Bilinguals’ Twisted Tongues: Frequency Lag or Interference?

Chuchu Li¹, Matthew Goldrick², Tamar H. Gollan¹

¹University of California, San Diego
²Northwestern University

Acknowledgement: This research was supported by grants from the National Institute on Deafness and Other Communication Disorders (011492), the National Institute of Child Health and Human Development (050287; 077140; 079426; 051030), and the National Science Foundation (BCS0846147; BCS1457159). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NIH or NSF. We thank Mackenzie Young and Michael Blasingame for error coding.

Please send correspondence to:

Dr. Chuchu Li
Department of Psychiatry
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0948, USA
Email: chl441@ucsd.edu
Abstract

Though bilinguals know many more words than monolinguals, within each language bilinguals exhibit some processing disadvantages, extending to sub-lexical processes specifying the sound structure of words (Gollan & Goldrick, 2012). The current study investigated the source of this bilingual disadvantage. Spanish-English bilinguals, Mandarin-English bilinguals, and English monolinguals repeated tongue twisters composed of English non-words. Twister materials were made up of sound sequences that are unique to the English language (non-overlapping) or sound sequences that are highly similar – yet phonetically distinct – in the two languages for the bilingual groups (overlapping). If bilingual disadvantages in tongue twister production result from competition between phonetic representations in their two languages, bilinguals should have more difficulty selecting an intended target when similar sounds are activated in the overlapping sound sequences. Alternatively, if bilingual disadvantages reflect the relatively reduced frequency of use of sound sequences, bilinguals should have greater difficulty in the non-overlapping condition (as the elements of such sound sequences are limited to a single language). Consistent with the frequency-lag, but not the competition account, both Spanish-English and Mandarin-English bilinguals were disadvantaged in tongue twister production only when producing twisters with non-overlapping sound sequences. Thus, the bilingual disadvantage in tongue twister production likely reflects reduced frequency of use of sound sequences specific to each language.

Key words: Bilingualism, Tongue twister, Speech error, Phonological processing
**Introduction**

Though bilinguals are not obviously different from monolinguals when they speak in a single language, carefully controlled laboratory studies reveal subtle but significant linguistic processing disadvantages for bilinguals compared with monolinguals (see Bialystok, 2009 for review). Bilingual disadvantages have been found in a variety of different tasks, even when tested exclusively in their dominant-language, including smaller vocabulary size (Oller & Eilers, 2002), reduced verbal fluency (Gollan, Montoya, & Werner, 2002), slower response times in picture naming (Gollan, Montoya, Cera, & Sandoval, 2008), and more tip-of-the-tongue states (i.e., TOTs, in which speakers temporarily fail to retrieve a well-known word; e.g., Gollan & Silverberg, 2001; Pyers, Gollan, & Emmorey, 2009).

Recent evidence suggests that bilinguals also have more difficulty processing sub-lexical representations. In contrast to lexical representations, which refer to whole-words, sub-lexical representations correspond to parts of words such as segments. The word *dog* has a single lexical representation, but at the sub-lexical level is associated with multiple elements (e.g., the three segments /d/ /a/ /g/). There is no lexical representation for the nonword *wug*, but speakers access three segments (/w/ /ʌ/ and /g/) at the sub-lexical level when they produce this nonword. Bilinguals have been shown to be disadvantaged relative to monolinguals in nonword repetition (Gibson et al., 2015; Summers, Bohman, Gillam, Peña, & Bedore, 2010), and in repeating both word and nonword tongue twisters (Gollan & Goldrick, 2012). Adopting the tongue twister paradigm, a tool for inducing sound-based speech errors (Wilshire, 1998, 1999), Gollan and Goldrick (2012) instructed Spanish-English bilinguals, Mandarin-English bilinguals, and English speaking
monolinguals to repeat English sound sequences made of real words (e.g., *dirt bus boot dose*) or nonwords (e.g., *deat bock dote bart*). Although bilinguals in Gollan and Goldrick (2012) had acquired English at an early age, were highly proficient in the target language, and fully immersed in an English-majority culture, they still produced significantly more errors during tongue twister production than English speaking monolinguals. Additionally, Spanish-English bilinguals were most disadvantaged for nonword twister production, and Mandarin-English bilinguals were equally disadvantaged for nonword and word twister production. Thus, the bilingual disadvantage in twister production could not be attributed to lexical processing deficits; if anything, the disadvantage was enhanced without lexical support (as nonwords lack lexical representations), suggesting that it arises at a sub-lexical processing level that specifies sound structure. Though these results suggested that the bilingual disadvantage in tongue twister production arises at the sub-lexical level, the mechanism of the disadvantage remains unclear. In the present study, we consider the source of these sub-lexical processing difficulties in light of two explanations that have been widely considered to account for bilingual disadvantages in verbal tasks: language competition and frequency lag.

**Competition between languages**

On one view, bilinguals are disadvantaged because the intention to speak activates representations in both languages, and this co-activation creates greater competition for selection relative to that faced by monolinguals (Green, 1998; Kroll, Bobb, Misra, & Guo, 2008). Supporting this interference hypothesis, considerable evidence suggests that even when bilinguals intend to speak in just one language, the non-target language is
nevertheless activated automatically—and competes with the target language. For example, Hermans, Bongaerts, De Bot, and Schreuder (1998) found that Dutch-English bilinguals took longer to name the picture in English (e.g., *mountain*, “berg” in Dutch) when it appeared with a Dutch word phonologically related to the target's translation (e.g., *berm*, “verge” in English) than when it was accompanied by a control (e.g., *kaars*, “candle” in English; referred to as phono-translation interference). The above observation has been replicated in Spanish-Catalan bilinguals as well (Costa, Colomé, Gómez, & Sebastián-Gallés, 2003; but see Costa, Miozzo, & Caramazza, 1999 who showed significant facilitation when the picture was accompanied by a translation equivalent). Furthermore, many studies have suggested that bilinguals sometimes produce words in a language they did not intend to speak (Gollan, Sandoval, & Salmon, 2011; Poulisse, 1999), and language mixing and switching often impairs production of the dominant language significantly more than the non-dominant language (Gollan, Schotter, Gomez, Murillo, & Rayner, 2014; Guo, Liu, Misra, & Kroll, 2011; Kroll et al., 2008; Philipp, Gade, & Koch, 2007; Philipp & Koch, 2009).

Language competition also occurs at the phonetic level (Amengual, 2012; Goldrick, Runnqvist, & Costa, 2014). For sounds that are phonologically similar, the phonetic systems of the two languages will interact so that bilinguals’ production of these sounds are different from monolinguals of either language (Flege, 1987, 1991, 1995, 2002; MacKay, Flege, Piske, & Schirru, 2001). For example, voice-onset time (VOT) refers to the time between the release of the consonant’s constriction and the onset of periodicity signaling modal vocal-fold vibration. Many languages contrast voiced and voiceless consonant contrasts using VOTs, but do so in different ways. For instance, in
Spanish this contrast is realized by negative vs. short positive VOT (i.e., vibration occurs first vs. constriction release occurs first). While English has a similar contrast, it is realized differently -- by a short vs. long positive VOT (Lisker & Abramson, 1964). As a result, when speaking English, Spanish-English bilinguals tended to produce shorter and more prevoiced VOTs than English monolinguals (Flege, 1991). This effect is not simply due to mis-learning English sounds, as the effect is magnified when producing an English word after having recently produced a Spanish word (Balukas & Koops, 2015; Goldrick et al., 2014; Olson, 2013). Interestingly, these effects of cross-language interference are enhanced for lexical items that induce strong cross-linguistic activation – cognates (translation equivalents that are similar; e.g., the Spanish word for tiger is tigre), Amengual (2012) found that in cognates Spanish-English bilinguals produced Spanish voiceless stops with longer VOTs (i.e., more English-like) than they do when producing non-cognates (see Jacobs, Fricke, & Kroll, 2015, for complementary effects in English learners of Spanish). These studies suggest that interference between languages arises at the phonetic level during bilingual speech production, particularly when the non-target language is strongly activated (e.g., producing cognates or recent production in the non-target language). In such contexts, bilinguals tend to produce sounds with phonetic properties similar to that of the non-target language.

Importantly, although bilinguals sometimes process cognates more easily than noncognates (e.g., Costa, Caramazza, & Sebastián-Gallés, 2000; Gollan & Acenas, 2004; Hoshino & Kroll, 2008), these effects likely reflect lexical level facilitation effects, which could not possibly arise when bilinguals process nonwords. Additionally, even with lexical level support, cognates do not always elicit facilitation effects, and in some cases
instead elicit interference effects. For example, interference can arise when cognates are written similarly but pronounced differently across languages (Schwartz, Kroll, & Diaz, 2007). Interference may also occur in a language-switching context when reading aloud written cognate words (Filippi, Karaminis, & Thomas, 2014), or sometimes in picture-naming (see English-dominant bilinguals in Broersma, Carter, & Acheson, 2016; but see Christoffels, Firk, and Schiller (2007) and Verhoef, Roelofs, and Chwilla (2009) for the opposite results, i.e., consistent faster naming latency in cognates than noncogantes).

Finally, interference has also been found for typed translation of written words, probably due to language competition at the orthographic level (Muscalu, 2007). If such competition effects can arise even with lexical level support, they might well be considerably stronger when cross-language overlap arises at a sub-lexical level, specifically when bilinguals must produce subtle language-specific differences without lexical support. Thus, during tongue twister production, between-language competition effects for nonwords with overlapping sounds might be particularly strong.

**Frequency-lag**

An alternative explanation is that bilingual disadvantages arise due to a frequency-lag (Gollan, Slattery, Goldenberg, Van Assche, Duyck, & Rayner, 2011; a.k.a. weaker links, Gollan et al., 2008), an emergent property of bilingual language use. On this view, bilinguals have more difficulty with lexical access because they speak each of their languages relatively less frequently than monolinguals. Because frequency has robust effects on lexical accessibility, this leads to relative difficulty in accessing lexical items (Gollan & Acenas, 2004; Gollan et al., 2008; Gollan, Slattery et al., 2011).

Consistent with this view, bilinguals name pictures more slowly than monolinguals, but
especially if those pictures have low-frequency names (Gollan et al., 2008; Gollan, Slattery, et al., 2011; Ivanova & Costa, 2008). Note that although frequency lag and notions of competition between languages are sometimes mutually compatible hypotheses (see Runnqvist, Strijkers, Sadat, & Costa, 2011 for a review), some aspects of bilingual performance can only be explained by one or the other account (for discussion see Kroll & Gollan, 2014).

**The current study**

The present study aimed to address the source of bilingual disadvantages at the sub-lexical level by asking bilinguals to repeat two different types of English nonword tongue twisters (i.e., sound sequences that are not real English words but consist of sounds that can be found in English and follow the phonotactic constraints of English). To distinguish between the frequency-lag and interference hypotheses, we manipulated whether the twisters consisted of *overlapping sound sequences* (e.g., /puni/) that can be found both in English and the other language spoken by the bilinguals tested here (Spanish or Mandarin), or *non-overlapping sound sequences* (e.g., /spæv/), that can be found in English, but not in the bilinguals’ other language. (Note that the nonword materials used in our previous work (Gollan & Goldrick, 2012) are an uncontrolled mixture of these sequence types.)

If the bilingual disadvantage in twister production is caused by cross-language interference at the level of sub-lexical retrieval, it could result from competition from co-activated language-specific sound structure representations. As reviewed above, such interference effects have been found even in tasks conducted entirely in a monolingual context (e.g., Amengual, 2012). In such cases, similar, but phonetically distinct,
language-specific representations disrupted the articulation of speech sounds. In the present study such competition would be possible for twisters with overlapping sound sequences, but not for twisters with non-overlapping sound sequences. For example, the initial sound of the nonword /puni/, /p/, could be pronounced with long VOT (i.e. English-like) or short VOT (i.e., Spanish-like). If representations of both sounds are activated in Spanish-English bilinguals, competition may occur. In contrast, because the sound sequences in the non-overlapping condition are unique to English, the absence of highly similar forms across the two languages would prevent or reduce co-activation, allowing bilinguals to produce tongue twisters as accurately as monolinguals.

Conversely, if the bilingual disadvantage in twister production is caused by frequency lag, the opposite pattern should be found. Bilinguals will have had more cumulative practice producing overlapping than non-overlapping sound sequences, because the shared elements of cross-linguistically similar sounds would be produced in both languages. Given that frequent exposure facilitates processing, this will result in better performance on tongue twisters with overlapping sound sequences, and worse performance on lower frequency, non-overlapping sequences. For example, in the overlapping sound sequence /puni/, although the /p/ is pronounced slightly differently across languages (long vs. short VOT in English vs. Spanish), the frequency of this segment type is upped by production in both languages. As a result, pronouncing this type of sound is easier pronouncing the /sp/ sequence, /æ/ sound, or /v/ sound in the non-overlapping sound sequence /spæv/, given that the all the three sound sequences (i.e., /sp/, /æ/, and /v/) are absent in Spanish. One might think production of individual sounds should not exhibit frequency effects because all sounds would be used very frequently
and therefore at maximum levels of proficiency. However, low frequency segments elicit more speech errors than high frequency segments (Goldrick & Rapp, 2007; Kupin, 1982). Such effects arise independent of the phonetic complexity of the sound sequences. The same sound can elicit different error rates depending on experiment-specific frequency distributions (Goldrick & Larson, 2008). Thus, the fact that sound sequences unique to English are practiced less frequently by bilinguals than they are by monolinguals should lead bilinguals to have greater processing difficulty relative to sound sequences found in both languages.

Methods

Participants

Ninety-six undergraduates at University of California, San Diego in each of three language groups (32 Spanish-English bilinguals, 32 Mandarin-English bilinguals, and 32 English speaking monolinguals) participated for course credit. Table 1 shows self-reported participant characteristics and MINT (a Multilingual Naming Test) in English, and for bilinguals also in Spanish or Mandarin (the bilinguals’ native language). Most participants in the two bilingual groups were English-dominant (higher MINT scores in English vs. native language; 29 Spanish-English and 23 Mandarin-English). Both bilingual groups reported acquiring English later, lower self-rated English proficiency, and had significantly lower English MINT scores than monolinguals.

Materials and procedure

Forty-eight tongue twisters composed of English nonsense words were designed. Nonwords in each tongue twister were identical except for the stop consonants in word-initial onsets which contrasted in place of articulation. In the overlapping condition, each
of the twenty-four tongue twisters consisted of four nonsense words with a CVCV (consonant-vowel-consonant-vowel) structure (e.g., *tuni puni puni tuni*). In each nonsense word, the initial consonant was */t/, */p/, or */k/, the second consonant */m/ or */n/, vowels were */u/ or */i/. All these sounds occur in English, Spanish and Mandarin, and all the CV syllables can be found in all three languages as well. In the non-overlapping condition, each of the twenty-four tongue twisters also had four nonsense words, but with CCVC structure (e.g., *staev spaev spaev staev*). For each of these, the initial consonant was */s/, the second was */t/, */p/, or */k/. The final consonant was */v/ or */ŋ/, and the vowel was either */æ/ or */ə/. These initial consonant clusters do not occur in Mandarin or Spanish, while */v/ and */ŋ/ as final consonants do not occur in Spanish. For vowels, */æ/ and */ə/ do not occur in Spanish and occur only in restricted phonological environments in Mandarin.

Note that phonological structures that are less commonly found across languages also tend to also be less frequent within a language (see Goldrick & Larson, 2008, for discussion). This holds true for our tongue twisters. Using word position-specific counts from Baayen, Piepenbrock, and Gulikers, (1995), the mean frequency of syllables in our overlapping tongue twisters is 3309/million vs. 8/million for syllables in our non-overlapping twisters. The non-overlapping syllables are therefore predicted to be more difficult than overlapping syllables for both monolingual and bilingual speakers. Critically, the frequency lag account predicts that this frequency effect will be significantly larger for the bilingual vs. for the monolingual speakers.

Stimuli were presented using PsyScope X software (Cohen, MacWhinney, Flatt, & Provost, 1993, http://www.psy.ck.sissa.it) on an iMac 7 computer with a 20-inch color monitor. For each trial, the four nonsense words were presented in written form in the
center of the screen. Participants were instructed in English to read aloud the four nonsense words sequentially 4 times paced by a 100BPM metronome. The twenty-four twisters in each condition were evenly divided between ABBA (e.g., tuni puni puni tuni) and ABAB (e.g., tuni puni tuni puni), and the order of the two conditions was counterbalanced between participants.

**Results**

Productions were transcribed by two annotators. Reliability was assessed using 4008 nonsense words across 6 participants (3 monolingual, 3 bilingual). Agreement between the transcribers was 97%.

The annotators identified a total of 6498 errors. Figure 1 shows the average number of errors that each participant made within each trial (i.e., the total number of additions, substitution, segmental or syllable deletions produced by a participant for each tongue twister sequence). The total counts of errors were analyzed using mixed effect negative binomial regressions (using the R package lme4; Bates, Maechler, Bolker, & Walker, 2015). These models are appropriate for count data such as number of errors, relative to alternatives such as linear regressions (which analyze continuous variables) and logistic regressions (used to model proportional data). Separate regressions compared Spanish-English bilinguals to monolinguals and Mandarin-English bilinguals to monolinguals. Contrast-coded fixed effects included language group (bilingual vs. monolingual), condition (overlapping vs. non-overlapping), and the interaction of these factors. Subject and twister/trial were served as two random intercepts (models with random slopes failed to converge). The significance of each fixed effect was assessed via likelihood ratio tests (Barr, Levy, Scheepers, & Tily, 2013).
Like many statistical models, our regressions assume that events are independent. This assumption may be particularly problematic within each trial; if an error occurs on one syllable within a particular tongue twister, it may increase the probability that errors will occur on subsequent syllables (violating independence). We therefore also used logistic mixed-effects regressions (Jaeger, 2008) to analyze accuracy at the level of whole trials (i.e., whether the trial was completely correct vs. contained one or more, that is, any number of errors; see Figure 2). The only difference in model structure relative to the first analysis was the addition of correlated random slopes. Note this analysis has reduced power relative to the first analysis, as it obscures differences in degree of difficulty within a trial.

**Spanish-English Bilinguals**

In the negative binomial regressions, Spanish-English bilinguals produced significantly more errors than monolinguals overall (mean 1.91 vs. 1.07 errors; $\beta = .51; SE \beta = .23; \chi^2 (1) = 4.55, p = .033$). While there was no overall effect of overlap (mean 1.58 errors for non-overlapping vs. 1.39 errors for overlapping; $\beta = -.09; SE \beta = .13; \chi^2 (1) = .51, p = .477$), there was a significant interaction between participant group and overlap ($\beta = -.34; SE \beta = .12; \chi^2 (1) = 7.87, p = .005$). Consistent with the predictions of the frequency-lag account, bilinguals were disadvantaged only for non-overlapping twisters. In follow-up regressions within each group, bilinguals produced significantly more errors than monolinguals in the non-overlapping condition (mean 2.15 vs. 1.01 errors; $\beta = .71; SE \beta = .24; \chi^2 (1) = 7.91, p = .005$), while the two groups performed similarly in the overlapping condition (mean 1.66 vs. 1.13 errors; $\beta = .32; SE \beta = .26; \chi^2 (1) = 1.47, p = .225$). In addition, bilinguals produced more errors in the non-overlapping
condition than the overlapping condition (mean 2.15 vs. 1.66 errors; $\beta = -.27$; $SE\ \beta = .13$; $\chi^2 (1) = 3.89$, $p = .048$), while monolinguals performed similarly in the two conditions (mean 1.01 vs. 1.13 errors; $\beta = .076$; $SE\ \beta = .15$; $\chi^2 (1) = .25$, $p = .614$). The logistic mixed-effects regressions with trial-level accuracy as the dependent variable showed similar results, except that the overall bilingual disadvantage (i.e., the main effect of language group) was only marginally significant ($\beta = .52$; $SE\ \beta = .27$; $\chi^2 (1) = 3.5$, $p = .062$).

To consider if the bilingual disadvantage in twister production might be attributable to the difficulty of producing tongue twisters in a non-dominant language we asked if similar results are found when excluding all bilinguals whose MINT score was higher in Spanish than in English. When we repeated our analysis including only the 29 English dominant bilinguals, the results were qualitatively similar. Critically, the interaction between language group and condition was significant ($\beta = -.36$; $SE\ \beta = .13$; $\chi^2 (1) = 8.33$, $p = .004$), and follow-up comparisons showed that English-dominant bilinguals produced more errors than monolinguals in the non-overlapping condition (mean 2.23 vs. 1.01 errors; $\beta = .74$; $SE\ \beta = .25$; $\chi^2 (1) = 8.25$, $p = .004$), but not in the overlapping condition (mean 1.68 vs. 1.13 errors; $\beta = .33$; $SE\ \beta = .27$; $\chi^2 (1) = 1.46$, $p = .227$), and English-dominant bilinguals produced more errors in the non-overlapping than in the overlapping conditions (mean 2.23 vs. 1.68 errors; $\beta = -.29$; $SE\ \beta = .14$; $\chi^2 (1) = 4.28$, $p = .039$). Similar effects were found in logistic mixed-effects regressions examining trial-level accuracy.
Mandarin-English Bilinguals

In the negative binomial regressions, bilinguals tended to produce more errors than monolinguals overall, but this difference was just marginally significant (mean 1.25 vs. 1.07 errors; $\beta = .33; SE \ \beta = .19; \chi^2 (1) = 3.15, p = .076$). As reported above, overall, speakers produced equivalent numbers of errors in the overlapping and non-overlapping conditions (mean 1.18 errors for non-overlapping vs. 1.14 errors for overlapping; $\beta = -.060; SE \ \beta = .12; \chi^2 (1) = .25, p = .620$), but there was a significant interaction between participant group and condition ($\beta = -.31; SE \ \beta = .13; \chi^2 (1) = 6.09, p = .014$). As found for Spanish-English bilinguals, and as predicted by the frequency-lag account, Mandarin-English bilinguals were more disadvantaged for non-overlapping than for overlapping twisters. In follow-up regressions within each condition, bilinguals produced significantly more errors than the monolinguals in the non-overlapping condition (mean 1.36 vs. 1.01 errors; $\beta = .50; SE \ \beta = .20; \chi^2 (1) = 5.93, p = .015$), while the two groups performed similarly in the overlapping condition (mean 1.15 vs. 1.13 errors; $\beta = .14; SE \ \beta = .23; \chi^2 (1) = .35, p = .554$). In addition, bilinguals produced marginally ($\beta = -.23; SE \ \beta = .13; \chi^2 (1) = 3.01, p = .083$) more errors in the non-overlapping than the overlapping conditions (mean 1.36 vs. 1.15 errors), whereas monolinguals exhibited no such trends (i.e., they performed similarly in the two conditions; mean 1.01 vs. 1.13 errors; $\beta = .076; SE \ \beta = .15; \chi^2 (1) = .25, p = .614$). Similar results were found in logistic mixed-effects regressions, except that the interaction between language group and condition was only marginally significant ($\beta = -.42; SE \ \beta = .22; \chi^2 (1) = 3.5, p = .060$).

Fewer of the Mandarin-English bilinguals were English-dominant (relative to the Spanish-English bilinguals). Analyses including only the 23 English-dominant Mandarin-
English bilinguals revealed qualitatively similar results. Critically, the interaction between language group and condition was significant ($\beta = -.27; SE \ \beta = .16; \chi^2(1) = 3.90, p = .048$) in the negative binomial regressions. Possibly due to the reduction in power, the follow-up comparison of the two language groups within the non-overlapping condition was just marginally significant, with English-dominant Mandarin-English bilinguals producing more errors than monolinguals (mean 1.31 vs. 1.01 errors; $\beta = .43; SE \ \beta = .23; \chi^2(1) = 3.36, p = .067$). Additionally, the interaction between participant group and condition trended in the right direction in the logistic mixed-effects regressions, but was not significant ($\beta = -.39; SE \ \beta = .24; \chi^2(1) = 2.57, p = .109$).

**Vocabulary Matched Bilinguals and Monolinguals**

In order to investigate whether the bilingual disadvantage is a result of lower English proficiency, we kept all the 32 monolinguals and selected a group of 31 bilinguals (18 Mandarin-English bilinguals, 13 Spanish-English bilinguals) whose English MINT scores (Mean= 64.5, SD= 2.0) matched with 32 monolinguals’ (Mean=65, SD= 2.0). Of interest, this vocabulary matched subgroup of bilinguals elicited a similar pattern of results; that is, the interaction between language background and tongue twister type was significant ($\beta = -.39; SE \ \beta = .12; \chi^2(1) = 10.1, p = .001$). The high-vocabulary bilinguals produced significantly more errors than monolinguals in the nonoverlapping condition (mean 1.80 errors vs. 1.01 errors; $\beta = .57; SE \ \beta = .24; \chi^2(1) = 5.51, p = .019$) but not in the overlapping condition (mean 1.38 vs. 1.13 errors; $\beta = .12; SE \ \beta = .26; \chi^2(1) = .22, p = .640$), and the high-vocabulary bilinguals produced more errors in the non-overlapping than the overlapping conditions (mean 1.80 errors vs. 1.38 errors; $\beta = -.30; SE \ \beta = .14; \chi^2(1) = 4.40, p = .036$), while the monolinguals performed similarly in the
two conditions (mean 1.01 vs. 1.13 errors; $\beta = .076$; $SE \beta = .15$; $\chi^2 (1) = .25$, $p = .614$). The logistic mixed-effects regressions showed similar results.

**Error Types**

Even though our analyses imply similar mechanisms underlying the bilingual disadvantage in Spanish-English and Mandarin-English bilinguals (i.e., both bilingual groups were more disadvantaged with non-overlapping than with overlapping twisters), it is possible that further investigation will reveal subtle differences between groups. Figure 3 illustrates an exploratory division of our data into error types, including segment substitution (e.g., saying spaev for skaev), syllable deletion (e.g., failing to say the whole syllable skaev), segment deletion (e.g., saying kaev for skaev), and addition (e.g., saying skaevi or skskaev for skaev).

We did not conduct statistical analysis of each error type, given that power is greatly reduced (due to smaller numbers of observations and the need to correct for multiple comparisons). However, inspection of the figure suggests some striking differences that could be pursued in future work. The pattern we observed above (collapsing across error types) is a higher error rate for non-overlapping vs. overlapping sequences, with the difference largest for bilingual vs. monolingual speakers. While deletions of individual segments follow this pattern for both groups of bilingual speakers, deletions of whole syllables suggest possible differences across groups; Spanish-English but not Mandarin-English show more syllable deletions for non-overlapping vs. overlapping sequences.

Segment substitutions show a quite unexpected pattern; the monolinguals and Spanish-English bilinguals show higher error rates on overlapping sequences. This could
reflect distinct patterns of difficulty with overlapping vs. non-overlapping sequences; there may be differing opportunities for minor errors like substitutions vs. majors distortions like whole syllable deletion for these two types of twisters. When production of a non-overlapping sequence is disrupted, the result may frequently be a severe distortion of the target forms, with major alterations to their phonological structure (i.e., deletions, or reductions of consonant clusters, and omitting entire syllables). This precludes the possibility of producing a substitution error. In contrast, when production of an overlapping twister is disrupted, the relative ease of processing might allow for the possibility of minor distortions of the target twister (i.e., substitution errors - which largely preserve the overall syllabic and prosodic structure of the twister).

What these descriptive analyses make clear is the need for future investigation of these patterns: further exploration of the differences in processing of these two different sequence types, and a more detailed examination of how different types of bilingual experience shape the structure of errors made during sublexical processing.

Discussion

The present study compared two groups of bilinguals (Spanish-English bilinguals and Mandarin-English bilinguals) to English monolinguals in their ability to produce English non-word tongue twisters. The twisters either consisted of sound sequences and syllables that are found in both languages spoken by the bilinguals (overlapping) or that are unique to English (non-overlapping). Both bilingual groups performed similarly to monolinguals in the overlapping condition, but produced significantly more errors than monolinguals in the non-overlapping condition. In addition, both bilingual groups, but not monolinguals, produced more errors in the non-overlapping than in the overlapping
condition. The results were consistent in both bilingual groups, suggesting that our findings were not an artifact of one population or the demographic features that are associated with each group. Instead, they provide reliable and broadly generalizable results that will likely apply to bilingual speakers of all language combinations. In addition, we showed that similar results were found when the analysis was restricted to English-dominant bilinguals, and even to vocabulary matched groups of bilinguals and monolinguals. This suggests that the bilingual disadvantage does not simply reflect relatively low English proficiency, and is not an artifact of production in a non-dominant language. Further investigation is needed to determine why vocabulary matching – which might eliminate any effects of frequency lag at the lexical level – did not also eliminate lag effects at the sublexical level.

The present study replicated and extended the findings in Gollan and Goldrick (2012), which showed bilingual disadvantages at the sub-lexical level but did not investigate the mechanism of these disadvantages. The present results suggest that the bilingual disadvantage in non-word tongue twister production is not likely to be caused by between-language interference. Under this account, bilinguals should have shown the highest accuracy in the non-overlapping condition, which would minimize cross-linguistic interference. Instead, the results are consistent with the frequency-lag account; non-overlapping sound sequences are less practiced for bilinguals because they are used in one but not the other language spoken by bilinguals. Thus, their relatively low frequency appears to result in processing difficulty. Note that our results did not suggest that competition at the phonetic level never occurs. Instead, such competition may be more likely to arise in situations where there is uncertainty about which language will be
used, such as a language-switching task (Goldrick, et al., 2014), and further research is needed to reveal when competition between languages can arise at the phonetic level.

A further question that merits additional inquiry is why monolinguals did not exhibit robust differences between conditions in the present study. Based on frequency lag, weaker effects of twister type should be found with monolinguals, as the non-overlapping syllables have lower frequency in English than the overlapping syllables. However, monolingual performance may be too close to ceiling in our task, to produce significant frequency effects. At faster speaking rates, which would elicit higher error rates, a more robust frequency effect should emerge for monolinguals (see Goldrick & Larson, 2008 in which monolinguals repeated tongue twisters at a rate of 210 BPM).

Another issue is why monolinguals and Mandarin-English bilinguals performed similarly in the overlapping condition while Spanish-English bilinguals made more errors than the other two groups (see Figure 1). This could arise because the overlapping sequence does not occur exactly equally often for all the three groups—the critical comparison in the present study is simply that overlapping sound sequences are more frequent than the non-overlapping ones, not that they have a particular frequency level. It is also possible is that the considerable overlap in writing systems between English and Spanish could have allowed our written materials to induce a more bilingual mode (see Grosjean, 1994) in Spanish-English relative to Mandarin-English bilinguals. Alternatively, differences between groups might reflect other aspects of their exposure to and use of two languages. For example, Spanish has a much more prominent presence in the linguistic and cultural environment of San Diego than Mandarin; this (rather than
orthography) may have led to a more bilingual mode for the Spanish-English vs. Mandarin-English bilinguals.

Beyond frequency lag, difficulty with acquisition of two different and contrasting phonological systems (see Bosch & Sebastián-Gallés, 2003; Werker, 2012 for simultaneous acquisition, and Best & Tyler, 2007 for sequential acquisition) may also contribute to the bilingual disadvantage in sub-lexical processing. Like frequency lag, this account predicts a greater disadvantage for bilinguals with production of non-overlapping twisters. Properties of one language can facilitate the acquisition of the other language for shared sounds (e.g., those found in overlapping sound sequences). In contrast, acquiring unshared sounds is more difficult since they belong to two distinct phonological systems and do not allow for positive transfer between languages (e.g., those found in non-overlapping sequences). Supporting this view, previous studies found that bilingual children have higher production accuracy for shared sounds compared to unshared sounds (Fabiano-Smith & Goldstein, 2010). Note that the difficulty of acquisition account is different from an age of acquisition effect since acquiring two distinct phonological systems is more difficult than acquiring one even for young language learners. It is difficult to differentiate the frequency lag and difficulty of acquisition accounts, as the factors underlying each account are highly correlated. As noted in the methods section, frequency within a language tends to correlate with frequency across languages. Unshared sounds are likely to be less frequent cross-linguistically, and therefore less frequent within a particular language; this general correlation presents a challenge to discriminating these accounts.
In summary, the present study suggests the bilingual disadvantage in speech production at the sub-lexical level that specifies sound structure is likely due to frequency-lag, or difficulty of acquisition of two distinct phonological systems, or both. Conversely, the bilingual disadvantage appears not to be due to difficulty with selecting between competing sound representations in each language, and also not by lower proficiency or production in a non-dominant language. Instead, bilinguals’ increased production of tongue twisters errors appears to reflect reduced accessibility of sound structures unique to one of the bilinguals’ two languages. The findings of the present study imply that caution is needed when developing materials for future work with bilinguals, as overlapping sounds may cancel out bilinguals’ disadvantages in some language production tasks. More broadly, these results provide additional support for the view that sub-lexical processes are sensitive not only to the inherent articulatory complexity of different sound sequences but also to the relative frequency with which different speakers encounter them (Goldrick & Larson, 2008).
References


Muscalu, L. (2007). *The role of phonological activation in bilinguals' orthographic realization of words* (doctoral dissertation). Claremont Graduate University, California, USA.


http://doi.org/10.1080/09541440600758812


http://doi.org/10.1037/a0016376


http://doi.org/10.1016/j.cognition.2009.04.007


Table 1. Means and standard deviations of participant characteristics

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Age</td>
<td>20.4</td>
<td>2.5</td>
<td>20.1</td>
<td>1.4</td>
<td>20.3</td>
<td>1.6</td>
</tr>
<tr>
<td>Age of Acquisition of English</td>
<td>3.8</td>
<td>2.8</td>
<td>4.0</td>
<td>3.4</td>
<td>0.5</td>
<td>1.2</td>
</tr>
<tr>
<td>Self-rated spoken English proficiency&lt;sup&gt;a&lt;/sup&gt;</td>
<td>6.5</td>
<td>0.7</td>
<td>6.5</td>
<td>0.8</td>
<td>6.8</td>
<td>0.8</td>
</tr>
<tr>
<td>Percent of English use during childhood</td>
<td>52.5</td>
<td>16.5</td>
<td>49.6</td>
<td>21.8</td>
<td>94.5</td>
<td>8.8</td>
</tr>
<tr>
<td>Current percent of English use</td>
<td>78.8</td>
<td>20.1</td>
<td>81</td>
<td>19.5</td>
<td>98</td>
<td>5.9</td>
</tr>
<tr>
<td>Primary caregiver English proficiency&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3.6</td>
<td>1.8</td>
<td>4.6</td>
<td>1.5</td>
<td>6.4</td>
<td>1</td>
</tr>
<tr>
<td>Secondary caregiver English proficiency&lt;sup&gt;a&lt;/sup&gt;</td>
<td>4.0</td>
<td>2.2</td>
<td>4.7</td>
<td>1.6</td>
<td>6.2</td>
<td>1.3</td>
</tr>
<tr>
<td>Primary caregiver education level (years of education)</td>
<td>10.9</td>
<td>4.6</td>
<td>16</td>
<td>3.9</td>
<td>15.9</td>
<td>3.5</td>
</tr>
<tr>
<td>Secondary caregiver education level (years of education)</td>
<td>11.2</td>
<td>4.6</td>
<td>15</td>
<td>4.6</td>
<td>15.7</td>
<td>3.4</td>
</tr>
<tr>
<td>Years lived in non-English speaking country</td>
<td>1.6</td>
<td>3.6</td>
<td>6.5</td>
<td>6.7</td>
<td>0.2</td>
<td>0.6</td>
</tr>
<tr>
<td>MINT score in English&lt;sup&gt;b&lt;/sup&gt;</td>
<td>60.2</td>
<td>3.8</td>
<td>60.1</td>
<td>7.6</td>
<td>65.0</td>
<td>2.0</td>
</tr>
<tr>
<td>MINT score in other language&lt;sup&gt;b&lt;/sup&gt;</td>
<td>48.3</td>
<td>7.9</td>
<td>45.6</td>
<td>12.6</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

For all comparisons not marked with *or *** p > .14.

* Significant difference at p < .05 level.

*** Significant difference at p < .001 level.

<sup>a</sup> Proficiency-level self-ratings were obtained using a scale from 1 (little to no knowledge) to 7 (like a native speaker).

<sup>b</sup> The maximum possible MINT score is 68
Figure 1. Mean number of errors within each trial for each condition and language group.
Figure 2. Mean percent of trials with errors for each condition and language group.
Figure 3. Mean number of errors within each trial for each condition and language group according to error type.