuis & Berent (2015) for ASL and Bosworth et al. (2021) for both ASL and English. These findings indicate that deaf ASL-English bilinguals rapidly and automatically access the meaning of both ASL signs and written English words. The present study was the first to implement a between-language condition for deaf bilinguals in which the response language did not match the language of the stimuli, revealing the presence of between-language Stroop effects in bimodal bilinguals and replicating the existence of between-language Stroop effects previously found in unimodal bilinguals (Rosselli et al., 2002; Tse & Altarriba, 2012).

Surprisingly, the between-language (English-ASL) Stroop effect was also *larger* than the within-language (ASL-ASL) Stroop effect (in both RTs and errors), i.e., Stroop effects were larger when deaf participants signed responses to English compared to ASL stimuli. This pattern of results has not been seen in previous studies of unimodal bilinguals who either showed

equivalent Stroop effects within- vs. between-languages, or *smaller* between-language than within-language Stroop effects (Fang et al., 1981; Tzelgov et al., 1990). Bosworth et al. (2021) did not test between-language Stroop effects – participants in that study responded with button presses for both the ASL and English stimuli. However, Stroop effects were larger with English than with ASL stimuli, which resembles the pattern of results reported here. A key difference between unimodal and bimodal bilinguals that might explain the differences in the relative size of Stroop effects with ASL vs English stimuli may be that processing written words – which are presented as a whole – is faster than processing signs which unfold more slowly over time. Indeed, ASL stimuli elicited slower responses overall compared to English stimuli. Recognition times for spoken words, which also unfold over time, are also typically longer than for printed words (Carver, 1990). Dynamic ASL signs may allow more time to control interference effects when responding compared to static English words, thus leading to weaker within-language Stroop effects.

Larger Stroop effects in English than in ASL may also be due to increased competition between activated lexical items. Previous studies have shown that deaf signers automatically access ASL signs when reading printed English (Meade et al., 2017; J. P. Morford et al., 2011; Saunders et al., 2025). Thus, in the English-ASL condition participants may be suppressing competition from both the incongruent printed English word *and* its ASL translation. In contrast, for the within-language (ASL-ASL) condition, participants only need to inhibit producing the ASL color sign that was presented. Additionally, even if a deaf individual's first language is ASL, their first *written* language is typically English, and all of their subsequent reading and writing is in English. This situation may create a different type of proficiency profile for deaf bilinguals compared to hearing bilinguals. In fact, several studies comparing skilled deaf and

hearing readers have found that deaf signers are more efficient in processing print (N. N. Bélanger & Rayner, 2015; Cooley et al., 2025; Schotter et al., 2024), and make faster word and non-word decisions (Fariña et al., 2017; Lee et al., 2022; J. Morford et al., 2015). Thus, we speculate that deaf bilinguals make more errors and have longer RTs with English stimuli because they process English words faster than ASL signs, and this leads to early co-activation of ASL color sign translations which compete with the ASL signs that must be retrieved to name the ink color.

Finally, a critical difference between the current study and previous studies investigating Stroop performance in unimodal bilinguals is that the English task in the current study also represents a modality difference between the stimulus and the response, making this condition both a cross-modal and a between-language task. Thus, the differences seen in Stroop effects for the within-language (ASL-ASL) and between-language (English-ASL) conditions could be related to the modality difference in the stimuli (as argued above) or to the language modality difference for the between-language condition. A future study could control for modality by testing Stroop performance in deaf signers who are bilingual in two sign languages (e.g., ASL and DGS - German Sign Language). The between-language condition (ASL-DGS) for these bilinguals would be more comparable to unimodal between-language conditions, and the differences related to static text vs dynamic videos would be removed.

Aging Effects on Stroop Performance

Older participants showed larger Stroop effects in the analysis of both raw RTs and baseline adjusted RTs compared to younger participants, aligning with previous research that has identified a robust aging related increase in Stroop effects (Nicosia et al., 2021). As in previous studies with hearing participants, aging effects survived a control for baseline response speed,

and unique to the present study, aging effects were larger in English than in ASL, possibly for some of the same reasons as noted above to explain why Stroop effects were larger in English. Finally, it is notable that older participants did not produce more errors on incongruent trials than younger participants. Stroop errors have been found to be particularly sensitive for predicting conversion to AD in hearing participants (Balota et al., 2010). The lack of aging effects on error rates in the current sample presents raises further questions to explore in future studies comparing patients to healthy controls to see if Stroop errors are predictive of cognitive decline in this population.

Proficiency and AoA Effects on Stroop Performance

Though previous studies with unimodal bilinguals (Rosselli et al., 2002; Tse & Altarriba, 2012) and bimodal bilinguals (Bosworth et al., 2021) have found an association between language proficiency and the magnitude of the Stroop effect in reaction times, there was no evidence for this relationship in the current group of ASL-English bilinguals. Importantly, we used the same ASL proficiency measure as Bosworth et al. (2021), and we included a similar (or wider) range in ASL-CT scores. Their scores ranged from 67-97% ($M=85\%$; $SD=31$) whereas our participants' scores ranged from 36%-93% ($M=79\%$; $SD=12$). The present study also had a much larger number of participants with ASL-CT scores ($n=108$ versus $n=14$, one participant from Bosworth et al. was missing an ASL-CT score). The absence of a relationship between proficiency level and Stroop effects is consistent with Tzelgov et al.'s (1990) suggestion that Stroop effects vary little in magnitude once a minimal level of proficiency level is reached as participants in the present study regularly used both ASL and English to communicate in their daily lives.

Age of acquisition (AoA) of ASL increased steadily across age groups from 3.3 years for the young signers to 9.8 years for the older signers (see Table 3.1). This difference most likely reflects changes in US education and in the acceptance of ASL as a language between the 1940s and 1950s when the oldest signers attended school and the 1990s/2000s when the youngest signers began pre-school (Humphries, 2013; Lane, 1984). Bilingual pre-school and kindergarten programs were not available for older signers who typically learned ASL from deaf signing peers when they entered elementary or junior high school. Importantly, we found no effect of AoA on any of our Stroop effect measures (see Table 3.2), suggesting no cognitive impact from potential language deprivation that can occur with late AoA. However, AoA was correlated with ASL proficiency – those with a later AoA performed worse on the ASL Comprehension Test. This result is consistent with many studies that report poorer ASL skill with later AoA, even when acquisition occurs during childhood (see Mayberry & Kluender, 2018, for review).

Limitations

Although the current sample size is larger than those of any previous study investigating Stroop effects in ASL-English bilinguals, it is still relatively small and more research is necessary to determine if these findings hold across deaf participants with more varied backgrounds. The current sample was relatively homogenous in terms of education (i.e., primarily individuals with a high school education and some college education), race, and ethnicity (i.e., primarily individuals who identified as white). Additionally, while both measures of language proficiency are established, well-normed tasks, they are not directly comparable to each other, i.e., picture naming (English MINT Sprint 2.0) vs. sentence comprehension (ASL-CT).

Conclusions

The results of the current study present a more in-depth understanding of deaf ASL-English bilinguals' performance on Stroop tasks in each of their languages. Overall response times were slower with ASL stimuli, likely due to modality differences between dynamic sign videos and static printed words. However, participants demonstrated larger Stroop interference with English stimuli, in both reaction time and error rates. There was also an effect of aging on Stroop performance (reaction times) for both English and ASL, but the aging effect was larger in English, the non-dominant language for these participants, than it was in ASL. This pattern of results is unexpected given that standard recommended practice for unimodal bilinguals is to administer tests in the dominant language, which will generally be more sensitive for detecting group differences (Gollan et al., 2024). Understanding how typical aging influences Stroop performance is a critical first step in determining whether a Stroop task can be used for identifying signs of cognitive decline within populations of deaf ASL-English bilinguals. Our finding that English stimuli elicited larger Stroop effects than ASL stimuli could be useful in clinical practice. Currently, English Stroop tasks are more easily accessible in that many clinicians already administer tasks with English Stroop stimuli, and the materials are easier to create given that they only require presenting static English words while ASL Stroop stimuli require presenting dynamic videos of signs. The current study suggests that, so long as participants are allowed to respond in ASL, it may be possible to detect aging effects using the Stroop task in English with deaf participants. However, to know which version of the Stroop test will be sensitive to cognitive decline in older deaf ASL-English bilinguals requires both longitudinal and patient studies.

Acknowledgements

Chapter 3 has been submitted for publication of the material and is currently under revised peer review. Camilliere, S., Emmorey, K., Hauser, P. C., Contreras, J., McKee, M. M., & Gollan, T. H. (Under review). Bigger between-language than within-language Stroop effects in young and older deaf American Sign Language-English bilinguals. *Journal of International Neuropsychological Society*. The dissertation author was the primary investigator and author of this paper. This research was supported by NIH grant R01-AG074928. The authors have no conflicts of interest to disclose. We would like to thank Lucinda Farnady and Julie Fisher for their help with participant recruitment and data collection; Chehally Soto, Emma Kane, Hayden Orr, and Gigi Zheng for their help with data collection; and Caitlin Pearson for her help with data coding. We would also like to thank all of our deaf participants.

References

- Balota, D. A., Tse, C.-S., Hutchison, K. A., Spieler, D. H., Duchek, J. M., & Morris, J. C. (2010). Predicting conversion to dementia of the Alzheimer's type in a healthy control sample: The power of errors in stroop color naming. *Psychology and Aging, 25*(1), 208–218. <https://doi.org/10.1037/a0017474>
- Bélanger, N. N., & Rayner, K. (2015). What Eye Movements Reveal About Deaf Readers. *Current Directions in Psychological Science, 24*(3), 220–226. <https://doi.org/10.1177/0963721414567527>
- Bélanger, S., Belleville, S., & Gauthier, S. (2010). Inhibition impairments in Alzheimer's disease, mild cognitive impairment and healthy aging: Effect of congruency proportion in a Stroop task. *Neuropsychologia, 48*(2), 581–590. <https://doi.org/10.1016/j.neuropsychologia.2009.10.021>
- Belleville, S., Fouquet, C., Hudon, C., Zomahoun, H. T. V., Croteau, J., & Consortium for the Early Identification of Alzheimer's disease-Quebec. (2017). Neuropsychological Measures that Predict Progression from Mild Cognitive Impairment to Alzheimer's type dementia in Older Adults: A Systematic Review and Meta-Analysis. *Neuropsychology Review, 27*(4), 328–353. <https://doi.org/10.1007/s11065-017-9361-5>
- Belleville, S., Rouleau, N., & Van der Linden, M. (2006). Use of the Hayling task to measure inhibition of prepotent responses in normal aging and Alzheimer's disease. *Brain and Cognition, 62*(2), 113–119. <https://doi.org/10.1016/j.bandc.2006.04.006>
- Bosworth, R. G., Binder, E. M., Tyler, S. C., & Morford, J. P. (2021). Automaticity of lexical access in deaf and hearing bilinguals: Cross-linguistic evidence from the color Stroop task across five languages. *Cognition, 212*, 104659. <https://doi.org/10.1016/j.cognition.2021.104659>
- Botting, N., Jones, A., Marshall, C., Denmark, T., Atkinson, J., & Morgan, G. (2017). Nonverbal Executive Function is Mediated by Language: A Study of Deaf and Hearing Children. *Child Development, 88*(5), 1689–1700. <https://doi.org/10.1111/cdev.12659>
- Brown, T. L., Joneleit, K., Robinson, C. S., & Brown, C. R. (2002). Automaticity in Reading and the Stroop Task: Testing the Limits of Involuntary Word Processing. *The American Journal of Psychology, 115*(4), 515–543. <https://doi.org/10.2307/1423526>
- Camilliere, S., Emmorey, K., Hauser, P. C., Contreras, J., McKee, M. M., & Gollan, T. H. (2026). Influence of testing language and aging on verbal list memory in deaf American Sign Language–English bilinguals. *Neuropsychology, 40*(4), 402–417. <https://doi.org/10.1037/neu0001065>
- Carver, R. P. (1990). *Reading rate: A review of research and theory* (pp. x, 514). Academic Press.

- Cohn, N. B., Dustman, R. E., & Bradford, D. C. (1984). Age-related decrements in stroop color test performance. *Journal of Clinical Psychology*, 40(5), 1244–1250. [https://doi.org/10.1002/1097-4679\(198409\)40:5%3C1244::AID-JCLP2270400521%3E3.0.CO;2-D](https://doi.org/10.1002/1097-4679(198409)40:5%3C1244::AID-JCLP2270400521%3E3.0.CO;2-D)
- Comalli Jr., P. E., Wapner, S., & Werner, H. (1962). Interference Effects of Stroop Color-Word Test in Childhood, Adulthood, and Aging. *The Journal of Genetic Psychology*, 100(1), 47–53. <https://doi.org/10.1080/00221325.1962.10533572>
- Cooley, F. G., Emmorey, K., Saunders, E., Sinclair, G., Stringer, C., & Schotter, E. R. (2025). Identifying text-based factors that contribute to the superior reading efficiency of skilled deaf readers: An eye-tracking study of length, frequency, and predictability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 51(7), 1178–1189. <https://doi.org/10.1037/xlm0001415>
- Dupuis, A., & Berent, I. (2015). Signs are symbols: Evidence from the Stroop task. *Language, Cognition and Neuroscience*, 30(10), 1339–1344. <https://doi.org/10.1080/23273798.2015.1069363>
- Emmorey, K., Petrich, J. A. F., & Gollan, T. H. (2012). Bilingual processing of ASL–English code-blends: The consequences of accessing two lexical representations simultaneously. *Journal of Memory and Language*, 67(1), 199–210. <https://doi.org/10.1016/j.jml.2012.04.005>
- Fang, S.-P., Tzeng, O. J. L., & Alva, L. (1981). Intralanguage vs. Interlanguage Stroop effects in two types of writing systems. *Memory & Cognition*, 9(6), 609–617. <https://doi.org/10.3758/BF03202355>
- Fariña, N., Duñabeitia, J. A., & Carreiras, M. (2017). Phonological and orthographic coding in deaf skilled readers. *Cognition*, 168, 27–33. <https://doi.org/10.1016/j.cognition.2017.06.015>
- Gollan, T. H., Garcia, D. L., Murillo, M., Vargas, J., Pulido, B., & Salmon, D. P. (2024). Sprinting in two languages: Picture naming performance of older Spanish–English bilinguals on the Multilingual Naming Test Sprint 2.0. *Neuropsychology*, 38(7), 653–664. <https://doi.org/10.1037/neu0000958>
- Hall, W. C., Levin, L. L., & Anderson, M. L. (2017). Language deprivation syndrome: A possible neurodevelopmental disorder with sociocultural origins. *Social Psychiatry and Psychiatric Epidemiology*, 52(6), 761–776. <https://doi.org/10.1007/s00127-017-1351-7>
- Hauser, P. C., Paludnevičienė, R., Riddle, W., Kurz, K. B., Emmorey, K., & Contreras, J. (2016). American Sign Language Comprehension Test: A Tool for Sign Language Researchers. *The Journal of Deaf Studies and Deaf Education*, 21(1), 64–69. <https://doi.org/10.1093/deafed/env051>

- Humphries, T. (2013). Schooling in American Sign Language: A paradigm shift from a deficit model to a bilingual model in deaf education. *Berkeley Review of Education*, 4(1). <https://doi.org/10.5070/B84110031>
- Lane, H. (1984). *When the mind hears: A history of the deaf*. Random House.
- Lee, B., Martinez, P. M., Midgley, K. J., Holcomb, P. J., & Emmorey, K. (2022). Sensitivity to orthographic vs. phonological constraints on word recognition: An ERP study with deaf and hearing readers. *Neuropsychologia*, 177, 108420. <https://doi.org/10.1016/j.neuropsychologia.2022.108420>
- MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, 109(2), 163–203. <https://doi.org/10.1037/0033-2909.109.2.163>
- Marschark, M. (1988). Automaticity in Word & Sign Recognition by Deaf Adults & Children. *Sign Language Studies*, 58(1), 1–19. <https://doi.org/10.1353/sls.1988.0000>
- Marschark, M., & Shroyer, E. H. (1993). Hearing Status and Language Fluency As Predictors of Automatic Word and Sign Recognition. *American Annals of the Deaf*, 138(4), 370–375. <https://doi.org/10.1353/aad.2012.0381>
- Mayberry, R. I. (2002). Cognitive development in deaf children: The interface of language and perception in neuropsychology. In S.J. Segalowitz & I. Rapin (Eds), *Handbook of neuropsychology*, 8(Part II), pp. 71-107; Elsevier.
- Mayberry, R. I., & Kluender, R. (2018). Rethinking the critical period for language: New insights into an old question from American Sign Language. *Bilingualism* (Cambridge, England), 21(5), 938–944. <https://doi.org/10.1017/S1366728918000585>
- Meade, G., Midgley, K. J., Sevcikova Sehyr, Z., Holcomb, P. J., & Emmorey, K. (2017). Implicit co-activation of American Sign Language in deaf readers: An ERP study. *Brain and Language*, 170, 50–61. <https://doi.org/10.1016/j.bandl.2017.03.004>
- Merchán, A., Fernández García, L., Gioiosa Maurno, N., Ruiz Castañeda, P., & Daza González, M. T. (2022). Executive functions in deaf and hearing children: The mediating role of language skills in inhibitory control. *Journal of Experimental Child Psychology*, 218, 105374. <https://doi.org/10.1016/j.jecp.2022.105374>
- Mitchell, R. E., & Karchmer, M. A. (2004). Chasing the Mythical Ten Percent: Parental Hearing Status of Deaf and Hard of Hearing Students in the United States. *Sign Language Studies*, 4(2), 138–163.
- Morford, J., Occhino, C., Piñar, P., Wilkinson, E., & Kroll, J. (2015). The time course of cross-language activation in deaf ASL–English bilinguals. *Bilingualism: Language and Cognition*, 20, 337–350. <https://doi.org/10.1017/S136672891500067X>

- Morford, J. P., Wilkinson, E., Villwock, A., Piñar, P., & Kroll, J. F. (2011). When deaf signers read English: Do written words activate their sign translations? *Cognition*, 118(2), 286–292. <https://doi.org/10.1016/j.cognition.2010.11.006>
- Naylor, L. J., Stanley, E. M., & Wicha, N. Y. Y. (2012). Cognitive and Electrophysiological Correlates of the Bilingual Stroop Effect. *Frontiers in Psychology*, 3. <https://doi.org/10.3389/fpsyg.2012.00081>
- Neveu, A., Garcia, D. L., Escobedo, B., Vazquez, P. E., Mejia, M., Hoversten, L. J., & Gollan, T. H. (2025). Predicting proficiency. *Bilingualism: Language and Cognition*, 1–18. <https://doi.org/10.1017/S1366728925000367>
- Nicosia, J., Cohen-Shikora, E. R., & Balota, D. A. (2021). Re-examining age differences in the Stroop effect: The importance of the trees in the forest (plot). *Psychology and Aging*, 36(2), 214–231. <https://doi.org/10.1037/pag0000599>
- Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 51(1), 195–203. <https://doi.org/10.3758/s13428-018-01193-y>
- Rosselli, M., Ardila, A., Santisi, M. N., Arecco, M. D. R., Salvatierra, J., Conde, A., & Lenis, A. B. (2002). Stroop effect in Spanish–English bilinguals. *Journal of the International Neuropsychological Society*, 8(6), 819–827. <https://doi.org/10.1017/S1355617702860106>
- Saunders, E., Mirault, J., & Emmorey, K. (2025). Activation of ASL signs during sentence reading for deaf readers: Evidence from eye-tracking. *Bilingualism: Language and Cognition*, 28(1), 208–216. <https://doi.org/10.1017/S1366728924000336>
- Schotter, E. R., Stringer, C., Saunders, E., Cooley, F. G., Sinclair, G., & Emmorey, K. (2024). The role of perceptual and word identification spans in reading efficiency: Evidence from hearing and deaf readers. *Journal of Experimental Psychology: General*, 153(10), 2359–2377. <https://doi.org/10.1037/xge0001633>
- Sehyr, Z., & Emmorey, K. (2022). Contribution of lexical quality and sign language variables to reading comprehension. *Journal of Deaf Studies and Deaf Education*.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18(6), 643–662. <https://doi.org/10.1037/h0054651>
- Suarez, P. A., Gollan, T. H., Heaton, R., Grant, I., Cherner, M., & the HNRC Group. (2014). Second-Language Fluency Predicts Native Language Stroop Effects: Evidence from Spanish–English Bilinguals. *Journal of the International Neuropsychological Society*, 20(3), 342–348. <https://doi.org/10.1017/S1355617714000058>
- Tse, C.-S., & Altarriba, J. (2012). The effects of first- and second-language proficiency on conflict resolution and goal maintenance in bilinguals: Evidence from reaction time distributional analyses in a Stroop task. *Bilingualism: Language and Cognition*, 15(3), 663–676. <https://doi.org/10.1017/S1366728912000077>

Tzelgov, J., Henik, A., & Leiser, D. (1990). Controlling Stroop interference: Evidence from a bilingual task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(5), 760–771. <https://doi.org/10.1037/0278-7393.16.5.760>

CHAPTER 4: Semantic and Letter Fluency in Young and Older Deaf ASL-English Bilinguals:
Effects of Aging and Testing Language

Sadie Camilliere¹, Karen Emmorey², Peter C. Hauser³, Jessica Contreras³, Michael M. McKee⁴,
David P. Salmon⁵, and Tamar H. Gollan⁵

¹ Joint Doctoral Program in Language and Communicative Disorders, San Diego State
University & University of California, San Diego, CA USA

² School of Speech, Language, and Hearing Science, San Diego State University, San Diego, CA
USA

³ National Technical Institute for the Deaf, Rochester Institute of Technology, Rochester, NY
USA

⁴ Department of Family Medicine, University of Michigan, Ann Arbor, MI USA

⁵ Department of Psychiatry, University of California, San Diego, La Jolla, CA USA

Correspondence concerning this article should be addressed to Sadie Camilliere, Laboratory for
Language and Cognitive Neuroscience, 6495 Alvarado Rd Suite 200, San Diego, CA 92182

Email: sadiecamilliere@gmail.com

Abstract

Objective: The present study explored verbal fluency performance in ASL and English within a wide age range of deaf ASL-English bilinguals.

Method: One hundred nine deaf ASL-English bilinguals between the ages of 18-88 completed six verbal fluency tasks testing ASL and English semantic category fluency and English letter fluency. Participants were instructed to name as many items as they could in the given category in 60 seconds. For the ASL fluency task, participants signed their response. For English fluency tasks, participants were given the option to fingerspell or speak their responses.

Results: When comparing ASL and English category fluency, younger participants produced more correct responses in ASL than English while older participants performed similarly in both languages, and the number of correct responses was negatively correlated with age for ASL category fluency. For English verbal fluency, participants did not show any task differences between category and letter fluency, and better English fluency performance was predicted by education and English proficiency levels (but showed no significant effects of aging).

Conclusions: The findings from this study provide preliminary normative data on verbal fluency performance in deaf ASL-English bilinguals across a wide age range and suggest that ASL category fluency may be more sensitive to age-related declines compared to English category and letter fluency. Additional work is needed to determine which fluency tasks will be most sensitive for detecting cognitive impairment in this population.

Introduction

Tests of verbal fluency are widely used in neuropsychological assessment because they are sensitive to changes in lexico-semantic knowledge, lexical retrieval, executive function, and language production processes, making them effective in detecting a variety of language and cognitive disorders. Verbal fluency tasks typically require individuals to say the name of as many different items as they can from a particular semantic (e.g., animals or fruits) or phonemic/letter (e.g., words starting with F- or S-) category in 60 seconds. Given the brevity, efficiency and clinical utility of these tests, studies have begun to develop and characterize semantic category fluency tests in American Sign Language (ASL) and other sign languages, and have even made attempts to create signed-based phonemic fluency tasks, e.g., using handshape categories. However, very little information is available about normative performance on fluency tasks in ASL by deaf individuals, and no study to date has compared semantic fluency test performance in deaf ASL-English bilinguals in both their signed and spoken languages or compared English semantic and letter fluency in this population. Furthermore, no study has explored how typical aging may impact semantic and letter fluency in the unique population of deaf bimodal bilinguals who use a sign language as their primary form of communication and read and write in a spoken language (though typically do not speak it).

Aging and Verbal Fluency

A number of studies have shown that verbal fluency declines with aging. This decline is often attributed to age-related slowing (Salthouse, 1996), reduced efficiency of lexical retrieval, or decline in executive functioning (Braver & West, 2008; but see Verhaeghen, 2011 for a meta-analysis discussing alternate findings). An additional factor could be the age-related degradation of semantic knowledge that has been documented by reduced picture naming abilities

(Verhaegen & Poncelet, 2013) and an increased rate of tip of the tongue (TOT) states (Burke et al., 1991).

Semantic fluency is likely to be more sensitive than letter fluency to age-related semantic degradation, whereas letter fluency may require more executive control abilities than semantic fluency because words are not strongly organized in memory by letter, and search strategies may require more conflict resolution and greater response monitoring to suppress semantic-based retrieval (Luo et al., 2010), processes likely to decline in older adults (Braver & West, 2008). While both semantic and letter fluency tasks engage multiple underlying cognitive processes, including some that are susceptible to age-related decline, there generally appears to be greater age-related decline in semantic fluency than in letter fluency. However, evidence that semantic and letter fluency are differentially impacted by aging is mixed.

Vaughan et al. (2016) provided normative data on the difference between semantic and letter fluency scores in cognitively healthy older adults by calculating discrepancy scores for 5780 adults between the ages of 50 and 98 years. Participants typically obtained higher scores with semantic (i.e., animals) than letter categories (i.e., F-), with average discrepancy scores of around 9.2 words, in alignment with previous suggestions that semantic fluency is an easier task than letter fluency (but see Azuma et al., 1997 for a discussion of how category size can influence fluency performance). Semantic and letter fluency scores were negatively correlated with age; however, the advantage of semantic over letter fluency was maintained across aging (Vaughan et al., 2016). This contrasts with Alzheimer's disease which typically produces the

opposite pattern with worse performance and faster decline on semantic than letter fluency tests (Salmon et al., 1999).

Weintraub et al. (2018) provided normative data for verbal fluency tasks from a group of 3602 cognitively healthy older adults who completed the neuropsychological test battery of the National Alzheimer's Coordinating Center (NACC) Uniform Data Set (UDS) (Beekly et al., 2007). Cross-sectional analyses showed significant age-related decline in both semantic and letter fluency scores of these cognitively healthy older adults, such that fluency scores decreased with each ten-year age band. The authors did not directly compare age-related decline in semantic and letter fluency, but an examination of the difference between scores for the youngest (< 60 years) and the oldest groups (> 90 years) suggests that aging effects are larger for semantic than letter fluency (e.g., 2.7 items for the letter *F*- fluency and 6.6 for *animal* fluency; Weintraub et al., 2018).

Gordon et al. (2017) also found a significant effect of age on both letter and category fluency test performance, and a more dramatic age-related decline for category fluency. These investigators examined the impact of processing speed on this effect by calculating the proportion of total responses that a participant produced in 15 second time bins across a 90 second response period and creating a linear slope as a measure of naming latency. Regression analyses revealed an interaction between word retrieval times and age for semantic fluency tasks; younger participants with faster retrieval times (indicated by less steep slopes) generated more items, while older participants with faster retrieval times did not show this advantage. In addition, the number of items generated in letter fluency tasks was better predicted by vocabulary and digit span performance (as measured by the WAIS-IV vocabulary and digit span subtests; Wechsler, 2008) than lexical retrieval time or age (i.e., higher vocabulary scores and

better digit span was associated with more words produced). Given these results, the authors proposed that disproportionate age-related decline in semantic fluency may be related to decline in lexical retrieval speed, while letter fluency may be closely related to crystallized vocabulary knowledge that remains relatively stable with age (Gordon et al., 2017).

Verbal Fluency in Unimodal and Bimodal Bilinguals

Assessment of verbal fluency becomes more complicated when evaluating bilingual individuals since characteristics such as language proficiency and dominance can impact fluency performance. Several studies in hearing, spoken-language “unimodal” bilinguals (e.g., Spanish-English speakers) have shown they produce lower category fluency scores than their monolingual peers (Gollan et al., 2002; Portocarrero et al., 2007; Rosselli et al., 2000; Sandoval et al., 2010), although like monolinguals, they tend to generate more exemplars from semantic than from letter categories, even when tested in their less dominant language (Gollan et al., 2002). In contrast, Luo et al. (2010) found a bilingual advantage on a letter fluency task when they compared bilinguals and monolinguals matched for vocabulary level, possibly due to bilinguals having an increased ability to manage executive control demands.

Blumenfeld et al. (2016) assessed letter and category fluency in monolingual English speakers and English-dominant Spanish-English bilinguals to investigate the impact of language dominance and task type on verbal fluency performance. Monolinguals and bilinguals performed similarly, and both groups showed an advantage for category over letter fluency when tested in their dominant language (i.e., English). However, bilinguals had lower scores in Spanish than in

English and displayed no advantage for category over letter fluency when tested in Spanish, their nondominant, less proficient language (Blumenfeld et al., 2016).

Only a few studies have investigated verbal fluency in bimodal bilinguals who use a signed and a spoken/written language. Marshall et al. (2014) found that 30 deaf British Sign Language (BSL) users, aged 21-60 years, generated a similar number of BSL responses in semantic category fluency tasks compared to normative data from spoken language users. The signers in this study exhibited similar retrieval strategies as hearing individuals (e.g., thematic clustering) as well as sign-specific strategies such as clustering of manual homonyms (i.e., signs that are manually articulated the same way, but disambiguated by different mouthings). Sehry et al. (2018) directly compared performance of hearing monolingual English speakers and deaf early-acquisition ASL signers and found similar fluency scores for the two groups when fingerspelled responses were included for ASL trials. Fingerspelling is a system found in many sign languages in which handshapes are used to represent the letters of a written alphabet. Furthermore, Sehry et al. (2018) found an effect of language dominance (i.e., better performance in English, the dominant language) for hearing ASL-English bilinguals who acquired ASL either early (from deaf parents) or later in adulthood, an effect that parallels the dominance effect observed in unimodal bilinguals. These studies highlight that semantic fluency can be compared across signed and spoken languages, but that modality specific retrieval strategies must be considered when evaluating performance on signed fluency tasks.

While semantic category fluency performance appears to be comparable across signed and spoken language modalities, this is less clear for letter fluency since a correlate for the letter fluency task in the signed modality is difficult to develop. Marshall et al. (2014) attempted to create a sign-based phonemic fluency task for BSL in which deaf signers were asked to generate

as many signs as possible with a particular handshape (e.g., “I”, “G”, or “claw 5”), location (e.g., “above the shoulders”, “on the non-dominant hand”), or movement (e.g., “two hands, both hands moving”). However, phonological fluency in BSL on this task was not directly comparable to spoken language letter fluency. Participants generated fewer items and showed much more variability between different categories on the BSL phonemic fluency task compared to norms for spoken letter fluency. Marshall et al. (2014) hypothesized that this may be due, in part, to the lack of a writing system associated with BSL (or any sign language), leading to less meta-linguistic awareness of BSL phonology compared to awareness of phonology in individuals who become literate in a language with a corresponding orthographic system. The authors also found a high rate of semantic clustering in the BSL phonemic fluency task, which may suggest that signed phonological fluency tasks rely more on semantic knowledge and connections than classic spoken-language letter fluency tasks and index different underlying search mechanisms.

Taken together, the results of these studies suggest that differences in lexical retrieval and the organization of signed and spoken lexicons makes comparison of semantic and phonemic/letter fluency performance in the signed modality difficult. Evaluation of these two types of signed verbal fluency may not lead to the same diagnostic predictions that are commonly made with spoken verbal fluency performance.

Current Study

A primary goal of the current study is to obtain preliminary normative data for ASL versions of widely used and highly effective verbal fluency tasks in deaf individuals. ASL is the most accessible language for much of the American deaf population, but there are limited resources available for neuropsychological evaluation in ASL for use in clinical practice.

Additional goals are to investigate semantic category fluency in both ASL and English, and to compare English category and letter fluency in deaf bilinguals.

To better understand the verbal fluency performance of deaf ASL-English bilinguals in both of their languages we administered six semantic fluency tasks in ASL and English as well as letter fluency in English. Previous studies investigated either BSL or ASL fluency in deaf signers (Marshall et al., 2014; Sehyr et al., 2018) or spoken English fluency in hearing signers (Giezen & Emmorey, 2017; Sehyr et al., 2018). For English responses in the current study, participants could either say or fingerspell the words. Note that many deaf ASL-English bilinguals access English primarily through reading and writing, and do not use speech in their everyday life, making it an unnatural response mode for verbal fluency tasks. In contrast, fingerspelling is a natural part of ASL discourse (Padden, 1998) and provides a link between ASL and written English, making it a good candidate for completing English fluency tasks. The lack of a conventionalized writing system in ASL may also lead to imbalances in the predicted language dominance profiles. Many deaf individuals consider ASL their dominant language as it is the primary language used in their everyday life, yet they read and write only in English, leading to an asymmetry in their language use contexts which may have implications for their performance on language tests. Additionally, educational backgrounds and language proficiency can vary greatly among deaf populations, likely impacting verbal fluency performance.

The current study is the first to characterize and compare the performance of deaf ASL-English bilinguals across a wide age range on both ASL and English versions of verbal fluency tests. Understanding the performance of cognitively healthy deaf bilinguals across the lifespan is

a necessary first step in evaluating whether verbal fluency tasks may be used in the clinical assessment of neurocognitive disorders in this population.

Methods

Participants

One hundred twelve deaf ASL-English bilinguals between the ages of 18-88 years completed six tasks probing ASL semantic category fluency, English semantic category fluency, and English letter fluency as part of a larger neuropsychological test battery that also included, among other tests, the MINT Sprint 2.0, ASL Comprehension Test, and the Trail-Making Test Parts A and B. Three participants were excluded from analyses due to self-reported difficulties with memory and struggles to complete many of the cognitive tasks presented on the day of examination. Data from 4 participants were unavailable for one or both English fluency tasks due to loss of video recording. Data from the remaining 109 participants are reported here. Forty-four of the 109 participants were tested in San Diego or Riverside, CA, and 65 were tested in Rochester, NY. Participants were recruited through collaboration with Deaf senior citizen groups, presentations at Deaf community events, and through local recruitment networks. Based on a background questionnaire, the breakdown of self-reported race was 83% white (90/109), 8.3% black (9/109), 2.8% Asian (3/109), 3.7% more than one race (4/109), 0.9% American Indian/Alaskan Native (1/109) and 1.8% declined to answer (2/109). The self-reported ethnicity breakdown was 10% Hispanic or Latino (11/109), 84% not Hispanic or Latino (92/109), and 5% unknown or declined to answer (6/109). All participants reported using ASL as a primary language in their everyday life. Participants were divided into four age groups: Young (18-45 years; $n = 31$); Middle (46-64 years; $n = 12$); Young-Old (65-70 years; $n = 34$); and Older (71-88 years; $n = 32$). Cutoffs for the Young and Middle age groups followed those reported in our

previous study on list memory performance in this population (Camilliere et al., 2026). Cutoffs for the older age groups were decided based on previous literature showing that semantic picture naming ability declines around age 70 (Connor et al., 2004; Feyereisen, 1997; Verhaegen & Poncelet, 2013; Wen & Dong, 2023). Table 4.1 summarizes participant demographic characteristics by age group.

Informed Consent

Research protocols were approved by the University of California, San Diego (UCSD) Institutional Review Board in accordance with the Helsinki Declaration. Participants gave informed consent and were given monetary compensation for their participation.

Procedure, Stimuli, and Scoring

Participants were tested individually by fluent ASL signers. Video instructions for each of the three fluency tasks were presented in ASL by a deaf signer. All responses were videotaped for later transcription and analysis. The semantic fluency categories requiring ASL responses were *Animals* and *Fruits*, while the semantic fluency categories requiring English responses were *Clothing* and *Vegetables*. The English letter fluency trials used the letters *F* and *S*. The decision to not counterbalance semantic categories across response languages was motivated in part by an overarching goal to provide preliminary normative data on verbal fluency of deaf signers in ASL (the most accessible language) for use in clinical practice with this population. *Animals* was chosen as the point of focus because it is a prototypical semantic category used to assess verbal fluency in neuropsychological evaluations (Lezak, 1995), it is a relatively large semantic category and thus more vulnerable to decline (Azuma et al., 1997), and it is sensitive to aging and cognitive decline (Bayles et al., 1989; Butters et al., 1987).

For the semantic category fluency tasks requiring ASL responses, participants were instructed to produce as many items as possible in the category until the experimenter instructed them to stop. They were given 1 minute to produce signs in each category. Following Sehyr et al. (2018), participants were told that fingerspelled words were allowed, but that fingerspelling should not replace known ASL signs. For the semantic category fluency tasks requiring English responses, participants were again instructed to produce as many items as possible in the category, but to respond in English either by fingerspelling or speaking their responses. Participants were further instructed not to use the same word with different endings. They were given 1 minute to produce words in each category. Finally, for the letter fluency tasks requiring English responses, participants were instructed to produce as many English words that start with the target letter by fingerspelling or speaking, without using proper nouns or repeating the same word with different endings. They were given 1 minute to produce words in each category.

All signed and fingerspelled responses were transcribed from video recordings by a deaf native signer, and all spoken English responses were transcribed by a hearing native English speaker. ASL responses were coded for fingerspelled (FS) responses (e.g., kiwi (FS)) and combinations of lexical signs and fingerspelled responses (e.g., water (FS) + MELON). All fingerspelled and combination responses were scored as correct in accordance with Sehyr et al. (2018). Participants were given one point for every unique, correct answer following scoring protocols used in the Alzheimer Disease Centers' Uniform Data Set (UDS) administration and scoring guidelines (Weintraub et al., 2018). Responses were scored for accuracy and error type. Errors were classified into several different categories: repetitions of the same word; intrusions, including morphological variants of previous responses (e.g., *sing* and *singing*) and words not belonging to the target category (e.g., *laces* in the category clothing, or *peach* in the category

vegetables); nonwords, including neologisms and words or signs articulated in an incomprehensible manner; proper nouns; and cross language intrusions (e.g., when an ASL sign was produced during the English category or letter fluency tasks).

Table 4.1. Participant Characteristics

	Younger (<i>n</i> = 31)		Middle (<i>n</i> = 12)		Younger-Old (<i>n</i> = 34)		Older (<i>n</i> = 32)	
Characteristic	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Age	29.8	9.2	57.4	5.2	67.5	1.8	77.4	5.4
Years Education	14.7	2.5	15.4	2.5	16.1	2.7	15.3	2.5
ASL AoA ^A	5.6	6.1	4.8	6.3	9.3	10.1	9.1	7.2
ASL CT score ^B	23.8	2.9	24.3	3.3	21.9	3.9	20.4	4.4
English MINT Sprint 2.0 ^C	64.1	10.6	68.4	5.9	68.0	8.5	69.3	6.1
Trail-making A ^D	27.7	7.3	41.7	21.6	46.1	20.3	51.6	19.2
Trail-making B ^D	74.4	23.5	99.6	34.0	111.7	46.6	136.2	53.1

Note. ^AAoA = reported age of acquisition of ASL. ^BASL CT = ASL Comprehension Test. ^CTotal number of pictures named correctly in English (out of 80) on the Sprint 2.0 version of the Multilingual Naming Test (MINT). ^DTrail Making Test, Parts A and B = completion time in seconds.

Results

Response Mode

Participants used fingerspelling for about 19% of responses during the ASL category fluency task, similar to the fingerspelling rate in the Sehyr et al. (2017) study of category fluency in ASL. The proportion of fingerspelled responses differed between the *animals* (10% of responses) and *fruits* (32% of responses), paralleling the pattern reported in Beal-Alvarez & Figueroa Quintana (2017) for ASL. Participants were given a choice of fingerspelling or

speaking their responses for the English fluency tasks, and many more chose to fingerspell ($n = 92/107$) than to speak ($n = 15/107$).

Fluency performance and associations with demographic features

Correlation coefficients from preliminary analyses exploring associations between fluency outcomes and key demographic characteristics are shown in Table 4.2. There was a significant negative correlation between ASL category fluency and age, but no significant correlation between age and English category or letter fluency (see Figure 4.1). Additionally, ASL proficiency (measured by the ASL Comprehension Test; Hauser et al., 2016) was correlated with ASL category fluency ($r = .351, p < .001$), as well as with English category ($r = .221, p < .05$) and letter ($r = .192, p < .05$) fluency. English proficiency (measured by the English MINT Sprint 2.0; Gollan et al., 2024) was correlated with both English category ($r = .418, p < .001$) and letter ($r = .363, p < .001$) fluency, but not with ASL category fluency. Age of acquisition of ASL was not significantly correlated with verbal fluency performance (all $ps > .10$). The means and standard deviations for the number of correct responses for each fluency task are presented as a function of age group in Table 4.3 and illustrated in Figure 4.2.

Table 4.2. Correlation matrix of fluency performance and demographics

	Age	Edu.	ASL AoA ^A	ASL CT ^B	English MINT ^C	Trails A ^D	Trails B ^D	ASL Category	English Category
Age									
Education	0.19								
ASL AoA	0.18	0.06							
ASL CT	-0.27**	0.32***	-0.26**						
English MINT	0.29**	0.45***	0.07	0.25*					
Trails A	0.38***	-0.22	-0.03	-0.35**	-0.11				
Trails B	0.45***	-0.29**	0.05	-0.43***	-0.09	0.52***			
ASL Category	-0.31***	0.20*	-0.11	0.35***	0.13	-0.21	-0.38***		
English Category	0.14	0.43***	0.11	0.22*	0.42***	-0.33**	-0.39***	0.46***	
English Letter	-0.02	0.23*	0.09	0.19*	0.30**	-0.05	-0.16	0.42***	0.36***

Note. ^AAoA = reported age of acquisition of ASL. ^BASL CT = ASL Comprehension Test. ^CTotal number of pictures named correctly in English (out of 80) on the Sprint 2.0 version of the Multilingual Naming Test (MINT). ^DTrail Making Test, Parts A and B = completion time in seconds. *** $p \leq 0.001$, ** $p \leq .01$, * $p < .05$, † $p \leq .10$.

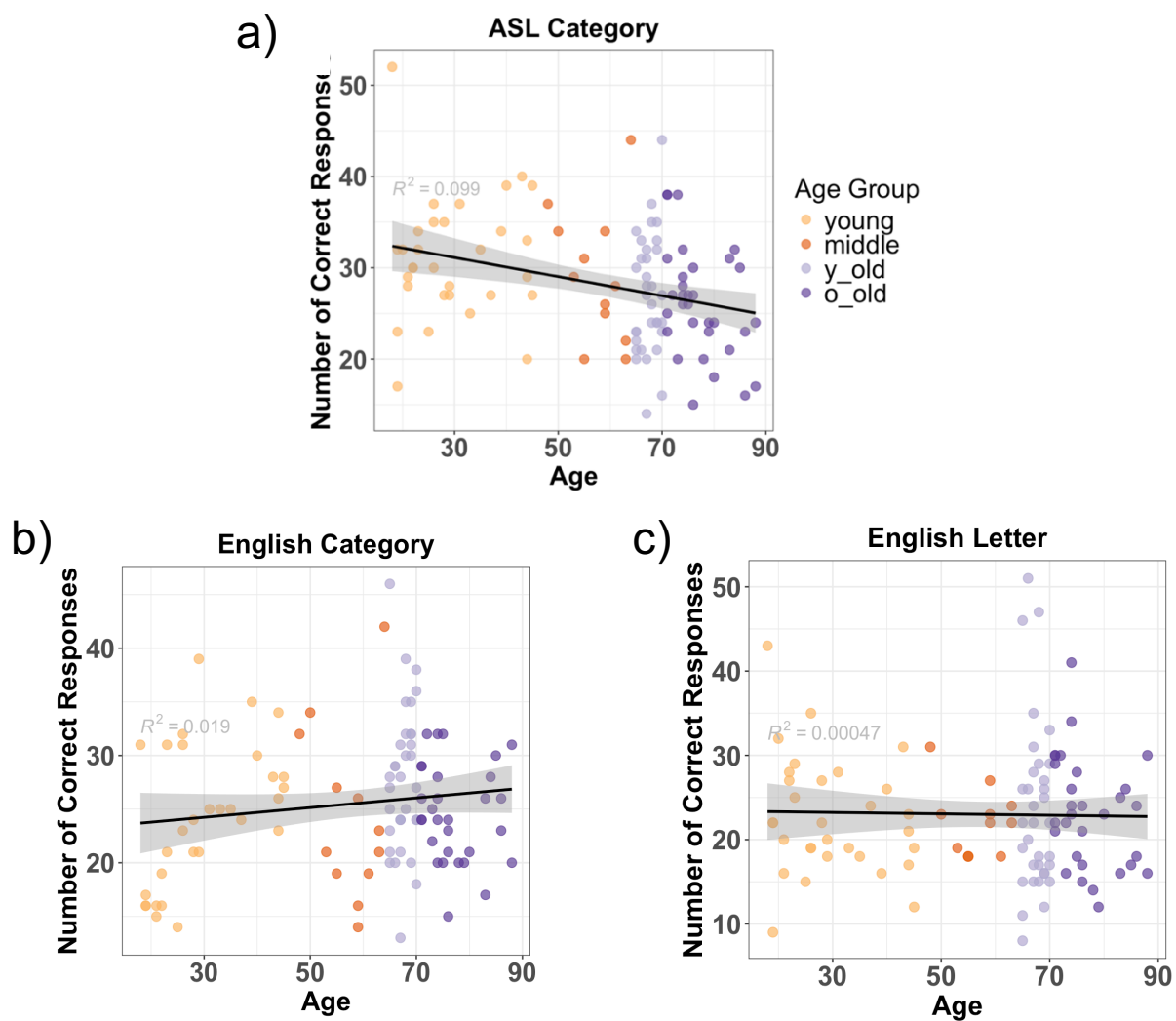


Figure 4.1. Scatterplots illustrating the significant relationship between ASL category fluency and age (a) and the null relationship between English category and English letter fluency and age (b-c)

Table 4.3. Means (and SDs) for fluency tasks by age group

	Younger (<i>n</i> = 31)		Middle (<i>n</i> = 12)		Younger-Old (<i>n</i> = 34)		Older (<i>n</i> = 32)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
ASL Category								
Animals	19.1	4.6	17.3	4.8	15.6	4.0	15.3	3.8
Fruits	12.0	3.7	11.9	3.9	11.5	3.1	10.8	2.9
Total	31.1	6.7	29.2	7.3	27.0	6.4	26.1	6.0
English Category								
Clothing	14.3	4.1	13.1	4.9	15.2	3.8	13.8	2.7
Vegetables	10.2	3.2	11.4	4.3	12.5	4.1	10.8	2.5
Total	24.4	6.6	24.5	8.2	27.7	6.8	24.6	4.5
English Letter								
F	10.6	4.0	10.2	3.1	10.9	5.3	10.8	3.6
S	12.1	3.7	12.6	2.1	12.6	5.2	12.3	3.6
Total	22.6	7.0	22.3	4.1	23.5	10.1	23.1	6.6

ASL vs English Semantic Category Fluency

The impact of testing language and aging on category fluency performance was investigated using linear regression modeling with the number of correct responses (combining scores from the two semantic categories used with each language) as the dependent variable and language (ASL vs English), age, and their interaction as predictors. Education level was included as a noninteracting covariate. This model significantly predicted category fluency scores, $F(4, 210) = 13.6, p < .001$, adj. $R^2 = .19$. There were significant effects of age ($\beta = -.13, p < .001$), language ($\beta = -11.38, p < .001$), and education level ($\beta = .88, p < .001$) such that younger participants produced more correct responses than older participants, participants produced more correct responses in ASL than in English, and participants with higher education level produced

more correct responses. A significant interaction between language and age ($\beta = .15, p < .001$) revealed that aging effects were larger for ASL than English category fluency tests (see Figure 4.2).

To investigate which factors contributed most to category fluency performance in each language, follow-up regressions examining ASL and English category fluency were run including age, education, and proficiency (measured by the ASL-CT and English MINT Sprint 2.0) as predictors. For ASL category fluency, this model significantly predicted fluency scores, $F(3, 105) = 8.61, p < .001, \text{adj. } R^2 = .18$. Older participants produced fewer correct ASL responses ($\beta = -1.9, p = 0.004$) and participants with higher levels of ASL proficiency produced more correct responses ($\beta = 1.4, p = 0.3$). Education level just marginally predicted ASL category fluency when ASL proficiency was included in the model ($\beta = 1.2, p = 0.06$). The follow-up model for English category fluency also significantly predicted fluency scores, $F(3, 101) = 11.2, p < .001, \text{adj. } R^2 = .23$. As found in the above models, age did not significantly predict English category fluency performance ($\beta = -0.04, p = 0.95$). However, education ($\beta = 1.9, p = 0.002$) and English ($\beta = 1.8, p = 0.006$) proficiency both significantly predicted English category fluency – participants with higher levels of education and English proficiency produced more correct English responses.

English Category vs Letter Fluency

Performance on English semantic category and letter fluency tasks was compared using linear regression modeling with the number of correct responses produced in English (combining scores from the two categories used in each condition) as the dependent variable and task type (category vs letter), age, and their interaction as predictors. Education level was included as a noninteracting covariate. This model significantly predicted fluency scores, $F(4, 207) = 8.2, p <$

.001, adj. $R^2 = .12$; however, there was no difference between category and letter fluency, ($\beta = .74, p = .80$), no main effect of age ($\beta = .02, p = .51$), and no interaction between age and task type ($\beta = -.05, p = .24$). There was a significant effect of education ($\beta = .92, p < .001$) as participants with more education produced more correct responses.

To investigate which factors contribute most to letter fluency performance, an additional follow-up regression examining was run including age, education, and English proficiency (measured by the English MINT Sprint 2.0) as predictors. This model significantly predicted English letter fluency scores, $F(3, 101) = 4.5, p = .005$, adj. $R^2 = .09$. As found in the initial model, there was no significant effect of age ($\beta = -0.96, p = .20$). Individuals with better English proficiency produced more correct responses for letter fluency, a significant main effect of proficiency ($\beta = 2.14, p = .011$), and education was not a significant predictor when English proficiency was included in the model ($\beta = 1.07, p = .20$).

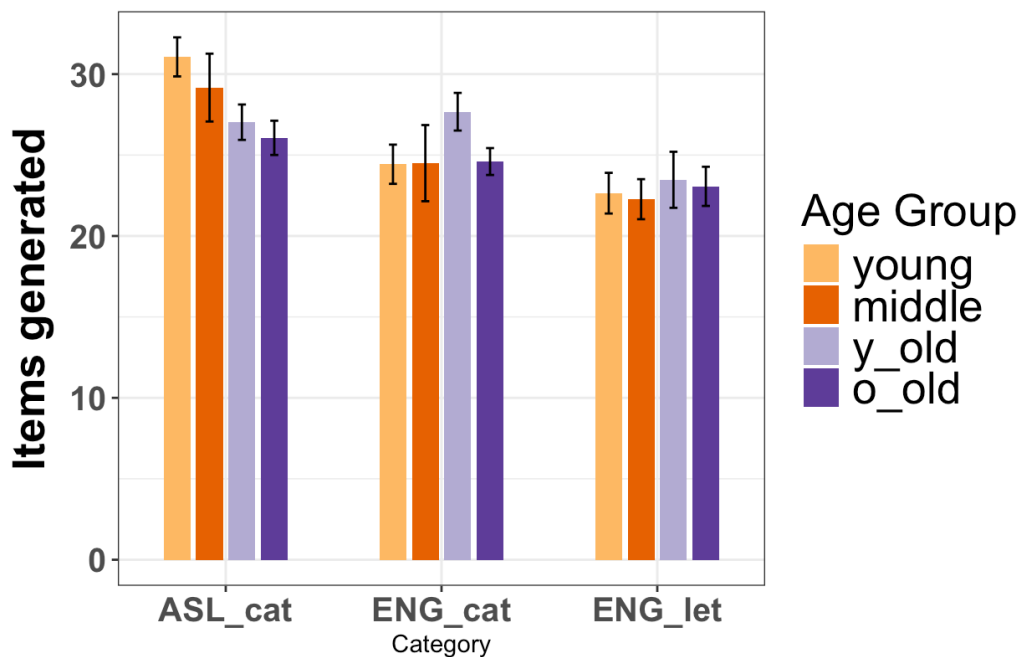


Figure 4.2. Performance on ASL and English category fluency and English letter fluency by age group

Discussion

Summary of Key Results

The present study investigated verbal fluency in 109 deaf ASL-English bilinguals aged 18-88 years old using ASL and English category fluency tasks and an English letter fluency task. When comparing ASL and English category fluency, younger participants produced more correct responses in ASL than English while older participants performed similarly in both languages (a significant interaction between language and age; see Table 4.3; Figure 4.2). This relationship was also apparent when examining the correlation between age and category fluency within each language; older signers performed worse on ASL category fluency, but there was no statistically significant relationship between age and English category fluency (see Figures 4.1). When comparing English category and letter fluency performance, participants produced a similar number of correct responses for both English category and letter fluency tasks (see Figure 4.2). Separate analyses investigating the contribution of demographic factors to fluency performance in each language and task type revealed that for ASL category fluency, age and ASL proficiency significantly predicted performance, while education was not significant (after controlling for age and ASL CT scores). By contrast, age was not a significant predictor for English category or letter fluency, but individuals with higher education level and higher English proficiency (as measured by English picture naming scores) produced more correct responses. Finally, the only significant predictor for English letter fluency was English proficiency (again measured by English picture naming scores).

Task and Testing Language Effects on Verbal Fluency

The finding that deaf ASL-English bilinguals performed worse on English than ASL category fluency and similarly on English category and letter fluency is in line with previous

studies that found unimodal bilinguals performed worse in their nondominant than dominant language and showed no advantage for category over letter fluency when tested in their nondominant language (Blumenfeld et al., 2016). This pattern of results suggests that many of the participants in the current study are ASL dominant which aligns with participants' reports of their preferred language (i.e., 98/109 reported ASL as their preferred language; 7/109 preferred Pidgin Signed English or Signed Exact English⁷; 3/109 preferred spoken English; and 1/109 reported "other" as their preferred language). Language dominance effects have been observed previously in hearing ASL-English bilinguals who produced more correct responses in an English category fluency task (their dominant language) than in an ASL category fluency task (Sehyr et al., 2018). The pattern of results we observed could be indicative of smaller vocabulary sizes for English compared to ASL. Another possibility is that participants experienced more cross-language interference during English tasks. This possibility is supported by the occurrence of cross-language errors in the English category and letter fluency tasks. These errors typically occurred when a participant produced a lexical sign during an English fluency task, followed by the English fingerspelled word, e.g., a participant would produce the sign PANTS, then fingerspell P-A-N-T-S. Research from hearing bilinguals suggests that cross-language interference errors may be more likely to occur during semantic than letter fluency tasks, and are more frequent when the verbal fluency task is performed in the nondominant language (Rosselli et al., 2000; Sandoval et al., 2010). This type of error may reflect use of a translation-based strategy for some participants completing the English fluency tasks or may signal executive

⁷ Pidgin Signed English (PSE) refers to a contact form of signing that has strong influences from English (Cokely, 1983); Signed Exact English (SEE) refers to the artificially constructed signed system based on English grammar.

control decline in some older participants (Gollan et al., 2011), however this type of error was very infrequent (less than 1.8% of all responses).

Aging Effects in Deaf Bilinguals

Investigation of the effect of aging on verbal fluency in this population of deaf ASL-English bilinguals revealed that the number of correct responses decreased with increasing age only for ASL semantic category fluency. This finding is in accordance with previous literature demonstrating that age-related decline is most apparent in semantic fluency tasks (Gordon et al., 2018; Vaughan et al., 2016), and that category fluency may be particularly sensitive to aging effects. It is also possible that aging effects are most obvious in the ASL category task in the current study because this was the language in which the category *animals* was presented. Animals is a large semantic category and may be more susceptible to age-related decline. The lack of an association between age and performance in the English category or letter fluency tasks may be because English verbal fluency was more strongly associated with English language proficiency than with age in deaf ASL-English bilinguals. One possible explanation for this pattern is that most deaf individuals acquire English as an L2 through print and receive much of their education in English. This may be particularly true for older deaf individuals who only had access to oral education due to U.S. policies that limited or forbade the use of sign language in education (Moore, 2010). As a result, more years of education leads to higher levels of English proficiency and larger vocabulary which influences the number of correct responses produced on English fluency tasks. Also notable was that the oldest-old participants scored highest on average on the English picture naming test (see Table 1) but did not obtain the highest semantic fluency scores in English (see Table 3). Thus, higher English proficiency may have offset aging effects on semantic fluency in English in this population.

Limitations

Several studies have pointed out that category size can impact the number of correct items produced in verbal fluency tasks (Azuma et al., 1997), and some have even shown that certain categories (e.g., larger semantic categories such as *animals*) are more vulnerable to cognitive decline (Diaz et al., 2004). A primary goal of the study was to provide normative data on verbal fluency performance in ASL, the most accessible language for many deaf individuals. As a result, categories were not counterbalanced across language in order to maximize data collected on the category *animals* for ASL. Therefore, it is possible that differences in category size may have impacted verbal fluency performance in ASL compared to English.

Another limitation in the present study was the inclusion of just two categories per task (2 semantic in each language, and 2 in each type of English fluency task). This limited our ability to compare semantic to letter fluency with respect to overall difficulty in deaf participants, which could be of interest to examine in future studies because of the potential sensitivity to Alzheimer's disease (Salmon et al., 1999). The relative difficulty of semantic versus letter fluency is also determined in part by category size (Azuma et al., 1997; Gollan et al., 2002). For example, college aged Spanish-English bilinguals in the study by Gollan et al., (2002) produced fewer *animals* in English than did young participants in the present study in ASL (16.8 versus 19.1 respectively), but the same number of *fruits* (12.0 in both cases), and virtually the same number of *clothing* (14.0 and 14.3 respectively) and *vegetables* responses even though deaf bilinguals completed these categories in their nondominant language (whereas participants in Gollan et al were tested in their dominant language). By contrast, the same deaf participants produced fewer English *F* and *S* words (10.6 and 12.1) compared to Spanish-English bilinguals (12.0 and 13.9 respectively). Direct comparison across studies is difficult for many reasons, in

addition to the difference in which language was dominant (ASL in the present study versus English for the Spanish-English bilinguals), young participants in the present study were about 10 years older than young bilinguals in Gollan et al., (2002). However, the presence of higher semantic category scores but lower letter fluency scores in deaf ASL-English bilinguals versus in Spanish-English bilinguals in the previous study, could suggest that letter fluency is a more difficult task for deaf than for hearing participants.

Conclusions

The findings from the current study provide preliminary normative data on verbal fluency performance in deaf ASL-English bilinguals across a wide age range and offer some insights into factors that may contribute to this population performs on different types of verbal fluency tasks. ASL category fluency appears to be more sensitive to age-related declines compared to English category and letter fluency. This result may reflect larger vocabularies and semantic networks in ASL compared to English, which would result in poorer ASL category fluency for older participants because larger semantic categories are more vulnerable to decline. English verbal fluency tasks were closely related to English proficiency and education level, factors that vary across the deaf population. Levels of English proficiency and education level should be taken into account when using these tasks clinically with this population. Future longitudinal and patient studies are needed to confirm which categories are most sensitive to cognitive decline over and above the effects of typical aging.

Acknowledgements

Chapter 4 is being prepared for submission for publication of the material. Camilliere, S., Emmorey, K., Hauser, P. C., Contreras, J., McKee, M. M., Salmon, D.P., & Gollan, T. H. Semantic and Letter Fluency in Young and Older Deaf ASL-English Bilinguals: Effects of Aging and Testing Language (Manuscript in preparation). The dissertation author was the primary investigator and author of this paper. This research was supported by NIH grant R01-AG074928. The authors have no conflicts of interest to disclose. We would like to thank Julie Fisher, Lucinda Farnady, Chehally Soto, Emma Kane, Hayden Orr, and Gigi Zheng for their help with data collection; and Lucinda Farnady and Julie Fisher for their help with data coding. We would also like to thank all of our deaf participants.

References

- Azuma, T., Bayles, K. A., Cruz, R. F., Tomoeda, C. K., Wood, J. A., McGeagh, A., & Montgomery, E. B. (n.d.). *Comparing the Difficulty of Letter, Semantic, and Name Fluency Tasks for Normal Elderly and Patients With Parkinson's Disease*.
- Bayles, K. A., Salmon, D. P., Tomoeda, C. K., Jacobs, D., Caffrey, J. T., Kaszniak, A. W., & Tröster, A. I. (1989). Semantic and letter category naming in Alzheimer's patients: A predictable difference. *Developmental Neuropsychology*, 5(4), 335–347.
<https://doi.org/10.1080/87565648909540443>
- Beal-Alvarez, J. S., & Figueroa Quintana, J. (2017). *Generation of Signs Within Semantic and Phonological Categories: Data from Deaf Adults and Children Who Use American Sign Language*. Retrieved April 23, 2026, from <https://dx.doi.org/10.1093/deafed/enw075>
- Beekly, D. L. et al. The National Alzheimer's Coordinating Center (NACC) database: the uniform data set. *Alzheimer Dis. Associated Disord.* 21, 249–258 (2007).
- Blumenfeld, H. K., Bobb, S. C., & Marian, V. (2016). The role of language proficiency, cognate status and word frequency in the assessment of Spanish–English bilinguals' verbal fluency. *International Journal of Speech-Language Pathology*, 18(2), 190–201.
<https://doi.org/10.3109/17549507.2015.1081288>
- Braver, T. S., & West, R. (2008). Working Memory, Executive Control, and Aging. In *The Handbook of Aging and Cognition*. Psychology Press.
- Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults? *Journal of Memory and Language*, 30(5), 542–579. [https://doi.org/10.1016/0749-596X\(91\)90026-G](https://doi.org/10.1016/0749-596X(91)90026-G)
- Butters, N., Granholm, E., Salmon, D. P., Grant, I., & Wolfe, J. (1987). Episodic and Semantic Memory: A Comparison of Amnesic and Demented Patients. *Journal of Clinical and Experimental Neuropsychology*, 9(5), 479–497.
<https://doi.org/10.1080/01688638708410764>
- Camilliere, S., Emmorey, K., Hauser, P. C., Contreras, J., McKee, M. M., & Gollan, T. H. (2026). Influence of testing language and aging on verbal list memory in deaf American Sign Language–English bilinguals. *Neuropsychology*. <https://doi.org/10.1037/neu0001065>
- Connor, L. T., Spiro, A., III, Obler, L. K., & Albert, M. L. (2004). Change in Object Naming Ability During Adulthood. *The Journals of Gerontology: Series B*, 59(5), P203–P209.
<https://doi.org/10.1093/geronb/59.5.P203>
- Diaz, M., Sailor, K., Cheung, D., & Kuslansky, G. (2004). Category size effects in semantic and letter fluency in Alzheimer's patients. *Brain and Language*, 89(1), 108–114.
[https://doi.org/10.1016/S0093-934X\(03\)00307-9](https://doi.org/10.1016/S0093-934X(03)00307-9)

- Feyereisen, P. (1997). A Meta-Analytic Procedure Shows an Age-Related Decline in Picture Naming. *Journal of Speech, Language, and Hearing Research*, 40(6), 1328–1333. <https://doi.org/10.1044/jslhr.4006.1328>
- Giezen, M. R., & Emmorey, K. (2017). Evidence for a bimodal bilingual disadvantage in letter fluency. *Bilingualism: Language and Cognition*, 20(1), 42–48. <https://doi.org/10.1017/S1366728916000596>
- Gollan, T. H., Garcia, D. L., Murillo, M., Vargas, J., Pulido, B., & Salmon, D. P. (2024). Sprinting in two languages: Picture naming performance of older Spanish–English bilinguals on the Multilingual Naming Test Sprint 2.0. *Neuropsychology*, 38(7), 653–664. <https://doi.org/10.1037/neu0000958>
- Gollan, T. H., Montoya, R. I., & Werner, G. A. (2002). Semantic and letter fluency in Spanish–English bilinguals. *Neuropsychology*, 16(4), 562–576. <https://doi.org/10.1037/0894-4105.16.4.562>
- Gordon, J. K., Young, M., & Garcia, C. (2018). Why do older adults have difficulty with semantic fluency? *Aging, Neuropsychology, and Cognition*, 25(6), 803–828. <https://doi.org/10.1080/13825585.2017.1374328>
- Hauser, P. C., Paludneviene, R., Riddle, W., Kurz, K. B., Emmorey, K., & Contreras, J. (2016). American Sign Language Comprehension Test: A Tool for Sign Language Researchers. *The Journal of Deaf Studies and Deaf Education*, 21(1), 64–69. <https://doi.org/10.1093/deafed/env051>
- Lezak, M.D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.
- Luo, L., Luk, G., & Bialystok, E. (2010). Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition*, 114(1), 29–41. <https://doi.org/10.1016/j.cognition.2009.08.014>
- Marshall, C., Rowley, K., & Atkinson, J. (2014). Modality-Dependent and -Independent Factors in the Organisation of the Signed Language Lexicon: Insights From Semantic and Phonological Fluency Tasks in BSL. *Journal of Psycholinguistic Research*, 43(5), 587–610. <https://doi.org/10.1007/s10936-013-9271-5>
- Moore, D. F. (2010). The history of language and communication issues in deaf education. *The Oxford Handbook of Deaf Studies, Language, and Education*, 2, 17–30.
- Padden, C. A. (1998). The ASL lexicon. *Sign Language & Linguistics*, 1(1), 39–60. <https://doi.org/10.1075/sll.1.1.04pad>

- Portocarrero, J. S., Burright, R. G., & Donovan, P. J. (2007). Vocabulary and verbal fluency of bilingual and monolingual college students. *Archives of Clinical Neuropsychology*, 22(3), 415–422. <https://doi.org/10.1016/j.acn.2007.01.015>
- Rosselli, M., Ardila, A., Araujo, K., Weekes, V. A., Caracciolo, V., Padilla, M., & Ostrosky-Solís, F. (2000). Verbal Fluency and Repetition Skills in Healthy Older Spanish-English Bilinguals. *Applied Neuropsychology*, 7(1), 17–24. https://doi.org/10.1207/S15324826AN0701_3
- Salmon, D. P., Heindel, W. C., & Lange, K. L. (1999). Differential decline in word generation from phonemic and semantic categories during the course of Alzheimer's disease: Implications for the integrity of semantic memory. *Journal of the International Neuropsychological Society*, 5(7), 692–703. <https://doi.org/10.1017/S1355617799577126>
- Salthouse, T. A. (1996). *The Processing-Speed Theory of Adult Age Differences in Cognition*.
- Sandoval, T. C., Gollan, T. H., Ferreira, V. S., & Salmon, D. P. (2010). What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. *Bilingualism: Language and Cognition*, 13(2), 231–252. <https://doi.org/10.1017/S1366728909990514>
- Sehyr, Z. S., Giezen, M. R., & Emmorey, K. (2018). Comparing Semantic Fluency in American Sign Language and English. *The Journal of Deaf Studies and Deaf Education*, 23(4), 399–407. <https://doi.org/10.1093/deafed/eny013>
- Vaughan, R. M., Coen, R. F., Kenny, R., & Lawlor, B. A. (2016). Preservation of the Semantic Verbal Fluency Advantage in a Large Population-Based Sample: Normative Data from the TILDA Study. *Journal of the International Neuropsychological Society*, 22(5), 570–576. <https://doi.org/10.1017/S1355617716000291>
- Verhaegen, C., & Poncelet, M. (2013). Changes in Naming and Semantic Abilities With Aging From 50 to 90 years. *Journal of the International Neuropsychological Society*, 19(2), 119–126. <https://doi.org/10.1017/S1355617712001178>
- Verhaeghen, P. (2011). Aging and Executive Control: Reports of a Demise Greatly Exaggerated. *Current Directions in Psychological Science*, 20(3), 174–180. <https://doi.org/10.1177/0963721411408772>
- Weintraub, S., Besser, L., Dodge, H. H., Teylan, M., Ferris, S., Goldstein, F. C., Giordani, B., Kramer, J., Loewenstein, D., Marson, D., Mungas, D., Salmon, D., Welsh-Bohmer, K., Zhou, X.-H., Shirk, S. D., Atri, A., Kukull, W. A., Phelps, C., & Morris, J. C. (2018). Version 3 of the Alzheimer Disease Centers' Neuropsychological Test Battery in the Uniform Data Set (UDS). *Alzheimer Disease & Associated Disorders*, 32(1), 10. <https://doi.org/10.1097/WAD.0000000000000223>
- Wen, H., & Dong, Y. (2023). The effect of ageing on confrontation naming in healthy older adults: A three-level meta-analysis. *Journal of Cognitive Psychology*, 35(4), 480–508. <https://doi.org/10.1080/20445911.2023.2184745>

CHAPTER 5: General Discussion

This dissertation investigated the cognitive and linguistic outcomes of aging in a group of deaf ASL-English bilinguals by testing performance on three widely used neuropsychological tasks. Study one examined list memory performance in both ASL and English and found that participants showed different recall patterns based on age. Older bilinguals exhibited a unique memory pattern: they were able to recall more English words during immediate learning but also showed higher rates of forgetting in English compared to ASL. Study two investigated Stroop performance and found larger Stroop effects when participants responded to English rather than ASL stimuli, a difference that was magnified in older bilinguals. Study three explored ASL and English verbal fluency and found that ASL category fluency was most sensitive to aging effects, while English category and letter fluency were more closely linked to education and English proficiency levels. These findings together demonstrate that deaf bimodal bilinguals' performance on cognitive and linguistic tasks often deviates from what might be expected in hearing, unimodal bilinguals. In addition, these studies indicate testing language and educational background need to be considered when interpreting test results from this population - arguably even more so than in hearing bilinguals of two spoken languages. Finally, the three studies show that it is possible to develop culturally and linguistically appropriate neuropsychological tests for use with deaf ASL-English bilinguals, and that aspects of these tests are sensitive to aging effects.

Several common themes were apparent across the findings of these studies. Many of the results found in this dissertation diverged from what would be expected based on literature surrounding neuropsychological evaluation in hearing, unimodal bilinguals. While the common recommendation for bilingual evaluation is to test in the dominant language to maximize

sensitivity and performance (Gollan et al., 2024), the deaf bilinguals here showed the opposite pattern in some cases, such as in showing better initial learning for English list memory or larger Stroop effects to English compared to ASL stimuli. On the other hand, some observed effects align with what is expected for unbalanced bilinguals, such as the finding that category fluency in English (non-dominant language) was worse than category fluency in ASL (dominant language, and that there was no advantage for category fluency over letter fluency when tested in English (nondominant language). However, for other tasks, such as list memory, participants' self-ratings and recall performance were more reflective of how balanced bilinguals might perform. Taken together, this pattern of results suggests that deaf bimodal bilinguals have language dominance profiles that differ from unimodal bilinguals due to variable levels of language proficiency in both their first and second languages, as well as the vastly different contexts in which deaf signers learn and use English which sets them apart from hearing L2 learners.

Some of the differences observed here may be attributed to modality differences between signed and written languages. For example, distinct processing speeds for signed vs. written stimuli likely contributed to the language asymmetry observed in Stroop effects, and the potential for speech-based encoding strategies may have benefited English list learning. Interestingly, some of these modality differences are only apparent when examining the performance of older deaf bilinguals: the heightened learning and forgetting in English list memory, larger differences between English and ASL Stroop effects, and worse ASL than English category fluency performance. In contrast, some findings were consistent across older and younger bilinguals (e.g., similar rates of Stroop error production across ages).

The question of which aspects of performance are able to best predict early signs of cognitive decline remains to be investigated, but these findings presented here are foundational for understanding typical aging patterns in this population. The tasks in this dissertation provide several opportunities for identifying decline either through performance that shows exaggerated decline over and above the effects of typical aging (e.g., delayed list memory and ASL category fluency) or through worse performance on tasks that seemed to show no aging effects in the current sample which was living independently and had no specific cognitive or memory complaints (e.g., Stroop errors).

As seen throughout this set of studies, differences in the educational and language backgrounds of deaf signers impact neuropsychological testing. Educational experiences for deaf individuals in America vary widely, and older deaf individuals in particular are likely to have had oralist educations that focused on speech and English instruction while limiting or forbidding access to sign language. Some of the cohort differences observed in performance on English tasks may be attributable to these educational differences. Even for younger deaf individuals who were much more likely to have access to sign-based education, the experience of learning English through print is quite distinct from that of hearing individuals learning English as an L2. These differences may have lasting effects on the ways in which deaf individuals process written language.

Age of acquisition (AoA) of ASL is also an important consideration, as it can be variable in deaf populations, and can impact some cognitive and linguistic tasks. In the current studies, we found no significant effects of AoA on performance on list memory, Stroop, or verbal fluency tasks. However, there was a significant association between AoA and ASL proficiency, and ASL proficiency impacted some aspects of test performance. Further work should

investigate other variables, such as the quality of early language input, the degree of sign access deaf individuals had in different contexts growing up, or the amount of interaction with deaf signing peers to add to the understanding of how different experiences of language acquisition may influence performance on these tests. These results demonstrate that it is necessary to take into account unique circumstances of modality, language access, and educational backgrounds when studying deaf bimodal bilinguals, and suggest that comparisons to the performance of hearing, unimodal bilinguals or monolinguals may not be appropriate.

Finally, through the development and implementation of the tasks developed in this dissertation, it became apparent that there are important considerations that must be made when attempting to use these neuropsychological tests with deaf signers. The findings highlight the caution needed when choosing the language of test administration for this population, as several of the results would typically be interpreted as signs of cognitive decline (e.g., rapid forgetting in English list memory) whereas testing in both languages may reveal different patterns. Response mode should also be considered when administering neuropsychological tests to this population to ensure that appropriate options are provided (e.g., allowing fingerspelling, signing, or writing (when possible), rather than requiring speech). Clinicians should also ensure they are not inhibiting any natural strategies that signers may utilize (e.g., using list buoys during recall or fingerspelling during ASL category fluency).

The studies presented in this dissertation constitute a starting point for understanding how typical aging influences performance on cognitive and linguistic tasks which is critical for the later evaluation of these tasks' efficacy in assessing for cognitive decline. Future studies should compare performance of deaf ASL-English bilinguals with and without cognitive decline to determine which tasks are most sensitive to decline, while controlling for typical aging effects,

education, and language proficiency. The development of these potentially clinically applicable tasks as well as the normative data provided here represent a first step in addressing the gaps in research and assessment tools that are currently available for describing and identifying cognitive decline in deaf bilingual populations.

References

- Corina, D. P., Farnady, L., LaMarr, T., Pedersen, S., Lawyer, L., Winsler, K., Hickok, G., & Bellugi, U. (2020). Effects of age on American Sign Language sentence repetition. *Psychology and Aging, 35*(4), 529–535. <https://doi.org/10.1037/pag0000461>
- Dean, P. M., Feldman, D. M., Morere, D., & Morton, D. (2009). Clinical Evaluation of the Mini-Mental State Exam with Culturally Deaf Senior Citizens. *Archives of Clinical Neuropsychology, 24*(8), 753–760. <https://doi.org/10.1093/arclin/acp077>
- Gasquoine, P. G., & Gonzalez, C. D. (2012). Using Monolingual Neuropsychological Test Norms with Bilingual Hispanic Americans: Application of an Individual Comparison Standard. *Archives of Clinical Neuropsychology, 27*(3), 268–276. <https://doi.org/10.1093/arclin/acs004>
- Gollan, T. H., Garcia, D. L., Murillo, M., Vargas, J., Pulido, B., & Salmon, D. P. (2024). Sprinting in two languages: Picture naming performance of older Spanish–English bilinguals on the Multilingual Naming Test Sprint 2.0. *Neuropsychology*.
- Gollan, T.H., Salmon, D.P., Montoya, R.I., Da Pena, E. (2010). Accessibility of the nondominant language in picture naming: A counterintuitive effect of dementia on bilingual language production. *Neuropsychologia, 48*, 1356-1366.
- Gollan, T. H., Stasenko, A., & Salmon, D. P. (2023). Which language is more affected in bilinguals with Alzheimer’s disease? Diagnostic sensitivity of the Multilingual Naming Test. *Neuropsychology, 37*(5), 595–606. <https://doi.org/10.1037/neu0000893>
- Grosjean, F. (2010). Bilingualism, biculturalism, and deafness. *International Journal of Bilingual Education and Bilingualism, 13*(2), 133–145. <https://doi.org/10.1080/13670050903474051>
- Humphries, T., Kushalnagar, P., Mathur, G., Napoli, D. J., Padden, C., Rathmann, C., & Smith, S. (2016). Avoiding Linguistic Neglect of Deaf Children. *Social Service Review, 90*(4), 589–619. <https://doi.org/10.1086/689543>
- Luna, S., Joubert, S., Blondel, M., Cecchetto, C., & Gagné, J.-P. (2021). The Impact of Aging on Spatial Abilities in Deaf Users of a Sign Language. *The Journal of Deaf Studies and Deaf Education, 26*(2), 230–240. <https://doi.org/10.1093/deafed/ena034>
- Mayberry, R. I., Lock, E., & Kazmi, H. (2002). Linguistic ability and early language exposure. *Nature, 417*(6884), 38–38. <https://doi.org/10.1038/417038a>
- McKee, M. M., & Paasche-Orlow, M. K. (2012). Health Literacy and the Disenfranchised: The Importance of Collaboration Between Limited English Proficiency and Health Literacy Researchers. *Journal of Health Communication, 17*(sup3), 7–12. <https://doi.org/10.1080/10810730.2012.712627>

- McKee, M. M., Paasche-Orlow, Michael K., Winters, Paul C., Fiscella, Kevin, Zazove, Philip, Sen, Ananda, & Pearson, T. (2015). Assessing Health Literacy in Deaf American Sign Language Users. *Journal of Health Communication*, 20(sup2), 92–100. <https://doi.org/10.1080/10810730.2015.1066468>
- Mitchell, R. E., & Karchmer, M. A. (2004). Chasing the Mythical Ten Percent: Parental Hearing Status of Deaf and Hard of Hearing Students in the United States. *Sign Language Studies*, 4(2), 138–163.
- Moore, D. F. (2010). The history of language and communication issues in deaf education. *The Oxford Handbook of Deaf Studies, Language, and Education*, 2, 17–30.
- Moreno, D. A. (2013). Methodological Issues Associated with Sign-Based Neuropsychological Assessment. *Sign Language Studies*, 14(1), 8–20.
- Padden, C. A. (1996). From the Cultural to the Bicultural: The Modern Deaf Community. In I. Paraginis (Ed.), *Cultural and Language Diversity and the Deaf Experience* (pp. 79–98). Cambridge University Press. <https://doi.org/10.1017/CBO9781139163804.006>
- Paludneviciene, R., Hauser, P. C., Daggett, D. J., & Kurz, K. B. (2012). Issues and Trends in Sign Language Assessment. In D. Moreno & T. Allen (Eds.), *Assessing Literacy in Deaf Individuals: Neurocognitive Measurement and Predictors* (pp. 191–207). Springer. https://doi.org/10.1007/978-1-4614-5269-0_11
- Peña, E. D. (2007). Lost in Translation: Methodological Considerations in Cross-Cultural Research. *Child Development*, 78(4), 1255–1264. <https://doi.org/10.1111/j.1467-8624.2007.01064.x>
- Roberson, L., & Shaw, S. (2015). Reflections on Deaf Education: Perspectives of Deaf Senior Citizens. *Educational Gerontology*, 41(3), 226–237. <https://doi.org/10.1080/03601277.2014.951194>
- Zorzi, G., Aristodemo, V., Giustolisi, B., Hauser, C., Donati, C., & Cecchetto, C. (2023). Assessing Lexical and Syntactic Comprehension in Deaf Signing Adults. *The Journal of Deaf Studies and Deaf Education*, 28(4), 373–386. <https://doi.org/10.1093/deafed/enad022>