Psycholinguistic, cognitive, and neural implications of bimodal bilingualism*

(Received: December 21, 2014; final revision received: February 18, 2015; accepted: February 21, 2015; first published online 23 April 2015)

Bilingualism provides a valuable tool for understanding language processing and its underlying neurocognitive mechanisms. However, the vast majority of bilingual research has involved the study of spoken language users – ‘unimodal’ bilinguals. ‘Bimodal’ bilinguals have acquired a spoken and a signed language, and they offer a unique window into the architecture of the language processing system because distinct sensory-motor systems for comprehension and production. Differences between unimodal and bimodal bilinguals have implications for how the brain is organized to control, process, and represent two languages. Evidence from code-blending (simultaneous production of a word and a sign) indicates that the production system can access two lexical representations without cost, and the comprehension system must be able to simultaneously integrate lexical information from two languages. Further, evidence of cross-language activation in bimodal bilinguals indicates the necessity of links between languages at the lexical or semantic level. Finally, the bimodal bilingual brain differs from the unimodal bilingual brain with respect to the degree and extent of neural overlap for the two languages, with less overlap for bimodal bilinguals.

Keywords: bimodal bilingualism, language mixing, cross-language activation, language control, cognitive control


Karen Emmorey Laboratory for Language and Cognitive Neuroscience 6495 Alvarado Road suite 200 San Diego CA 92120 USA
kemmorey@mail.sdsu.edu

Karen Emmorey
School of Speech, Language and Hearing Sciences, San Diego State University
Marcel R. Giezen
San Diego State University
Tamar H. Gollan
University of California San Diego, Department of Psychiatry

* This research was supported by NIH grant HD047736 to Karen Emmorey and SDSU, Rubicon grant 446-10-022 from the Netherlands Organization for Scientific Research to Marcel Giezen, and NIH grant DC011492 to Tamar Gollan. We would like to thank Jennifer Petrich for her tremendous help in running many of the studies discussed in this paper, and all our research participants.

1 Jemina Napier recently suggested the acronym People from Deaf Families (PDF) to replace Coda (typically applied to adults) and Koda (typically applied to children) (http://lifeinlincs.wordpress.com/2014/10/28/whats-in-a-name/).
because this is the language of schooling and commerce in the broader sociolinguistic context (e.g., there are no communities with hearing monolingual signers).

**How the production system manages simultaneous language output**

Investigations of language mixing by bimodal bilingual adults (Bishop, 2010; Emmorey, Borinstein, Thompson & Gollan, 2008) and children (Lillo-Martin, de Quadros, Chen Pichler & Fieldsteel, 2014; Petitto, Katerelos, Levy, Gauna, Tétreault & Ferraro, 2001) reveal a strong preference for code-blending over code-switching and indicate that the majority of code-blends involve semantically congruent forms (translation equivalents).

Most studies only consider an utterance to contain a code-blend if the speech is voiced or whispered (Petroj, Guerrera & Davison, 2014), rather than silently mouthed (work by Van den Bogaerde & Baker (2005) and Baker & Van den Bogaerde (2008) is an exception). Code-blending is distinct from ‘simultaneous communication’, also known as SimCom, Total Communication, or Sign Supported Speech. Simultaneous communication is usually produced with mixed audiences of deaf and hearing people, requires continuous simultaneous speaking and signing, often includes invented signs for grammatical morphemes, and the syntactic structure always follows the spoken language. Below are examples of natural code-blends produced by adult ASL–English bilinguals in which either English is the matrix (or ‘base’) language (1, 2) or ASL is the matrix language (3, 4), providing the syntactic structure for the utterance. By convention, signs are written in capitals as English glosses (the nearest translation equivalent; hyphens indicate a multiword gloss), and brackets indicate the co-occurring speech.

1. She didn’t [even sign] until fourteen.

   **SIGN**

   (Emmorey, et al., 2008, p. 48)

2. I’m pretty [sure] they were [hearing].

   **SURE** **HEARING**

   (Bishop, 2010, p. 225)

3. [Wonder what do]

   **WONDER** **DO-DO**

   Translation: “(He) wonders what to do”

   (Emmorey, et al., 2008, p. 50)

4. [That hearing way]

   **THAT** **HEARING** **WAY**

   Translation: “That is the way hearing people tend to be”

   (Bishop, 2010, p. 226)

Emmorey, Borinstein et al. (2008) argued that the overwhelming preference for code-blending over code-switching indicates that a) the locus of lexical selection for bilingual language production is relatively late – a single lexical representation need not be selected for production at the conceptual or lemma level (otherwise code-blending would not occur). By implication, this result suggests that dual lexical retrieval is less costly than language inhibition, which is assumed to be required for code-switching (e.g., Green, 1998). Evidence for inhibition in unimodal bilinguals is that code-switching costs are asymmetric, such that switching into the easier (dominant) language is more costly than switching into the more difficult (non-dominant) language (Meuter & Allport, 1999; but see Verhoeof, Roelofs & Chwilla, 2009). In bimodal bilinguals asymmetries are found in the pattern of observed code-blend types. All studies (that we are aware of) report examples of single-sign code-blends (examples 1–2), but none report single-word code-blends in which the sign language is the matrix language of the utterance, and a single spoken word is produced, e.g., saying only the word “hearing” in example 4. Emmorey, Borinstein et al. (2008) suggest that bimodal bilinguals do not produce single words while signing because the dominant spoken language has been strongly inhibited when sign language is selected as the matrix language; continuous code-blending (example 3–4) can occur because speech has been completely released from inhibition. Single-sign code-blends (examples 1–2) can occur because the non-dominant sign language is not strongly inhibited.

Further evidence that English must be strongly suppressed in ASL-dominant contexts comes from the study of bimodal bilingual children who frequently whisper spoken words when signing with deaf interlocutors for whom whispered speech is inaccessible (Petroj et al., 2014). This phenomenon has not been studied for adult fluent bimodal bilinguals, but adult ASL learners are known to frequently whisper when signing, thus getting around the typical ‘voice off’ instruction from their deaf ASL teachers. Petroj et al. (2014) hypothesize that whispering in ASL contexts (e.g., when ASL is the matrix language) eases the pressure to fully suppress English, at least at the lexical level (the syntactic structure of English can conflict with ASL and must be inhibited when signing). Although whispered signing occurs in secretive contexts (e.g., producing reduced signs low and out of view of others), Petroj et al. (2014) did not report the use of whispered signs while speaking, which suggests that ASL may be easier to suppress than English and provides further support for asymmetric inhibition effects for bimodal bilinguals (although it should be noted that these authors were not coding or looking for whispered ASL).

Data from natural code-blending suggests that dual lexical retrieval is less costly than inhibiting one language, but is it ‘cost free’? Emmorey, Petrich and Gollan (2012) investigated this question by asking ASL–English bilinguals to name pictures in ASL alone, English alone, or with an ASL–English code-blend. Code-blending did not slow lexical retrieval for ASL: response times for ASL produced alone were not different than for ASL in
a code-blend. Kaufmann (2014) recently replicated this result with German and German Sign Language (DGS) for unbalanced bilinguals. Further, Emmorey et al. (2012) found that code-blending actually facilitated retrieval of low frequency ASL signs. Facilitation may occur via translation priming (e.g., Costa & Caramazza, 1999), such that retrieving the English word rendered the ASL sign more accessible. However, for English, code-blending delayed the onset of speech because bimodal bilinguals preferred to synchronize the onsets of words and signs (i.e., delaying speech onset until the hand reaches the target location of the sign). Similar synchronization is observed for co-speech gestures (McNeill, 1992). The overall pattern suggests that code-blend production does not incur a lexical processing cost despite the fact that two lexical items must be retrieved simultaneously. This result indicates that the bilingual language production system is capable of turning off inhibition between languages in order to allow the simultaneous production of translation equivalents.

However, there is some evidence for an indirect form of lexical competition between signs and words for bimodal bilinguals. Pyers, Gollan and Emmorey (2009) found that ASL–English bilinguals, like Spanish–English bilinguals, experienced more tip-of-the-tongue (TOT) states compared to monolingual English speakers when naming pictures in English. Increased rates of TOTs for unimodal bilinguals could be caused by a) competition between languages at the phonological level, b) competition between languages at a lexical level, or c) less frequent language use because bilinguals divide their language use between two languages (Gollan, Montoya, Cera & Sandoval, 2008; Gollan, Slattery, Goldenberg, Van Assche, Duyck & Rayner, 2011). The finding that bimodal bilinguals exhibited more TOTs than monolinguals rules out phonological blocking as the primary source of increased TOTs for bilinguals because speech and signs do not overlap in phonology. Although Pyers et al. (2009) argued for a frequency-of-use explanation over lexical-semantic competition, a recent study by Pyers, Emmorey and Gollan (2014) suggests that production of ASL signs during a TOT can impair the ability to retrieve English words. Pyers et al. (2014) found that bimodal bilinguals spontaneously produced ASL signs when attempting to retrieve English words in a picture-naming task, but fewer TOTs were resolved when they did so. A follow-up study manipulated whether bilinguals were allowed to sign (and gesture) or were prevented from using their hands when naming pictures. Significantly more TOTs occurred when bilinguals were allowed to sign, which contrasts with the results from monolingual speakers who experience fewer TOTs when they are allowed to gesture (e.g., Rauscher, Krauss & Chen, 1996). In addition, there was a trend for fewer TOTs to be resolved when bilinguals produced a sign translation. These new findings suggest that lexical alternatives may compete for selection, leading to lexical-semantic blocking of English by the spontaneous production of ASL signs.

The Pyers et al. (2014) results are surprising given that simultaneous production of English words facilitated (rather than blocked) access to low frequency ASL signs in a code-blend (Emmorey et al., 2012) and that experimental presentation of English translations speeded ASL sign production in the picture–word interference (PWI) paradigm (Giezen & Emmorey, 2015; discussed in more detail below). Translation facilitation is also commonly found in the PWI paradigm for unimodal bilinguals (e.g., Costa & Caramazza, 1999). Interestingly, this particular puzzle is not unique to bimodal bilinguals because Gollan, Ferreira, Cera and Flett (2014) recently reported that translation primes increased, rather than decreased, TOT rates for unimodal bilinguals. To explain this result, they suggested that translation equivalents may facilitate partial access to target words which effectively turns full retrieval failures (i.e., “don’t know” responses) into a partially successful lexical retrieval (i.e., a TOT). Thus, for all bilinguals translation equivalents may facilitate access to target lexical items, with different apparent effects in different paradigms.

However, it remains a mystery why sign retrieval does not facilitate TOT resolution once bimodal bilinguals are stuck in a failed retrieval attempt. One speculative hypothesis is that there is an important distinction between internally and externally generated translation equivalents. Self-generated translations may suppress access to the target word, whereas externally provided translations may facilitate access of the target word. In fact, data from TOTs in monolinguals support this hypothesis. Retrieval of a semantically related word during a TOT state is associated with delayed resolution of TOTs (e.g., Burke, MacKay, Worthley & Wade, 1991), but experimental provision of semantically related cues facilitates correct retrievals (Meyer & Bock, 1992). Further research is needed to determine whether internally and externally generated translation equivalents yield the same pattern of interference and facilitation effects in bilinguals.

Although the majority of naturally produced code-blends and all of the experimentally produced code-blends in the studies described above involve translation equivalents, there are nonetheless relatively frequent examples of naturally occurring code-blends that do not involve simultaneous production of translation equivalent signs and words: 16% in the Emmorey, Borinstein et al. (2008) study and 26–56% of the utterances produced by three hearing bimodal bilingual children at age three in a longitudinal study by Baker and Van den Bogaerde (2008) (their ‘code-blended mixed’ category). Below are examples from child and adult data sets (all Codas) that illustrate a wide variety of types of naturally produced code-blends which contain simultaneously produced signs and words that are not translation equivalents.
Examples 5 and 6 involve simultaneous production of semantically-related lexical items in which either the word is more specific (Tweety vs. BIRD) or the sign is more specific (SHOOT vs. kill). Similarly, in Example 7, a more precise adverbial phrase (ten years ago) is simultaneously produced with a single adverbial sign (LONG-TIME-AGO), while in Example 8 the deictic pronoun this co-occurs with the sign CONNECTION that specifies the deictic referent. These examples have semantic parallels with co-speech gesture because speech can be more specific than a gesture (e.g., pointing upward while saying “God”), and a gesture can also provide more specific information than speech (e.g., saying “killed him” while producing a stabbing gesture). However, signs – unlike gestures — have lexical representations with semantic, morphosyntactic, and phonological specifications that are retrieved during language production, rather than spontaneously created on the fly. These code-blend examples illustrate that dual lexical selection is not constrained to translation equivalents and that the production system must be capable of simultaneously selecting lexical items that differ in semantic specificity and in syntactic category (as in 8).

Examples 9 and 10 provide evidence that the production system must be able to distribute distinct syntactic constituents across languages. In the ASL–English example 9, the adverbial phrase all of a sudden and the ASL verb LOOK-AT must be integrated to create a coherent utterance. Similarly, in Example 10 spoken Italian provides the indirect object (Snow white) while LIS provides the (post-verbal) subject (HUNTER), and the utterance is only interpretable by integrating the simultaneously produced constituents. To determine whether these constructions were indeed interpretable, Donati and Branchini (2013) conducted a comprehension task with three of their Coda participants (ages 6–8 years old) in which an adult Coda produced these types of non-linearized constituents, and the children were asked targeted comprehension questions. The authors reported that the children had no difficulty understanding these types of utterances.

Examples 11–12 illustrate the possibility of producing two distinct word orders in the same utterance. To assess the acceptability of these types of utterances, Donati and Branchini (2013) asked three bilingual children to repeat such code-blends (produced by an adult bilingual), along with other types of code-blends, as well as monolingual sentences. One child who was exposed to Signed Italian at his bilingual school changed the LIS word order to match the Italian word order for most of the code-blends. However, the two other children faithfully reproduced the majority (70%–80%) of the code-blended utterances, producing a distinct word order for each language. LIS–Italian bilinguals may be more likely to produce such code-blends because LIS is a head final language, while Italian is a head initial language. Examples like 11–12 are particularly interesting with respect to theories that posit shared bilingual lexical-syntactic representations in which lexical representations of translation equivalents are linked to the same combinatorial node that specifies syntactic structure. Hartsuiker, Pickering and Veltkamp (2004) argue for such shared syntactic representations in bilinguals based on their finding of syntactic priming for similar structures between languages. Verbs in LIS and Italian in the examples above link to distinct syntactic structures, and yet presumably both structures (combinatorial nodes) must be activated in order for each language to be simultaneously produced.

Zeshan (2014) also reports this phenomenon, which he terms ‘sign-speaking’ for hearing interpreters/teachers who simultaneously produce two different word orders in spoken Hindi and Indian Sign Language.
If comprehension and production share representational structure, then it might be possible to observe cross-linguistic syntactic priming between different word orders for bimodal bilinguals, in contrast to what has been found for unimodal bilinguals (e.g., Bernolet, Hartsuiker & Pickering, 2007).

Much more experimental and corpus-based research is needed to account for the patterns of code-blending illustrated in 5–12. Currently, there is no agreed upon terminology for categorizing such examples, and different authors have used different terms, e.g., mixed code-blending, independent blending, semantically non-equivalent code-blends. Without agreement, it may be difficult to determine the relative frequency of different code-blend types across data sets. It is also not clear whether these types of code-blends are more costly to produce or comprehend compared to code-blends that involve translation equivalents. Further, the constraints on such code-blending behavior are unknown, e.g., what causes these examples to be less frequent than code-blends involving translation equivalents and what would constitute an impossible (unacceptable) code-blend? We suggest that bimodal bilingual code-blends provide a rich, new source of data for investigating the syntactic and semantic constraints on intra-sentential language mixing and for the architecture of the language system (e.g., Lillo-Martin, Quadros, Koulidobrova & Chen Pichler, 2012).

In sum, the data thus far indicate that the bilingual language production system must allow for a late locus of lexical selection and be capable of turning off inhibition between languages in order to produce speech-sign code-blends. Evidence for inhibition effects in bimodal bilinguals comes from asymmetric patterns of code-blending in adults (single-sign vs. single-word code-blends) and whispering in bilingual children (whispered words vs. whispered signs). The preference for code-blending over code-switching implies that dual lexical retrieval is less costly than inhibiting one language, and evidence from picture-naming tasks suggests that simultaneous production of translation equivalents does not incur a processing cost. Indirect lexical retrieval costs are observed when bimodal bilinguals are in a TOT state and produce a sign translation, but the cause of this effect requires further investigation. Finally, the code-blending corpus data indicate that the production system must be at least capable of simultaneously selecting lexical items that are not translation equivalents, as well as simultaneously producing distinct syntactic structures, although these code-blend types are dispreferred.

Turning languages ‘on’ and ‘off’: Are there switch-costs for code-blending?

Although production of a single code-blend (i.e., the simultaneous production of a word and a sign) does not appear to incur a processing cost (Emmorey et al., 2012; Kaufmann, 2014), bimodal bilinguals must nonetheless switch into and out of a code-blend, as illustrated in Examples 1 and 2 above. This raises the question of whether there are switch costs for code-blending that parallel the switch-costs observed for unimodal bilinguals. When unimodal bilinguals switch between languages, they must simultaneously turn one language “off” (i.e., inhibit that language) and turn “on” the other language, releasing that language from inhibition (Green, 1998). The existence of code-blends allows us to examine these processes separately. For example, switching from speaking (English alone) into an ASL–English code-blend involves just turning ASL ‘on’ – English is produced in the code-blend and therefore is not inhibited. Switching out of an ASL–English code-blend into English alone involves only turning ASL ‘off’.

Emmorey, Petrich and Gollan (2014) investigated the nature of switch costs associated with code-blending by requiring ASL–English bilinguals to name pictures in two conditions: switching between speaking and code-blending or between signing and code-blending (two sequentially presented pictures were named in each language condition: English, ASL, and Code-blend). As in previous studies, response time (RTs) for English were measured from voice onset, and RTs for ASL were measured from a manual key-release. RTs for code-blend switch trials were not longer than for stay (repeat) trials for either ASL or English. This pattern of results indicates that ‘turning an additional language on’ is not costly; or, put differently, that releasing a language from inhibition does not incur a processing cost. This finding is consistent with evidence that a non-target language is easily and automatically activated even when bilinguals speak in just one language (e.g., Kroll, Bobb, Misra & Guo, 2008). In contrast, RTs for the switches out of a code-blend (i.e., switches into English or ASL alone out of a code-blend on the previous trial) were longer than for stay trials. This suggests that ‘turning a language off’ (inhibiting a language) is costly and suggests that what makes language switching difficult is applying inhibition to the recently selected language. For bimodal bilinguals, releasing an inhibited language appears to be ‘cost free’, but applying inhibition to a non-target language is not.

The study by Emmorey, Petrich and Gollan (2014) also contained a condition that was parallel to code-switching in unimodal bilinguals in which ASL–English bilinguals switched between ASL and English. Unlike unimodal code-switching, however, the articulators for ASL and English do not compete for production, and thus it is possible that switch costs might not be observed for bimodal bilinguals. However, the results revealed a significant switch cost for English (longer RTs for switch than stay trials), but surprisingly there was no switch cost for ASL. In fact, RTs for ASL on switch
trials were slightly, but significantly, faster than on stay trials, a very surprising result. Emmorey and colleagues conducted a follow-up study to examine whether this ‘switch advantage’ was a lexical effect specific to bilingual language production, or a modality effect specific to speaking and signing. To this end, hearing non-signers named Arabic digits (1–5) with English, with one-handed gestures (holding up the relevant number of fingers), or switched back and forth between English and gesture. The results of this experiment revealed a parallel pattern, with significant switch costs for speech and a switch advantage for gestures. The explanation for this finding is not completely clear, but one speculative hypothesis is that a lack of competition between articulators may allow for the anticipatory preparation of a manual response. That is, a manual response (either a sign or gesture) may be partially prepared on a preceding speech trial. In contrast, speech preparation may be blocked by the covert or overt mouthing of the English word produced on the preceding manual trial.

In sum, the language mixing possibilities for bimodal bilinguals are more complex than those for unimodal bilinguals, and these differences allow us to examine different aspects of language control (see also Kaufmann, 2014). As with unimodal bilinguals, we find evidence that inhibiting a language incurs a processing cost, as when bimodal bilinguals switch out of a code-blend into a single language (i.e., ‘turning a language off’) or when switching into the dominant spoken language from ASL alone. However, the evidence also indicates that switching from a single language into a code-blend is not costly (i.e., ‘turning a language on’), and that switching from speech to either sign or gesture does not incur a switch cost (in fact, a switch advantage is observed).

**How the comprehension system manages simultaneous language input**

In contrast to code-blend production, much less is known about the psycholinguistic processes involved in comprehending code-blends. Emmorey et al. (2012) asked ASL–English bilinguals to perform a semantic categorization task (“is it edible?”) for ASL signs, audio-visually presented English words, and code-blends (all produced by the same Coda model). Response times were faster in the code-blend condition than for either language alone. Thus, code-blending appears to speed lexical access during language comprehension for both the dominant language (English) and the non-dominant language (ASL). This result is somewhat analogous to findings from visual word recognition studies in unimodal bilinguals that reported facilitation when both languages were relevant to the task. Specifically, Dijkstra, Van Jaarsveld and Ten Brinke (1998) found that when Dutch–English bilinguals were asked to make lexical decisions to strings that could be words in either Dutch or English, response times were facilitated by interlingual homographs (e.g., *brand*, which exists in both languages but has different meanings) compared to words that exist only in one language. Based on frequency data, Dijkstra et al. (1998) argued that bilinguals made their decision based on whichever lexical representation became available first. Similarly, it is possible that bimodal bilinguals made their semantic decision based on whichever lexical item (the sign or the word) was recognized first during the code-blend.

Emmorey et al. (2012) contrasted this ‘race’ model of code-blend comprehension with a second account that involves the integration of lexical material from both languages, which can facilitate lexical recognition and semantic processing. Specifically, integration could occur at the phonological and/or semantic level. At the form level, lexical cohort information from both languages could be combined to constrain lexical recognition for each language. For example, the English onset “ap” activates several possible words (*apple, aptitude, apathy, appliqué, . . .*), but only *apple* is consistent with the onset of the ASL sign *APPLE*. Similarly, the initial handshape and location of the ASL sign *APPLE* activates several possible ASL signs (*APPLE, COOL/NEAT, RUBBER, . . .*), but only *APPLE* is consistent with the onset of the English word *apple*. In addition, integration could occur at the semantic level, with both languages providing congruent and confirmatory information which facilitates the semantic decision in a code-blend. The integration account predicts that code-blend facilitation should always be observed for both languages, whereas the race explanation predicts that facilitation should only occur for one language in a code-blend (i.e., the language that ‘loses’ the race). To tease apart these two explanations, Emmorey et al. (2012) determined which language should ‘win’ the race based on RTs for the same lexical item presented in English and ASL alone. Code-blend facilitation was found even for those items that had ‘won’ the race (i.e., had faster RTs in the single language condition), which is predicted only by the integration account. Thus, code-blend facilitation does not appear to be solely due to recognition of one language before the other. Rather, the results suggest that lexical information from each language can be integrated either at the form level by reducing the uniqueness point for signs and words or at the semantic level because translation equivalents converge on the same semantic concept.

Weisberg, McCullough and Emmorey (under review) recently replicated code-blend facilitation effects for ASL–English bilinguals in an fMRI study using the same semantic categorization task. Weisberg et al. reported that comprehension of code-blends recruited a neural network that resembled a combination of brain regions active for each language alone. Unlike comprehension of auditory code-switches in unimodal...
bilinguals, code-blend comprehension did not recruit cognitive control regions (e.g., the anterior cingulate; Abutalebi, Brambati, Annoni, Moro, Cappa & Perani, 2007). This result is consistent with the behavioral results and supports the hypothesis that there is little or no competition between languages when translation equivalents are perceived simultaneously. Rather, code-blend facilitation effects were reflected by reduced neural activity in language-relevant sensory cortices: auditory regions when code-blends were compared to English and visual regions when compared to ASL. In addition, prefrontal regions associated with semantic processing demonstrated reduced activity during code-blend comprehension compared to ASL alone, but not compared to English alone. This pattern is consistent with stronger facilitation effects for ASL than for English in the behavioral measures in both the Emmorey et al. (2012) and the Weisberg et al. studies. Overall, the Weisberg et al. results suggest that code-blend integration effects can be observed at both the form level (with reduced neural activity in sensory cortices) and at the semantic level (with reduced neural activity in left frontal regions for ASL).

What is completely unknown at this point, however, is how the comprehension system might deal with simultaneous input of the sort illustrated in Examples 5–12 where the simultaneously perceived signs and words are not translation equivalents. It is quite possible that such code-blends incur a processing cost related to integrating different semantic and/or syntactic representations. On the other hand, it is possible that under some conditions, this type of code-blend could facilitate processing in the same way that adding a disambiguating gesture facilitates processing of an ambiguous spoken word (e.g., Holle, Gunter, Rüschemeyer, Hennenlotter & Iacoboni, 2008). Exploring how bilinguals comprehend such ‘incongruous’ code-blends can provide novel insight into the architecture of the bilingual language system by identifying how different levels of representation are integrated without competition at a sensory level.

**Does language co-activation during comprehension require perceptual overlap?**

When unimodal bilinguals listen to speech, they simultaneously access information in both languages; that is, lexical access is largely language non-selective (for discussion, see e.g., Kroll, Bogulski & McClain, 2012; Kroll & Dussias, 2013; Wu & Thierry, 2010). Cross-language activation is highly automatic and is even observed in tasks that do not require explicit language processing (Wu, Cristino, Leek & Thierry, 2013). One likely source of co-activation during spoken word comprehension in unimodal bilinguals is sub-lexical activation in phonological input from the two languages (e.g., Dijkstra & Van Heuven, 2002; Shook & Marian, 2013). Indeed, many studies assess cross-language activation through phonological overlap between word pairs in different languages (e.g., Ju & Luce, 2004; Marian & Spivey, 2003a, b; Weber & Cutler, 2004). This raises the question whether bilinguals, whose languages involve non-overlapping phonological systems, exhibit cross-language activation during language comprehension, or whether such automatic cross-language interaction is modality-specific and only observed when the two language systems overlap in phonological structure.

Morford, Wilkinson, Villwock, Piñar and Kroll (2011) addressed this question by asking deaf ASL–English bilinguals to judge written English word pairs for semantic relatedness. A subset of the English word pairs had ASL translation equivalents that overlapped in sign phonology. They found that semantically-related word pairs (e.g., *bird* and *duck*) were judged more quickly (i.e., the “yes” responses were faster) when their ASL sign translation equivalents overlapped in sign phonology (see Figure 1A). In contrast, semantically unrelated word pairs (e.g., *movie* and *paper*) were judged more slowly (i.e., the “no” responses were slower) when their ASL sign translation equivalents overlapped in sign phonology (see Figure 1B). This pattern of results provides clear evidence of cross-language activation between a signed language and (the written form of) a spoken language, suggesting that perceptual overlap between phonological systems is not required for language co-activation to occur (see also Kubus, Villwock, Morford & Rathmann, 2014; Ormel, Hermans, Knoors & Verhoeven, 2012 for related findings in DGS (German Sign Language) and NGT, respectively).

However, it is possible that deaf bimodal bilinguals develop particularly strong associations between lexical signs and lexical orthographic representations during reading acquisition in bilingual education programs (e.g., Hermans, Knoors, Ormel & Verhoeven, 2008; Hermans, Ormel, Knoors & Verhoeven, 2008). Hearing bimodal bilinguals are unlikely to have established similar associations with written word forms because of their full access to spoken language phonology. Shook and Marian (2012) investigated co-activation of signs during spoken word recognition in a visual world paradigm eye-tracking study with hearing ASL–English bilinguals. Participants were presented with spoken words while looking at displays with four images, while their eye movements were monitored. Critical displays included an image of the target object as well as an image of the target object as well as an image of a cross-language ASL phonological competitor, e.g., an image of *paper* in a trial with the English target word *cheese*. The ASL signs *PAPER* and *CHEESE* share a similar phonological relationship as the signs *PAPER* and *MOVIE* in Figure 1B. ASL–English bilinguals...
Karen Emmorey, Marcel R. Giezen and Tamar H. Gollan

Figure 1. Examples of phonologically-related pairs of ASL signs from Morford et al. (2011) used in the A) semantically-related condition and B) semantically-unrelated condition.

...looked longer at the picture of the cross-language competitor than at the two unrelated pictures, indicating co-activation of ASL signs during English auditory word recognition. These findings strongly suggest that cross-language activation can also occur through connections between signs and spoken word forms. Furthermore, the Shook and Marian (2012) results show that the non-dominant language (ASL for hearing bimodal bilinguals) is activated during comprehension of the dominant language (English).

Recently, Morford, Kroll, Piñar and Wilkinson (2014) compared language co-activation in hearing ASL–English bilinguals who learned ASL as a late second language and deaf bilinguals using the same paradigm as Morford et al. (2011). Both groups of bilinguals showed an interference effect for semantically unrelated English word pairs (“no” responses) with phonologically-related ASL translations (as in Figure 1B). However, hearing bilinguals did not show a facilitation effect for semantically-related word pairs (“yes” responses) with phonologically-related ASL translations (as in Figure 1A). This weaker co-activation pattern for the hearing bilinguals could arise because ASL was their L2; it is well established that cross-language activation of the L2 when comprehending the L1 is weaker and strongly dependent upon L2 proficiency (e.g., Van Hell & Dijkstra, 2002; Van Hell & Tanner, 2012). However, Morford et al. (2014) suggest a second possible explanation, namely that deaf bilinguals may co-activate ASL signs directly through English orthography, whereas hearing bilinguals may first activate English phonology before co-activating ASL signs. Such an indirect route may have reduced language co-activation effects for ASL for hearing bilinguals.

Thus far, all studies have indicated activation of a signed language during written or auditory comprehension of a spoken language. Is there any evidence for activation of a spoken language during sign language comprehension? To our knowledge, there are only two preliminary reports indicating activation of a spoken language when comprehending a sign language, a sign-picture verification experiment with hearing NGT–Dutch bilinguals (Van Hell, Ormel, Van der Loop & Hermans, 2009) and an ERP signed sentence processing study with deaf DGS–German bilinguals (Hosemann, Altvater-Mackensen, Hermann & Mani, 2013). In the Van Hell et al. (2009) study, verification times were affected by whether the spoken translation equivalents of the sign and the picture name were phonologically- and orthographically-related in Dutch. Hosemann et al. (2013) recorded ERPs from deaf native signers of DGS while viewing signed sentences containing a prime and target sign whose German translations rhymed and were also...
orthographically similar, e.g., *LAST WEEK MY HOUSE, THERE MOUSE HIDE*. Their preliminary EEG analysis indicated a priming effect (weaker N400) for target signs that were form-related in spoken German. These preliminary findings provide evidence for activation of a spoken language when comprehending a sign language, and together with the above results indicate the existence of bidirectional co-activation in bimodal bilinguals, even when the spoken language is less dominant, as for the deaf signers in the Hosemann et al. (2013) study.

In sum, the absence of perceptual overlap through shared phonological systems between languages does not prevent co-activation of the non-target language during word or sign comprehension in either deaf or hearing bimodal bilinguals. Because cross-language activation cannot occur at the phonological level, the presence of language co-activation in bimodal bilinguals provides unequivocal evidence for the contribution of lexical connections and/or shared semantics to cross-linguistic interaction in bilingual comprehension.

**Implications of bimodal bilingualism for the architecture of the bilingual lexicon**

The finding that language co-activation can readily occur between two languages with fully distinct phonological systems in different modalities has important theoretical implications for the architecture of the bilingual lexicon. Most models of bilingual word recognition such as the Bilingual Interactive Activation model (Dijkstra & Van Heuven, 2002), the Bilingual Model of Lexical Access (Grosjean, 1988, 2008) or the Bilingual Language Interaction Network for Comprehension of Speech (Shook & Marian, 2013), assume that the primary source of cross-language activation in unimodal bilinguals is at a shared phonological level (consisting of speech sounds and/or phonological features). To account for co-activation in bimodal bilinguals these models need to be adjusted to allow modality-specific phonological information from each language to independently feed forward to the respective lexical levels for spoken and signed items. Co-activation can then arise at the lexical level, either through direct lateral links or through feed-forward and feedback connections between the lexical representations of each language and shared semantic/conceptual representations (Shook & Marian, 2009; 2012).

Evidence that deaf bimodal bilinguals can activate L1 (signed) translation equivalents without necessarily having to first activate the semantics of the corresponding L2 written words was recently found in a study with LIS–Italian bilinguals (Navarrete, Caccaro, Pavan, Mahon & Peressotti, 2015). They compared cumulative semantic costs in picture naming and printed word translation in deaf Italian signers. Participants had to name lists of pictures in LIS or translate printed Italian words to LIS signs. The lists consisted of related items drawn from the same semantic category or unrelated items. In monolinguals and unimodal bilinguals, longer picture naming latencies are typically observed for semantically-related lists than unrelated lists (e.g., Howard, Nickels, Coltheart & Cole-Virtue, 2006; Kroll & Stewart, 1994; Runnqvist, Strijkers, Alario & Costa, 2012). Indeed, Navarrete et al. (2015) found semantic costs in the picture-naming condition, but no semantic interference was observed for the printed word condition (word translation), suggesting that deaf signers directly activated translation equivalents during L2 to L1 translation without first activating semantics of the L2 words. Critically, although objective reading assessments were not available, the authors report that their participants were proficient readers of Italian, suggesting that the absence of direct semantic access of the L2 stimuli was not a consequence of low L2 proficiency (see below).

The question of the relative contributions of lexical links versus semantic connections to language co-activation effects is not unique to bimodal bilinguals and is currently widely debated in the unimodal literature (e.g., Brysbaert & Duyck, 2010; Kroll, Van Hell, Tokowicz & Green, 2010; Van Hell & Kroll, 2013). Much of this debate has been initiated by the Revised Hierarchical Model (RHM) of bilingual word acquisition (Kroll & Stewart, 1994). According to this model, during the early stages of L2 acquisition, the bilingual lexicon contains stronger links from L2 to L1 than from L1 to L2, and stronger word-to-concept mappings for L1 than L2. The explanation for this asymmetry is that L2 semantic access is dependent on L1 mediation during the early stages of L2 acquisition, while L1 words can directly access their meaning. Over time, the ability to directly access the semantics of L2 words increases and the dependency on L2–L1 translation links decreases. In recent years a growing body of research has suggested that lexical connections between languages are not only available to beginning L2 learners, but also to proficient bilinguals (e.g., Schoonbaert, Duyck, Brysbaert & Hartsuiker, 2009), and that depending on the acquisition context even beginning L2 learners can readily activate L2 word meanings (e.g., Comesaña, Perea, Piñeiro & Fraga, 2009). Whether similar observations hold for bimodal bilinguals is currently unknown, but the comparison between deaf and hearing bilinguals may be particularly revealing in this regard because these groups differ in how spoken language is acquired and represented. Deaf bilinguals may be more likely to establish strong lexical links between signs and orthographic representation of words during reading instruction and literacy development. In contrast, hearing bilinguals often access spoken words via phonological representations, and co-activation for sign language may
be more likely to be mediated by semantics than by direct lexical connections.

Bimodal bilinguals can also provide insight into the nature of non-target language activation during production. Although there is abundant evidence for phonological activation of the non-target language during speech production, e.g., from cognate production studies (Costa, Caramazza & Sebastián-Gallés, 2000; Gollan & Acenas, 2004; Hoshino & Kroll, 2008; Strijkers, Costa & Thierry, 2009) and picture–word interference (PWI) studies with cross-language phonological distractors (Costa, Miozzo & Caramazza, 1999; Herman, Bongaerts, De Bot & Schreuder, 1998), the relative contributions of sub-lexical and lexical connections between languages to these observed co-activation effects are unclear (e.g., Costa et al., 1999; Hermans, 2004). For example, although translation equivalents as distractor words in picture–word interference studies have consistently been found to facilitate target picture naming times (e.g., Costa & Caramazza, 1999), distractor words that are phonologically related to the target picture through translation into the target language (or non-target language) have yielded mixed results and appear to be relatively weak, if observed at all (Hall, 2011). Examples of such translation-mediated phonological effects would be if, compared to unrelated words, Spanish–English bilinguals were faster to name a picture of dog in English in the context of the English distractor word lady, which is dama in Spanish (phonologically related through translation into the non-target language), or in the context of the Spanish distractor word muñeca, which is doll in English (phonologically related through translation into the target language). The relatively weak evidence for translation-mediated phonological effects in PWI studies suggests that unimodal bilinguals may not automatically activate the translation equivalents of distractor words or translation activation does not spread to the phonological level.

A recent study by Giezen and Emmorey (2015) investigated whether bimodal bilinguals exhibit translation-mediated phonological effects in picture naming. If cross-language phonological effects in PWI studies exclusively occur through sub-lexical input-to-output connections between languages, then translation-mediated phonological effects should also not be observed in bimodal bilinguals. Alternatively, translation activation may not (automatically) spread to the phonological level in unimodal bilinguals because of the presence of (stronger) sub-lexical connections between the languages. In that case, bimodal bilinguals may actually be more likely to exhibit translation-mediated phonological effects than unimodal bilinguals, because of the absence of sub-lexical connections between the two languages. Using the PWI paradigm, hearing ASL–English bilinguals named pictures in ASL (e.g., PAPER) while they were presented with English auditory distractor words that were either 1) unrelated to the ASL target name (e.g., stamp), 2) English translation equivalents of the ASL target names (i.e., the word paper), or 3) phonologically related to the target sign through its ASL translation (e.g., the word cheese; the ASL signs PAPER and CHEESE have a similar phonological relationship as PAPER and MOVIE in Figure 1B). Giezen and Emmorey (2015) found that translation distractors facilitated ASL production, extending reports of translation priming in unimodal bilinguals (e.g., Costa & Caramazza, 1999) to translation equivalents in different language modalities. Critically, significant facilitation was also observed in the phonologically-related-through-translation condition (PAPER-cheese). Furthermore, despite the fact that it concerned an indirect effect, translation-mediated phonological facilitation was not significantly different from direct translation facilitation. This finding indicates that automatic activation of ASL translation equivalents of English distractor words spreads to the phonological level and influences lexical access of target ASL signs. Furthermore, although this remains speculative, the robustness of the translation-mediated phonological effect in this study suggests that automatic activation of translation equivalents may be more likely to spread to the phonological level when there are no sub-lexical connections between the two languages and thus no possibility for direct cross-language activation at the phonological level.

In sum, although both unimodal and bimodal bilinguals may activate translation equivalents through direct lexical connections or through conceptual links, unimodal bilinguals, hearing bimodal bilinguals and deaf bimodal bilinguals may differ in the relative strength and availability of either route. Furthermore, the complete absence of sub-lexical overlap between the two languages may promote phonological activation of translation equivalents in bimodal bilinguals, and as we will see later, may also have important consequences for the time course of recruitment of non-linguistic inhibitory control mechanisms to suppress non-target language activation.

Code-blending in monolingual contexts: A unique window into cross-language activation

Although both spoken languages may be active within the mind of a unimodal bilingual, code-switches in monolingual contexts are very rare (Gollan, Sandoval & Salmon, 2011; Poulisse, 1999). In contrast, bimodal bilinguals have been shown to produce ASL–English code-blends when communicating with interlocutors who do not know ASL. Casey and Emmorey (2009) reported that in such monolingual contexts nine out of thirteen native ASL–English bilinguals produced at least one ASL sign when retelling the cartoon story Canary Row (range:
0–12 signs, with a mean of 4.15 signs). In a survey study, Casey, Emmorey and Larrabee (2012) found that many ASL learners spontaneously reported that after learning ASL, they sometimes produced ASL signs when talking with others (63%; 39/62), and an experimental study revealed that a small proportion of ASL learners in fact produced at least one ASL sign when re-telling the Canary Row cartoon to a non-signer (24%; 5/21), whereas none of the Romance language learners (N = 20) produced a code-switch to their other language. Similar to code-blends produced in bilingual contexts, the ASL signs produced in monolingual contexts expressed the same meaning as the accompanying speech (see Examples 13–14). Critically, many of the signs were non-iconic and unlikely to convey meaning to the conversation partner. This finding suggests that these signs were unintentional intrusions of ASL, possibly resulting from failed suppression processes.

13. He’s [lookin’ around].

14. Ah you know [the train] wires . . .

Further, Pyers and Emmorey (2008) found evidence that ASL is not always suppressed at the syntactic level when speaking English in a monolingual situation. In ASL, as in many sign languages, facial expressions mark a variety of syntactic structures. In particular, raised eyebrows (and a slight head tilt) co-occur with conditional clauses in ASL, while furrowed brows co-occur with WH-questions. Pyers and Emmorey (2008) asked a group of bimodal bilinguals and monolingual non-signers to tell a sign-naive English listener what they would do in certain hypothetical situations (in order to elicit conditional clauses) and to obtain certain pieces of information, e.g., number of siblings, home town (in order to elicit WH-questions). Although non-signers produced conditionals with raised brows in 48% of their utterances, bimodal bilinguals did this significantly more often (79%). Most importantly, only the bilinguals synchronized their raised brows with the onset of the conditional clause, indicating that they were producing a grammatical marker rather than a facial gesture. With respect to furrowed brows, non-signers almost never produced a WH-question with furrowed brows (5%), while bimodal bilinguals did this in 37% of their utterances. These results show that distinct morphosyntactic elements from two languages can be simultaneously produced, which provides support for shared syntactic representations in bilinguals (e.g., Hartsuiker et al., 2004). In addition, they suggest a dual-language architecture in which grammatical information is integrated and coordinated at all processing levels, including phonological implementation.

Pyers and Emmorey (2008) attribute the difference in the proportion of ASL appropriate facial expressions for WH-questions and conditional clauses to the fact that furrowed brows tend to convey negative affect, while raised brows tend to express openness to interaction. Bimodal bilinguals may be more likely to suppress ASL facial grammar when it conflicts with conventional facial gesture (WH-questions) than when it does not (conditional clauses). This pattern of results also indicates that inhibition of grammatical elements of the non-selected language is relatively difficult, given that bimodal bilinguals did not completely inhibit the production of furrowed brows with WH-questions, even when this facial gesture could communicate pragmatically inappropriate or misleading information.

The fact that bimodal bilinguals produce both ASL signs and grammatical facial expressions in monolingual contexts, while unimodal bilinguals almost never produce code-switches with monolingual speakers, suggests that the language control demands are greater for unimodal than for bimodal bilinguals. Code-switches can disrupt communication in monolingual situations, while code-blends clearly do not because the relevant information is present from the speech and gestures accompanying speech are common. Further, given articulatory constraints, unimodal bilinguals ultimately must select one language for production (even if in error), while bimodal bilinguals are not so constrained. Given these differences in language control demands for unimodal versus bimodal bilinguals, we next explore how the study of bimodal bilingualism can help elucidate the nature of language and cognitive control in bilinguals.

**Implications of bimodal bilingualism for language control**

So far, we have discussed considerable evidence that because their two languages are produced and perceived in different modalities, bimodal bilinguals may not need to control activation of the non-target language to the same extent as unimodal bilinguals. Bimodal bilinguals frequently produce elements of both languages at the same time, which is ‘cost free’ for the producer and even incurs processing advantages for the comprehender. This is in sharp contrast to the costs typically associated with code-switching for unimodal bilinguals. Furthermore, although turning a language off (switching from code-blending to just speaking or signing) is costly, turning a language on (switching from just speaking or signing to code-blending) is not. Finally, although bimodal bilinguals exhibit language co-activation at the lexical and/or semantic level during language comprehension (and possibly also production), they do not experience direct sub-lexical perceptual and articulatory competition between their two languages.
For unimodal bilinguals, it has been hypothesized that managing attention to the target language, avoiding interference from the non-target language, and more generally monitoring activation in different languages engages domain-general executive control processes such as inhibition, updating and shifting (e.g., Bialystok, Craik, Green & Gollan, 2009). Furthermore it has been argued that such specific links between bilingual language processing and cognitive control might provide explanatory mechanisms for the widely reported – and currently equally widely debated – bilingual advantages in executive control abilities (e.g., Blumenfeld & Marian, 2011, 2013; Linck, Swieter & Sunderman, 2012; Prior & Gollan, 2011). Bilinguals’ continuous use of executive control processes for language control may enhance their efficiency and yield performance advantages on tasks that do not require language processing. The evidence for and against bilingual advantages in cognitive control is discussed in great detail elsewhere, and we will not repeat those discussions here (e.g., Baum & Titone, 2014 and commentaries; Hilchey & Klein, 2011; Paap & Greenberg, 2013; Paap & Sawi, 2014; Valian, 2015 and commentaries). We note only that bilingual advantages are not always found, and there may be a number of possible factors that influence whether advantages are observed or not (for discussion see e.g., Duñabeitia, Hernández, Antón, Macizo, Estévez, Fuentes & Carreiras, 2014; Kroll & Bialystok, 2013).

Aspects of the bimodal bilingual experience can provide unique insight into the nature of the relationship between cognitive control and bilingual language processes. Emmorey, Luk, Pyers and Bialystok (2008) compared the performance on a conflict resolution task (an Eriksen flanker task) by hearing ASL–English bilinguals who had learned ASL from an early age (Codas), unimodal bilinguals who learned two spoken languages from an early age, and English monolinguals. The unimodal bilinguals performed better than the other two groups, but the bimodal bilinguals did not differ from the monolinguals. These results suggest that bimodal bilinguals may not experience the same advantages in cognitive control as unimodal bilinguals. Enhanced executive control observed for unimodal bilinguals might thus stem from the need to attend to and perceptually discriminate between two spoken languages (also see Kovács & Mehler, 2009), whereas perceptual cues to language membership are unambiguous for bimodal bilinguals. Furthermore, because bimodal bilinguals can code-blend, less monitoring might be required to ensure that the correct language is being selected.

However, evidence from two other studies suggests that cognitive control may nevertheless guide some aspects of processing in bimodal bilinguals. Kushalnagar, Hannay and Hernández (2010) found similar performance for balanced and unbalanced deaf ASL–English bilinguals on a selective attention task, but better performance for the balanced bilinguals on an attention-switching task, suggesting that there might be enhancements in attention shifting for bimodal bilinguals who are highly proficient in both languages. MacNamara and Conway (2014) tested ASL interpreting students on a battery of cognitive tests at the beginning of their program and two years later. In these two years, the students improved on measures of task switching, mental flexibility, psychomotor speed, and on two working memory tasks that required the coordination or transformation of information. While these findings suggest that interpreting experience can modulate the cognitive system (cf. Yudes, Macizo & Bajo, 2011 for unimodal bilinguals), they may not generalize to bimodal bilinguals who are not training to be interpreters. Furthermore, unfortunately, neither the Kushalnagar et al. (2010) study or the MacNamara and Conway (2014) study included comparison samples of unimodal bilinguals and/or monolinguals, to rule out the possibility that the improvements came about for reasons other than increased language proficiency or increased experience with bilingual language management demands.

Instead of assessing the consequences of bimodal bilingualism on cognitive functions through group comparisons, Giezen, Blumenfeld, Shook, Marian and Emmorey (2015) directly investigated the relationship between individual differences in cognitive control and cross-language activation within a group of hearing ASL–English bilinguals. They used a visual world eye-tracking paradigm adapted from Shook and Marian (2012) to assess co-activation of ASL signs during English auditory word recognition. In addition, participants completed a non-linguistic spatial Stroop task as an index of inhibitory control (cf. Blumenfeld & Marian, 2011). Bimodal bilinguals with better inhibitory control exhibited fewer looks to the cross-language competitor ~150–250 ms post word-onset, i.e., positive correlations were observed between the Stroop effect and the proportion of competitor looks, suggesting that bimodal bilinguals with better executive control abilities experienced reduced cross-language competition and/or resolved such competition more quickly. These results parallel recent findings with Spanish–English bilinguals using the same design (Blumenfeld & Marian, 2013). Together, these results strongly suggest that inhibitory control during word recognition does not depend upon perceptual competition.

Unlike Giezen et al., however, Blumenfeld and Marian (2013) found negative correlations with the Stroop effect 300–500 ms post word-onset, and positive correlations 633–767 ms post word-onset. That is, Spanish–English bilinguals with better inhibitory control experienced increased competition early in the time course and reduced competition later in the time course. The positive correlations for the bimodal bilinguals in the
Giezen et al. study were observed much earlier than for unimodal bilinguals, despite similar onsets of language co-activation effects in both studies. Giezen et al. argued that the early negative correlations between Stroop performance and lexical competition in unimodal bilinguals may be driven by perceptual ambiguity between targets and cross-language competitors that occurred before the uniqueness point in the auditory stream. Because the overlap between targets and cross-language competitors in the Giezen et al. study was based on translation similarity in ASL, their bimodal participants did not experience similar perceptually-based ambiguity. Instead, the incoming phonological information provided unambiguous cues to a single target item, which may have allowed bimodal bilinguals to immediately engage their inhibitory control to resolve competition at the lexical level.

These findings suggest that, similar to unimodal bilinguals, bimodal bilinguals engage non-linguistic inhibitory control processes during auditory word recognition. Of course, it is still possible that they may not do so to the same extent as unimodal bilinguals. Only a few studies so far have investigated possible cognitive consequences of bimodal bilingualism and given that the evidence for advantages in cognitive control in unimodal bilinguals is mixed, variable findings in studies with bimodal bilinguals should similarly be interpreted with caution. Furthermore, any variation in the bilingual experience that may contribute to variation in cognitive consequences for unimodal bilinguals would be expected to similarly cause variation for bimodal bilinguals.

However, the bilingual experience for bimodal bilinguals is distinct because one of their languages is a visual-spatial language, which has independent cognitive consequences. Experience with a sign language for both deaf and hearing signers has been found to impact various non-linguistic visual-spatial abilities. For example, signers exhibit enhanced abilities to generate and transform mental images and discriminate facial expressions (see Emmorey & McCullough, 2009, for review). Furthermore, some studies have found that signers exhibit superior spatial memory abilities compared to non-signers (e.g., Geraci, Gozzi, Papagno & Ceccheto, 2008; Wilson, Bettger, Niculae & Klima, 1997). We suggest that such visual-spatial cognitive advantages should be considered effects of acquiring a visual language rather than effects of being bilingual per se. Nonetheless, changes in visual-spatial cognition in bimodal bilinguals could interact with more general cognitive effects of bilingualism. For example, many non-linguistic cognitive control tasks used to assess the effects of bilingualism utilize visual-spatial stimuli precisely to eliminate verbal processing or verbal strategies as much as possible. However, such tasks may tap into visual-spatial abilities that are sensitive to visual language experience in bimodal bilinguals. For example, inhibitory control tasks such as the spatial Stroop or the Eriksen flanker task require inhibition of visual information (location of the arrow; direction of the surrounding flankers), and it is possible that performance on these tasks could benefit from enhanced visual-spatial processing.

In sum, because of the absence of perceptual and articulatory competition at the phonological level, bimodal bilinguals may experience weaker demands on language control than unimodal bilinguals and as a result, they may not develop enhanced executive control abilities. However, bimodal bilinguals do have competing language representations at lexical levels and appear to engage the same cognitive control mechanisms as those of unimodal bilinguals when resolving this competition. Finally, it is quite possible that bimodal bilinguals exhibit advantages on other cognitive tasks that tap into aspects specific to sign language processing or to bimodal language processing; for example, monitoring attention cross-modally or integrating information cross-modally.

The bimodal bilingual brain

One longstanding and critical question raised by bilingualism is the degree to which the neural systems that support language processing in each language are separate or overlapping. For unimodal bilinguals, there is general consensus that a high degree of neural convergence exists between two spoken languages, particularly for proficient and early bilinguals (see Buchweitz & Prat, 2013, for a recent review). More than two decades of lesion-based and neuroimaging research has also identified a shared neural network that supports both signed and spoken language processing (see MacSweeney, Capek, Campbell & Woll, 2008, and Emmorey, 2002, for reviews). Comprehension of both spoken and signed language recruits bilateral superior temporal cortex – including auditory regions, despite the fact that sign languages are perceived visually. Further, damage to left posterior temporal cortex leads to comprehension deficits and fluent sign language aphasia, parallel to fluent aphasia in spoken language. Sign and speech production recruit a similar left-lateralized fronto-temporal network, and damage to left inferior frontal cortex results in a non-fluent aphasia for both signers and speakers. Most of these findings are based on comparisons between deaf native signers and hearing monolingual speakers, but these core similarities are also supported by studies that compared hearing bimodal bilinguals with each of these two groups separately (e.g., MacSweeney, Woll, Campbell, McGuire, David, Williams, Suckling, Calvert & Brammer, 2002; Neville, Bavelier, Corina, Rauschecker, Karni, Lalwani, Braun, Clark, Jezzard & Turner, 1998) or that conducted
within-subject conjunction analyses with hearing bimodal bilinguals (e.g., Braun, Guillemin, Hosey & Varga, 2001).

Nonetheless, important differences between the neural representation of spoken and sign languages are beginning to emerge (Corina, Lawyer & Cates, 2013), and direct contrasts between speech and sign in bimodal bilinguals reveal striking differences between the sensory-motor neural resources required to produce and comprehend languages in different modalities (Emmorey, McCullough, Mehta & Grabowski, 2014). For unimodal bilinguals, both languages engage auditory regions during comprehension, whereas for bimodal bilinguals, speech comprehension activates bilateral auditory regions (superior temporal cortex) to a much greater degree than sign comprehension. Further, sign comprehension activates bilateral occipitotemporal visual regions more than audiovisual speech comprehension (e.g., Söderfeldt, Ingvar, Rönnberg, Eriksson, Serrander & Stone-Elander, 1997). With respect to production, unimodal bilinguals engage essentially the same motor regions for both languages, whereas bimodal bilinguals recruit bilateral parietal and left posterior middle temporal cortex (LpMTG) significantly more when signing than speaking (e.g., Zou, Abutalebi, Zinszer, Yan, Shu, Peng & Ding, 2012a). Further, sign language comprehension also engages parietal cortices to a greater extent than speech comprehension in bimodal bilinguals (Emmorey, McCullough et al., 2014). Thus, the bimodal bilingual brain differs from the unimodal bilingual brain with respect to the degree and the extent of neural overlap when processing their two languages, with more neural overlap for unimodal than bimodal bilinguals.

Consistent with evidence of language co-activation reviewed above, neuroimaging studies suggest activation of sign-related brain regions when bimodal bilinguals name pictures with spoken nouns (Zou et al., 2012a) or name spatial relations with spoken prepositions (Emmorey, Grabowski, McCullough, Ponto, Hichwa & Damasio, 2005). Compared to monolingual Chinese speakers, late bimodal bilinguals fluent in Chinese and Chinese Sign Language (CSL) exhibited greater activation in the right superior occipital gyrus (RSOG; a visual region) and also showed increased functional connectivity between the RSOG and the LpMTG during picture naming. Given that these two brain regions were linked to sign language processing, Zou et al. (2012a) interpreted their results as reflecting automatic activation of CSL during the production of spoken Chinese. Emmorey et al. (2005) reported a similar result with early ASL–English bilinguals who engaged right parietal cortex when naming spatial relations in English, in contrast to monolingual English speakers (Damasio, Grabowski, Tranel, Ponto, Hichwa & Damasio, 2001). Given that expressing spatial relations with ASL classifier constructions recruits right parietal cortex, Emmorey et al. (2005) hypothesized that ASL–English bilinguals encode spatial relationships in ASL, even when the task is to produce English prepositions.

Because their two languages rely on neural systems that are only partially overlapping, Zou, Ding, Abutalebi, Shu and Peng (2012b) suggested that bimodal bilinguals may have to co-ordinate more complex neural networks than unimodal bilinguals, which could involve more control processes. However, no study has yet directly compared the neural regions engaged during language switching for bimodal and unimodal bilinguals. Nonetheless, Zou et al. (2012b) found that switching between CSL and Chinese compared to naming pictures only in Chinese (non-switching blocks) engaged the left caudate nucleus (LCN) and the anterior cingulate, and behavioral switch costs for Chinese were observed when the task was performed outside the scanner. In addition, they reported that the grey matter volume of the LCN was positively correlated with the functional activity of the switching effect. Unimodal bilinguals also engage the anterior cingulate and the caudate when switching between languages, which Abutalebi, Annoni, Zimine, Pegna, Seghier, Lee-Jahnke, Lazyras, Cappa and Khateb (2008) attribute to the need to inhibit inappropriate responses (i.e., the non-target language) and to select distinct actions plans for each language. Using fNIRS (which cannot detect subcortical activations) and a picture-naming task, Kovelman, Shalinsky, White, Schmitt, Berens, Paymer and Pettito (2009) reported increased activation in a left posterior temporal region and a sensory-motor region (postcentral and precentral gyri) when early bimodal bilinguals mixed their languages (their ‘bilingual mode’) compared to producing each language alone (their ‘monolingual mode’). However, the ‘bilingual mode’ combined both code-blending and code-switching, and thus it is difficult to interpret whether the neural differences reflect cognitive control processes (associated with switching) or dual lexical production processes (associated with code-blending). Kovelman et al. (2009) reported no difference in neural activation between the code-blending and switching conditions, but this could be due to low power, with only five bimodal bilinguals included in the study.

Finally, anatomical changes in brain structure have been linked to bilingual experiences (see Li, Legault & Litcofsky, 2014, for a recent review). Bimodal bilinguals can provide novel insight into the mechanisms that underlie such neural plasticity by teasing apart anatomical changes that are linked to the acquisition and use of two languages (general effects of bilingualism), changes that are linked to the need to engage the same sensory-motor system for both languages (i.e., effects specific to unimodal bilinguals), and changes that are specific to acquiring a signed language (i.e., effects specific to bimodal bilinguals). With respect to general effects, Zou
et al. (2012b) found that CSL-Chinese bilinguals had greater grey matter volume in the left caudate nucleus compared to monolingual Chinese speakers, paralleling results with unimodal bilinguals. This finding suggests that neural plasticity within the caudate is not specifically linked to experience with two spoken languages and may be associated with the more general demands of controlling two languages. Unimodal bilinguals have also been shown to exhibit increased grey matter density in the left inferior parietal lobe (IPL) (e.g., Mechelli, Crinion, Noppeney, O’Doherty, Ashburner, Frackowiak & Price, 2004), but two voxel-based morphometry studies comparing hearing bimodal bilinguals with monolingual speakers failed to detect differences in grey matter volume in the left or right IPL (Olulade, Koo, LaSasso & Eden, 2014; Zou et al., 2012b). Thus, it is possible that changes in IPL in unimodal bilinguals are related to accessing two phonological systems within the same modality, given that IPL has been associated with phonological working memory (e.g., Awh, Jonides, Smith, Schumacher, Koeppe & Katz, 1996). Similarly, it is likely that the increased size of primary auditory cortex (Heschl’s gyrus) in unimodal bilinguals (Ressel, Pallier, Ventura-Campos, Diaz, Roessler, Avila & Sebastian-Galles, 2012) is specific to those acquiring two spoken languages. An anatomical difference that may be specific to bimodal bilinguals is a change in the asymmetry of the motor hand area related to life-long experience with sign language (Allen, Emmorey, Bruss & Damasio, 2013). Another morphometry study by Allen, Emmorey, Bruss and Damasio (2008) found a difference in white matter volume in the right insula in hearing (and deaf) bimodal bilinguals compared to non-signers that they attributed to enhanced connectivity resulting from an increased reliance on cross-modal sensory integration (e.g., tactile, visual, and proprioceptive integration during sign production).

In sum, the bimodal bilingual brain is likely to differ from the unimodal bilingual brain due to differences in the convergence of neural networks that control their two languages, with more sensory-motor overlap for unimodal bilinguals. However, much more research is needed to identify the neuroanatomical consequences of bimodal bilingualism, e.g., there are no studies (thus far) examining whether structural changes can be induced by short-term learning of a sign language as a second language (cf. Schlegel, Rudelson & Tse, 2012).

Conclusions and future directions

The fact that bimodal bilinguals can and do produce elements from their two languages at the same time provides a novel testing ground for theories of bilingual language representation and processing. Thus far, we know that the bilingual language production system must be able to simultaneously access two lexical representations (translation equivalents) without cost and that the comprehension system must be able to integrate lexical information from two languages simultaneously. However, the existence of code-blends that do not involve translation equivalents or that contain different syntactic constituents across languages raises interesting questions for how lexical and syntactic representations are integrated for bilinguals. In addition, with respect to language mixing, bimodal bilinguals allow us to tease apart the cognitive processes involved in releasing a language from inhibition (turning a language ‘on’) and inhibiting a language (turning a language ‘off’).

Because the phonological systems of signed and spoken languages involve non-overlapping features, bimodal bilinguals provide a means for investigating the importance of perceptual or phonological competition in bilingual language processing. Evidence from bimodal bilinguals illustrates the necessity of links between languages at the lexical or semantic level and that perceptual overlap is not required for language co-activation. A phenomenon that we have not touched upon, however, is the existence of mouthings from spoken language words that are often produced silently and simultaneously with signs (e.g., Boyes Braem & Sutton-Spence, 2001). If mouthings are part of the lexical representation of a sign (and this is debated, see Vinson, Thompson, Skinner, Fox & Vigliocco, 2010), then there is the possibility of phonological overlap between a signed and a spoken language. Currently, the impact of mouthings on bimodal bilingual language processing is completely unknown.

In addition, bimodal bilinguals provide unique insight into the nature of language and cognitive control processes because their languages do not compete for the same articulatory output system and perceptual cues to language membership are unambiguous. Despite these differences, evidence from natural language mixing, psycholinguistic experiments, and neuroimaging studies indicates that bimodal bilinguals, like unimodal bilinguals, must engage cognitive control mechanisms to manage their languages. However, to the extent that we observe cognitive enhancements or cognitive control effects only in unimodal bilinguals, it will suggest that perceptual or articulatory competition between languages gives rise to these effects.

Finally, many open questions remain with respect to the neural systems that underlie bimodal bilingual language processing, the nature of bimodal bilingual language acquisition in both deaf and hearing children (and L2 acquisition by adults), and possible cognitive effects of bimodal bilingualism across the life-span. The study of bimodal bilingualism has much to offer to our understanding of the bilingual mind and promises to enrich models of language processing more generally.
References


Emmorey, K., Petrich, J. A. F., & Gollan, T. H. (2012). Evidence from bimodal bilinguals indicates “turning a language ON” is not costly but “turning a language OFF” is. Presented at the 55th Annual Meeting of the Psychonomic Society, Long Beach.


Karen Emmorey, Marcel R. Giezen and Tamar H. Gollan


