Face Recognition and Bilingual Lexical Access: Familiarized faces prime performance in a written language-selection task

David Abugaber (dabuga2@uic.edu)
Department of Hispanic and Italian Studies, University of Illinois at Chicago
601 S. Morgan Street
Chicago, IL 60607 USA

Abstract

The unique nature of face perception (Bruce & Young, 1986) suggests that faces may have a strong capacity to prime language. Indeed, a recent line of studies (Molnar, Ibáñez-Molina, & Carreiras 2015; Woumans et al., 2015; Martin, Molnar, & Carreiras, 2016) shows that bilingual language activation can be modulated by visual information from videos of speakers as they produce a test word. However, these studies' use of video stimuli leaves the possibility that the face-priming effects were driven by the viewers' knowledge of speaker-specific phonetic features. To avoid this potential confound, this study used face prime images that were temporally separated from written test words. In the experiment, English native speakers with self-reported high L2 Spanish proficiency were first familiarized with eight people from either a Spanish-speaking or English-speaking background, via the co-presentation of face images with self-introductory texts in the corresponding language. Participants then performed a timed language-decision task on individually-presented Spanish/English words. Critically, each word was preceded by the brief presentation of a familiarized face. Congruence between the language of the face primes and of the words had a significant effect on response times for English words but not for Spanish words. These findings suggest that non-linguistic stimuli that typically co-occur with a specific language (in this case, familiar faces) may modulate access to that language in bilinguals, but that this effect may depend on proficiency. Such results indicate the psycholinguistic importance of all of a multilingual's experiences in a community of speech, and not just strictly linguistic ones.

Keywords: bilingualism; face perception; priming

Introduction

Linguistic processing involves competition between similar candidate forms at various levels of language (Vitevitch & Luce, 1998; Grainger, O'Regan, Jacobs, & Segui, 1989; Garrett, 1980). For bilinguals, overcoming such competition would seem especially difficult because they must reject not only words with the wrong meaning but also words with the right meaning in the wrong language. Such *parallel activation* of the non-target language has been found at the level of orthography (Bijeljac-Babic, Biardeau, & Grainger, 1997), phonology (Jared & Kroll, 2001), semantics (Elston-Güttler & Williams, 2008), and syntax (Loebell & Bock, 2003). Although bilinguals are generally successful in avoiding unintentional mixing (Poulisse & Bongaerts, 1994), their performance in language selection tasks can be affected by contextual factors such as whether experimenters present themselves as monolinguals or bilinguals (Canseco-González et al., 2010) or by the language of an unrelated film shown before the task (Elston-Güttler, Gunter, & Kotz, 2005).

Several models have been proposed to account for such parallel activation, including the Bilingual Interactive Activation+ model (BIA+; Dijkstra & Van Heuven, 2002), the Adaptive Control hypothesis (Green & Abutalebi, 2013), the Bilingual Interactive Model of Lexical Access (BIMOLA; Léwy & Grosjean, 2008), and the Bilingual Language Interaction Network for Comprehension of Speech (BLINCS; Shook & Marian, 2013). These models generally posit that features for a bilingual's different languages are stored in the same network, but that separate sub-networks eventually emerge for each language. Additionally, higher-level contextual information can facilitate selection of a certain language, either by inducing top-down activation/ inhibition or by adjusting language decision thresholds. The idea that higher-level contextual information can affect bilingual processing seems more plausible in light of eye-tracking evidence that visual information is quickly incorporated in language comprehension (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995) and that culturally-biased images can prime bilinguals towards an associated language (Zhang, Morris, Cheng, & Yap, 2013).

There are several reasons why faces may be particularly powerful sources of bilingual priming effects. As producers of speech, faces usually accompany language in a way that inanimate objects don't. Additionally, speakers might be more attentive in general to faces than to objects/scenes, as faces are important sources of information such as sex, age, health status, and emotion. Furthermore, speakers must often consider the identity of the addressee, e.g., when adjusting their speech style for a certain interlocutor.

Further suggesting the priming potential of faces, a wide body of psychology research indicates that face perception is a specialized, automatic cognitive process that is qualitatively different from the recognition of objects or scenes (Bruce & Young, 1986). Face perception has been associated with a specific brain region (the fusiform gyrus; Von Kriegstein, Kleinschmidt, Sterzer, & Giraud, 2005), with a specific ERP response (the N170; Kanwisher, 2000), and with a specific impairment (prosopagnosia; Bauer, 1984). A strong link between face perception and speech perception is suggested by findings of altered phonological perception based on a face's place of articulation (McGurk & McDonald, 1976); activation of the fusiform gyrus by voices (Von Kriegstein et al., 2005); facilitation of voice-learning by co-occurring faces

(Sheffert & Olson, 2004); and faster ERP responses to speech when simultaneously viewing a speaking face (van Wassenhove, Grant, & Poeppel, 2005).

A recent line of experiments investigated the link between interlocutor identity and bilingual language selection. In the first of these (Molnar, Ibáñez-Molina, & Carreiras, 2015), native speakers of Spanish with either low or high Basque proficiency were shown videos in which interlocutors gave brief self-introductions in either Basque, Spanish, or a mixture of the two languages. Afterwards, participants carried out a lexical decision task on video clips in which those same interlocutors produced either Basque words, Spanish words, Basque-like pseudowords, or Spanish-like pseudowords. Both low- and high-proficiency bilinguals showed slower reaction times to trials from interlocutors who had presented themselves as bilinguals (by mixing both languages in their self-introductions); the experimenters attributed this to the fact that participants could not make a prediction in favor of either language for the upcoming word. Furthermore, high-proficiency bilinguals showed faster response times for congruent trials (in which "monolingual" interlocutors produced a word in the same language as their self-introduction) vs. incongruent trials. However, low-proficiency bilinguals did not show any language-congruency effect whatsoever. The study also included trials in which the voice of the interlocutor was artificially replaced with a different voice; slower response times to these trials confirmed that the participants were incorporating visual information in their language processing, rather than just the voice of the interlocutor.

A follow-up study with the same design used ERP methods to compare processing when the test item was produced by a bilingual vs. a monolingual (Martin, Molnar, & Carreiras 2016). Differences were found in the P3 component after video onset but before hearing the test item (suggesting pre-speech task preparation after seeing a monolingual's face) as well as in the N1 and N400 components after hearing the test item (reflecting more efficient word discrimination in the monolingual case, due to priming for that language).

In a related study involving language production (Woumans et al., 2015), bilinguals were familiarized with certain faces and their corresponding language through simulated videochat conversations. In a test phase, participants were asked to produce verbs associated with nouns uttered by these faces in either the familiarized or a different language. Production was faster for congruent trials, an effect that disappeared when that face was discovered to be bilingual (i.e., after several incongruent trials from that face). These results suggest that face-priming effects only occur when the face is a reliable cue for a specific language; as such, bilingual faces would not serve as strong primes.

Although these three studies provide evidence for a language priming effect from faces, they all share a potential confound: because participants viewed videos of interlocutors' faces as they spoke, it is possible that the effects were driven by the extraction of speaker-specific phonetic information. Previous studies in phonetics indicates that familiarity with a certain voice can affect speech processing in subsequent tasks (Sheffert & Olson, 2004; Smith & Hawkins 2012). As such, it is possible that the facilitation in bilingual language selection in the above studies was due to general speech processing mechanisms from monolinguals, rather than to non-linguistic contextual information related to the "linguistic identity" of the faces.

Methods

This study sought to test whether familiarized faces could prime a specific language in bilinguals by inducing them to access information about the person's linguistic background. Importantly, the study sought to avoid any potential confounds from speaker-specific phonetic information by using static face images that were temporally separate from test words that were presented in a written modality.

Participants

Eighteen English native speakers with self-reported advanced Spanish proficiency were recruited from a British university community. Following a procedure common for bilingual processing studies (e.g., Dijkstra et al., 2010; Elston-Güttler & Williams, 2008) and previously found to be reliable (Marian, Blumenfeld, & Kaushanksaya, 2007), a questionnaire was conducted (in English) to gauge L2 experience. All participants claimed at least four years of Spanish study and rated themselves as "advanced" or higher. Twelve participants had worked or studied in a Spanish-speaking country, and four of the remaining ones had travelled in a Spanish-speaking country since beginning their Spanish studies. Although participants were not full, "balanced" bilinguals, their extensive experience as Spanish learners would make their performance in this experiment relevant for investigating the nature of L2 processing.

Materials

Face stimuli were acquired from the Psychological Image Collection at Stirling (PICS; http://pics.psych.stir.ac.uk/). From this database, eight different individuals were selected (four males and four females), each with four associated images including two different facial expressions and two different poses. This variation in pose and expression allowed for a distinction between actual face recognition and simple low-level association with particular images. The photographs were all in black-and-white and had the same composition, size (269x369 pixels), and plain black background. The eight faces were divided into two sets (balanced for gender), to be associated with either English or Spanish via co-presentation with a short (35-50 word) self-introductory text written in the corresponding language. Such straightforward biographical presentation has been shown to facilitate face recall and modulate ERP responses to faces (Tsujimoto, Yokoyama, Noguchi, Kita, & Kakigi, 2011).

Word stimuli in English and Spanish (40 for each language) were prepared using the CLEARPOND database (Marian, Bartolotti, Chabal, & Shook, 2012). These were matched for orthographic length; frequency; bigram token frequency;

imageability; and number of higher-frequency orthographic and phonological neighbors. Independent-samples t-tests found no significant differences for any of these variables between the two languages (all *p*<.05), and none of the words appeared in the self-introductory texts from the familiarization phase. No words contained diacritics or had orthographic or phonological neighbors in the other language (where a "neighbor" is a word that can be constructed by the substitution, addition, or deletion of a single letter or phoneme). The 40 words in each language were divided into two lists, to be presented with either English or Spanish faces; independent samples t-tests found no significant differences for any of the variables described above between items in the same language across the two lists (all *p*<.05). Finally, the association between the presentation lists and the language of the co-presented face stimuli was counterbalanced between participants as per a Latin squares design.

Procedure

The experiment consisted of a familiarization phase, a test phase, and a language-association test. For the familiarization phase, participants first read instructions in English saying that they would be introduced to some people about whom they would answer a few questions. They were then presented with a face image (for three seconds) and then the same face along with a corresponding self-introductory text (for ten seconds); order of presentation was randomized between participants. After all eight individuals had been presented once, there was a short assessment in which each face was shown along with a yes-or-no question (in English) about the person, based on the biographical information from the texts. Participants were instructed to respond with either the left or right key on a response pad (key assignment was counterbalanced between participants). After each response, the relevant face and biographical self-introduction were presented again for four seconds. There were three such cycles of questions, with different questions for each cycle.

For the test phase, participants indicated via button response whether the presented word was an English or Spanish word. Such a language-selection task has been found to induce L2 activation even in lower-proficiency bilinguals (Dijkstra et al., 2010; Casaponsa, Carreiras, & Duñabeitia, 2014), suggesting that it was appropriate for our purposes. Critically, participants were also instructed to pay attention to face images preceding the words, as they would have to answer questions about them. Each trial consisted of a fixation cross at the center of the screen (for 500 milliseconds), followed by either an English or Spanish face (1000ms), and then by either an English or Spanish word (presented until participant response, up to a 2500ms maximum). Words were presented over a black background in white lowercase letters in 20-point Arial font. Each participant viewed 20 English words and 20 Spanish words after English faces, and 20 English words and 20 Spanish words after Spanish faces. There were also 24 catch trials in which a word was followed by a yes-or-no question about the previously-shown face, providing an incentive for participants to recall the face stimuli during the critical naming trials. A short practice session was included at the beginning of the test phase, as well as a break after every 40 trials. Order of presentation was randomized, and key assignments were counterbalanced between participants.

A language-association phase was included at the end of the experiment to ensure that participants had truly associated the faces with their corresponding language. Participants were instructed to indicate via button response whether the presented face was that of an English speaker or a Spanish speaker. Each trial consisted of a fixation cross (500ms) and then a face (presented until participant response, up to a 2500s maximum). All 32 face images (four for each of the eight individuals) were shown during this phase. The order of presentation was randomized, and the response key assignments were counterbalanced between participants.

Results

One-sample t-tests on responses to the biographical questions, catch trials, and language-association questions confirmed that participants' performance was significantly above chance (all *p*<.001). Analyses of reaction times to the word stimuli excluded erroneous responses and responses more than 2.5 standard deviations from the mean (calculated per participant per language). Table 1 below shows reaction time means and standard deviations as well as error rates for each condition.

Table 1: Reaction time means and standard deviations (in milliseconds) and error rates for different trial types.

Trial type	Response Time	Error Rate
English face, Eng. word	729ms (113)	2.78%
English face, Span. word	725ms (129)	3.61%
Spanish face, Span. word	731ms (121)	2.50%
Spanish face, Eng. word	772ms (154)	0.83%

A mixed-design Analysis of Variance (ANOVA) by participants was performed on average reaction times to word stimuli, using word language and congruence with the co-occurring face as within-subject factors and presentation list in the Latin squares design as a between-subject factor. No significant effects were found for presentation list, $F(1, 16)=0.68$, $p=.420$; word language, $F(1, 16)=2.17$, $p=.160$; or congruence, $F(1, 16)=2.02$, $p=.175$. A significant interaction was found between language and congruence, $F(1, 16)=6.98$, $p=.018$. Follow-up paired-items t-tests found that response times were significantly faster for congruent vs. incongruent words in English, $t(17)=-2.56$, $p=.020$, but not in Spanish $t(17)=0.39$, $p=.699$.

For a by-item analysis, a mixed-design ANOVA was performed on each word's average response times across congruent and incongruent conditions, with congruence as a within-item factor and word language and presentation list as between-item factors. No significant effects were found for language, $F(1, 76)=3.24$, $p=.076$, or presentation list, $F(1, 76)=0.13$, $p=.718$. A significant main effect was found for congruence, $F(1, 76)=4.03$, $p=.048$, wherein congruent words showed faster responses. A significant effect was also found for the interaction of congruence and word language, $F(1, 76)=5.02$, $p=.028$. Follow-up paired-samples t-tests found significantly faster responses for congruent than for incongruent words in English, $t(39)=-2.52$, $p=.016$, but not in Spanish, $t(39)=0.16$, $p=.876$.[1]

Discussion

The results suggest that face stimuli did have some priming effect on language, as measured by faster language-decision response times to English words following English-associated faces than to English words following Spanish-associated faces. This would indicate that linguistic processing is sensitive to non-linguistic contextual effects from face stimuli, and that previously-reported face-priming effects were not driven by viewers' knowledge speaker-specific phonetic information (Molnar et al., 2015; Woumans et al., 2015; Martin et al., 2016). In our study, however, face primes only led to significant response time effects for words in English (the L1) but not Spanish (the L2), despite similar accuracies for both languages on measures of face/language association (i.e., on the biographical questions, catch trials, and explicit language-association questions). Although sample size limitations would temper our confidence in the robustness of this finding, a tentative exploration of the possible reasons behind such an asymmetry is given below.

The fact that only native-language words were primed by faces aligns with previous psycholinguistic studies indicating that parallel activation is asymmetrically biased towards the more dominant language (e.g., de Groot, 1995). For our study, participants' non-native proficiency in Spanish may have reduced the accessibility of lexical items in that language, making it more difficult for face primes to induce a change in activation levels strong enough to manifest itself as a significant response time differences. Note, however, that such an account would conflict with findings that priming is stronger when targets are less accessible in the first place, e.g., due to low frequency (Yap, Tse, & Balota, 2009), or degraded visual presentation (Thomas, Neely, & O'Connor, 2010), a discrepancy worth further exploration.

A similar account for the lack of a face priming effect for Spanish words might involve a proficiency-based asymmetry in language-switching costs. Previous research has shown that bilinguals perform worse in language-switching tasks when switching from their weaker to their stronger language than when switching in the reverse direction (Costa & Santesteban, 2004). This is attributed to a greater magnitude of inhibition for the L1, which is more highly activated by default and thus requires stronger suppression to allow a bilingual to use only the weaker L2. For the current study, it is possible that seeing Spanish faces induced participants to strongly inhibit their native English lexicon in preparation for L2 Spanish, leading to slower response times in the incongruent condition for English words. Meanwhile, seeing a face associated with English would not lead to a strong inhibition of Spanish, as this was a weaker language that would not need to be suppressed as strongly.

An alternative explanation for the asymmetry in face priming effects across languages might invoke the higher level of cognitive effort required to process an L2. Processing a non-native language has been shown to impose larger demands on working memory (Kroll, Bobb, Misra, & Guo, 2002), suggesting one explanation for our results: that proficiency differences between English and Spanish led to a reduced ability to integrate information from visual stimuli in the environment during online processing of Spanish due to the higher cognitive demands involved. Meanwhile, processing English words would not impose such cognitive demands, allowing for priming from face stimuli to occur.

How could the different models of bilingual language processing account for our results? Dijkstra and van Heuven's Bilingual Interactive Activation+ model (2002) would categorize the face stimuli as elements of non-linguistic context that affect a broader task schema system above the level of word identification. As such, the bilingual's task schema system would lower the decision threshold for word selection for the face-congruent language while raising the threshold for the face-incongruent language. However, if contextual factors had no direct effect on the activation level of individual lexical items within a language but instead only affected a more general, language-neutral process of task control, then both English and Spanish words should have shown a clear congruency effect. One possible explanation may lie in some limitation on identifying and utilizing relevant contextual information presented on a short time scale during processing of the more cognitively-demanding L2; after all, our experiment presented the face stimuli only for 1000ms immediately before the target words. This contrasts with studies which either introduced relevant contextual manipulations well before the task (e.g., Elston-Güttler et al., 2005); co-presented faces and test words (e.g., Molnar et al., 2015; Woumans et al., 2015; Martin et al., 2016); or maintained a single manipulation throughout the task (e.g., Dijkstra, Van Jaarsveld, & Ten Brinke, 1998).

Under the Adaptive Control model (Green & Abutalebi, 2013), language-relevant face information would be picked up by a cognitive process of "salient cue detection," which would trigger processes of "task disengagement" (of the incongruent language) and "task engagement" (of the congruent language). This would alter the parameters in a neural network to facilitate access to the target language. For the current study, the reason that only English words were sensitive to this effect may lie in some difference in the neural network's representations of more vs. less dominant

[1]In exploring the possibility of a speed-accuracy tradeoff, a by-participants mixed-design ANOVA on arcsine-transformed error rates (compensating for the high number of participants with zero errors) found no significant effects for language, $F(1, 16)=1.77$, $p=.187$, congruence, $F(1, 16)=0.15$, $p=.698$, or their interaction, $F(1, 16)=3.04$, $p=.086$. However, we note the possibility of a floor effect given the low error rates overall (35 errors in 1,440 trials).

languages, making more highly-accessible nodes in the network more susceptible to task disengagement. Alternately, it is possible that the more cognitively-demanding nature of the L2 led to difficulties in implementing salient cue detection for identity-specific information from faces.

Under the BIMOLA model, knowledge about the person shown in the face prime would provide "global language information," which, under Léwy and Grosjean's (2008) definition, captures the base language of the interlocutor. This would pre-activate lexical items within the subsystem for the target language, facilitating their recognition in a language decision task. To account for the language asymmetries in this study's findings, the BIMOLA approach might posit that participants had not yet formed a full, separate subnetwork for their L2 Spanish. Rather, their Spanish lexicon would be based on unstable, underdeveloped networks, resulting in weaker cross-activation within and between different levels of language. This would lower the capacity that global language information tied to the face stimuli would have for directly pre-activating Spanish lexical items. Meanwhile, comprehension of words in (native) English would be based on a robust network that would allow for full top-down effects from global language information.

The BLINCS model involves increases in the resting activation of semantic representations for items that the visual-input module indicates are currently visible (Shook & Marian, 2013). For our study, the face stimuli might be assumed to activate certain nodes corresponding to the language congruent to the person; this activation would feed into other nodes for that language at the ortho-lexical and phono-lexical levels, facilitating the selection of lexical items appropriate to the linguistic identity of the person being seen. Meanwhile, activation would be shifted away from the incongruent language, leading to potential inhibition effects. The observed language asymmetry in priming effects may be attributed to a less developed network for Spanish in our participants, who may not have had enough experience with their L2 to have built up robust connections between certain kinds of visual information and Spanish-specific nodes. However, the BLINCS model to date only links visual input to the phonological and semantic levels, offering no obvious mechanism for how non-linguistic visual information could trigger activation specifically for words in one language.

Future studies along this line of inquiry could further explore the link between face perception and bilingual language activation by systematically testing learners with differing levels of L2 proficiency. Other studies might manipulate the strength of the priming effect, e.g., by changing the amount of time during which participants form face-language associations or by altering stimulus presentation so as to make it more/less similar to the kind of language that often co-occurs with faces (for instance, by using auditorily rather than visually presented words). Another approach might disentangle proficiency from behavioral ecology of language use, i.e., the community context in which a multilingual speaker uses their different languages (Green, 2011). For the participants in this study, this would refer to the fact that they lived in an English-speaking area and used English as their primary means of communication at the time of data collection, meaning that they would have few opportunities to establish interlocutor-language associations for Spanish words than proficiency-matched speakers in a dual language environment who would frequently have to switch between their two languages. This difference in experience linking faces to their corresponding languages may in turn affect sensitivity to face primes over and above any effects from sheer proficiency. Other studies could also test whether a face can prime for particular words associated with that person (rather than just at a "global," language-wide level), and for lower-level features in phonology and orthography. Neuroimaging techniques could be applied to evaluate face-priming effects more directly, thus avoiding the drawbacks of behavioral methodologies (e.g., reactivity). Additionally, studies on prosopagnosic (i.e., 'face-blind') bilinguals may evaluate whether face-priming can occur at a subconscious level. Further experiments might also test for correlations between the magnitude of face-priming effects and general measures of cognitive ability such as working memory. With the growing proportion of bilinguals in the world, it would only become more important to understand how their ability to manage two distinct languages can be affected by the environment around them.

Acknowledgments

The author would like to thank John Williams for his assistance with this project, which was completed at the University of Cambridge and supported by the Gates Cambridge Trust. Any mistakes are the author's own.

References

Bauer, R. (1984). Autonomic recognition of names and faces in prosopagnosia: a neuropsychological application of the Guilty Knowledge Test. *Neuropsychologia, 22*, 457–69.

Bijeljac–Babic, R., Biardeau, A., & Grainger, J. (1997). Masked orthographic priming in bilingual word recognition. *Memory and Cognition, 25*, 447–457.

Bruce, V. & Young, A. (1986). Understanding face recognition. *British Journal of Psychology, 77*, 305–27.

Canseco–Gonzalez, E., Brehm, L., Brick, C., Brown–Schmidt, S., Fischer, K., & Wagner, K. (2010). Carpet or Carcel: the effect of age of acquisition and language mode on bilingual lexical access. *Language and Cognitive Processes, 25*, 669–705.

Casaponsa, A., Carreiras, M., & Duñabeitia, J. (2014) Discriminating languages in bilingual contexts: The impact of orthographic markedness. *Frontiers in Psychology, 5*, 424.

Costa, A. & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language, 50*, 491–511.

de Groot, A. (1995). Determinants of bilingual lexicosemantic organization. *Computer Assisted Language Learning, 8*, 151–180.

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010) How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language, 62*(3), 284-301.

Dijkstra, T. & Van Heuven, W. (2002). The architecture of the bilingual word recognition system: from identification to decision. *Bilingualism: Language and Cognition, 5*, 175– 197.

Dijkstra, A., Van Jaarsveld, H., & Ten Brinke, S. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language and Cognition, 1*, 51–66.

Elston-Güttler, K., Gunter, T., & Kotz, S. (2005). Zooming into L2: Global language context and adjustment affect processing of interlingual homographs in sentences. *Cognitive Brain Research, 25*, 57–70.

Elston-Güttler, K. & Williams, J. (2008). First language polysemy affects second language meaning interpretation: evidence for activation of first language concepts during second language reading. *Second Language Research, 24*, 167–187.

Garrett, M. (1980). Levels of processing in sentence production. In B. Butterworth (ed.), *Language Production*. New York: Academic Press.

Grainger, J., O'Regan, J., Jacobs, A., & Segui, J. (1989). On the role of competing word units in visual word recognition: The neighborhood frequency effect. *Perception and Psychophysics, 45,* 189–195.

Green, D. (2011). Language control in different contexts: The behavioral ecology of bilingual speakers. *Frontiers in Psychology, 2*, 1–4.

Green, D. & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology, 25*, 515–530.

Jared, D. & Kroll, J. (2001). Do bilinguals activate phonological representations in one or both of their languages when naming words? *Journal of Memory and Language, 44*, 2–31.

Kanwisher, N. (2000). Domain specificity in face perception. *Nature Neuroscience, 3*, 759–63.

Léwy, N. & Grosjean, F. (2008). The Léwy and Grosjean BIMOLA model. In F. Grosjean (Ed.), *Studying bilinguals*. Oxford, UK: Oxford University Press.

Loebell, H. & Bock, K. (2003). Structural priming across languages. *Linguistics, 41*, 791–824.

Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross–Linguistic Easy–Access Resource for Phonological and Orthographic Neighborhood Densities. *PLoS ONE, 7*, 43230.

Marian, V., Blumenfeld, H., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP–Q): Assessing Language Profiles in Bilinguals and Multilinguals. *Journal of Speech, Language, & Hearing Research, 50*, 940–967.

Martin, C., Dering, B., Thomas, E., & Thierry, G. (2009). Brain potentials reveal semantic priming in both the 'active' and the 'non–attended' language in early bilinguals. *NeuroImage, 47*, 326–333.

McGurk H. & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*, 746–8.

Molnar, M., Ibáñez-Molina, A., & Carreiras, M. (2015). Interlocutor identity affects language activation in bilinguals. *Journal of Memory and Language, 81*, 91–104.

Poulisse, N. & Bongaerts, T. (1994). First language use in second language production. *Applied Linguistics, 15*, 36–57.

Sheffert, S. & Olson, E. (2004). Audiovisual speech facilitates voice learning. *Perception and Psychophysics, 66*, 352–362

Shook, A. & Marian, V. (2013). The bilingual language interaction network for comprehension of speech. *Bilingualism: Language and Cognition, 16*, 304–324.

Smith, R. & Hawkins, S. (2012). Production and perception of speaker–specific phonetic detail at word boundaries. *Journal of Phonetics, 40*, 213–233.

Tanenhaus, M., Spivey–Knowlton, M., Eberhard, K., & Sedivy, J. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*, 1632–1634.

Thomas, M., Neely, J., & O'Connor, P. (2010). When word identification gets tough, retrospective semantic processing comes to the rescue. *Journal of Memory and Language, 66*, 623–643.

Tsujimoto, S., Yokoyama, T., Noguchi, Y., Kita, S., & Kakigi, R. (2011). Modulation of neuromagnetic responses to face stimuli by preceding biographical information. *European Journal of Neuroscience, 34*, 2043–2053.

van Wassenhove, V., Grant, K., & Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proceedings of the National Academy of Science, 102*, 1181–86.

Vitevitch, M. & Luce, P. (1998). When words compete: Levels of processing in spoken word perception. *Psychological Science, 9*, 325–329.

Von Kriegstein, K., Kleinschmidt, A., Sterzer, P., & Giraud, A. (2005). Interaction of face and voice areas during speaker recognition. *Journal of Cognitive Neuroscience, 17*, 367–376

Woumans, E., Martin, C. D., Bulcke, C. V., Van Assche, E., Costa, A., Hartsuiker, R. J., & Duyck, W. (2015). Can faces prime a language? *Psychological Science*, *26*, 1343–1352.

Yap, M., Tse, C., & Balota, D. (2009). Individual differences in the joint effects of semantic priming and word frequency revealed by RT distributional analyses: The role of lexical integrity. *Journal of Memory and Language, 61*, 303–325.

Zhang, S., Morris, M., Cheng, C., & Yap, A. (2013). Heritage–culture images disrupt immigrants' second–language processing through triggering first–language interference. *Proceedings of the National Academy of Sciences, 110*, 11272–77.