Title
One fish, uh, two fish: Effects of fluency and bilingualism on adults' novel word learning

Permalink
https://escholarship.org/uc/item/32k167x0

Journal
Psychonomic Bulletin & Review, 30(2)

ISSN
1069-9384

Authors
Libersky, Emma
Neveu, Anne
Kaushanskaya, Margarita

Publication Date
2023-04-01

DOI
10.3758/s13423-022-02189-8

Peer reviewed

HHS Public Access

One fish, uh, two fish: Effects of fluency and bilingualism on adults' novel word learning

Emma Libersky[a,1], **Anne Neveu**[a], **Margarita Kaushanskaya**[a]
[a]Waisman Center, University of Wisconsin-Madison

Abstract

Listeners utilize speech disfluencies to anticipate novelty in spoken language, but little is known about how speech disfluencies impact novel word learning. We investigated how monolingual and bilingual adults learn novel words under disfluent speech conditions, focusing on fillers such as *uh* and *um*. If fillers highlight novelty, they might be an especially potent cue during word learning; however, because fillers also signal uncertainty, listeners may be less willing to learn in a disfluent condition. We also tested whether an effect of fillers on word learning would be moderated by bilingual experience, expecting that bilinguals would be affected differently because their exposure to distributional information within each language is reduced relative to monolinguals. In Experiments 1 and 2, where participants were exposed only to novel words, we found that participants learned words equally well in fluent and disfluent conditions, and that this effect was not moderated by bilingual experience. In Experiment 3, when novel words were embedded within a larger set of known words, we observed a bilingualism by condition interaction, wherein bilinguals benefited from fluency, but monolinguals performed equally well across conditions. These findings suggest that monolinguals' word learning—unlike word processing—may be robust to variations in speaker fluency, but that language experience may moderate the effect of fluency on learning.

Disfluent filled pauses or *fillers* (*uh* and *um* in American English) occur in predictable locations in the speech stream (Schnadt & Corley, 2006; Beattie & Butterworth, 1979; Clark & Fox Tree, 2002), and listeners can harness these statistics to make predictions about upcoming novelty. Studies of adults' processing of fillers have repeatedly indicated that listeners expect speakers to refer to uncommon or discourse-novel objects following fillers (Arnold et al., 2007). Moreover, listeners make assumptions about speakers' topic knowledge based on fillers (Brennan & Williams, 1995), and speakers' disfluency rates are negatively correlated with how credible they are rated by listeners (Carpenter, 2012). These

assumptions may affect learning, as children have been found to prefer to learn from more fluent—i.e., more credible—speakers, when given a choice (White et al., 2020).

In addition to their effects on online language processing, fillers have been found to impact memory for spoken passages (Fraundorf & Watson, 2011; Bosker et al., 2015). Fraundorf and Watson (2011) tested adults' memory for audio-recorded passages that were interrupted by fillers or coughs. They found that participants had better recall for events that were interrupted by fillers than for those that were fluent or interrupted by coughs. They replicated their findings in a second experiment in which fillers occurred in non-predictive locations, suggesting that fillers aid memory by increasing listener attention rather than by enabling specific predictions about upcoming novelty.

While the effects of fillers on processing and memory are becoming well established, investigations of word learning in disfluent speech contexts remain extremely limited. In the only published study of disfluent word learning that we are aware of, White and colleagues (2020) investigated the impact of fillers on children's word learning. Across two experiments, the authors found that children preferred fluent speakers but learned words equally well when presented in a fluent or disfluent frame. The authors suggested that the hypothesized facilitative (i.e., predictive utility and/or increased attention) and detrimental (i.e., reduced speaker credibility) effects of disfluencies were both at play in the study, effectively canceling one another out.

However, it is possible that, because children have less exposure to fillers than adults (Kidd et el., 2011; Newport et al., 1977), they utilize fillers in word learning scenarios differently than adults. To our knowledge, no existing studies have investigated adults' word learning in the context of fillers, and it is possible that, unlike children, adults would be sensitive to fillers during word learning. The first goal of the current study therefore was to examine the impact of fillers on adult word learning. Our second goal was to investigate whether the hypothesized relationship between fillers and word learning would be moderated by bilingualism.

Bilingual listeners and disfluencies

Although there has been some work investigating how native listeners respond to disfluencies produced by nonnative speakers (Bosker et al., 2014) and variation in fillers across languages is well-documented (Clark & Fox Tree, 2002), there is very little work examining how bilinguals respond to fillers in their native vs. non-native language (Morin-Lessard & Byers-Heinlein, 2019 is one exception). Where monolinguals experience all language input in a single language, bilinguals split their exposure across languages. This distributed exposure means relatively less exposure to each language and therefore a reduced exposure to distributional information within each language. Distributed exposure has been used to explain interactions between bilingual experience and word frequency in picture naming tasks (Gollan et al., 2008; Gollan et al., 2011). By similar logic, bilinguals should have less exposure to the specific realizations of fillers within each of their languages and less experience with specific fillers cueing discourse-novel or infrequent words. As a result, bilingual listeners might be hypothesized to be less sensitive to fillers during word learning.

Alternatively, disfluencies could impact bilinguals and monolinguals similarly. In a study of online processing of French and English fillers, Morin-Lessard and Byers-Heinlein (2019) found that monolingual and bilingual listeners used disfluencies predictively whether they followed the French or English phonetic realization. Although this finding needs to be evaluated for language pairs other than French and English, where fillers are particularly similar in form, it may suggest that both bilinguals and monolinguals could rely on common features that mark fillers across languages. Another reason we may expect to see similar performance across bilingual and monolingual participants concerns the mechanism underlying a potential effect. If fillers affect word learning through an attentional rather than predictive mechanism (Fraundorf & Watson, 2011), relative familiarity with fillers and their distributional patterns may not matter in the context of word learning.

Current study

We investigated the effects of fillers on word learning in a paired-associate word learning task. Participants were taught nonword labels for unfamiliar fish in three conditions: disfluent, fluent, and cough. If the benefits of fillers extend beyond sentence processing and memory for passages, we would expect better learning in the disfluent condition. The cough condition was included to investigate the mechanism underlying the effect of fluency, should one be observed. If performance in the disfluent condition was significantly different from performance in the fluent condition but comparable to performance in the cough condition, the pattern of results would indicate that fillers affect learning through an attentional mechanism, where any interruption to fluent speech helps learning. In Experiments 1 and 2, we assessed performance in these conditions within a pool of novel stimuli, only. In Experiment 3, we added filler trials of fluent familiar stimuli, exploring the possibility that effects of condition would only appear when disfluency was predictive of novelty.

We also tested a moderating effect of bilingualism, hypothesizing that bilinguals and monolinguals would be differentially affected by disfluency, as a consequence of bilinguals' distributed language exposure. We ran three experiments. In Experiment 1, we compared monolinguals to bilinguals who were largely native speakers of English. In Experiment 2, we compared monolinguals to bilinguals who were non-native speakers of English and therefore exposed to relatively less English language input. In Experiment 3, we compared monolinguals to non-native English-speaking bilinguals on a task that integrated novel and familiar stimuli. If distributed language exposure impacts how bilinguals process fillers, we would expect a stronger moderating effect of bilingualism in Experiments 2 and 3.

Experiment 1

Methods

Participants—Participants were recruited through undergraduate courses within the University of Wisconsin-Madison (n = 36) and through Prolific.co (n = 102; Palan & Schitter, 2018). For undergraduates, our goal was to recruit as many participants as possible during the Fall 2020 semester. For participants recruited via Prolific, we set our recruitment target to 50 participants per language group; two participants who experiences technical difficulties during the sound check were sent new links for the study and included in the

dataset. All participants gave informed consent through a protocol approved by the UW-Madison IRB and reported normal hearing, age between 18-40, monolingual or bilingual status, and residence in the United States.

Bilingual status was based on self-identification, and participants were not required to report a minimum level of fluency in their non-dominant language(s). The undergraduate sample mostly consisted of women (women = 32; men = 3; nonbinary = 1) but had a roughly equal number of bilingual (n = 16) and monolingual (n = 20) participants. Prolific recruitment was targeted such that we recruited a roughly equal number of monolingual men (n = 23), monolingual women (n = 24), bilingual men (n = 28), and bilingual women (n = 25), as well as two bilingual participants who identified as nonbinary.

Self-reported language experience and ability across English and other languages was collected via the Language Experience and Proficiency Questionnaire (LEAP-Q, Marian et al., 2007), and English language proficiency was assessed using a speeded version of the Lexical Test for Advanced Learners of English (LexTALE; Lemhöfer & Broersma, 2012; Poort & Rodd, 2019) and a spelling task (Fidler et al., 2011). Participant characteristics, including language characteristics, are described in Table 1.

Undergraduate and Prolific participants were demographically similar, except Prolific participants were significantly older than the undergraduate group ($t(136) = -5.36$, $p <$ .001), and bilingual participants recruited via Prolific reported a greater variety of language pairings than the undergraduate group, who were largely Spanish-English bilinguals. We ran all statistical models on undergraduate and Prolific groups separately and combined, to ensure that patterns of results were consistent across groups. Because the pattern of results was consistent across recruitment groups, we report combined models.

Procedure—Participants were tested remotely, via Gorilla Experiment Builder (Anwyl-Irvine et al., 2019). After consenting, participants completed a sound check (James, 2019) before continuing to the experiment. The tasks were completed as follows: LEAP-Q, experimental task, spelling task, and LexTALE. The procedure took approximately 30 minutes to complete.

Materials—Stimuli consisted of nine nonword-image pairs embedded in sentences and presented in a random order. The condition in which the nonword-image pair appeared (fluent, disfluent, cough) was counterbalanced across participants, and each participant learned three words in each condition. Nonwords were taken from Gupta et al. (2004). All nonwords consisted of two syllables and five phonemes and had no neighbors in English (Marian et al., 2012). Words were matched on syllable stress and English biphone frequencies (Marian et al., 2012). The images were all illustrations of fish taken from Renard (1754) and were selected based on realism and novelty. The images were matched on objective visual complexity, proxied by JPG file size (Bates et al., 2003). The sentences were based on entries in the National Geographic Photo Ark (Sartore, n.d.) to sound realistic without describing specific fish. Each sentence gave a "fact" about the fish without reference to its appearance (e.g., "The gagek lives in warm streams in sunny areas"). Sentences were matched on average SubtleX frequency (Brysbaert & New, 2009) as well as semantic

richness (proxied by propositional idea density, Brown et al., 2008). Match (van Casteren & Davis, 2007) was used to create stimulus lists.

Teaching stimuli were recorded by a female native speaker of American English. A version of each sentence was created for each of the three conditions: fluent, disfluent, and cough. Each version consisted of three parts that were spliced together in Audacity (Audacity Team, 2018). The first part consisted of the determiner, which varied by condition. The fluent condition had a fluent "the" (pronounced "thuh") preceded by silence (to match the length of the other conditions). The disfluent condition consisted of a disfluent "the" (pronounced "thee") followed by a filler (either *uh* or *um*). The cough condition had the same fluent "the" used in the fluent condition, followed by a cough. Disfluencies and coughs occurred at a rate of 3.9 per 100 words, which is somewhat higher than filler rates of approximately 2 per 100 words observed in naturalistic data by Bortfield et al. (2001) and implemented in an experimental setting by Fraundorf and Watson (2011). The second part of the sentence was the nonword. The rest of the sentence (e.g., "lives in warm streams in sunny areas") was included in the last segment. The second and third segments were the same across conditions (i.e., the same recordings were used) and only the first segment differed across conditions. Segments were normalized to 70.0dB in Praat (version 6.1.16; Boersma & Weenink, 2020). We included two attention checks that asked participants to click "next" to continue with the experiment. These were presented after one third and two thirds of the stimuli had been presented. Participants were excluded if their response time to either check was beyond 2.5 standard deviations from the group's mean.

At test, participants saw three fish (one from each condition) and heard a fluent singleword label by a new speaker (another female speaker of American English, from the supplementary materials recorded by Gupta et al., 2004). During each test trial, we presented a fish from each condition to equalize difficulty across trials, so that participants would not experience competition from same-condition alternatives. The novel speaker was used to test participants' ability to generalize their learning across speakers. Participants had to select the fish that the speaker was referring to. Each fish was tested twice, in a single block. Test trials were presented in a pseudo-randomized order, such that participants were tested on all fish-label pairs before being tested a second time. The procedure is described in Figure 1.

Analyses—We examined word learning accuracy at the item-level using mixed-effects logistic regression models in R Studio, version 1.2.5042, using the lme4 package (Bates et al., 2015). To avoid including trials where participants were randomly clicking or not engaged with the task, we excluded trials based on response time. We excluded 356 trials and retained 2128 trials (14% excluded; see exclusionary criteria in Supplementary Materials). In accordance with the "keep it maximal" approach (Barr et al., 2013), the models converged and singularity issues were resolved when retaining random intercepts for participants and items and removing random slopes. Model assumptions in the generalized linear model were tested using the DHARMa package (Hartig, 2022).

Fixed effects included condition (fluent, cough, or disfluent, coded using planned non-orthogonal contrasts), bilingual status (monolingual or bilingual, coded as –0.5 and 0.5) and interactions between condition and bilingual status. Because many prior studies have

linked socioeconomic status (Fernald et al., 2013; Maguire et al., 2018) and language ability (Hill & Wagovich, 2020) to word learning performance, we proxied these variables through years of education and LexTALE score, respectively, and included them as subject-level covariates.

Results

All participants (including those missing data for covariates) learned at above chance levels (> 33%) in the fluent ($M = 57\%$, $SD = 50\%$; $t(717) = 13.04$, $p < .0001$), cough ($M = 55\%$, $SD = 50\%$; $t(716) = 12.04$, $P < .001$), and disfluent conditions ($M = 57\%$, $SD = 50\%$; $t(725) = 12.91$, $P < .0001$) . The model did not reveal any significant main effects of condition, bilingual status, or their interactions when controlling for years of education (mean centered) and LexTALE score (mean centered) (see Figure 2) (see Tables 1a and 1b in supplementary materials for the full regression model).

Discussion

Experiment 1 findings suggest that fillers did not influence learning. This may be because there is truly no effect of fillers on word learning (consistent with White et al.'s, 2020, findings in children) or because of our specific design (a possibility we take up in Experiment 3, and in the General Discussion). We also did not observe an effect of bilingualism. This may be due to the characteristics of our bilingual group, which consisted mostly of English-dominant bilinguals exposed largely to English (mean percent English exposure = 89%, median = 95%).

Because our predictions regarding the interactive effect of fluency condition and bilingualism were based on consequences of distributed exposure, we reasoned that a bilingual group with relatively less familiarity with the statistics of English filled pauses might perform differently than the monolingual group. Therefore, in Experiment 2, we replicated our experiment with a new group of bilingual participants with English as a second language (L2 participants).

Experiment 2

In Experiment 1 we found no effect of bilingualism on adults' ability to learn novel words in fluent and disfluent conditions. However, bilingual participants were largely English dominant and most of their regular language exposure was in English. In Experiment 2, we tested a group of L2 English bilinguals in the same paradigm, predicting that bilinguals with less exposure to English would be more impacted by English-language disfluencies than the participants included in Experiment 1.

Methods

Participants—Forty-nine L2 participants were recruited through Prolific. Inclusionary criteria were similar to those of Experiment 1, except that all participants reported a native language other than English. As in Experiment 1, we set our recruitment target to 50 participants, but one participant was excluded because their LEAP-Q revealed they were actually monolingual. Participant characteristics are described in Table 1.

Procedure and materials—Experiment 2 followed the same protocol as Experiment 1.

Analyses—Monolingual English speakers (tested in the first wave of data collection) and bilingual participants with L2 English were compared (coded as −0.5 and 0.5). After excluding 287 trials (14% of the data), 1819 trials were analyzed (see Supplementary Materials for details on excluded trials). The statistical analysis approach was the same as for the first experiment, except that, because the L2 group was significantly older that the monolinguals, age was included as a covariate.

Results

All participants (including those missing data for covariates) learned at above chance levels (> 33%) in the fluent ($M = 56\%$, $SD = 50\%$; $t(619) = 11.77$, $p < .001$), cough ($M = 57\%$, $SD = 50\%$; $t(601) = 11.87$, $p < .001$), and disfluent conditions ($M = 58\%$, $SD = 49\%$; $t(631) = 12.94$, $p < .001$) . The model did not reveal any significant main effects of condition, language group or their interactions when controlling for years of education (mean centered), LexTALE score (mean centered), and age (mean centered) (see Figure 2). (See Tables 2a and 2b in supplementary materials for the full regression model.)

To further investigate the role of bilingual experience in disfluent word learning, we ran a follow-up analysis collapsing groups across both experiments (including monolinguals and bilinguals from both experiments) and examining accuracy across the continuous variable of percent exposure to English. A total of 2858 trials were analyzed (526 or 16% excluded; see Supplementary Materials for full details). Once again, the model did not reveal any significant main effects of condition, percent English exposure, or their interactions, when controlling for years of education (mean centered), LexTALE score (mean centered), and age (mean centered). See Tables 3a and 3b in supplementary materials for the full regression model. See Figure 3 for predicted performance by condition and percent exposure.

Discussion

Despite experiencing more distributed language exposure than the bilingual group in Experiment 1, the bilingual group in Experiment 2 performed similarly to monolinguals, and neither group was sensitive to the fluency manipulations. When bilingual experience was included as a continuous predictor (percent English exposure), our results remained null.

One limitation to the design use in both Experiment 1 and 2 is that participants were learning novel words on *every* trial. If fillers affect word learning through distributional cues, our manipulation of fluency would not be sensitive to this effect, since participants knew that they would be presented with novel words on every trial (regardless of fluency condition). Moreover, fillers occurred at an unnaturally high rate, which may have affected listeners' perception of the speaker and the information presented. To address both limitations, we ran a third experiment in which we embedded filler trials of familiar fish into the learning phase of the previous experiments.

Experiment 3

In Experiments 1 and 2, we found no effects of fluency or bilingualism on adults' ability to learn novel words. However, the materials of Experiments 1 and 2 differed from spontaneous language in that disfluencies occurred at above average rates and all trials presented novel information. Disfluencies are predictive of novelty in spontaneous speech (Arnold et al., 2003), but this was not true within the experimental context. To address this limitation, we ran a third experiment in which novel words were integrated into a larger pool of known words.

Methods

Participants—Forty-seven monolingual and 49 L2 participants were recruited through Prolific. Inclusionary criteria for monolinguals were the same as in Experiment 1 and inclusionary criteria for L2 bilinguals were the same as in Experiment 2. We set our recruitment target to 50 participants per language group, but one participant was excluded from the L2 group because the LEAP-Q revealed they were monolingual. Three monolinguals were recruited but timed out on the task or experienced technical difficulties. Eleven self-identified monolinguals reported some exposure to a second language, but their self-reported spoken language proficiency was low ($M = 4.00$, $SD = 1.41$) and they were included in the final sample. Monolingual participants were significantly older than bilingual participants ($p < .05$). Participant characteristics are described in Table 2.

Procedure—The procedure was the same as in Experiments 1 and 2, except that participants completed a familiarity check for the familiar and novel fish names immediately following the testing phase.

Materials—Stimuli consisted of the nine nonword-image pairs used in Experiments 1 and 2, which appeared in the same counterbalanced conditions (e.g., participants learned 3 novel words in each of the conditions - fluent, disfluent, and cough). Nine fluent filler trials describing familiar fish were randomly integrated with these novel trials. The familiar fish all had two-syllable names and were reported to be known by at least 90% of participants tested by Brysbaert et al. (2019). Following the learning and test phases, we presented participant with a written list of the familiar and novel fish names and asked them to "indicate if [they] had heard of each fish before participating in [the] study." Within our sample, the proportion of familiar fish known to participants was high ($M = 0.86$, $SD = 0.35$), although monolinguals reported knowing a significantly greater proportion of the familiar fish ($M = 0.93$, $SD = 0.26$) than bilinguals ($M = 0.79$, $SD = 0.40$; $p < .01$). Images of the fish were taken from Renard (1754); if the exact fish had not been illustrated by Renard, we chose an illustration of a similar looking fish. The fish names were embedded in sentences created in the same manner described in Experiment 1, which were recorded by the same speaker. These sentences conveyed realistic but not necessarily factual information about the familiar fish. The familiar trials were used as fillers to make disfluencies and coughs predictive cues to novelty; therefore, participants were not tested on these trials in the recognition phase. With the introduction of the familiar stimuli, the likelihood of a fluent trial cueing a novel fish was reduced to 3/12, or 25%. The likelihood of a disfluent

or cough trial cueing a novel fish remained at 100%. The addition of the fluent trials also served to lower the experiment-wide disfluency rate to 1.8 disfluencies per hundred words, which is consistent with rates from Bortfield et al. (2001) and Fraundorf and Watson (2011). Attention checks were included after one third and two thirds of the stimuli had been presented, regardless of whether those stimuli were familiar or novel.

The test phase was exactly the same as in Experiments 1 and 2. Immediately following the test phase, participants were asked to "indicate if [they] had heard of each fish before participating in [the] study." Both real and pseudo names were presented in a list and participants provided a yes or no response to each item. The procedure is described in Figure 1.

Analyses—Monolingual and bilingual participants were compared (coded as −0.5 and 0.5). After excluding 527 trials, 2965 trials were analyzed (see Supplementary Materials for details on excluded trials). The statistical analysis approach was the same as for the first two experiments, and we controlled for age as the bilingual group was significantly younger than the monolingual group.

Results

All participants (including those missing data for covariates) learned at above chance levels (> 33%) in the fluent ($M = 53\%$, $SD = 50\%$; $t(1001) = 12.67$, $p < .0001$), cough ($M = 49\%$, $SD = 50\%$; $t(1011) = 10.18$, $p < .0001$), and disfluent conditions ($M = 52\%$, $SD = 50\%$; t(986) = 11.86, $p < .0001$). The model revealed a significant interaction between the contrast comparing the fluent and disfluent conditions and bilingualism ($B = −0.42$, $SE = 19$, $z = −2.16$, $p = .031$), such that bilinguals were more likely than monolinguals to show an advantage for learning in the fluent condition. When the disfluent condition was made the reference group (to obtain the last contrast comparing the cough and disfluent conditions), we observed a significant effect of bilingual status ($B = −0.43$, $SE = z = −2.14$, $p = 0.032$), reinforcing the conclusion that bilinguals show poorer performance in the disfluent condition relative to monolinguals (with the caution that the fluent condition was our a priori reference group). See Figure 4 for visualization of performance. See Tables 4a and 4b for full model output.

Discussion

In Experiment 3, we embedded novel stimuli within a larger set of familiar stimuli. This change resulted in coughs and disfluencies being predictive of novelty as well as meaningful attention getters. Despite the utility of these cues, bilinguals performed *better* in the fluent condition relative to the disfluent condition and monolinguals performed similarly across conditions. This finding indicates that interruptions to the speech stream, even when they have predictive utility, may be detrimental to learners who are relatively less exposure to the target language.

General Discussion

This study investigated the effects of filled pauses and bilingualism on adults' word learning across three experiments. Experiments 1 and 2, where only novel words were presented

at exposure, revealed no impact of disfluencies or coughs on participants' word learning, whether participants were monolingual English speakers, L1 English bilinguals, or L2 English bilinguals. Experiment 3 indicated similar patterns for monolinguals when novel stimuli were intermixed with familiar stimuli at exposure (rendering disfluencies predictive of novelty, as they are in spontaneous speech). However, in this experiment, bilingual participants performed better in the fluent condition relative to the disfluent condition. These findings suggest that bilinguals may benefit from fluency, but only in certain circumstances, and it is unclear why such an effect would only appear in Experiment 3. One possibility is that, because disfluency rates were higher than typical in Experiments 1 and 2, bilingual participants had enough exposure to disfluencies that they were able to find strategies to mitigate any detrimental effects on learning. Alternatively, because bilingual participants were less familiar with the real fish included in Experiment 3, the findings could stem from a difference in the relative utility of disfluencies or difficulty or the task across participants. Because bilinguals were less familiar with the real fish, some participants would have been exposed to novel-to-them fish within fluent contexts, meaning that disfluencies had relatively less predictive power for these participants. Moreover, bilingual participants who did not know the English labels for the real fish might assign resources to learning both familiar and novel fish, making the task more difficult overall.

Our findings indicate that bilinguals learning in their L2 demonstrate reduced performance when the speech stream is interrupted by disfluencies, whereas monolinguals learn equally well across fluent and interrupted learning conditions. We localize this difference to the amount of exposure participants have to English, with L2-English bilinguals experiencing significantly less English exposure than monolinguals, and therefore having less experience dealing with English-specific interruptions to English speech (i.e., *uh* and *um*).

In future studies, it will be important to examine the role of language experience and testing paradigm in a variety of ways. To fully understand the fluency by bilingualism interaction observed in Experiment 3, future studies should test bilingual adults with variable levels of L2 proficiency and children (including beginning second language learners) and quantify language exposure in a more fine-grained manner. Moreover, future studies should examine word learning performance via other methods, perhaps by assessing free- or cued-recall, response times, and lexical recognition, and manipulate the degree to which participants expect the test phase. These measures may be more sensitive to fluency effects in participants with high exposure to the target language than the recognition task used here. Lastly, future studies should manipulate the placement and type of disfluencies to make the learning procedure less repetitive and more naturalistic.

In conclusion, our results extend previous findings in children (White et al., 2020), suggesting that monolingual adults are not sensitive to fillers in a word learning context. However, bilingual adults may be, in specific circumstances. Unlike monolinguals, who were expected to benefit from disfluency during learning, bilinguals benefited from fluency. These findings contrast with the beneficial effects of fillers in online processing (Arnold et al., 2007; Bosker et al., 2015) and memory (Fraundorf & Watson, 2011), perhaps suggesting that the utility of fillers varies according to timescale, listener characteristics, and listener goals. Future research should further explore the bilingualism by fluency interaction

observed here, emphasizing the role of language experience in recovering from interruptions to the speech stream.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements & Funding:

Availability of data and materials:

The datasets and used in the current study are available at https://osf.io/ym8jw/?view_only=58deee976ade4f019e667130a05a2487

References

Anwyl-Irvine A, Massonnié J, Flitton A, Kirkham N, & Evershed J (2018). Gorilla in our Midst: An online behavioral experiment builder. Behavior Research Methods, 52, 388–407. 10.3758/sl3428-019-01237-x

Arnold JE, Fagnano M, & Tanenhaus MK (2003). Disfluencies signal theee, um, new information. Journal of Psycholinguistic Research, 32(1), 25–36. 10.1023/A:1021980931292 [PubMed: 12647561]

Arnold JE, Hudson Kam CL, & Tanenhaus MK (2007). If you say thee uh you are describing something hard: the on-line attribution of disfluency during reference comprehension. Journal of Experimental Psychology, 33(5), 914–930. 10.1037/0278-7393.33.5.914 [PubMed: 17723069]

Audacity-Team. (2018). Audacity(R): Free audio editor and recorder (Version 2.3.3). https://www.audacityteam.org/

Barr DJ, Levy R, Scheepers C, & Tily HJ (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. Journal of Memory and Language, 68(3), 255–278. 10.1016/j.jml.2012.ll.001

Bates D, Mächler M, Bolker BM, & Walker SC (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 67(1), 1–48. 10.18637/jss.v067.i01

Bates E, D'Amico S, Jacobsen T, Székely A, Andonova E, Devescovi A, Herron D, Lu CC, Pechmann T, Pléh C, Wicha N, Federmeier K, Gerdjikova I, Gutierrez G, Hung D, Hsu J, Iyer G, Kohnert K, Mehotcheva T, … Tzeng O (2003). Timed picture naming in seven languages. Psychonomic Bulletin and Review, 10(2), 344–380. 10.3758/BF03196494 [PubMed: 12921412]

Beattie GW, & Butterworth BL (1979). Determinants of pauses and errors in spontaneous speech. Language and Speech, 22(3), 201–211.

Boersma P, & Weenink D (2020). Praat: doing phonetics by computer (Version 6.1.16). https://www.fon.hum.uva.nl/praat/

Bosker HR, Quené H, Sanders T, & de Jong NH (2014). Native 'urn's elicit prediction of low-frequency referents, but non-native 'urn's do not. Journal of Memory and Language, 75, 104–116. 10.1016/jjml.2014.05.004

Bosker HR, Tjiong J, Quené H, Sanders TJM, & De Jong NH (2015). Both native and non-native disfluencies trigger listeners' attention. In Lickley R & Eklund R (Eds.), Proceedings of Disfluency in Spontaneous Speech 2015 (pp. 5–8). https://www.mpi.nl/publications/item_2156905

Brennan SE, & Williams M (1995). The feeling of another's knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. Journal of Memory and Language, 398(3), 383–398. 10.1006/jmla.1995.1017

Brown C, Snodgrass T, Kemper SJ, Herman R, & Covington MA (2008). Automatic measurement of propositional idea density from part-of-speech tagging. Behavior Research Methods, 40(2), 540–545. 10.3758/BRM.40.2.540 [PubMed: 18522065]

Brysbaert M, & New B (2009). Moving beyond Ku era 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(4), 977–990. 10.3758/BRM.4L4.977 [PubMed: 19897807]

Carpenter CJ (2012). A meta-analysis and an experiment investigating the effects of speaker disfluency on persuasion. Western Journal of Communication, 76(5), 552–569. 10.1080/10570314.2012.662307

Clark HH, & Fox Tree JE (2002). Using uh and um in spontaneous speaking. Cognition, 84(1), 73–111. 10.1016/S0010-0277(02)00017-3 [PubMed: 12062148]

Fernald A, Marchman VA, & Weisleder A (2013). SES differences in language processing skill and vocabulary are evident at 18 months. Developmental Science, 16(2), 234–248. 10.1111/desc.12019 [PubMed: 23432833]

Fidler LJ, Plante E, & Vance R (2011). Identification of adults with developmental language impairments. American Journal of Speech-Language Pathology, 20(1), 2–13. 10.1044/1058-0360(2010/09-0096) [PubMed: 20739630]

Fraundorf SH, & Watson DG (2011). The disfluent discourse: Effects of filled pauses on recall. Journal of Memory and Language, 65(2), 161–175. 10.1016/jjml.2011.03.004 [PubMed: 21765590]

Gollan TH, Montoya RI, Cera C, & Sandoval TC (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. Journal of Memory and Language, 58(3), 787–814. doi: 10.1016/j.jml.2007.07.001 [PubMed: 19343088]

Gollan TH, Slattery TJ, Goldenberg D, Van Assche E, Duyck W, & Rayner K (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. Journal of Experimental Psychology: General, 140(2), 186–209. 10.1037/a0022256

Gupta P, Lipinski J, Abbs B et al. Space aliens and nonwords: Stimuli for investigating the learning of novel word-meaning pairs. Behavior Research Methods, Instruments, & Computers 36, 599–603 (2004). 10.3758/BF03206540

Hill MS, & Wagovich SA (2020). Word learning from context in school-age children: relations with language ability and executive function. Journal of Child Language, 47(5), 1006–1029. 10.1017/S0305000919000989 [PubMed: 32252839]

James E (2019). Pre-experiment sound check. Gorilla Experiment Builder. https://app.gorilla.sc/openmaterials/42292

Kidd C, White KS, & Aslin RN (2011). Toddlers use speech disfluencies to predict speakers' referential intentions. Developmental Science, 14(4), 925–934. 10.1111/j.1467-7687.2011.01049.x [PubMed: 21676111]

Lemhöfer K, & Broersma M (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. Behavior Research Methods, 44(2), 325–343. 10.3758/sl3428-011-0146-0 [PubMed: 21898159]

Lickley RJ (2015). Fluency and Disfluency. In Redford MA (Ed.), The Handbook of Speech Perception. Wiley. 10.1002/9781118584156.ch20

Maguire MJ, Schneider JM, Middleton AE, Ralph Y, Lopez M, Ackerman RA, & Abel AD (2018). Vocabulary knowledge mediates the link between socioeconomic status and word learning in grade school. Journal of Experimental Child Psychology, 166, 679–695. 10.1016/jjecp.2017.10.003 [PubMed: 29103588]

Marian V, Bartolotti J, Chabal S, & Shook A (2012). Clearpond: Cross-linguistic easy-access resource for phonological and orthographic neighborhood densities. PLoS ONE, 7(8). 10.1371/journal.pone.0043230

Marian V, Blumenfeld HK, & Kaushanskaya M (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. Journal of Speech Language and Hearing Research, 50(4), 940. 10.1044/1092-4388(2007/067)

Morin-Lessard E, & Byers-Heinlein K (2019). Uh and euh signal novelty for monolinguals and bilinguals: Evidence from children and adults. Journal of Child Language, 46(3), 522–545. 10.1017/S0305000918000612 [PubMed: 30829567]

Newport EL, Gleitman H, & Gleitman LR (1977). Mother , I'd rather do it myself: Some effects and non-effects of maternal speech style. In Snow CE & Ferguson CA (Eds.), Talking to Children: Language Input and Acquisition (pp. 109–150). Cambridge University Press.

Owens SJ, Thacker JM, & Graham SA (2018). Disfluencies signal reference to novel objects for adults but not children. Journal of Child Language, 45(3), 581–609. 10.1017/S0305000917000368 [PubMed: 29141698]

Palan S, & Schitter C (2018). Prolific.ac—A subject pool for online experiments. Journal of Behavioral and Experimental Finance, 17, 22–27. 10.1016/jjbef.2017.12.004

Poort ED, & Rodd JM (2019). Towards a distributed connectionist account of cognates and interlingual homographs: evidence from semantic relatedness tasks. PeerJ—Life & Environment, 7, e6725. 10.7717/peeij.6725

Renard L (1754). Poissons, écrevisses et crabes, de diverses couleurs et figures extraordinaires, que l'on tronve autour des Isles Moluques, et sur les côtes des Terres Australes. Reinier & Josué Ottens.

Sartore J (n.d.). National Geographic Photo Ark. National Geographic. https://www.nationalgeographic.ors/projects/photo-ark/

Schnadt MJ, & Corley M (2006). The influence of lexical, conceptual and planning based factors on disfluency production. Proceedings of the Annual Meeting of the Cognitive Science Society, 28. https://escholarship.org/uc/item/9337x2hk

van Casteren M, & Davis MH (2007). Match: A program to assist in matching the conditions of factorial experiments. Behavior Research Methods, 39(4), 973–978. 10.3758/BF03192992 [PubMed: 18183914]

White KS, Nilsen ES, Deglint T, & Silva J (2020). That's thee, uuh blicket! How does disfluency affect children's word learning? First Language, 40(1), 3–20. 10.1177/0142723719873499
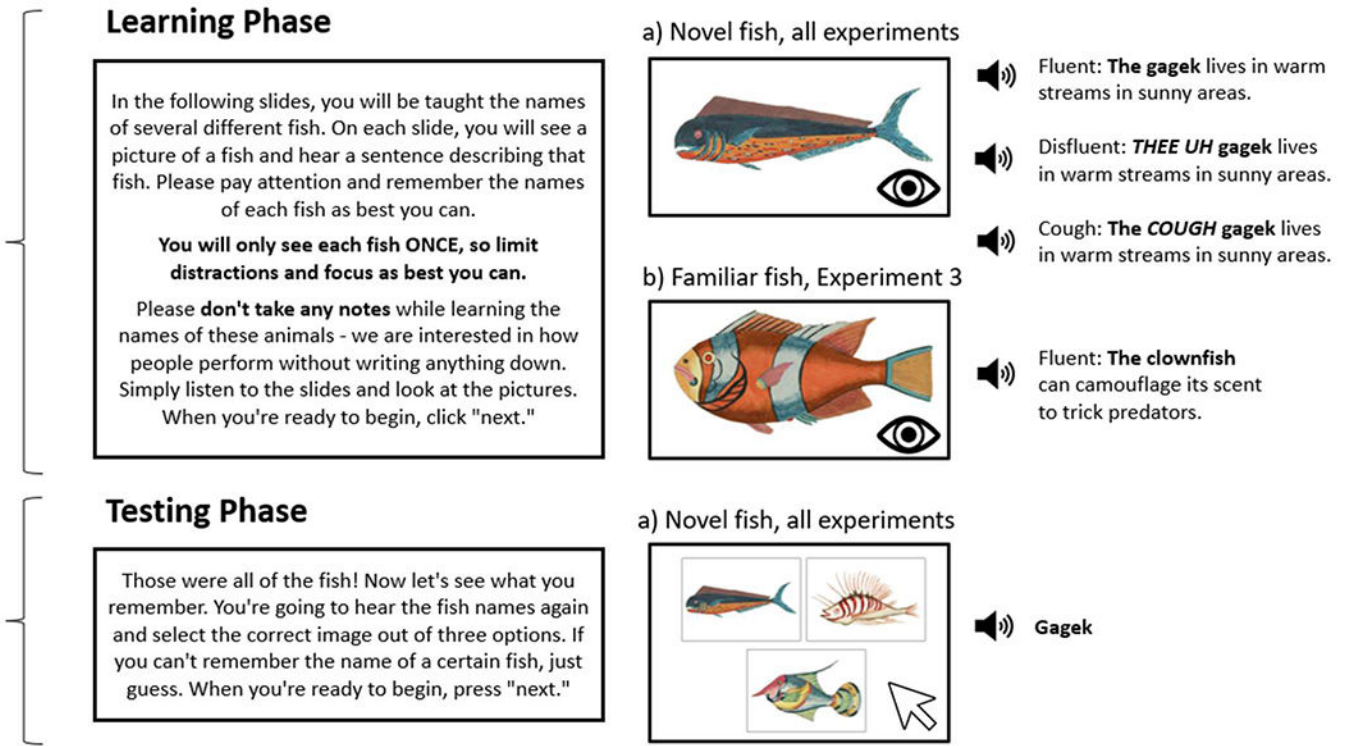
Figure 1.
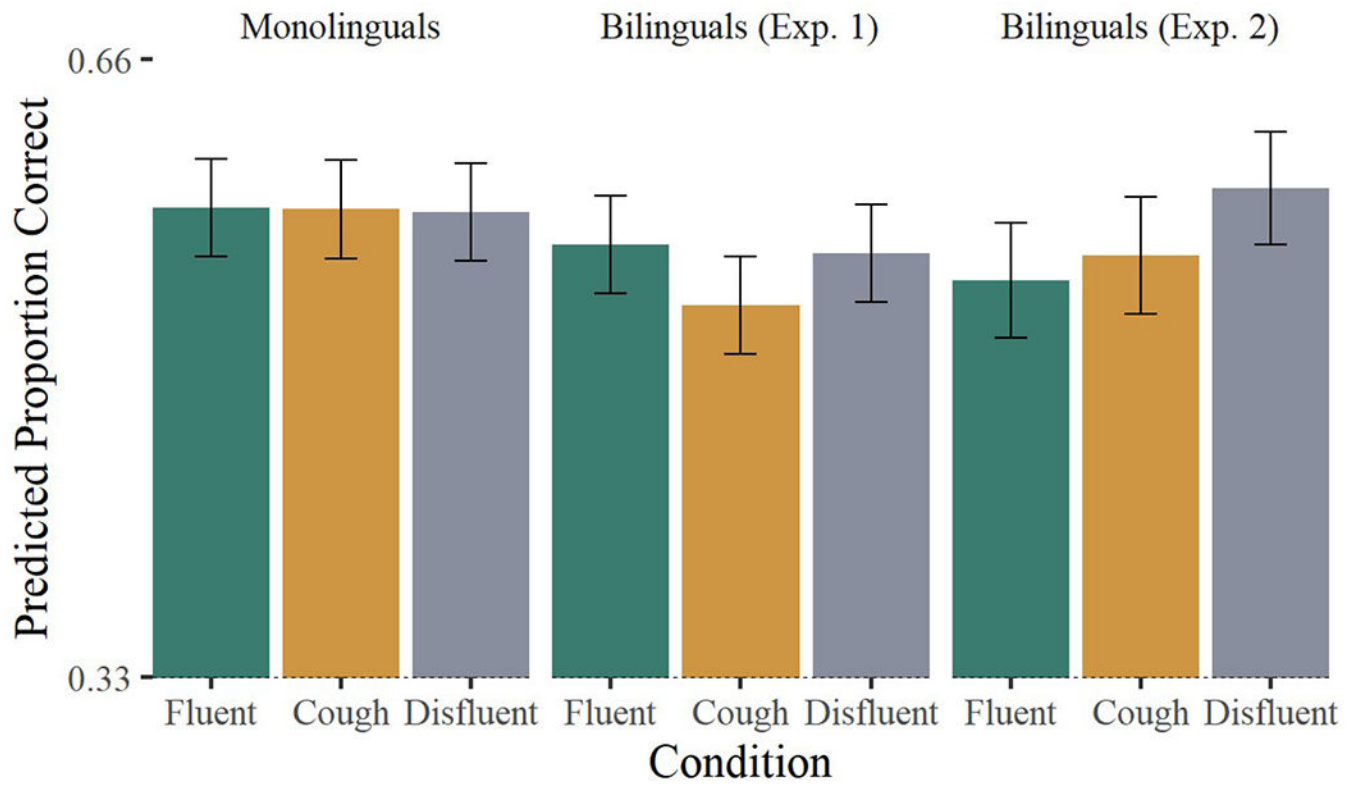Graphical depiction of learning and test phase, for all experiments.

Figure 2.
Predicted proportion correct by condition and group, with chance at .33. Error bars represent standard errors.
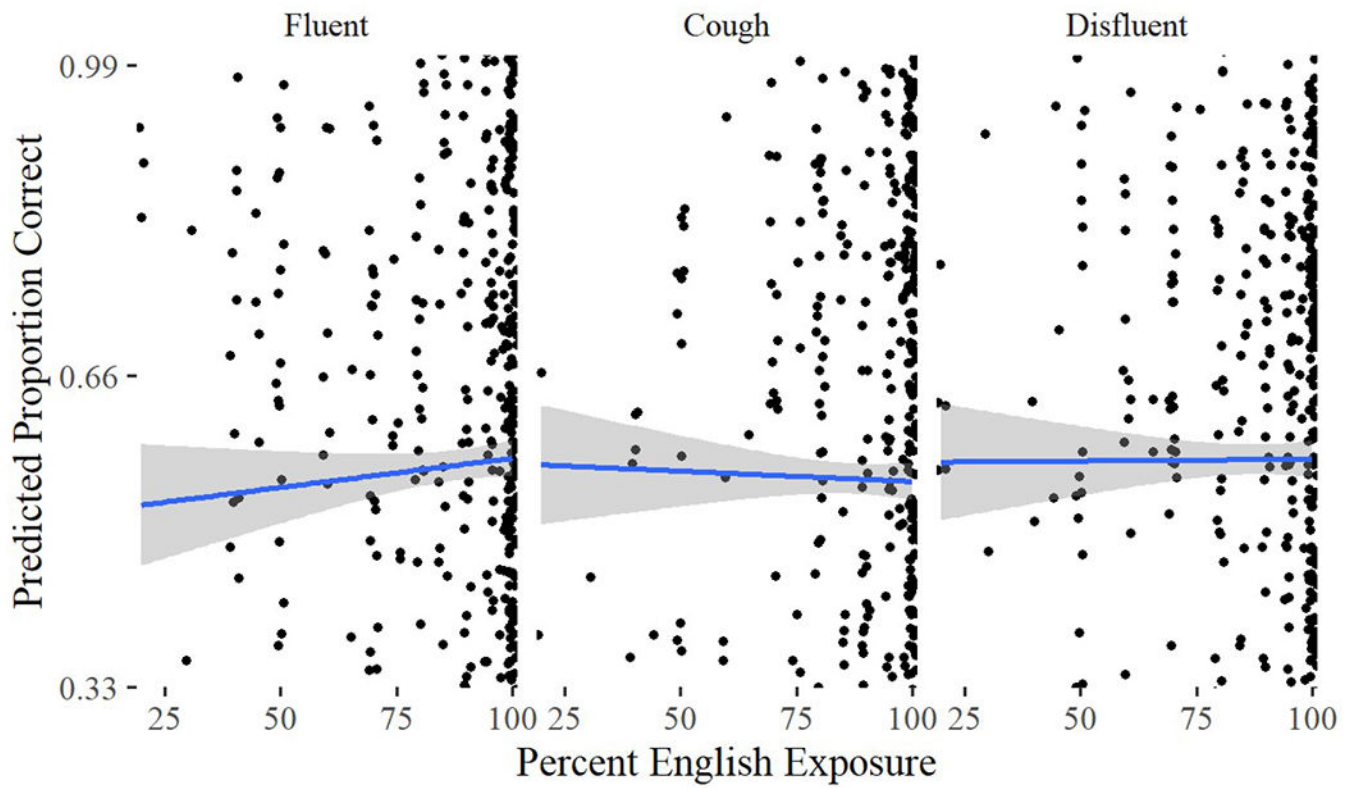
Figure 3.
Predicted proportion correct on the word learning task by percent English exposure and
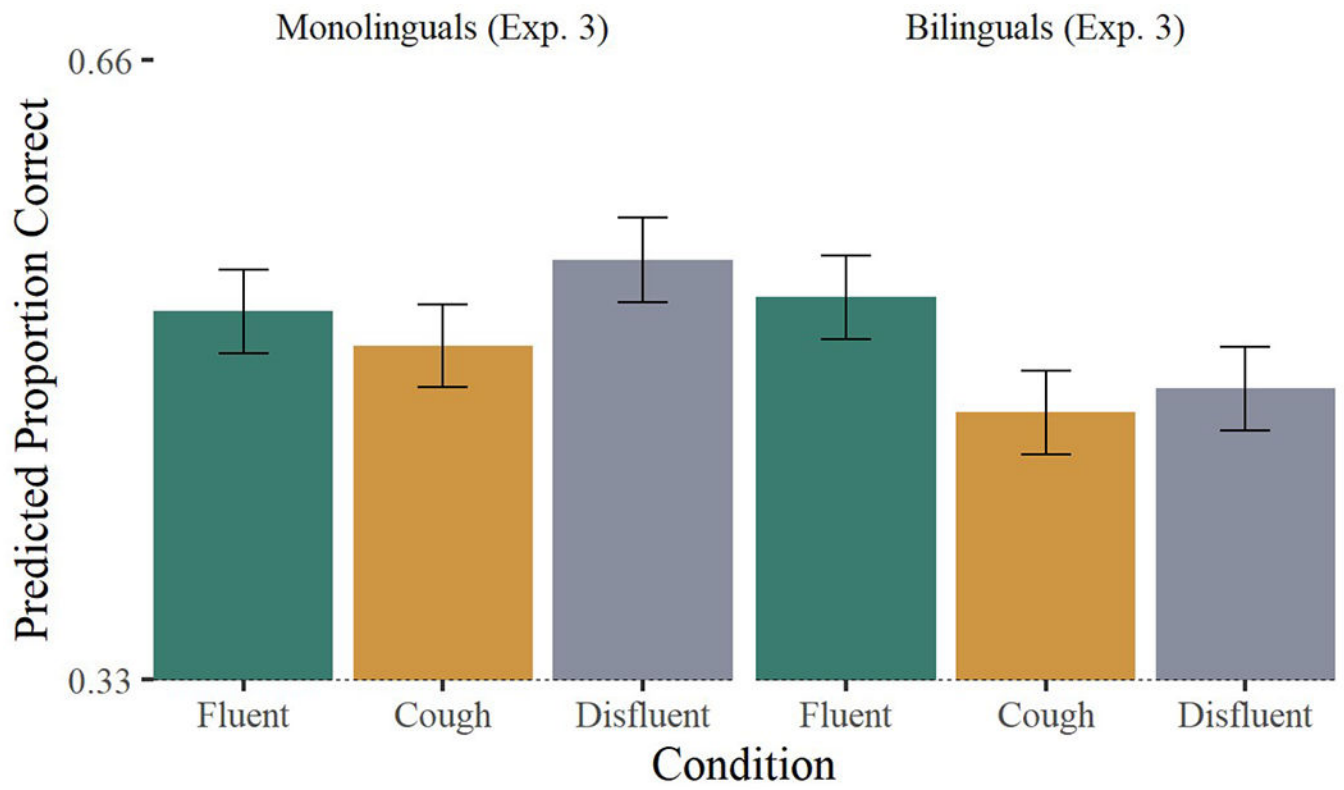condition, with chance level at .33. Grey ribbons represent standard errors.

Figure 4.
Predicted proportion correct by condition and bilingual status, with chance at .33. Error bars
represent standard errors.

Table 1.

Participant Characteristics.

	Monolinguals	Bilinguals (Experiment 1)	Bilinguals (Experiment 2)	*P*	*P*	*P*
	M(SD)	*M(SD)*	*M(SD)*	Mono. vs bi., Exp. 1	Mono vs bi., Exp. 2	Bi., Exp. 1 vs bi., Exp. 2
N	67	71	49	-	-	-
Age (years)	24.5(5.89)	24.5(6.13)	27.0(5.92)	.99	.032	.033
Years education	14.1(4.01)	14.4(3.26)	16.1(3.52)	.65	.005	.007
Age of acquisition, in years, for English	1.33(2.55)	1.28(2.22)	4.71(2.97)	.898	<.001	<.001
Self-rated proficiency for English (0-10)	9.65(0.648)	9.63(0.681)	9.18(0.950)	.914	…004	.006
LexTALE score (0-100)	83.5(11.3)	87.4(9.72)	82.1(11.0)	.033	.481	.007
Spelling score (0-15)	11.0(2.72)	11.6(2.73)	11.2(2.91)	.194	.681	.464
Percent English exposure	99.7(1.42)	88.6(14.2)	70.7(20.7)	<.001	<.001	<.001
Age of acquisition, in years, for other language	-	8.94(6.21)	1.20(0.957)	-		<.001
Self-rated proficiency for other language (0-10)	-	5.23(2.08)	7.51(2.30)	-	-	<.001

Table 2.

Participant characteristics for Experiment 3.

	Monolingual (Experiment 3)	Bilinguals (Experiment 3)	P
	M(SD)	M(SD)	
N	47	49	-
Age (years)	28.6(6.56)	25.9(5.37)	.035
Years of education	15.2(2.38)	15.9(2.74)	.213
Age of acquisition, in years, for English	1.33(2.86)	5.98(5.20)	<.001
Self-rated proficiency for English (0-10)	9.77(0.520)	9.17(0.975)	<0.001
LexTALE score (0-100)	85.5(16.3)	83.4(13.2)	.499
Spelling score (0-15)	11.0(2.79)	10.8(2.98)	.669
Percent English exposure	98.1(5.25)	69.3(16.4)	<.001
Age of acquisition, in years, for other language	-	2.06(5.00)	-
Self-rated proficiency for other language(0-10)	-	7.57(2.10)	-