Spring 3-2021

Automaticity of lexical access in deaf and hearing bilinguals: Cross-linguistic evidence from the color Stroop task across five languages

Rain G. Bosworth
Rochester Institute of Technology

Sarah C. Tyler
University of California, San Diego

Eli M. Binder
Colorado College

Jill P. Morford
University of New Mexico

Recommended Citation
Bosworth RG, Binder EM, Tyler SC, Morford JP. Automaticity of lexical access in deaf and hearing bilinguals: Cross-linguistic evidence from the color Stroop task across five languages. Cognition. 2021 Mar 31;212:104659. doi: 10.1016/j.cognition.2021.104659. Epub ahead of print. PMID: 33798950.

Automaticity of Lexical Access in Deaf and Hearing Bilinguals: Cross-linguistic Evidence from the Color Stroop Task across Five Languages

Rain G. Bosworth[1]
Eli M. Binder[2]
Sarah C. Tyler[3]
Jill P. Morford[4]

[1] National Technical Institute for the Deaf, Rochester Institute of Technology
[2] Department of Psychology, Colorado College
[3] Department of Psychology, University of California, San Diego
[4.] Department of Linguistics, University of New Mexico

* *Corresponding Author:*

Rain Bosworth, Ph.D.
Assistant Professor
National Technical Institute for the Deaf
Rochester Institute of Technology
52 Lomb Memorial Drive
Rochester, NY 14623
rain.bosworth@rit.edu

***Acknowledgments**: This research was supported in part by research grants from NSF (#1423500) and NIH (R01EY024623) to RGB. Many thanks to our participants and to Jiajun Yuan who contributed to data collection and analysis.

Highlights:

- Deaf bilinguals exhibit interference on the Color Stroop Task for dynamic ASL signs
- ASL, Chinese, Korean, Spanish and English Bilinguals were tested on the Stroop task
- Deaf bilinguals who don't use speech show strong Stroop interference for English
- Deaf bilinguals' visual processing of signs and written words is highly automated
- Cross-language orthographic similarity increases Stroop interference in bilinguals

Abstract

The well-known Stroop interference effect has been instrumental in revealing the highly automated nature of lexical processing as well as providing new insights to the underlying lexical organization of first and second languages within proficient bilinguals. The present cross-linguistic study had two goals: 1) to examine Stroop interference for dynamic signs and printed words in deaf ASL-English bilinguals who report no reliance on speech or audiological aids; 2) to compare Stroop interference effects in several groups of bilinguals whose two languages range from very distinct to very similar in their shared orthographic patterns: ASL-English bilinguals (very distinct), Chinese-English bilinguals (low similarity), Korean-English bilinguals (moderate similarity), and Spanish-English bilinguals (high similarity). Reaction time and accuracy were measured for the Stroop color naming and word reading tasks, for congruent and incongruent color font conditions. Results confirmed strong Stroop interference for both dynamic ASL stimuli and English printed words in deaf bilinguals, with stronger Stroop interference effects in ASL for deaf bilinguals who scored higher in a direct assessment of ASL proficiency. Comparison of the four groups of bilinguals revealed that the same-script bilinguals (Spanish-English bilinguals) exhibited significantly greater Stroop interference effects for color naming than the other three bilingual groups. The results support three conclusions. First, Stroop interference effects are found for both signed and spoken languages. Second, contrary to some claims in the literature about deaf signers who do not use speech being poor readers, deaf bilinguals' lexical processing of both signs and written words is highly automated. Third, cross-language similarity is a critical factor shaping bilinguals' experience of Stroop interference in their two languages. This study represents the first comparison of both deaf and hearing bilinguals on the Stroop task, offering a critical test of theories about bilingual lexical access and cognitive control.

Keywords**:** Lexical Access, Sign Language, Cross-linguistic, Stroop task, Bilingualism, Deafness

**Automaticity of Lexical Access in Deaf and Hearing Bilinguals:
Cross-linguistic Evidence from the Color Stroop Task across Five Languages**

The lexical processing abilities of deaf individuals have traditionally been investigated within a monolingual framework, focused either on the modality-specific features of sign recognition (e.g., Carreiras, Gutierrez-Sigut, Baquero & Corina, 2008; Grosjean, 1981) or on how deafness impacts the recognition of written words (e.g., Fariña, Duñabeitia, & Carreiras, 2017; McEvoy, Marschark, & Nelson, 1999). Recently, there has been a greater appreciation that deaf signers are not merely monolingual signers who are able to read, rather they have mastery of two full-fledged and distinct languages in different modalities (e.g., De Quadros, Lillo-Martin, & Pichler, 2015; Emmorey & McCullough, 2009; Kuntze, 1990; Plaza-Pust, 2016). With this increasing awareness of bimodal bilingualism, the need to understand lexical processing of both languages as they relate to each other is gaining scientific attention. The view that all readers, deaf or hearing, rely primarily on phonological decoding to recognize written words of a spoken language (Hanson, 1989; Mayer & Trezak, 2014; Wang, Trezek, Luckner, & Paul, 2008) has come under greater scrutiny due to investigations that control for the bilingual experience of the deaf participants (Barca, Pezzulo, Castrataro, Rinaldi, & Caselli, 2013) or that investigate how knowledge of a signed language may impact processing of written words in a spoken language (Meade, Midgley, Sevcikova, Holcomb, & Emmorey, 2017; Morford, Wilkinson, Villwock, Piñar, & Kroll, 2011; Quandt & Kubicek, 2018). Although bilingual lexical processing of both words and signs has been studied in hearing bimodal bilinguals (Emmorey, Borinstein, & Thompson et al., 2005; Emmorey, Petrich, & Gollan, 2012; Giezen, Blumenfeld, Shook, Marian, & Emmorey, 2015), no study to date has evaluated the automaticity of lexical access for words versus signs in deaf bilinguals in a single paradigm. The aim of this study was to investigate the automaticity of lexical access for both signs and written words in deaf bilinguals, and to probe the relationship between the two languages by comparing performance of deaf bilinguals to other bilinguals with comparable levels of proficiency in their two languages rather than comparing deaf bilinguals to hearing monolinguals whose proficiency is limited to a single spoken language. The question of automaticity is particularly pertinent due to the common assumption that deaf people are not able to gain high levels of reading proficiency due to their lack of full access to spoken phonology resulting in less detailed phonological representations (Geers, 2003;

Perfetti & Sandak, 2000; Treiman & Hirsh-Pasek, 1983). In the case of sign recognition, only a minority of deaf people begin acquiring a signed language from birth, leading to questions about whether automatic retrieval and efficient lexical access of signs can be achieved (Mayberry & Eichen, 1991). In the present study, we recruited deaf adults who use both American Sign Language (ASL) and written English on a daily basis, but who differed in their language learning history of each language, as is common in bilingual populations. In doing so, we are able to evaluate whether automaticity of lexical access is achieved in both languages, ASL and English, of deaf bilinguals despite language learning conditions that differ from the hearing population.

The well-documented Stroop effect, observed with the Stroop task (1935), has been used to probe the automaticity of visual word recognition in monolinguals and bilinguals. Participants complete two tasks: color naming and word reading. Stimuli for the Stroop task consist of color words (e.g., *blue, green, purple*). Typically, the words appear in a color congruent with the word meaning (e.g., *blue, green, purple*) or a color incongruent with the word meaning (e.g., *blue, green, purple*). In the color naming task, participants respond to the color of the ink, ignoring the meaning of the color word. In the word reading task, participants respond to the word meaning, and ignore the color of the ink. A consistent finding is that participants completing the color naming task experience high levels of interference from the meaning of an incongruent color word, indicating that visual word recognition is a highly automated skill that participants are not able to suppress. This inability to suppress is manifested as slower reaction time or greater number of errors for incongruent stimuli (i.e., slower to respond "green" for *blue,* or responding incorrectly "blue" instead of "green" for *blue)*. The Stroop effect is one of the most well replicated effects in psychology (see Brauer, 1998 for a review). While there is some controversy around the mechanism that best accounts for the pattern of results, the Stroop effect is widely considered to provide evidence of the highly automated nature of visual word recognition (Brown, Joneleit, Robinson, & Brown, 2002).[1]

More recently, the Stroop task has been administered to bilinguals to better understand how the two languages compete during lexical processing, and how Stroop effects are related to relative proficiency of the two languages. The Stroop interference effect has been observed in

[1] "Reverse" Stroop effects may or may not be seen when the task is to read the words and ignore colors; results vary depending on task implementation (Brauer, 1998; MacLeod, 2015). The classical Stroop effect is generated by the Color Naming Task that reflects interference generated by an inability to inhibit lexical processing.

bilinguals of many various languages in the world (Preston & Lambert, 1969; Tzelgov, Henik, & Leiser, 1990). Bilinguals typically display greater Stroop interference when they name the color of words from their dominant language compared to the weaker or more recently acquired language (Chen, 1986; Geukes, Gaskell, & Zwitserlood, 2015; Rosselli, et al., 2002; Mägiste, 1984). Hence, the more proficient first language (L1) causes stronger interference than the less proficient second language (L2). This finding indicates that proficiency and age of acquisition are factors influencing the magnitude of Stroop interference. In the present study, we measure the magnitude of Stroop interference to gauge the degree to which lexical access is automated for each language in deaf bilinguals. If deaf bilinguals achieve automated lexical access only in a signed language, we predict slower responses or higher error rates only for incongruent ASL stimuli. By contrast, if automated lexical access is achieved in both languages, we predict that deaf bilinguals will experience Stroop interference for incongruent stimuli presented in ASL and in English.

Coderre & Van Heuven (2014) recently proposed a hypothesis that the similarity of the orthographies of the two languages that a bilingual knows might impact how the two languages are mediated during word processing. For example, Chinese-English bilinguals who know one non-alphabetic language (Chinese) and one alphabetic language (English) might differ from bilinguals whose languages use the same orthography, such as Spanish-English bilinguals. Script similarity could *improve* performance on a Stroop task because bilinguals whose languages share the same orthography may have better cognitive control resulting from their lifelong experience inhibiting the non-target language. Alternatively, script similarity could lead to greater interference on a Stroop task due to the competition generated by similar word forms across the two languages. To test these hypotheses, Coderre & Van Heuven (2014) compared performance on a Stroop color task across groups whose languages share the same script (German-English bilinguals) and those who have very different scripts (Arabic-English bilinguals). Results demonstrated that different-script bilinguals showed the smallest Stroop interference effects in both the first and the second language, indicating that performance on the Stroop task may differ across different types of bilinguals depending on the relationship between their languages.

All prior studies of the Stroop effect in bilinguals have included participants who use two *spoken* languages. What has yet to be studied is whether the Stroop interference effect occurs for bilinguals who use two languages that have little shared articulatory, perceptual, or orthographic forms. Such is the case for deaf bilinguals who use ASL and English. ASL is a full-fledged language, possessing phonological and morpho-syntactic patterning typical of all natural languages of the world (Hill, Lillo-Martin, & Wood, 2018). Like many minority languages, it has no widely used orthography. ASL possesses a phonology that is not sound-based, but is based on manual combinations of handshape, movement, location and orientation. In comparison to hearing bilinguals who use two languages that share motor-articulatory and orthographic similarities, deaf bilinguals' experience with a signed and a spoken language might be expected to generate much less competition. In the present study, we address whether deaf bilinguals exhibit Stroop interference effects in both ASL and English, and further, whether the lack of orthographic similarity between ASL and English generates smaller Stroop effects relative to same-script and different-script hearing bilinguals.

A handful of studies have attempted to generate Stroop effects in ASL. The first studies used still photos of the hand producing ASL color signs (Marschark, 1988; Marschark & Schroyer, 1993; Wolff, Radecke, Kammerer, & Gardner, 1989). Evidence for a Stroop interference effect for ASL was inconsistent, possibly because there were several confounds in these early studies. First, proficiency of the participants was not measured. Second, the rate of presentation of the stimulus images was not controlled. And third, participants responded verbally, in either sign language or in spoken English. Given articulatory differences in the speed of production of these languages, it is difficult to compare naming times in ASL and spoken English directly (Emmorey et al., 2012; Wolff et al., 1989). The use of still photos as stimuli is also problematic. Signers develop their experience of sign recognition based on *dynamic* stimuli rather than a static image. Attempting to recognize a sign based on a still image amounts to a task in which participants are presented with degraded and atypical input, while the standard Stroop task with printed words taps a highly practiced skill.

To our knowledge, only one study to date by Dupuis & Berent (2015) developed a Stroop task with dynamic videos of ASL. Participants were ten deaf adults consisting of 4 adults who learned ASL from their deaf parents, 5 who began learning ASL by age 5 from hearing parents, and one deaf adult who started learning ASL at age 15. Participants' ASL proficiency was not assessed. Dupuis & Berent (2015) introduced an important methodological control in how participants responded. They compared performance when participants provided their response in ASL to performance when participants responded with a button press. The Stroop interference effect was replicated for both response types. Hence, when the signed signal is more comparable to standard linguistic input, signers exhibit the standard Stroop interference effect. They cannot ignore the ASL sign being produced when naming the color of a signer's hands.

The task developed by Dupuis & Berent (2015), however, would not be appropriate for comparing the processing of signs and written words. Their study overlooked an important point about perceptual features of the stimuli. They presented the video of a signer cropped from the waist up, with the whole body overlaid with a congruent or incongruent color. It is necessary to consider how processing of color and word information could be unduly affected if the stimulus size of the signer is larger than the size of the word font. Moreover, it is well known in vision science that the surface area of the colored region directly affects the speed of integration of that nonverbal color information, as greater surface area means greater retinal summation across many photoreceptors, and a stronger signal of wavelength information (Abramov, Gordon, & Chan, 1991; Atick, Li, & Redlich, 1992). Finally, it is also known that some colors are darker than others, causing less photoreceptor stimulation, than colors that emit more light (Boynton, 1979).

The current study introduces a novel methodology for assessing the Stroop color interference effect in ASL, with several added controls. First, the deaf participants completed an ASL proficiency test. The direct assessment of proficiency allowed us to evaluate the relationship between ASL fluency and interference effects on the ASL Stroop task. Second, in order to compare deaf bilinguals to hearing bilinguals on a comparable scale, all four bilingual participant groups completed proficiency self-assessments. The groups were carefully balanced in degree of fluency. Third, the current study used keyboard presses for responses, to prevent cross-language temporal and motor differences in participant responses. These controls allow us to compare performance for multiple languages. Similar to Dupuis & Berent (2015), dynamic stimuli were employed rather than still images of ASL signs. However, careful control of the visual characteristics of the signed stimuli ensured that all conditions were comparable across languages in terms of visual stimulation. The sign and word stimuli were the same size[2], and the

[2] We were able to equate the stimulus area of print and sign by cropping the visible portion of the signer to show only the dominant hand and forearm during word production. In the real world, signers typically fixate on the face during sign watching, but they can, and sometimes do, foveate on the articulating hand itself under natural conditions. This cropping was made possible because we selected color signs that are produced in neutral space (blue, green, yellow, purple) and do not contact the body.

luminance (brightness) across colors on the monitor was equated (See *Figure 1*). The sign and print stimuli were not comparable in their timing characteristics, with signs being dynamic, and taking more time to unfold, while static printed words were presented instantaneously. For this reason, when comparing deaf and hearing bilinguals we used a dependent variable that minimizes effects of the modality differences by comparing the *Stroop Interference Effect* which is calculated as the difference in performance in the incongruent and congruent conditions rather than directly reporting mean response latencies in each condition separately. This approach can mitigate the effects of modality differences since differences due to modality are present in both conditions but slowing due to interference should be above and beyond those modality differences.

Due to the lack of shared form similarity in signed and spoken languages, we predicted that deaf bilinguals would exhibit smaller Stroop interference effects than bilinguals whose languages were more similar (see Giezen et al., 2015 for a similar prediction for hearing bilingual signers). The bilingual comparison groups were selected based upon the degree of orthographic similarity of their two languages. Specifically, we selected Chinese-English bilinguals because the orthographic system of Chinese is non-alphabetic while the orthographic system of English is alphabetic. Chinese-English bilinguals learn two very distinct orthographies. By contrast, Korean-English bilinguals learn two orthographies that share no graphemes, but that are both alphabetic, which we call "moderate" similarity in the two scripts. Finally, we selected Spanish-English bilinguals who use the same alphabet for both of their languages for a group with a high degree of cross-language orthographic similarity. Thus, there were a total of four groups with increasing levels of cross-language orthographic similarity: ASL-English bilinguals (no similarity - single script), English-Chinese bilinguals (low script similarity), English-Korean bilinguals (moderate script similarity), English-Spanish bilinguals (high script similarity).

In sum, this study aims to investigate the following questions: 1) *Do deaf ASL-English bilinguals show Stroop interference in both languages?* If so, this result is indicative that visual recognition of both ASL signs and English words is highly automated in deaf ASL-English bilinguals who report no reliance on speech or audiological aids. 2) *Does cross-language script similarity impact the Stroop interference effect?* If so, we predict that ASL-English bilinguals will show the smallest interference effects since they do not have competing orthographic systems. We address these questions by first evaluating performance on the Stroop task in ASL and English in a group of deaf ASL-English bilinguals. Second, we compare Stroop color task performance in four groups of deaf and hearing bilinguals who vary in the cross-language script similarity of their two languages.

Figure 1. English and ASL Stimuli in all four colors, which were equated for differences in luminance in the actual stimuli. The ASL signs shown are still shots from the video stimuli of signs shown to participants (see *Supplementary Materials*). The corresponding English color words are presented in the font and colors used. For both English print and ASL still images shown here, the top row represents (a) congruent examples and the bottom illustrates (b) examples from the incongruent condition.

2. Experiment 1 Method

2.1 Participants

Fifteen deaf ASL-English bilinguals (47% female, age range 19 to 39, mean age 29.4 years; *SD* = 5.58) were recruited from the local community using social media and advertisements. All participants reported not using any audiological aids since age 10 years, not having a benefit with them, and using ASL daily as their primary language at home since exposure. Three participants had completed high school, 6 had completed some college or were currently enrolled in college, 3 had completed college, and 3 had completed graduate school. All were deaf from birth, except 3 who reported becoming deaf by the ages of 1, 2, and 3 years, respectively. Seven reported being native signers, with exposure to ASL since birth. Four reported exposure between ages 1-4, and 3 were exposed during late childhood between 9 and 15 years of age. The average age of ASL exposure was 3.5 years (*SD* = 4.9). Participants were compensated with a small monetary sum for their time.

The research protocol observed the tenets of the Declaration of Helsinki and was approved by the Institutional Review Board (IRB) of the University of California, San Diego. Informed consent was obtained from the participants prior to the experiment.

2.2 Materials

Stimuli consisted of the four ASL color signs BLUE, GREEN, YELLOW and PURPLE and their four corresponding English color words (See *Figure 1* and *Supplementary Materials* for video stimuli). These particular color signs were chosen because they all have the same location and movement pattern, and differ only in handshape (i.e., handshapes B, G, Y and P). Furthermore, because these signs are articulated in neutral sign space without contact on the body, we were able to crop the video to only the handshape in order to match sign and print in stimulus size.

Filming of each color sign was done against a black background, for maximum contrast. The hand through mid-forearm were cropped in iMovie, keeping exactly the same image size for all color signs. The sign model produced three repetitions of each sign. Then ASL videos were converted to greyscale and presented against a white background. For ASL congruent trials, the color of the hand matched the sign stimulus, e.g., the hand was colored blue for the sign BLUE (see *Figure 1a*). Incongruent trials were counterbalanced so that there were equal numbers of each incongruent color combination. For example, the ASL sign for BLUE was presented in equal numbers overlaid with the actual colors green, yellow, and purple (see *Figure 1b*). One third of all trials were congruent and the remaining were incongruent.

Each ASL video stimulus was shown at 30 frames per second. The production of the respective ASL signs for Blue, Green, Yellow, and Purple were 37, 36, 35, and 38 frames in duration, which was looped to create a 2 second video. The English color words were presented in Impact font and centered on the screen. As in the ASL trials, a third of all English trials were congruent (the English word *blue* was presented in blue font color, see *Figure 1*a) and the remaining incongruent trials were evenly divided between the other font colors (green, yellow and purple, see *Figure 1*b).

The RGB values for the four colors were as follows: Blue (0, 120, 255), Green (50, 205, 50), Yellow (230, 205, 0), Purple (153, 50, 204). The brightness of the four colors was controlled to be as close as possible to each other. This was confirmed with a PR650 photometer, with the resulting values: Blue (32.75 cd^2), Green (30 cd^2), Yellow (34 cd^2), Purple (27.75 cd^2). The size for the ASL stimuli was 5.4° x 10.4° and for English color words was 12.2° x 3.9° in width and height, keeping the stimulus area of each sign and word as similar as possible.

2.3 Procedure

All deaf participants were tested by a deaf experimenter, who was a native ASL user. Before beginning the Stroop task, participants completed a questionnaire about their background and their language history. Participants also completed a self-assessment of both ASL and English proficiency.

Questions probing age of acquisition, contexts of use, and proportion of day using each language were administered as well as a modified version of the Interagency Language Roundtable (ILR) self-assessment proficiency form (Interagency Language Roundtable, 2010; Stansfield, Gao, & Rivers, 2010) which had a total of 21 statements that participants answer with “yes” or “no”. The same form was used twice, once for English reading and once for ASL understanding. The 21 items each had a rank of 0, 1, 2, 3, or 4 in proficiency. Scores were based on the rank of the items to which the participant responded “yes”. For example, if a participant checked “yes” to all items in ranks 0, 1, and 2 for English, they were assigned a proficiency score of 2 for English. If half of the items were checked in rank 3, they were scored as a 2.5. This resulted in a possibility of one of 9 scores between 0 (low proficiency) and 4 (high proficiency).

They also completed an ASL proficiency self-assessment questionnaire taken from Marschark et al. (2015). Individuals give themselves scores between 0 (low proficiency) to 5 (high proficiency) for various statements about their own sign language fluency (such as “I know some signs or short phrases, and I can respond to basic questions signed to me but I very often have to ask for signs to be repeated or ask that something be signed in a different way" (1, low proficiency) to "I am able to have a very comfortable, in-depth conversation about social and school topics" (5, high proficiency).

Following the Stroop task, the participants completed the American Sign Language Comprehension Test (ASL-CT) developed by Hauser et al. (2016) which is an online multiple-choice test requiring comprehension of 30 words and simple sentences. ASL-CT scores for one participant were lost due to computer error. ASL-CT scores were collected to evaluate the relationship between ASL proficiency and the magnitude of the Stroop interference effect.

2.4 Color Stroop Task

Participants were seated 56 cm away from the monitor and positioned their head in a chin rest in a darkened room, with the monitor set to the maximum luminance value of 112 cd/m^2. Visual stimuli were presented on an LCD 24" monitor (1920 x 1080; 60 Hz) controlled by an Apple MacMini equipped with Matlab r2015b (Mathworks) and Psychtoolbox 3.0.8 (Brainard, 1997; Pelli, 1997).

Participants completed a total of 8 blocks (with 200 trials each) of the Color Stroop Task using a button press response. Four blocks were for English and 4 blocks were for ASL. For each language condition (ASL or English), there were 2 blocks of the color naming task and 2 blocks of the word/sign reading task. Order of blocks was randomized and counterbalanced.

At the beginning of each block, participants were informed which language (ASL or English) and which task (color naming or word/sign reading) they would be performing. They performed that task in the same language for the entire block. The responses were Blue, Green, Yellow and Purple. These were labeled on the keyboard with color patches on the keys *d*, *c*, *k*, and *m*, on a standard keyboard, which allowed the participant to comfortably position the index finger and middle finger of each hand on the keyboard. Each participant had a randomly paired color and key combination that stayed the same for the entirety of their own session, with which they gained practice before proceeding with the experiment. This color mapping reappeared on the monitor at the start of each block as a reminder for the participant.

Before the start of Block 1, the participant completed a practice session in order to familiarize themselves with the key mapping for the colors. The practice session presented the words and signs in grayscale for word/sign reading and colored squares for color naming.

Participants needed to correctly press the correct key for at least 16 consecutive practice trials before they could move on to initiate the experiment. Participants also had the option to revisit the practice session before each block if they wished, but this was rarely necessary.

For the practice and experiment, each stimulus was presented for 2 seconds. Participants were encouraged to respond as quickly as they could. If the participant waited longer than 2 seconds to respond, a reminder would appear on the screen of the key mapping. Once this reminder appeared, the previous trial was considered incorrect, for purposes of analysis. Only responses under 2 seconds were recorded as accurate responses. If participants got too many trials incorrect, they were required to rerun the practice session. Participants pressed the spacebar with their thumb to advance to the next trial. Feedback, as correct or incorrect, was provided after each trial. A green smiley face appeared following correct and a red X following incorrect responses for 500 milliseconds. At the end of the block, participants were alerted to take a mandatory break. Once the participant was ready to resume, the experimenter informed the participant what the instructions were for the next block. The experiment ended after 8 blocks. The Matlab code recorded the reaction time from when the stimulus first appeared to when the participants produced a keypress, and automatically calculated accuracy per trial.

2.5 Data Analysis

Only participants who performed better than 80% correct on the congruent word/sign reading blocks in both languages were included in the analysis to ensure that all participants had a high level of proficiency in both languages. No participants were eliminated. Average accuracy for the congruent ASL condition was 96.7% (*SD* = 3.4%) and average accuracy for the congruent English condition was 96.0% (*SD* = 3.3%).

The primary goal of Experiment 1 was to examine Stroop interference for dynamic signs and printed words of deaf ASL-English bilinguals who report no reliance on speech or audiological aids. We first investigated whether there was a Stroop effect for each language independently, using RT and accuracy as dependent variables. We analyzed reaction time and accuracy with 2 (Task: color naming, word reading) x 2 (Congruency: congruent, incongruent) within-subjects ANOVAs first to confirm that interference was seen within the incongruent condition, compared to the congruent condition. This was done separately for ASL and English, because watching dynamic signs and reading static print may not be not comparable in timing.

We subsequently used difference scores between the congruent and incongruent conditions to compare effects across languages. The Stroop Interference difference scores were calculated as follows for RT and accuracy:

1. Incongruent RT minus Congruent RT = Stroop Interference RT
2. Congruent accuracy minus Incongruent accuracy = Stroop Interference Accuracy

Stroop interference difference scores were analyzed with a 2 (Task: color naming, word reading) x 2 (Language: ASL, English) ANOVA. We conducted the ANOVA analyses with Block Order (first, second) as a factor, and with age as a covariate. Because these were not significant and did not impact the main results, they were not included in main analyses.

3. Experiment 1 Results

3.1 ASL - Effect of Color Incongruency on Reaction Time and Accuracy

A 2 (Task: color naming, word reading) x 2 (Congruency: congruent, incongruent) within-subjects ANOVA with repeated measures over both variables revealed a highly significant main effect of congruency on reaction time, F (1,14) = 35.80; $p < .0001$, $\eta_p^2 = .72$. Participants were significantly faster to name the hand color of an ASL sign or to identify the meaning of an ASL sign when the color was congruent with the lexical meaning of the target sign (M = 807.86 ms, SE = 37.82) than when it was incongruent with the target sign (M = 861.47 ms, SE = 35.48; see *Figure 2a*). Neither a main effect of task (F(1,14) = 1.12; p = .31, $\eta_p^2 = .07$) nor an interaction of task with congruency (F(1, 14) = 0.30; p = .59, $\eta_p^2 = .02$) was found.

A comparable analysis of the accuracy scores revealed a significant main effect of congruency on accuracy (F (1,14) = 30.70, $p < .0001$, $\eta_p^2 = .69$) such that congruent trials (M = 97.01%, SE = 0.69%) were significantly more accurate than incongruent trials (M = 93.09%, SE = 1.08%; see *Figure 2b*). As in the RT analysis, no main effect of task (F(1,14) = 2.17, p = .16, $\eta_p^2 = .13$) or interaction of task and congruency (F(1,14) = 0.21, p = .66, $\eta_p^2 = .015$) was found.

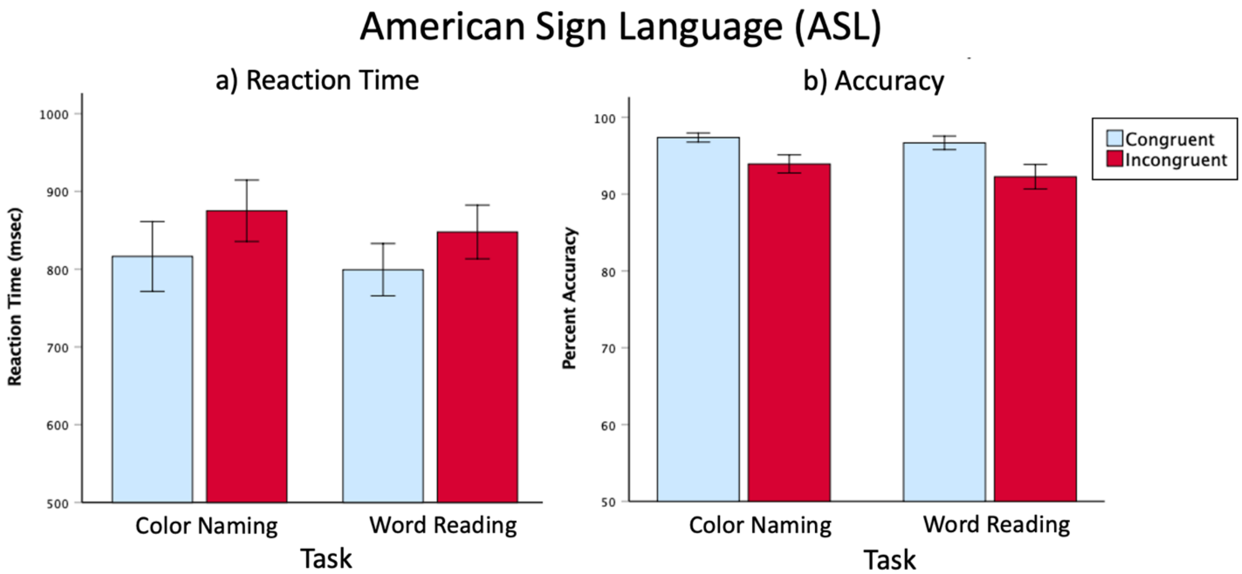

Figure 2. Reaction Time (a) and Accuracy (b) results for *ASL* from deaf ASL-English bilinguals for the color naming and word reading tasks. Averages are shown for congruent trials (light blue bars) and incongruent trials (dark red bars). Error bars denote standard error of the mean.

3.2 English - Effect of Color Incongruency on Reaction Time and Accuracy

A 2 (Task: color naming, word reading) x 2 (Congruency: congruent, incongruent) within-subjects ANOVA with repeated measures over both variables revealed a highly significant main effect of congruency on reaction time, $F(1,14) = 64.71$, $p < .0001$, $\eta_p^2 = .82$. Participants were significantly faster to name the font color of an English word, or to identify the meaning of an English word when it was presented in a color congruent with its meaning (M = 782.79 ms, SE = 30.25) than for words presented in an incongruent color (M = 836.85 ms, SE = 27.56). There was also a main effect of task ($F(1,14) = 7.05$; $p = .02$, $\eta_p^2 = .34$). Participants were significantly faster for word reading (M = 789.62 ms; SE = 28.87) than for color naming (M = 830.03 ms; SE = 30.58; See *Figure 3a*). There was no interaction of task with congruency ($F(1, 14) = 1.83$; $p = .20$, $\eta_p^2 = .12$).

A comparable analysis of the accuracy scores revealed a significant main effect of congruency on accuracy ($F(1,14) = 14.26$, $p = .002$, $\eta_p^2 = .51$) such that congruent trials (M = 96.53%, SE = 0.55%; See *Figure 3b*) were significantly more accurate than incongruent trials (M = 92.78%, SE = 1.08%). No main effect of task ($F(1,14) = 2.43$, $p = .14$, $\eta_p^2 = .15$) or interaction of task with congruency ($F(1,14) = 1.52$, $p = .24$, $\eta_p^2 = .10$) was found.

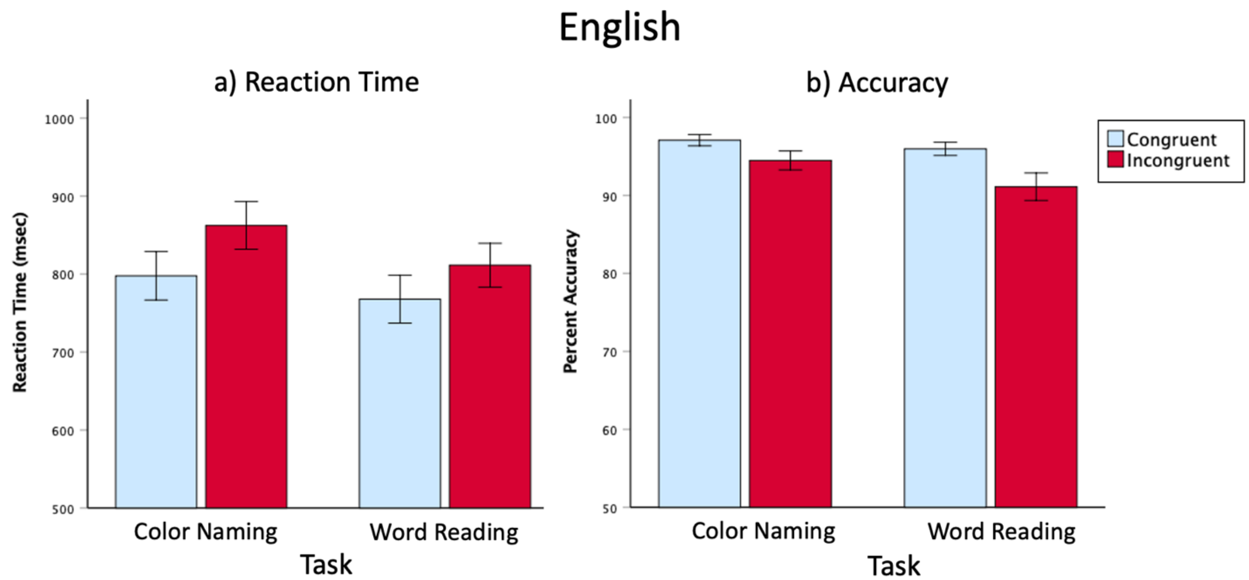

Figure 3. Reaction Time (a) and Accuracy (b) results for *English* from deaf ASL-English bilinguals for the English color naming and word reading tasks. Averages are shown for congruent trials (light blue bars) and incongruent trials (dark red bars). Error bars denote standard error of the mean.

3.3 Cross-language Comparison of Stroop Effects

A 2 (Task: color naming, word reading) x 2 (Language: ASL, English) within-subjects ANOVA with repeated measures over both variables revealed no significant effects of language, F (1,14) = 0.001; p = .97, η_p^2 = .05, or of task, F (1,14) = 1.57; p = .23, η_p^2 = .10, on Stroop Interference RT. In other words, the size of the Stroop interference effect did not differ significantly between ASL and English. Moreover, there was no Task x Language interaction, F (1, 14) = 0.22, p = .65, η_p^2 = .02 indicating that the interference effect was not significantly greater for one task or the other in either language.

A comparable analysis of the accuracy scores revealed no significant effects of language (F(1,14) = 0.10, p = .78, η_p^2 = .01) or task (F(1,14) = 0.82, p = .38, η_p^2 = .06) on Stroop Interference accuracy. No Task x Language interaction (F(1,14) =0 .55, p = .47, η_p^2 = .04) was found.

3.4 Relationship between ASL proficiency and ASL Stroop Interference Effect

In order to explore the relationship between proficiency in ASL and performance on the Stroop task, participants completed a test of ASL proficiency, the ASL-CT (Hauser et al., 2016). The average ASL-CT score was 85% (*SD* = 31%) with a range of 67% to 97%. We asked whether participants who were more proficient in ASL would have more difficulty suppressing their knowledge of ASL on the Stroop task (see *Figure 4*). ASL proficiency was not significantly correlated with Stroop Interference on the ASL word reading task which did not require inhibiting ASL lexical processing, $r = .02$, $p = .95$, 95% CI [-.52, .54], but it was for the color naming task ($r = .60$, $p = .02$, 95% CI [.10, .86]). The positive correlation between ASL proficiency and the size of the Stroop interference effect for color naming indicates that participants who were more proficient in ASL experienced more difficulty ignoring the meaning of the ASL signs when naming the color of the hand.

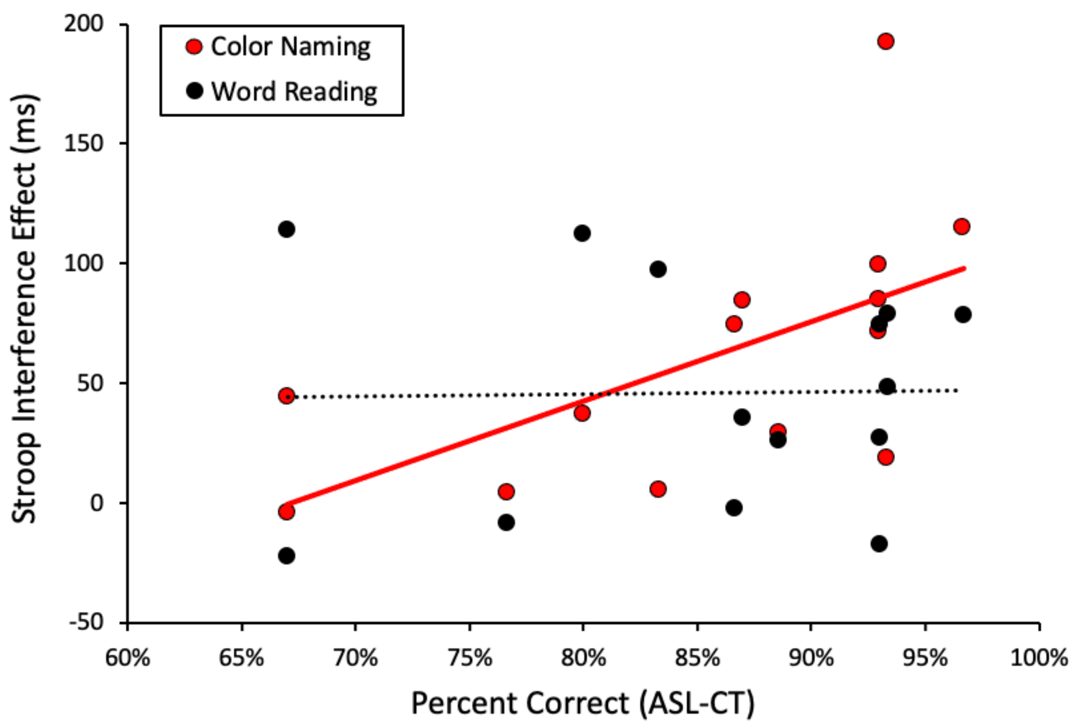

Figure 4. Correlation of ASL Proficiency and ASL Stroop Interference Effect for both tasks, the ASL color naming (*red*) and word reading (*black*). Percent correct on the ASL-CT proficiency test and Stroop Interference Effect (ms) values for each task in ASL were fit with a linear regression line. A significant positive relationship was seen for color naming. Note the correlation is still significant when the largest Stroop Interference RT is removed.

4. Experiment 1 Discussion

The aim of Experiment 1 was to establish whether the classic Stroop interference effect could be found for both ASL signs and English words for deaf ASL-English bilinguals, and to compare the relative magnitude of the Stroop interference effect in ASL and English in these bilinguals. This study is the first to test the Stroop task using dynamic ASL videos and English words equated in size and luminance, which is important because words and signs differ in size, and colors vary in luminance. Results confirmed strong Stroop interference in both languages of these participants. They were significantly slower to name the hand color of an ASL sign and the font color of an English word when the stimulus color conflicted with the lexical meaning of the sign or word.

Further, we found that within this population of highly skilled signers, individuals with the strongest ASL proficiency, as measured by the ASL-CT test (Hauser et al, 2016) which measures word and sentence comprehension, demonstrated the greatest ASL Stroop interference effects. Higher proficiency means more automatic retrieval of a word's meaning, or in this case, of a sign's meaning, and this generates greater interference with a participant's ability to ignore the irrelevant stimulus dimension. These results are consistent with prior findings from spoken language showing language proficiency is a primary determinant of the magnitude of Stroop effects (Chen, 1986; Geukes, Gaskell, & Zwitserlood, 2015; Mägiste, 1984; Marian, Blumenfeld, Mizrahi, Kania, & Cordes, 2013; Rosselli, et al., 2002).

The replication of the Stroop Task in ASL with increased controls and the parallel findings for signed and spoken language stimuli provides new evidence of the automaticity of lexical access for both ASL signs and English words in deaf bilinguals. Importantly, the participants recruited for this study did not all acquire ASL and English from birth, but reported using both languages on a daily basis, without reliance on or full access to sound-based phonological coding of print. Our approach to sampling from the deaf population was to include a representative sample of deaf bilinguals instead of including only the minority of bilinguals who have the privilege of early exposure to language. These initial findings can be further explored in future studies with larger populations of deaf bilinguals to investigate how more variation in proficiency and age of first exposure in each language is related to automaticity of lexical processing. For present purposes, the results demonstrate that Stroop effects are robust across bilingual learning contexts when the task controls for language proficiency, visual properties of the stimuli and uses dynamic videos of ASL signs.

The current results provide corroborating evidence that access to speech is not a necessary requirement to develop highly automated lexical access of written English words (see Hall, Hall & Caselli, 2019 for a discussion). Several recent studies also provide evidence disputing the assumption that deaf individuals must rely on the same phonological decoding strategies as hearing individuals to achieve high levels of reading proficiency. Mayberry, del Giudice & Lieberman (2011) showed that phonological decoding and awareness predicted only 11% of variability in reading proficiency in deaf readers. Likewise, Emmorey, McCullough, & Weisberg (2016) found that English reading comprehension in a group of deaf ASL-English bilinguals correlated with English vocabulary size and print exposure, but not with a measure of phonological awareness. While some deaf readers may use phonological decoding in comparable ways to hearing readers, accumulating evidence indicates that this is not the sole avenue for achieving reading fluency and automatic lexical access of written words.

No prior study of Stroop effects in deaf participants has compared the performance of deaf bilinguals to other populations of bilinguals. We now turn to a comparison of four groups of bilinguals who differ in the relationship between their two languages to investigate the hypothesis that Stroop effects in bilinguals depend on the orthographic similarity of their languages.

5. Experiment 2 Method

5.1 Participants

A total of 142 bilinguals were recruited from the University of California, San Diego (UCSD) student population for participation. We recruited individuals who identified themselves as being bilingual in English and either Chinese (n = 76), Korean (n=23), or Spanish (n=43). This diversity in spoken languages was made possible by the diverse student demographic population at UCSD. In order to select a subset of the spoken language bilinguals who best matched the deaf bilingual group, we included only bilinguals who were balanced in the proficiency of their two languages, or dominant in English. Balanced or English-dominant bilinguals were selected because all of the deaf bilinguals were raised in the US, and the deaf bilinguals all responded more quickly to the English stimuli than the ASL stimuli. For these two reasons, we included the hearing bilinguals who reported English proficiency as comparable to or greater than their proficiency in their other language.

All deaf and hearing participants completed a modified version of the Interagency Language Roundtable (ILR) reading assessment for each of their two languages and were scored according to the rubric (as described above for Experiment 1; Interagency Language Roundtable, 2010; Stansfield, Gao, & Rivers, 2010). There was no significant difference between the four groups in mean self-proficiency ratings for English reading, $F(1,3) = 2.61$; $p = .06$, $\eta_p^2 = .087$. Mean ratings are very similar for all groups, shown in *Table 1*. Likewise, there was no significant difference between the three hearing groups in mean self-proficiency ratings for L2 reading, $F(1,2) = 2.12$; $p = .13$, $\eta_p^2 = .06$. The groups also did not differ in their mean age at the time of test, or in their mean age when they learned English or their other language, all p's < 0.20.

All hearing participants completed the Bilingual Language Profile (BLP; Birdsong, Gertken, & Amengual, 2012).[3] The participants were sorted on the basis of their self-assessment of expressive and receptive language abilities on a 7-point scale (0 = not at all to 6 = very well) for both languages. For each participant, a language dominance score was computed as the sum of their self-assessment of speaking and understanding English minus the sum of their self-assessment of speaking and understanding the other language. A score of zero reflected balanced proficiency in the two languages, and a positive score reflected dominance in English. Only participants with a language dominance score of 0 or greater were included (Chinese, n = 23; Korean, n=15; Spanish, n = 36).

As described in the methods for Experiment 1, participants who performed below 80% accuracy on the congruent trials of the word reading task in either language were eliminated. Three participants (1 for Chinese, 2 for Spanish) were eliminated on this basis. The remaining 71 hearing bilinguals and the 15 deaf bilinguals from Experiment 1 were combined for a total of 86

[3] We began by collecting data on the deaf bilinguals. We directly assessed their proficiency in ASL, and used a reading assessment that is much more detailed than a simple Likert scale to assess their proficiency in English. We later added the hearing bilingual groups, and we were not able to directly assess proficiency in their L2, so we decided to use the BLP (i.e., self-assessment). The ILR measure of English reading proficiency used with the deaf bilinguals provides a better measure of English reading than the BLP self-rating, so we asked all hearing bilinguals to complete that as well.

participants whose data was included for analysis. The number of participants per group and demographics are presented in *Table 1***.**

Table 1. Participant Demographics

Bilingual Group	N	% Female	Mean Age (SD)	Age Range	Mean Self-Proficiency Reading English (0-4)	Mean Self-Proficiency Reading Language 2 (0-4)	BLP Speaking & Understanding Spoken English (0-6)	BLP Speaking & Understanding Language 2 (0-6)
Deaf ASL-English	15	47	29.4 (6.5)	19 to 39	3.67 (.52)	not applicable*	no speech used †	6.0 (0.0)
Hearing Chinese-English	22	86	19.6 (1.36)	18 to 23	3.55 (.72)	1.91 (1.38)	5.45 (1.39)	4.33 (1.65)
Hearing Korean-English	15	80	19.8 (.94)	18 to 21	3.80 (.41)	1.82 (.85)	5.93 (.18)	3.71 (1.75)
Hearing Spanish-English	34	76	20.2 (1.89)	18 to 25	3.91 (.31)	2.41 (.94)	5.85 (.38)	4.63 (1.24)

** ASL does not have orthography. † This is by subject selection/study design*

The research protocol observed the tenets of the Declaration of Helsinki and was approved by the UCSD Institutional Review Board (IRB). Informed consent was obtained from the participants prior to the experiment.

Table 2. Stimulus Words in each Language by Orthographic Similarity to English

Control Language	Single Script*	Low Similarity	Moderate Similarity	High Similarity
English	*ASL*	*Chinese*	*Korean*	*Spanish*
Blue		蓝	푸른 색	Azul
Green		绿	채색	Verde
Yellow		黄	노란색	Amarillo
Purple		紫	보라색	Morado

* Note that the ASL stimuli were presented as videos and these are screenshots.

5.2 Materials and Procedure

The stimuli and procedures were similar to Experiment 1, described above. In Experiment 1, the “other” language was ASL, while in Experiment 2, the “other” language was ASL, Chinese, Korean or Spanish (see *Table 2* for a list of all stimulus words). Written words were presented on a white background on which the graphemes were colored blue, green, yellow or purple. All written stimuli were matched as closely as possible in number of shaded vs. white pixels (by counting the number of pixels using Matlab code) and in the absolute area (width x height). Thus, the absolute size varied by only a few pixels. The height of the English, Korean, and Spanish font was 4.55 degrees of visual angle (viewed at 56 cm), while the height and width of the Chinese font was 4.86 degrees of visual angle. All incongruent trials were counterbalanced so that there was an even number of combinations of colors. One third of all trials were congruent and the rest were incongruent.

5.3 Color Stroop Task

As in Experiment 1, participants completed a total of 8 blocks (with 200 trials each) with four blocks for English, and four blocks for the participant’s other language. Two blocks of the word reading task and two blocks of the color naming task were completed in each of the participant’s two languages. Order of blocks was randomized and counterbalanced.

At the beginning of each block, the participant was informed which language and which task they would be performing. They performed the task in the same language for the entire block. The responses were Blue, Green, Yellow, and Purple. These were labeled on the keyboard with color patches on the keys *d*, *c*, *k*, and *m*, on a standard keyboard, which allowed the participant to comfortably position the index finger and middle finger of each hand on the keyboard. Each participant had a randomly paired color and key combination that stayed the same for the entirety of their own session. This color mapping reappeared at the start of each block as a reminder for the participant.

Before the start of Block 1, participants completed a practice session in order to familiarize themselves with the key mapping for the colors. The practice session presented the words and signs in grayscale for word/sign reading and colored squares for color naming. Participants needed to press the correct key for at least 16 practice trials in a row before they could initiate the first block of the experiment. The Matlab code recorded the reaction time from when the stimulus first appeared to when the participants produced a keypress and calculated accuracy per trial.

5.4 Data Analysis

As with Experiment 1, only participants who performed better than 80% correct on the congruent word/sign reading blocks in both languages were included in the analysis to ensure that all participants had a high level of proficiency in both languages. Because the aim of Experiment 2 was not to establish the Stroop Effect for all four groups, but to look for differences in the size of the Stroop effect relative to orthographic similarity and to minimize differences due to language-specific characteristics of the stimuli, the dependent measure selected was Stroop Interference difference scores, calculated as for Experiment 1.

Using difference scores as the dependent measure eliminated differences in reading time across the four groups (see *Supplementary Materials* for mean scores in each condition for each group). The Chinese-English bilinguals were the fastest readers of the four groups for both English and the other language, possibly due to their experience reading a language with a non-alphabetic orthography (Pasquarella, Chen, Gottardo, & Geva, 2015).

The study used a 4 (Bilingual Group) x 2 (Task) x 2 (Language) mixed design with Bilingual Group as a between-subjects variable, and Task (color naming, word/sign reading) and Language (English, Other Language) as within-subjects variables. ANOVAs were conducted for two dependent measures, Stroop Interference RT and Stroop Interference Accuracy. Age and Block Order (first, second) were examined as factors within the ANOVA model, and found not to impact results whatsoever, hence these variables were not included in the final ANOVA model.

Because the number of participants across Bilingual groups was unequal, it is necessary to check whether the variances of score distributions of the four groups are equal, a statistical assumption underlying the use of ANOVA. Levene's test was performed to test for homogeneity of variance of the four groups. Levene's test results indicated that the homogeneity of variance assumption was met ($F = .91$, $p = .44$), indicating that the four groups are homogeneous in terms of the distribution of variance in the dependent measure.

6. Experiment 2 Results

6.1 Effects of Orthographic Similarity on Stroop Interference Effects in Bilinguals

Table 3 presents the average Stroop Interference RTs for each group for each language. All four groups exhibited robust Stroop effects in both languages. That is, in no case was the Stroop Interference RT close to 0 as would be the case if performance on the incongruent trials was similar to performance on the congruent trials.

Table 3. Means (and Standard Errors) for Stroop Interference Effects on Reaction Time for the Color Naming and Word Reading Tasks for each Language by Bilingual Group.

	Color Naming Task		Word Reading Task	
Bilingual Group	**English**	**Other Language**	**English**	**Other Language**
Deaf ASL-English (N=15)	64.55 (13.13)	58.66 (13.02)	43.56 (12.03)	48.55 (13.42)
Hearing Chinese-English (N=22)	79.73 (10.84)	56.07 (10.75)	28.38 (9.93)	52.77 (11.08)
Hearing Korean-English (N=15)	89.90 (13.13)	54.14 (13.02)	28.27 (12.03)	54.58 (13.42)
Hearing Spanish-English (N=34)	115.05 (8.72)	91.27 (8.65)	45.72 (7.99)	45.94 (8.91)

The aim of the study was to determine whether bilinguals differ in the degree of Stroop Interference experienced relative to the similarity of the orthography of their two languages. A 4 (Bilingual Group) x 2 (Task: color naming, word reading) x 2 (Language: English, Other Language) mixed ANOVA revealed a significant main effect of Bilingual Group on Stroop Interference RT, $F(3, 82) = 3.56$, $p = .018$, $\eta_p^2 = .12$. There was also a significant main effect of Task on Stroop Interference RT, $F(1, 82) = 27.74$, $p = .0001$, $\eta_p^2 = .25$, and the Task x Group interaction was marginally significant, $F(3, 82) = 2.66$, $p = .05$, $\eta_p^2 = .09$. On average, participants exhibited a stronger Stroop interference effect for color naming ($M = 76.17$ ms, 95% CI[67.07, 85.27]) than for word reading ($M = 43.47$ ms, 95% CI [35.05, 51.89]).

Stroop effect means, which are shown in *Figure 5,* indicate that the significant main effect of Bilingual Group was driven by group differences for color naming and not for word reading, as predicted (see *Introduction*). To explore which groups were driving the significant main effect of Bilingual Group, pairwise comparisons with Bonferroni adjustments for multiple comparisons were conducted for the color naming task. Results indicated that the Spanish-English group had a significantly larger Stroop interference effect compared to the ASL-English ($p = .008$) and the Chinese-English ($p = .01$) groups, and compared to the Korean-English bilinguals, a trend was observed (p = .08). To put these differences in perspective, the Spanish-English group had Stroop interference that was, on average, 42, 35, and 31 ms *greater* than values observed in the ASL, Chinese, and Korean groups, respectively.

There was a significant Task x Language interaction, $F(1, 82) = 12.75$, $p = .001$, $\eta_p^2 = .14$. Pairwise comparisons with Bonferroni adjustments for multiple comparisons revealed that all four groups experienced more interference in English than in the other language (all p's < .01) for color naming, but not for word reading. These results suggest that English is the more highly automated language of the two used, across all groups, as would be expected due to the exclusion of participants who were not balanced or English-dominant.

No significant Language x Group interaction was found, $F(3,82) = 0.41$, $p = .75$, $\eta_p^2 = .015$, nor was there a significant three-way Task x Language x Group interaction, $F(3, 82) = 1.16$, $p = .33$, $\eta_p^2 = .04$.

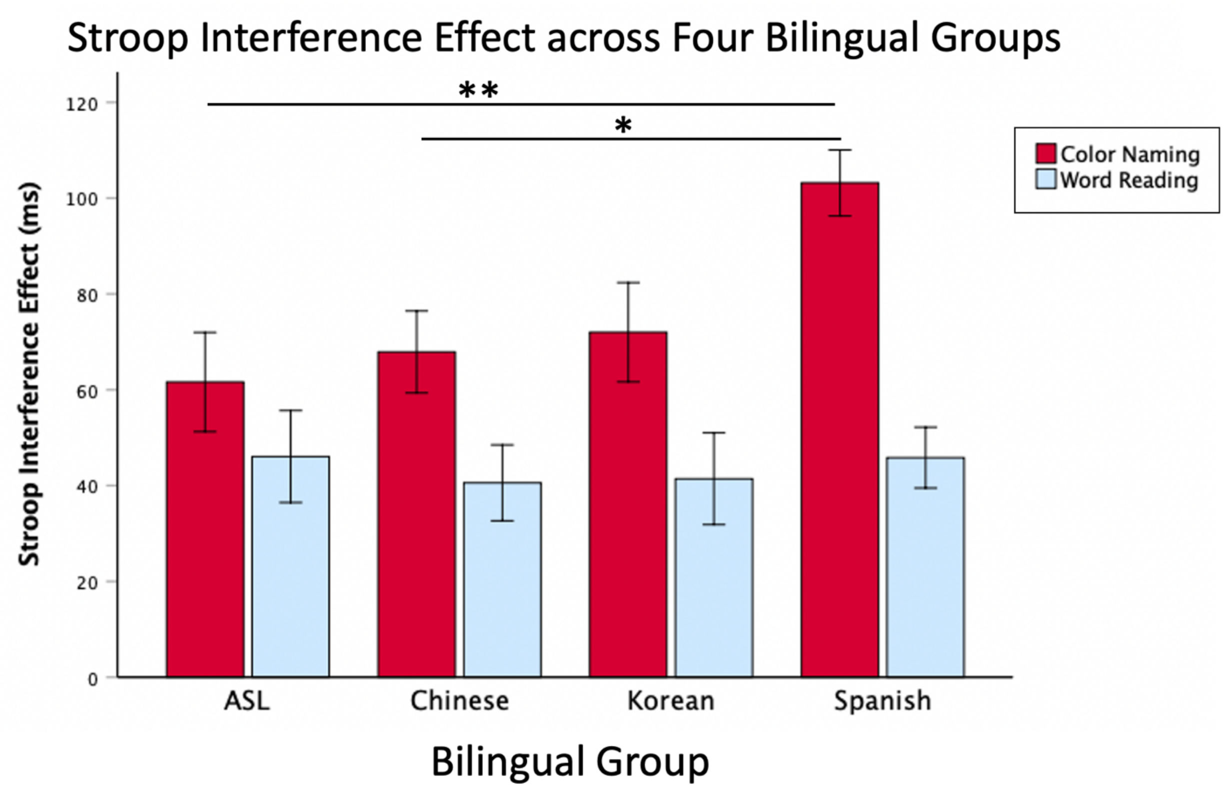

Figure 5. Mean Stroop Interference effects for reaction time for each Bilingual Group. Results are shown for color naming and word reading, collapsed across language condition. Comparisons with asterisks denote Bonferroni-adjusted pairwise comparisons between groups that were significant (** p < .01; * p < .05). Error bars denote standard error of the mean.

A comparable analysis of accuracy revealed no effects of Bilingual Group, $F(3,82) = 0.23$, $p = .87$, $\eta_p^2 = .01$, or Language, $F(1,82) = 1.52$, $p = .22$, $\eta_p^2 = .02$. The main effect for Task was significant, $F(1,82) = 7.02$; $p = .01$, $\eta_p^2 = .08$, driven by larger Stroop interference accuracy effects for Color naming (M = 4.6%, SE = 0.5) than word reading (M = 3.0%, SE = 0.5). None of the interactions reached significance (all *p*'s > .40).

7. Experiment 2 Discussion

The aim of Experiment 2 was to investigate the hypothesis that cross-language orthographic similarity may influence the ability of bilinguals to suppress language knowledge during the Stroop color naming task. We compared four types of bilinguals who differed in the similarity of the writing systems of their two languages. ASL-English bilinguals were predicted to show the smallest Stroop interference effect because the ASL-English bilinguals only use a writing system for English, and none for ASL. Chinese-English bilinguals, who know one alphabetic and one non-alphabetic language, were also predicted to show little interference. Korean-English bilinguals, who know two alphabetic languages, could potentially show larger interference effects. However, their two orthographies share no overlapping graphemes, so they were predicted to experience less interference than Spanish-English bilinguals, who were expected to show the greatest Stroop interference because they know two alphabetic languages that use the same alphabet. Importantly, all four bilingual groups exhibited the classic Stroop Interference effect; incongruent stimuli were more difficult to process than congruent stimuli for the color naming task. However, of the four groups, the Spanish-English bilinguals experienced the greatest interference on the color naming task, that is, when attempting to suppress word reading while naming the color of the font of words presented in English or Spanish. By contrast, there was no significant difference across the four bilingual groups on the word reading task.

Contrary to our predictions, the Stroop effect for color naming did not differ significantly between all groups. The dramatic difference between the Spanish-English group and the other three groups may be due not only to script similarity, as predicted by Coderre & Van Heuven (2014), but also to similarities in letter-sound mappings across Spanish and English. Both the Chinese-English and the Korean-English bilingual groups used writing systems with no shared graphemes. In the future, it would be informative to compare similar-script bilinguals whose languages differ contrastively in the transparency or alignment of the letter-sound mappings.

A main effect of Language on the color naming task and a marginal effect of Language on the word reading task indicated that the bilingual participants were balanced or English dominant. On the color naming task, when they had to inhibit language knowledge for successful task completion, they experienced more interference with English stimuli than with the other language. On the word reading task, when they had to inhibit attention to an irrelevant feature of the stimuli and activate their language knowledge, they experienced more interference with stimuli in the non-English language. These findings further support our conclusion that the bilingual participants were all highly proficient in English, including the deaf ASL-English bilinguals.

A number of researchers have investigated Stroop effects in bilinguals by manipulating whether the stimulus and response languages were the same (within language condition) or different (between language condition). These studies generally find greater Stroop interference effects within language than between language, with lower levels of between language Stroop interference for different script bilinguals (Fang, Tzeng & Alva, 1981). Although the current study did not include a between language condition, the results are generally consistent with prior claims that cross-language similarity leads to greater competition in mixed-language contexts. Although there is evidence that all bilinguals, regardless of language similarity, activate both languages during lexical processing (e.g., Morford et al., 2011 for ASL-English bilinguals; Thierry & Wu, 2007 for Chinese-English bilinguals; Lee, Nam & Katz, 2007 for Korean-English bilinguals), it is not yet clear to what degree word forms compete across languages with different phonetic or orthographic forms. The current results suggest that bilinguals whose languages are orthographically distinct are less likely to experience interference on the Stroop task even when there is no explicit language mixing included in the study design.

8. General Discussion

The results of both studies support three primary conclusions. First, Stroop interference effects are found for both signed and spoken languages. Second, deaf bilinguals who report no reliance on speech or audiological aids show highly automated lexical access of printed words as well as signed words. Third, cross-language script similarity is a critical factor shaping bilinguals' experience of Stroop interference in their two languages.

Although the Stroop task has often been used to study the automaticity of lexical access of printed words, few have done so with signs. The current study implemented a novel design for the ASL Stroop task by combining the use of dynamic video stimuli with careful control of the visual parameters of the signed stimuli to be matched to the visual parameters of printed words used for spoken language Stroop tasks. This allowed us to compare Stroop task performance across a variety of languages, including ASL, Chinese, Korean, Spanish and English. The results demonstrate that Stroop interference effects are not limited to spoken languages. All groups of bilinguals demonstrated Stroop interference effects regardless of the language used.

This study was also the first to directly assess ASL language proficiency of deaf participants completing an ASL Stroop task. The tendency to limit studies of ASL signers only to those who were exposed to ASL from birth has resulted in a preponderance of research findings that tell us about optimal performance, but not about performance in the deaf population as a whole. The current study revealed that ASL proficiency is correlated with the magnitude of the interference effect on the ASL color naming task. In other words, with greater proficiency in ASL, participants have more difficulty suppressing ASL knowledge when trying to name the color of the hand producing an ASL sign. Further, this study was the first to assess Stroop interference effects in both languages of deaf bilinguals. The pattern of results indicates that deaf bilinguals can achieve high levels of proficiency in both signed and spoken languages. Specifically, the study revealed Stroop interference effects on the color naming task in both ASL and English and no interaction of Stroop interference with language. Deaf bilinguals automatically accessed the meaning of the ASL signs and the written English words when trying to name the color of the stimulus, which slowed performance in the incongruent stimulus condition relative to the congruent stimulus condition. While these findings are entirely consistent with findings from the bilingual literature (Brauer, 1998; Coderre, & van Heuven, 2014; Tzelgov et al., 1990), they provide the first demonstration of automated lexical access in deaf bilinguals comparing across a signed and a spoken language. One implication of these findings is that it is essential to evaluate lexical processing in both sign and print in deaf bilinguals to inform best practices in bilingual literacy education for deaf students (Marschark & Hauser, 2008).

A central question about reading in the deaf population is whether orthographic forms automatically activate both phonological and lexico-semantic representations. Several recent studies provide evidence that deaf readers activate phonological forms automatically when viewing orthographic word forms, but that they appear to rely less on phonological processing than on more direct connections between orthography and lexico-semantics for reading comprehension (Glezer, Weisberg, Farnady, McCullough, Midgley, et al., 2018 for English; Gutierrez-Sigut, Vergara-Martínez & Perea, 2017 for Spanish). Bélanger, Mayberry & Rayner (2013) investigated whether extracting phonological information during parafoveal processing of text could differentiate stronger and weaker deaf readers of English (cf. Bélanger, Baum & Mayberry, 2012 for deaf readers of French). They found that only a hearing control group showed evidence of extracting phonological information during parafoveal processing while both groups of deaf readers as well as the hearing readers benefited from orthographic information in parafoveal preview. Evidence of strong orthography to lexico-semantic associations in deaf readers was provided by Gutierrez-Sigut, Vergara-Martínez & Perea (2019) who found that ERPs generated by deaf readers show an early sensitivity (150 ms post-stimulus) to orthographic differences such as upper- vs. lower-case print during masked identity priming regardless of whether they were viewing words or pseudowords, but by 250 ms post-stimulus, processing of the visual form of words and pseudowords diverges due to lexical semantic feedback. In other words, although deaf readers develop sufficiently strong associations between orthography and phonology for automatic spread of activation, these associations may play little role in the efficiency of visual word recognition; by contrast, orthography to lexico-semantic associations appear to influence reading for both weaker and stronger deaf readers as soon as 250 ms after a word is fixated in language-specific ways. The Stroop task cannot distinguish between a reliance on phonological or lexico-semantic associations during text processing, but the difficulty exhibited by deaf bilinguals to inhibit word recognition processing during color naming is another indication that lexico-semantics, in this case associated with color words, are rapidly activated by orthographic word forms.

There is robust evidence that lexical processing in bilinguals is language non-selective, i.e., that bilinguals activate words of both languages even in monolingual contexts for both deaf (Kubuş, Villwock, Morford, & Rathmann, 2015; Hosemann, Mani, Herrmann, Steinbach, & Altvater-Mackensen, 2020; Meade et al., 2017; Morford et al., 2011; Morford, Kroll, Piñar, &

Wilkinson, 2014; Morford, Occhino, Piñar, Wilkinson, & Kroll, 2017; Morford, Occhino, Zirnstein, Kroll, Wilkinson, & Piñar, 2019; Quandt & Kubicek, 2018) and hearing (Giezen et al., 2015; Marian & Spivey, 2003; Morford et al., 2014; Shook & Marian, 2012) bilingual signers. These findings have motivated comparisons of bilinguals and monolinguals on the Stroop task to investigate the hypothesis that bilinguals may exhibit greater executive control due to their experience inhibiting and selecting between multiple languages (Bialystok, Craik & Luk, 2008; Coderre, van Heuven, & Conklin, 2013). These investigations have been extended to hearing signers (Giezen et al., 2015), but not yet to deaf signers (but see Hauser, Lukomski & Hillman, 2008 for an overview of factors shaping executive function in deaf signing children). Giezen et al. (2015) compared inhibitory control on a non-linguistic spatial Stroop task to the degree of activation of cross-language ASL competitors during an English auditory word recognition task using the visual world paradigm. They found that hearing ASL-English bilinguals' ability to inhibit attention to the incongruent spatial dimension on the Stroop task was correlated with the ability to inhibit fixations on cross-language competitors during the spoken English word recognition task. They propose that both unimodal and bimodal bilinguals engage domain-general cognitive mechanisms to manage the competition resulting from the activation of cross-language competitors during lexical processing.

The current study does not address the question of differences in selective attention or executive control in monolinguals vs. bilinguals. Instead, we use the Stroop task to assess automaticity of lexical access and differences between bilingual populations. The Stroop task engages selective attention to one specific stimulus feature while inhibiting attention to a second stimulus feature (Algom & Chajut, 2019). The degree of cross-language activation in bilinguals is mediated to some extent by similarity in phonology, orthography and semantics as demonstrated, for example, by the finding of faster processing of cognates than of non-cognates by bilinguals (Cristoffanini, Kirsner, & Milech, 1986). In the current study, we asked whether cross-language orthographic similarity would influence Stroop interference effects. We argue that greater cross-language orthographic similarity is more likely to activate the non-target language in addition to activation of the irrelevant stimulus feature during the Stroop task. As a consequence, we predicted that Spanish-English bilinguals, who needed to inhibit both Spanish and English orthographic knowledge when presented with word forms in either language, would experience the greatest levels of Stroop interference. By consequence, ASL-English bilinguals, who use a single orthography, were predicted to experience the least Stroop interference due to a lack of competition between two orthographic systems.

Results supported our prediction that bilinguals who use the same orthography for both languages, in this case Spanish-English bilinguals, would experience the greatest Stroop interference (Coderre & Van Heuven, 2014). While there were nominal differences across the remaining bilingual groups – Korean-English, Chinese-English and ASL-English bilinguals – these differences did not reach significance. Hence, the data suggest that a shared orthography across the languages of bilinguals has the potential for increasing competition during the Stroop color naming task. Future studies could probe the possibility that variation in phoneme-grapheme correspondences across bilinguals who use the same orthography for multiple languages is related to variation in Stroop interference. For present purposes, we conclude that the evidence supports the view that the degree of cross-language cooperation and competition in bilinguals is mediated by the degree of similarity of various dimensions of the languages, and specifically, of cross-language orthographic similarity.

There are several limitations to our study that can be addressed in future work. An important question is how the sociolinguistic communities in which bilinguals are exposed to their languages influence their paths to proficiency and literacy. We recruited participants from the same region of the United States, but this does not mean that the participants experienced similar attitudes toward their languages (e.g., Hill, Lillo-Martin & Wood, 2018). Inclusion of a larger sample of deaf participants would also allow for additional questions about language learning history and its effects on lexical processing to be addressed. Finding a sample of deaf bilinguals with high levels of proficiency in both ASL and English but greater variation along other dimensions can be challenging. Moreover, at the time this study was carried out, there were few options for assessing ASL proficiency directly. Given the variation across languages, it is rare to find assessment measures that can control for language proficiency on a variety of languages. One strength of the current study was that the experimental task required very little linguistic knowledge, but nevertheless allowed for cross-group comparisons of bilingual lexical processing.

More generally, the Stroop Task has well-known limitations that should be taken into consideration. One is that there are few words presented over many trials, which may not realistically reflect lexical processing in contextually and grammatically rich discourse settings.

However, although generality is compromised, this can also be advantageous, because the stimuli are simple and the task can be administered quickly with minimal instructions, which may be ideal for studying children or populations where complicated instructions are not feasible. It is also easier to make cross-language comparisons when the stimuli do not vary along grammatical dimensions. The current results would be bolstered by corroboration using different methods.

In summary, our study has provided evidence of the high level of automaticity of lexical access for both ASL and English in a group of deaf bilinguals. Most explorations of bilingual lexical processing have excluded deaf bilinguals. The current study demonstrates that deaf and hearing bilinguals alike must manage cognitive demands on the language processing system, and that language proficiency and similarity modulate those demands for both signed and spoken languages.

References

Abramov, I., Gordon, J., & Chan, H. (1991). Color appearance in the peripheral retina: effects of stimulus size," *Journal of Optical Society of America A 8,* 404–414 (1991). doi: 10.1364/JOSAA.8.000404

Algom, D., & Chajut, E. (2019). Reclaiming the Stroop effect back from control to input-driven attention and perception. *Frontiers in Psychology, 10*, 1683. doi: 10.3389/fpsyg.2019.01683

Atick, J. J., Li, Z., & Redlich, A. N. (1992). Understanding retinal color coding from first principles. *Neural Computation, 4*, 559-572. doi:10.1162/neco.1992.4.4.559

Barca, L., Pezzulo, G., Castrataro, M., Rinaldi, P., & Caselli, M. C. (2013). Visual word recognition in deaf readers: Lexicality is modulated by communication mode. *PloS One, 8*, e59080. doi:10.1371/journal.pone.0059080

Bélanger, N. N., Baum, S. R., & Mayberry, R. I. (2012). Reading difficulties in adult deaf readers of French: Phonological codes, not guilty! *Scientific Studies of Reading*, *16*, 263–285. doi: 10.1080/10888438.2011.568555

Bélanger, N. N., Mayberry, R. I., & Rayner, K. (2013). Orthographic and phonological preview benefits: Parafoveal processing in skilled and less-skilled deaf readers. *Quarterly Journal of Experimental Psychology*, *66*, 2237–2252. doi: 10.1080/17470218.2013.780085

Bialystok, E., Craik, F.I.M., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 859–873. doi:10.1037/0278-7393.34.4.859

Birdsong, D., Gertken, L.M., & Amengual, M. (2012). *Bilingual Language Profile: An Easy-to-Use Instrument to Assess Bilingualism.* COERLL, University of Texas at Austin. Web. 20 Jan. 2012. https://sites.la.utexas.edu/bilingual/

Blumenfeld, H. & Marian, V. (2013). Parallel language activation and cognitive control during spoken word recognition in bilinguals. *Journal of Cognitive Psychology, 25,* 547-567. doi:10.1080/20445911.2013.812093

Boynton, R. M. (1979). *Human Color Vision.* New York: Holt, Rinehart and Winston.

Brauer, M. (1998). Stroop interference in bilinguals: the role of script similarity between two languages. In A. F. Healy & L. E. Bourne Jr. (Eds.). *Foreign Language Learning:*

Psycholinguistic Studies on Training and Retention (pp. 317-337). Mahwah, NJ: Erlbaum.

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*, 515–543. doi:10.2307/1423526

Carreiras, M., Gutierrez-Sigut, E., Baquero, S., & Corina, D. (2008). Lexical processing in Spanish Sign Language (LSE). *Journal of Memory and Language, 58*, 100–122. doi:10.1016/j.jml.2007.05.004

Coderre, E.L., & van Heuven, W. J. B. (2014). The effect of script similarity on executive control in bilinguals. *Frontiers in Psychology, 5*, 1070. doi: 10.3389/fpsyg.2014.01070

Coderre, E. L., van Heuven, W. J. B., & Conklin, K. (2013). The timing and magnitude of Stroop interference and facilitation in monolinguals and bilinguals. *Bilingualism: Language and Cognition, 16*, 420–441. doi:10.1017/S1366728912000405

Cristoffanini, P., Kirsner, K., & Milech, D. (1986). Bilingual lexical representation: The status of Spanish–English cognates. *The Quarterly Journal of Experimental Psychology, 38*, 367-393. doi:10.1080/14640748608401604

De Quadros, R. M., Lillo-Martin, D., & Pichler, D. C. (2015). Bimodal Bilingualism: Sign Language and Spoken Language. *The Oxford Handbook of Deaf Studies in Language, 181.* doi: 10.1093/oxfordhb/9780190241414.013.12

Dupuis, A., & Berent, I. (2015). Signs are symbols: evidence from the Stroop task. *Language, Cognition and Neuroscience*, *30*, 1339-1344. doi:10.1080/23273798.2015.1069363

Emmorey, K., Borinstein, H. B., & Thompson, R. (2005). Bimodal bilingualism: code-blending between spoken English and American Sign Language. In *Proceedings of the 4th International Symposium on Bilingualism* (pp. 663-673). Cascadilla Press Somerville, MA.

Emmorey, K., & McCullough, S. (2009). The bimodal bilingual brain: effects of sign language experience. *Brain and language*, *109*(2-3), 124-132. doi: 10.1016/j.bandl.2008.03.005

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*, 199–210. doi:10.1016/J.JML.2012.04.005

Fang, S. P., Tzeng, O. J., & Alva, L. (1981). Intralanguage vs. interlanguage Stroop effects in two types of writing systems. *Memory & Cognition*, *9*(6), 609-617.

Fariña, N., Duñabeitia, J. A., & Carreiras, M. (2017). Phonological and orthographic coding in deaf skilled readers. *Cognition*, *168*, 27-33. doi: 10.1016/j.cognition.2017.06.015

Geers, Ann E. (2003). Predictors of Reading Skill Development in Children with Early Cochlear Implantation, *Ear and Hearing, 24*, 59-68. doi: 10.1097/01.AUD.0000051690.43989.5D

Geukes, S., Gaskell, M. G., & Zwitserlood, P. (2015). Stroop effects from newly learned color words: Effects of memory consolidation and episodic context. *Frontiers in Psychology, 6*, 278. doi:10.3389/fpsyg.2015.00278

Glezer, L. S., Weisberg, J., Farnady, C. O., McCullough, S., Midgley, K. J., Holcomb, P. J. et al. (2018). Orthographic and phonological selectivity across the reading system in deaf skilled readers. Neuropsychologia, 117, 500-512.

Giezen, M. R., Blumenfeld, H. K., Shook, A., Marian, V., & Emmorey, K. (2015). Parallel language activation and inhibitory control in bimodal bilinguals. *Cognition*, *141*, 9-25. doi:10.1016/j.cognition.2015.04.009

Grosjean, F. (1981). Sign and word recognition: A first comparison. *Sign Language Studies, 32*, 195–220. doi:10.1353/sls.1982.0003.

Gutierrez-Sigut, E., Vergara-Martínez, M. & Perea, M. (2019). Deaf readers benefit from lexical feedback during orthographic processing. *Scientific Reports, 9.* 12321. doi: 10.1038/s41598-019-48702-3

Gutierrez-Sigut, E., Vergara-Martínez, M., & Perea, M. (2017). Early use of phonological codes in deaf readers: An ERP study. *Neuropsychologia, 106*, 261-279. doi: 10.1016/j.neuropsychologia.2017.10.006

Hall, M. L., Hall, W. C., & Caselli, N. K. (2019). Deaf children need language, not (just) speech. *First Language, 39*, 367–395. doi: 10.1177/0142723719834102

Hanson, V. L. (1989). Phonology and reading: Evidence from profoundly deaf readers. *Phonology and Reading Disability: Solving the Reading Puzzle*, *6*, 67-89.

Hauser, P. C., Lukomski, J. & Hillman, T. (2008). Development of deaf and hard-of-hearing students' executive function. In M. Marschark & P. C. Hauser (eds.), *Deaf Cognition: Foundations and Outcomes* (pp. 286-308). New York, NY: Oxford University Press.

Hauser, P. C., Paludneviciene, R., Riddle, W., Kurz, K. B., Emmorey, K., & Contreras, J. (2016). American Sign Language Comprehension Test: A Tool for Sign Language Researchers, *Journal of Deaf Studies and Deaf Education*, *21*, 64–69. doi:10.1093/deafed/env051

Hill, J. C., Lillo-Martin, D. C., & Wood, S. K. (2018). Language Attitudes. In *Sign Languages: Structures and Contexts.*: Routledge.

Hosemann, J., Mani, N., Herrmann, A., Steinbach, M., & Altvater-Mackensen, N. (2020). Signs activate their written word translation in deaf adults: An ERP study on cross-modal co-activation in German Sign Language. *Glossa: a journal of general linguistics 5*(1): 57. 1–25. doi:10.5334/gjgl.1014

Interagency Language RoundTable. (2010). *ILR interagency language roundtable Self-Assessment of foreign language Reading Ability.* Retrieved from: https://www.govtilr.org/Skills/readingassessment.pdf

Kubuş, O., Villwock, A., Morford, J. P., & Rathmann, C. (2015). Word recognition in deaf readers: Cross-language activation of German Sign Language and German. *Applied Psycholinguistics*, *36*, 831-854. doi:10.1017/S0142716413000520

Kuntze, M. (1990). Codeswitching in ASL and written English language contact. In K. Emmorey & H. Lane (Eds.), *The Signs of Language Revisited* (pp. 287-302). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Lee, C. H., Nam, K. & Katz, L. (2007). Nonselective access of spelling–sound knowledge for Korean-English bilinguals. *International Journal of Psychology, 40*, 407-415.

McEvoy, C., Marschark, M., & Nelson, D. L. (1999). Comparing the mental lexicons of deaf and hearing individuals. *Journal of Educational Psychology*, *91*(2), 312. doi: 10.1037/0022-0663.91.2.312

MacLeod, C. M. (2015) The Stroop Effect. In: Luo R. (ed.) *Encyclopedia of Color Science and Technology.* Berlin: Springer. doi: 10.1007/978-3-642-27851-8

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*, 163–203. doi:10.1037/0033-2909.109.2.163

Mägiste, E. (1984). Stroop tasks and dichotic translation: The development of interference patterns in bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*(2), 304–315. doi: 10.1037/0278-7393.10.2.304

Marian, V., Blumenfeld, H. K., Mizrahi, E., Kania, U. & Cordes, A. (2013). Multilingual Stroop performance: Effects of trilingualism and proficiency on inhibitory control. *International Journal of Multilingualism, 10*, 82-104. doi: 10.1080/14790718.2012.708037

Marian, V., & Spivey, M. J. (2003). Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition, 6*, 97–115.

Marschark, M. (1988). Automaticity in word and sign recognition by deaf adults and children. *Sign Language Studies, 58*, 1-19.

Marschark, M., & Hauser, P. C. (2008). Cognitive underpinnings of learning by deaf and hard of-hearing students. In M. Marschark & P. C. Hauser (Eds.), Deaf cognition: Foundations and outcomes (pp. 3–23). New York, NY: Oxford University Press.

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*, 370–375. doi:10.1353/aad.2012.0381

Marschark, M., Spencer, L. J., Durkin, A., Borgna, G., Convertino, C., Machmer, E., Kronenberger, W. G., & Trani, A. (2015). Understanding language, hearing status, and visual-spatial skills. *Journal of Deaf Studies and Deaf Education, 20*, 310–330. doi: 10.1093/deafed/env025

Mayer, C., & Trezek, B.J. (2014). Is reading different for deaf individuals? Reexamining the role of phonology. *American Annals of the Deaf, 159*, 359–371.

Meade, G., Midgley, K. J., Sevcikova Sehyr, Z., Holcomb, P. J., & Emmorey, K. (2017). Implicit coactivation of American Sign Language in deaf readers: An ERP study. *Brain and Language, 170*, 50–61. doi:10.1016/j.bandl.2017.03.004

Morford, J. P., Kroll, J. F., Piñar, P. & Wilkinson, E. (2014). Bilingual word recognition in deaf and hearing signers: Effects of proficiency and language dominance on cross-language activation. *Second Language Research, 30*, 251–271, doi:10.1177/0267658313503467

Morford, J. P., Occhino, C., Piñar, P., Wilkinson, E., & Kroll, J. F. (2017). The time course of cross-language activation in deaf ASL-English bilinguals. *Bilingualism: Language and Cognition, 20*, 337-350, doi:10.1017/S136672891500067X.

Morford, J. P., Occhino, C., Zirnstein, M., Kroll, J. F., Wilkinson, E., & Piñar, P. (2019). What is the source of bilingual cross-language activation in deaf bilinguals? *Journal of Deaf Studies and Deaf Education, 24*, 356–365. doi:10.1093/deafed/enz024

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*, 286–292. doi:10.1016/j.cognition.2010.11.006

Pasquarella, A., Chen, X., Gottardo, A., & Geva, E. (2015). Cross-language transfer of word reading accuracy and word reading fluency in Spanish-English and Chinese-English bilinguals: Script-universal and script-specific processes. *Journal of Educational Psychology, 107,* 96–110. doi:10.1037/a0036966.supp

Pelli, D. G. (1997). The Videotoolbox Software for Visual Psychophysics: Transforming Numbers into Movies. *Spatial Vision, 10*(4), 437-442.

Plaza-Pust, C. (2016). *Bilingualism and Deafness: On Language Contact in the Bilingual Acquisition of Sign Language and Written Language*. Berlin: De Gruyter. doi: 10.1515/9781501504990

Preston, M. S., & Lambert, W. E. (1969). Interlingual interference in a bilingual version of the Stroop color-word task. *Journal of Verbal Learning and Verbal Behavior*, *8*(2), 295-301.

Rosselli, M., Ardila, A., Santisi, M. N., Arecco, M. D. R., Salvatierra, J., Conde, A., & Lenis, B. (2002). Stroop effect in Spanish–English bilinguals. *Journal of the International Neuropsychological Society*, *8*(6), 819-827. doi: 10.1017.S1355617702860106

Shook, A., & Marian V. (2012). Bimodal bilinguals co-activate both languages during spoken comprehension. *Cognition, 124,* 314-324. doi:10.1016/j.cognition.2012.05.014

Stansfield, C. W., Gao, J., & Rivers, W. P. (2010). A concurrent validity study of self-assessments and the federal Interagency Language Roundtable oral proficiency interview. *Russian Language Journal*, *60*, 299-315. Retrieved December 4, 2020, from http://www.jstor.org/stable/43669189

Stroop, J. R. (1935). Studies of interference in serial verbal reaction. *Journal of Experimental Psychology, 18*(6), 643-662. doi:10.1037/h0054651

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign language comprehension. *Proceeding of National Academy of Sciences, 104*, 12530–12535. doi: 10.1073pnas.0609927104

Treiman, R., & Hirsh-Pasek, K. (1983). Silent reading: Insights from second-generation deaf readers. *Cognitive Psychology, 15*, 39-65. doi:10.1016/0010-0285(83)90003-8

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

Wang, Y., Trezek, B. J., Luckner, J. L., & Paul, P. V. (2008). The role of phonology and phonologically related skills in reading instruction for students who are deaf or hard of hearing. *American Annals of the Deaf, 153*, 396–407. doi:10.1353/aad.0.0061

Wolff, A. B., Radecke, D. D., Kammerer, B. L., & Gardner, J. K. (1989). Adaptation of the Stroop Color and Word Test for use with deaf adults. *Clinical Neuropsychologist, 3*, 369–374. doi:10.1080/13854048908401485

Table 1. Participant Demographics

Bilingual Group	N	% Female	Mean Age (SD)	Age Range	Mean Self-Proficiency Reading English (0-4)	Mean Self-Proficiency Reading Language 2 (0-4)	BLP Spea Underst Spoken En 6)
Deaf ASL-English	15	47	29.4 (6.5)	19 to 39	3.67 (.52)	not applicable*	no speech
Hearing Chinese-English	22	86	19.6 (1.36)	18 to 23	3.55 (.72)	1.91 (1.38)	5.45 (1
Hearing Korean-English	15	80	19.8 (.94)	18 to 21	3.80 (.41)	1.82 (.85)	5.93 (
Hearing Spanish-English	34	76	20.2 (1.89)	18 to 25	3.91 (.31)	2.41 (.94)	5.85 (

** ASL does not have orthography. †This is by subject selection/study design*

Table 2. Stimulus Words in each Language by Orthographic Similarity to English

Control Language	Single Script*	Low Similarity	Moderate Similarity	High Similarity
English	*ASL*	*Chinese*	*Korean*	*Spanish*
Blue		蓝	푸른 색	Azul
Green		绿	채색	Verde
Yellow		黄	노란색	Amarillo
Purple		紫	보라색	Morado

* Note that the ASL stimuli were presented as videos and these are screenshots.

Table 3. Means (and Standard Errors) for Stroop Interference Effects on Reaction Time for the Color Naming and Word Reading Tasks for each Language by Bilingual Group.

Table 3. Means (and SE) for Stroop Interference Effects on Reaction Time for the Color Naming and W Reading Task for each language by Bilingual Group.

	Color Naming Task		Word Reading Ta	
Bilingual Group	**English**	**Other Language**	**English**	**Other La**
Deaf ASL-English (N=15)	64.55 (31.40)	58.66 (52.48)	43.56 (46.62)	48.55 (
Hearing Chinese-English (N=22)	79.73 (59.63)	56.07 (52.37)	28.38 (38.32)	52.77 (
Hearing Korean-English (N=15)	89.90 (53.92)	54.14 (39.89)	28.27 (56.97)	54.58 (
Hearing Spanish-English (N=34)	115.05 (50.10)	91.27 (52.24)	45.72 (46.46)	45.94 (