In: Bilinguals: Cognition, Education and Language Processing ISBN 978-1-60741-710-1
Editor: Earl F. Caldwell, pp.

Chapter 2

LANGUAGE PROCESSING IN BIMODAL BILINGUALS

Anthony Shook and Viorica Marian
Northwestern University

ABSTRACT

Recent research suggests differences between bimodal bilinguals, who are fluent in a spoken and a signed language, and unimodal bilinguals, who are fluent in two spoken languages, in regard to the architecture and processing patterns within the bilingual language system. In this chapter, we discuss ways in which sign languages are represented and processed and examine recent research on bimodal bilingualism. It is suggested that sign languages display processing characteristics similar to spoken languages, such as the existence of a sign counterpart to phonological priming and the existence of a visual-spatial loop analogous to a phonological loop in working memory. Given the similarities between spoken and signed languages, we consider how they may interact in bimodal bilinguals, whose two languages differ in modality. Specifically, we consider the way in which bimodal bilingual studies may inform current knowledge of the bilingual language processing system, with a particular focus on top-down influences, and the fast integration of information from separate modalities. Research from studies looking at both production and perception suggests that bimodal bilinguals, like unimodal bilinguals, process their languages in parallel, with simultaneous access to both lexical and morphosyntactic elements. However, given the lack of overlap at the phonological level (the presumed initial locus of parallel activation in unimodal studies) in bimodal bilinguals' two languages, we conclude that there are key differences in processing patterns and architecture between unimodal and bimodal language systems. The differences and similarities between unimodal and bimodal bilinguals are placed in the context of current models of bilingual language processing, which are evaluated on the basis of their ability to explain the patterns observed in bimodal bilingual studies. We propose ways in which current models of bilingual language processing may be altered in order to accommodate results from bimodal bilingualism. We conclude that bimodal bilingualism can inform the development of models of bilingual language processing, and provide unique insights into the interactive nature of the bilingual language system in general.

> "The analytic mechanisms of the language faculty seem to be triggered
> in much the same ways, whether the input is auditory, visual, even tactual…"
> -Noam Chomsky (2000, p. 100-101)

1. Introduction to the Chapter

One of the most striking features of bimodal bilingualism (which refers to fluency in both a signed and a spoken language) is the total lack of phonological overlap between the two languages. Bimodal bilinguals are able to create distinct, meaningful utterances with two separate sets of articulators, and have two output channels, vocal and manual. In contrast, unimodal bilinguals, whose two languages are spoken, utilize only one modality for both input and output. Moreover, bimodal bilinguals are able to perceive distinct linguistic information in two domains, via listening to speech and via visually perceiving signs. Although unimodal bilinguals utilize visual information as well, it acts primarily as a cue to facilitate understanding of auditory input, rather than providing a source of visual linguistic input independent from the auditory signal.

Research on bimodal bilingualism carries implications for understanding general language processing. For example, one issue at the heart of conceptual modeling of language is the level of influence of bottom-up versus top-down processing. Specifically, when we process language, how much information do we gain from the signal itself (e.g. bottom-up input from phonological or orthographic features) and how much do we gain from higher-order knowledge (e.g., top-down input from background information, context, etc.)? While the existence of both top-down and bottom-up influences is universally acknowledged, the degree to which each holds sway over the language processing system is not entirely clear. Another important question is how bilinguals integrate auditory and visual information when processing language and whether that process differs between unimodal and bimodal bilinguals. To what extent does the ability to retrieve information from two separate modalities facilitate language comprehension? Is information from separate modalities accessed simultaneously or serially?

Given the alternate input/output structure found in bimodal bilinguals, and lack of linguistic overlap between signed and spoken languages, it is important to consider what studies about bimodal bilingualism can tell us about bilingual language processing in general. Since the vast majority of bilingual research is performed with unimodal bilinguals, it is somewhat unclear what similarities and differences exist between the two groups. Furthermore, comparing both the structural-linguistic and cognitive-processing aspects of unimodal and bimodal bilingual groups can illuminate the effects of modality on language processing. In the present chapter, we will review recent research on bimodal bilingualism and contrast it with results from unimodal studies in order to expand understanding of bilingual language processing.

To study the influence that research on bimodal bilingualism has on both the mechanisms of bilingual processing and the architecture of the underlying language system, we will outline several models of bilingual language processing and examine how well they account for the results seen in recent bimodal bilingual research. The present chapter consists of two main parts. The first part focuses on linguistic and cognitive aspects of sign languages in native signers and in bimodal bilinguals. Specifically, we will (a) compare and contrast how sign languages are represented linguistically, by examining previous work on the structural characteristics of sign languages, (b) discuss the cognitive patterns of sign language, in contrast to spoken language, by examining the similarities between phenomena found in spoken language research with those found in sign language research, and (c) examine results from studies looking first at language production and then at language perception in bimodal bilinguals, which will be directly contrasted with previous results from unimodal bilingual studies. In the second part of the chapter, we will introduce several models of bilingual language processing, focusing on models of both (a) language production and (b) language perception, and discuss them in light of the results from bimodal bilingual studies. We will conclude by suggesting that spoken and signed languages are represented similarly at the cognitive level and interact in bimodal bilinguals much the same way two spoken languages interact in unimodal bilinguals, and that bimodal bilingual research can highlight both the strengths and weaknesses of current models of bilingual language processing.

2. Representation and Processing of Sign Languages

2a. Structure of Sign Languages

Current models that explain the structure of sign languages are primarily based on the study of American Sign Language, and its contrast with spoken English. It is important to note that just as spoken languages differ in phonology, morphology and syntax, so do sign languages. For example, not only do American Sign Language (ASL) and British Sign Language (BSL) utilize separate lexicons, they display morphological distinctions as well, such as BSL's use of a two-handed finger-spelling system compared to ASL's one-handed finger-spelling system (Sutton-Spence and Woll, 1999). There are also phonological distinctions, in that phonological segments in one language do not necessarily exist in the other, much like in spoken languages (sign languages use handshape, location of the sign in space, and motion of the sign as phonological parameters). We will focus mainly on the phonological aspects of sign language structure, while briefly discussing certain morphosyntactic traits. Given that much of the current body of knowledge about sign language structure is based on American Sign Language, we will focus specifically on the relationship between the phonologies of ASL and spoken English, in order to highlight some of the fundamental differences between spoken and signed languages in general.

The most salient difference between signed and spoken languages is that signed languages exist in a spatial environment, and are expressed grammatically through the manipulation of the body, notably the hands and face, within a linguistic sign-space, which is a physical area centered around the front of the speaker's body. Like actors on a stage, the mechanics of grammar occur within this sign-space. This variance in articulatory location and modality results in interesting syntactic differences. While English uses prepositional information to determine the location and relation of object, ASL creates schematic layouts of objects within the sign-space to determine their relationship in space, as well as to show motion. For example, rather than describe the movement of some object from point A to point B, the lexical item is presented and then physically moved. Syntactically, movement of verbs within the sign-space can determine the objects and subjects of sentences. Consider, for instance, the sign "give." Whereas in English the difference between "I give you X" and "You give me X" is determined by word order and thematic role (subject "I" versus object "me"), ASL differentiates the two through variation in direction of movement of the verb. The two nominal entries are placed in the sign space, and the sentence structure is determined by whether the speaker moves the "give" sign from himself to his interlocutor, or vice versa (see Figure 1, top and middle panels). This system also allows for ASL to have a less strict word order than English, given that nominal entries in space can be manipulated freely (Bellugi, Poizner, and Klima, 1989).

Secondly, ASL marks tense lexically using temporal adverbs, rather than morphologically marking verbs as in English (i.e., whereas forming the past tense of "walk" in English involves adding a past tense morpheme to create "walked," in ASL "walk" becomes "walked" by combining the sign for "walk" with a temporal sign like "before" or "yesterday"). One could argue that this, too, is related to the nature of sign-language articulation. If we consider the sign-space as a stage on which lexical items can be placed and physically referenced, then we may consider the addition of temporal adverbs as setting-markers. This way, rather than consistently marking signs throughout an entire utterance, signers merely need to indicate the time to the listener once.

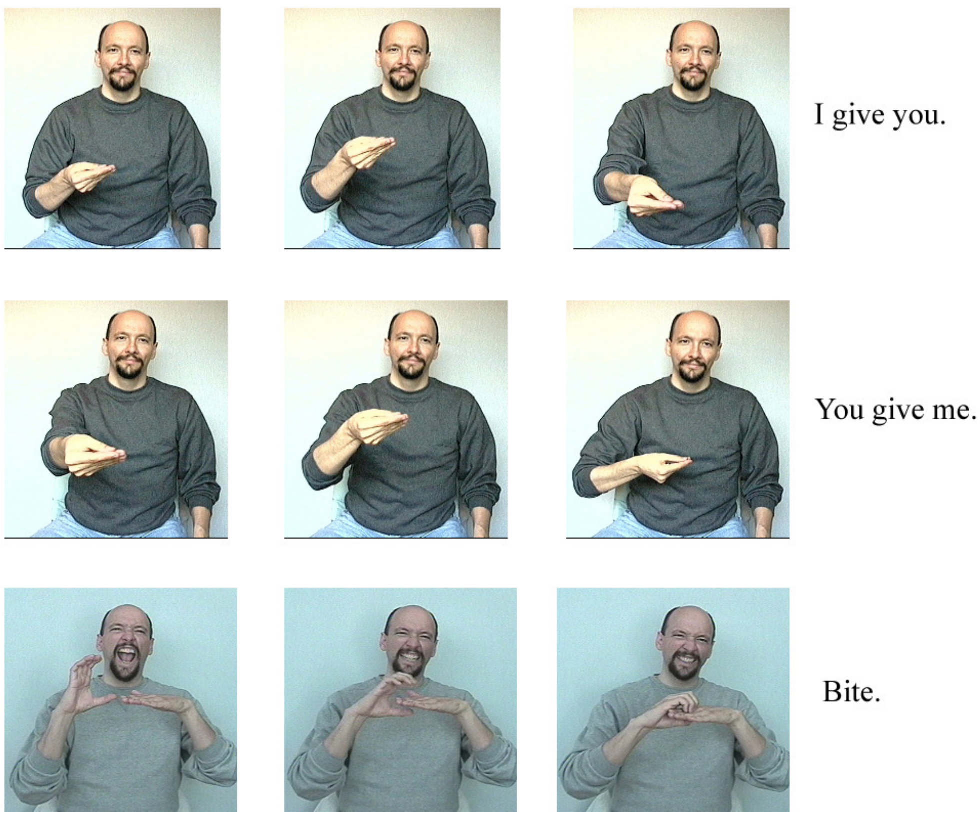

Figure 1. Signs for the ASL phrases "I give you" and "you give me," and the word "bite". All images © 2006, www.Lifeprint.com. Used by permission.

Another interesting difference between ASL and English is the use of facial expressions to mark certain morphosyntactic elements, like relative clauses, topicalization and conditionals (Liddell, 1980). For example, when producing conditional statements in ASL, the signs are accompanied by the raising of the eyebrows. The eyebrows are furrowed downward to accompany wh-questions. Often, combinations of manual and non-manual constructions are used in creating signs (Liddell, 1980) – for example, the sign for "bite" involves both a manual motion and a biting motion of the mouth (see Figure 1, lower panel). They can also act as suprasegmental features. A head shake can negate all or part of a sentence, even though a manual sign for negation exists as well.

Much like English, ASL contains a sublexical phonological system that combines segments according to combinatorial rules, but which uses manual rather than oral features. At the phonological level, current models of ASL recognize three main parameters, which are handshape, location of the sign in space, and movement of the sign (Brentari, 1998; Stokoe, 1960). Each of these parameters can be further broken down into a finite set of phonemes, but each sign contains at least one feature from all three parameters. In other words, a sign consists of a specific handshape that is held in a particular point in the sign space, and is then moved in a particular way. Each parameter can be varied independently of the others, which can result in pairs of signs that match in two parameters and vary only in the third.

This differs from spoken language phonology not only in modality of structure, but also in temporal relationships of the features. Since spoken languages unfold sequentially, the phonemes do not overlap within words. In ASL, and other sign languages, the features selected to create a sign do overlap temporally. The handshape used to form a sign occurs simultaneously with the location of the sign in space. As we will discuss later, this temporal overlap can provide a unique avenue into lexical and sublexical selection mechanisms during language processing.

2b. Cognitive Representation of Sign Languages

While comparisons of spoken and sign language often focus on the structural differences between the two languages, a number of studies have examined similarities in patterns of cognitive processing across both languages. Here, we discuss studies that compare psycholinguistic phenomena found in spoken language with those found in signed languages. In doing so, we show that there are similarities in the way that linguistic information is processed across signed and spoken languages, which allows us to view them as equally represented on a cognitive level, and therefore examine their interactions in bimodal bilinguals.

In order to inspect the way in which phonological information is handled in users of signed languages, Dye and Shih (2006) performed a study that examined the role of phonological priming in British Sign Language (BSL). The authors asked whether a sign could facilitate the activation of a phonologically similar sign. This notion was based on results that suggest spoken words can facilitate the activation of phonologically similar words (Goldinger, Luce, Pisoni and Marcario, 1992; but see Marslen-Wilson, 1990; Praamsta, Meyer and Levelt 1994). Dye and Shih tested the reaction times of monolingual native signers of BSL in a lexical decision task where sign targets were preceded by primes that shared none, one, two, or all three parameters of sign language phonology. They found that native signers were significantly faster at correctly naming lexical items when the preceding prime overlapped in location, as well as in location and movement (but not in other dimensions).

Interestingly, the authors used both signs and non-signs to test the priming effect and found that native BSL signers only showed priming in response to real signs. This implies that the locus of lexical priming is actually at the level of the lexicon in native signers – while in English, non-words are capable of priming words, this is not the case with native signers of BSL. Were the features themselves producing the lexical priming effect, one would expect to see priming due to heavily overlapping non-signs. However, the same paradigm used with non-native signers showed non-sign-to-sign priming, suggesting that the effect cannot be attributed exclusively to the lexicon. This suggests qualitatively different processes for handling input when considering the model of human speech perception. It is unclear whether the difference arises due to changes in processing patterns while maintaining standard architecture, or if the structure of the language processing system as a whole is different when developed around visual/gestural languages. Given models of lexical access like Marlsen-Wilson and Welsh's Cohort Theory (1978) that conceive of the build-up of phonological information to activate phonologically similar lexical items, an issue arises with the results yielded by native-signers. Specifically, auditory information is analyzed temporally, such that the input activates all lexical items that match, and over time as the system gets more information, the number of activated potential targets decreases. If this theory held true, native signers should show the same non-sign-to-sign priming effect as non-native signers, due to the fact that features that overlap with items in the signer's lexicon are presented. However, it appears that the featural information alone is not enough to promote the activation of overlapping lexical items.

It seems then that the nature of language acquisition influences the development of the language processing system. This raises several questions. First, to what extent is the system able to alternate between one lexical access mechanism versus the other – in other words, do non-native signers ever access lexical items the same way native signers do, or vice versa? Dye and Shih's non-native group was comprised of subjects who had learned BSL later in life, but were born profoundly deaf. It is, therefore, not the case that their processing system was influenced by spoken language processing. Still, the non-native signers' processing patterns matched those predicted by the Cohort theory better than the native signers.

The second question that arises is how bimodal bilinguals might access their lexicons. Age of acquisition of the two languages obviously plays a role, where native signers who learned to speak later in life might process more like Dye and Shih's native signers, and native speakers who learned a sign language later in life might process more like the non-native group. However, if a person learns both languages simultaneously, the predictions become less clear. One possibility is that lexical access could become task based, utilizing both processes in different circumstances. Slowiaczek and Pisoni (1986) tested monolingual English speakers in a lexical decision priming task and their results contradicted the predictions of the Cohort model – initial phonemes between prime and target actually caused inhibition of lexical access. However, initial phoneme overlap showed facilitation of response in identification-in-noise tasks. Phonological overlap affected the system differently across tasks. The same concept could be generalized to bimodal bilinguals, which could suggest that simultaneous bimodal bilinguals access their lexicon differently dependent on the nature of the input.

Another possibility is that the system simply uses both mechanisms for lexical access simultaneously. It is possible to think of the features of sign language phonology as analogous to certain features of spoken language phonology. Voice-onset time (VOT) can act as a cue to certain phonemes, and studies have shown that listeners are able to use fine-grained VOT information during lexical access to determine the word being spoken (McMurray, Tanenhaus, and Aslin, 2002). Dye and Shih (2006) showed that location and movement act as more salient cues to priming than handshape in non-native signers. So, perhaps the bimodal system is able to simultaneously process the featural information of signs, while simultaneously utilizing the lexical-level access shown by native signers in the study. To examine this, one could replicate the priming paradigm utilized by Dye and Shih and compare groups of Native-ASL speakers, Native-English speakers and ASL-English bilinguals.

Previous research has provided evidence for the existence of a visuospatial articulatory loop in working memory for sign languages. In users of a spoken language, the phonological loop in verbal working memory consists of a phonological storage buffer, which holds phonological information in working memory, and a rehearsal process, which refreshes the items in the storage buffer, preventing them from fading quickly (Gathercole and Baddeley, 1993). The evidence for the existence and subsequent separation of these two components comes from experimental effects such as the *phonological similarity effect*, the *word-length effect*, and *articulatory suppression*. If the same effects could be seen in users of sign-language, this would suggest that the cognitive system is capable of treating spatial sensorimotor information as it would language information from an auditory modality. In other words, specific linguistic experience shapes the kind of input that the phonological loop deems relevant, but does not necessarily change the way the system functions. So do the phonological similarity, word-length and articulatory suppression effects occur in sign-language users?

The *phonological similarity effect* refers to the phenomenon that words in a list that share phonological information are more difficult to recall than words that are phonologically diverse. Research has shown that lists of signs that contain the same handshape show worse recall than lists with diverse handshapes (Krakow and Hanson, 1985; Wilson and Emmorey, 1997). This provides evidence for a phonological similarity effect in ASL. The *word-length effect* shows that lists of long words are harder to recall than lists of short words. Wilson and Emmorey (1998) tested ASL signers with lists consisting of signs with long movement, and lists of signs with short, local movement. Since movement is a physical process, it requires more time to make large movements within a sign than it does to make small movements, thus increasing the temporal load on the listener. The results showed that lists of temporally long signs were recalled worse than those with short signs. Lastly, *articulatory suppression* is the effect in which repetition of phonemes or syllables that use relevant articulators disrupts the rehearsal mechanism of the phonological loop. The result is worse performance with suppression than without. This has also been shown in speakers of ASL (Wilson and Emmorey, 1997). When subjects were asked to make motoric hand movements (alternating fist and open hand) during list memorization, they showed worse recall than when they kept their hands still. In the same study, Wilson and Emmorey showed no interaction between articulatory suppression and phonological similarity, which is in accord with results from spoken language studies and suggests that the two tasks affect different aspects of the phonological loop.

There is also compelling evidence to suggest that children learning ASL as a native language develop much like children learning a spoken language. Children learning ASL reach the milestones of language development at about the same rate as those learning spoken languages (Bonvillian and Folven, 1993; Pettito and Marentette, 1991). One well-documented developmental phenomenon originates in Werker and Tees' (1984) study suggesting that infants under one-year of age were capable of discriminating sounds from their non-native language, but lost that ability as they grew older. This is an example of categorical perception, the phenomenon by which sounds are placed in phonemic categories and listeners are unable to distinguish between sounds that fall within the same category. The categories available to any given listener are based upon the relevant phonemes of the listener's native language. In regards to ASL signers, this raises two questions. First, does ASL display categorical perception? Emmorey, McCullough and Brentari (2003) examined whether native deaf signers displayed categorical perception based on differences in handshape. Much like in auditory studies, the experimenters developed continua of signs using still images, where the endpoints represented prototypical productions of some handshape. They found that native signers demonstrated categorical perception for the handshape stimuli, but a group of hearing non-signers did not. This result was corroborated by Baker, Idsardi, Golinkoff, and Petitto (2005), who also found that native ASL signers categorized handshapes linguistically, rather than on a purely perceptual basis. The second question is whether or not this ability develops similarly in signers and speakers. Baker, Golinkoff, and Petitto (2006) found that 4-month old hearing infants, who were not learning ASL, displayed categorical perception of handshape stimuli based on linguistic properties of the input, while 14-month infants failed to do so. This suggests a pattern of perceptual shift for ASL that is nearly identical to spoken languages, where infants initially have the capacity to linguistically categorize perceptual input, but lose this ability as their perceptual system becomes specialized to their own language.

Taken together, these studies suggest strong similarities between users of spoken languages and signed languages on a cognitive level. Sign languages show many of the same phenomena as spoken languages, such as lexical priming, categorical perception, and the presence of an articulatory loop, suggesting that spoken and signed languages may be processed by a similar language mechanism. In the next section, we examine the way two languages that differ in modality interact within bilinguals who are fluent in both.

2c. Processing Patterns in Bimodal Bilinguals

It has become commonly accepted in the field of bilingual study that unimodal bilinguals activate their two languages in parallel (Blumenfeld and Marian, 2007; Canseco-Gonzalez, et al., 2005; Marian and Spivey, 2003; Weber and Cutler, 2004). However, little work has been done to suggest that bimodal bilinguals activate their languages in parallel as well. Emmorey and colleagues (Casey and Emmorey, 2008; Emmorey, Borinstein, Thompson, and Gollan, 2008) have performed a series of experiments looking at how ASL might be active during production of English in bilingual users of both languages. In several studies, bimodal bilinguals were asked to tell a story to listeners, in English. Listeners were either known to be bilinguals as well (Emmorey et al., 2008), or their language background was unknown to the speaker (Casey and Emmorey, 2008). During English production, the experimenters recorded the hand gestures that were spontaneously created by the bilingual speakers. The results showed that bimodal bilinguals produced a significant number of what Emmorey and colleagues refer to as code-blends, which are semantically related signs inserted simultaneously with the related lexical item in speech. While bimodal bilinguals produced code-blends with both groups of listeners, they produced more code-blends when the listener was known to be a bimodal bilingual.

The findings from these studies suggest that bimodal bilinguals access their non-target language during production, even up to the point where semantically related signs are produced concomitantly with speech. This is in contrast with some previous work that suggests unimodal bilingual lexical access during speech is language-specific (e.g., Costa and Caramazza, 1999, but see Colomé, 2001). If lexical access during production is language-specific in unimodal bilinguals, but language-independent in bimodal bilinguals, does that imply that the bimodal bilingual processing system is, in some way, different from that of the unimodal bilingual, beyond the surface distinction of modality of input?

Emmorey et al. (2008) discuss one possible reason for the potential difference between unimodal and bimodal bilinguals in language production. While unimodal bilinguals are physiologically limited to using one language at a time (it is impossible to express a concept in both French and English simultaneously, for example), bimodal bilinguals face no such limitation. It is possible that the processing architecture is the same, but bimodal bilinguals are able to exploit the cross-modal nature of their languages to make use of a skill that unimodal bilinguals possess, but cannot access. Determining whether bimodal bilinguals process language differently than unimodal bilinguals or whether unimodal bilinguals and bimodal bilinguals process language similarly (i.e., the two groups employ the same mechanisms, but unimodal bilinguals are unable to produce both of their languages simultaneously due to biological constraints) has important implications for modeling the bilingual system.

There is also evidence to suggest that in addition to simultaneous production of lexical items, higher-order, ASL-specific morphosyntactic features can be found when bimodal bilinguals produce English. Pyers and Emmorey (2008) examined the interaction of ASL facial expressions with English grammatical constructions. ASL uses facial expressions to mark certain grammatical features of sentences (e.g., furrowed brows accompany wh-questions, raised eyebrows occur with conditionals, etc). The authors found that when bimodal bilinguals are speaking English, they produce grammatically relevant facial expressions simultaneously. Furthermore, the authors recorded the timing of facial productions and found that the facial expressions were very closely time-locked with the English grammatical constructions. This implies that bimodal bilinguals utilize a language system that integrates grammatical information from both languages at the same time, rather than separating the syntactical systems of the two languages (Hartsuiker, Pickering, and Veltkamp, 2004).

Another group of bimodal bilinguals, which Dufour (1997) refers to as sign-text bilinguals, can also provide a unique window into bilingual language processing. Sign-text bilinguals are those who are fluent in a sign language, as well as the written form of a spoken language. While sign-text bilinguals do not display the salient characteristic of both speaking and signing a language, they nevertheless process two languages with very different grammatical structures. Unfortunately, little work has been done to directly contrast the two language forms in sign-text bilinguals. However, as Dufour explains, there are studies that use text-based input as the stimuli or materials for studies with deaf signers. While these studies fail to control for the wide range of relevant variables found to influence bilingual status (such as age of acquisition, proficiency, etc.), they nevertheless provide an insight into the interaction between the two modalities.

Hanson (1982) examined the interaction between the signed and written modality. ASL-English sign-text bilinguals were presented with lists of words that were structurally similar (i.e., all the signs were similar), phonologically similar (the English translations had similar sounds) or orthographically similar (the English words were spelled similarly). The deaf participants were separated into two groups, where one received the word lists as signs, and the other received the lists as English written forms. Subjects performed a probe recall task, where an item from one of the lists was given, and subjects had to provide the item that followed it in the list. Hanson found that subjects who received the sign presentation showed worse recall to the phonologically and structurally similar lists, compared to control lists matched for frequency and item length. Signers who received the English lists showed significantly worse recall only for the phonologically similar list.

Hanson's study, from a bilingual standpoint, consists of two conditions – L1 recall of signs, and L2 recall of English word forms. If the L1 and L2 lexicons were separate (or at least privy to separate input), then we would expect to see signers who received the signed lists perform worse only with the structurally similar lists (akin to the *phonological similarity effect*), since the phonologically and orthographically related lists differed based on non-L1 features. However, this was not the case, which suggests that when native signers were presented with signs, they encoded the information in both L1 and L2 phonological structure. In other words, the perception and subsequent encoding of a sign activated structures in the L2, even though the two languages did not share modality. However, the results from the signers who received English word lists suggest something different – here, the input form only seems to disrupt recall in lists with relevant similarities, namely phonology of English.

Since Hanson was not intending to study the interaction between languages, she did not document characteristics of her subjects that could have affected the results, such as whether they were taught via total communication or oralist methodology[1]. The reason this is relevant is because it is unclear how the signers had access to the kind of distinctive phonological information found in English if they were unable to hear those distinctions. Previous work has focused on the functional equivalence of sign and speech in learning phonological information (Hanson, 1982; Leybaert, 2000; Miller, 2004), suggesting that underlying phonological representations can be equally tapped by sign or speech input. However, more recent work by McQuarrie and Parrila (2008) seems to suggest this is not the case; when rhyme judgments based on phonological information are contrasted with those based on visual or motor/tactile information, "phonological" facilitation is not seen.

There are, then, two possible explanations for Hanson's result. First, perhaps the participants in her study were trained in an oralist tradition, and did in fact have some motoric knowledge of the way in which to produce certain phonemes, and that phono-motor information was enough to map onto some underlying phonological structure. Another possibility is that the effects were caused by stimulus design – the list involving

[1] Total Communication, commonly referred to as Simultaneous Communication, refers to the practice of using various methods of communication (e.g., signing, writing, oral, etc.) to educate children. Oralism refers to the practice of teaching deaf children to pronounce spoken languages, and to understand spoken languages via lip-reading. Oralism, by default, involves teaching the child to produce spoken language, and Total Communication often involves an oral component.

phonologically similar items was simply more difficult by default. If we accept the latter explanation, we're forced to ignore the phonologically similar lists, and are left with the result that when the words are presented as signs, they are encoded as sign structures, and perhaps more troubling, that words presented orthographically result in *no specific encoding at all* (it is unclear why the orthographically related lists did not also show interference, given that they would appear to be the most salient). Regardless of which explanation is correct, we are left with a result that suggests single-language encoding in certain conditions, and therefore, separate lexicons.

Along with sign-text bilinguals, Dufour also mentions another understudied group – speech-sign bilinguals, who are more commonly referred to simply as bimodal bilinguals. Unfortunately, little work has been done to look at the processing of perceptual information in speech-sign bimodal bilinguals. In our lab, we are currently examining bimodal bilinguals who are fluent in ASL and spoken English to determine whether both languages are activated in parallel (Shook and Marian, in progress). More specifically, bimodal bilinguals are given a recognition task in which they are shown four images and asked to click on one of them while their eye movements are recorded. In experimental trials, two pictures (a target and a competitor) share three of the four phonological parameters in ASL (see Figure 2).

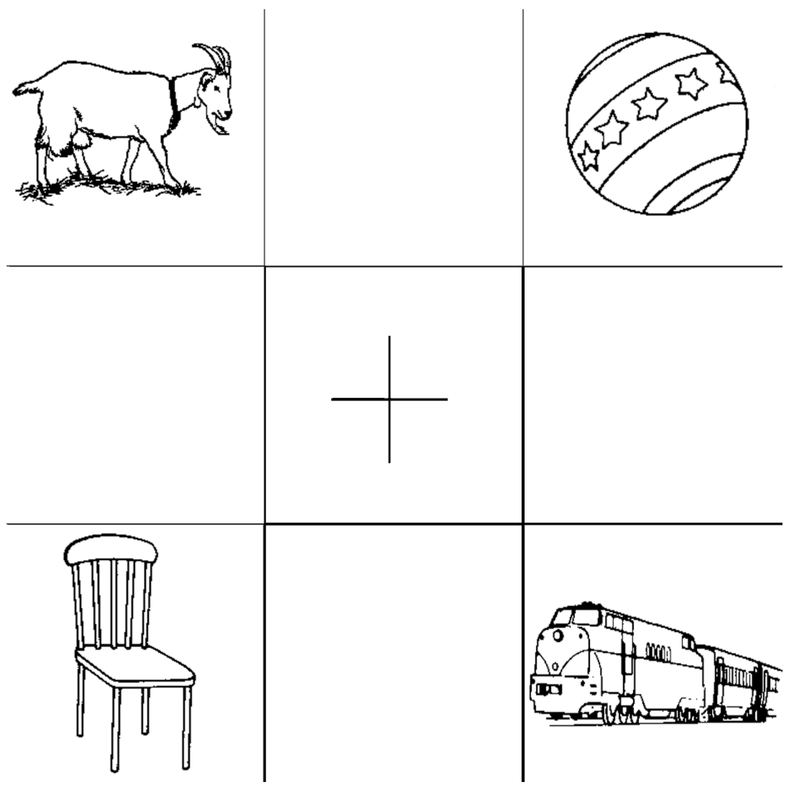

Figure 2. Example of eye-tracking display from Shook and Marian (in progress). The target item is "chair" which shares three of four parameters (handshape, location, orientation) with the sign for "train," differing only in motion.

Pilot data suggest that bimodal bilinguals look at competitor items that share phonological similarity to the target more than at semantically and phonologically unrelated control items. This implies that bimodal bilinguals activate both languages simultaneously during language comprehension. One interesting aspect of this result is the prediction that it makes regarding the mechanism with which the non-target language (ASL) is activated. Rather than relying on bottom-up information, the lack of phonological overlap between the two languages necessitates the inclusion of a top-down pathway in order to explain how ASL may be activated during a purely English task (see Figure 3).

The finding that bimodal bilinguals activate both languages simultaneously during comprehension is not surprising if we consider how audio and visual information is integrated generally, including in monolinguals and unimodal bilinguals. Likely the most famous account of audio-visual integration in monolingual speech recognition involves the McGurk effect (McGurk and MacDonald, 1976) in which an auditory signal, /ba/, is presented with a video of a speaker articulating the syllable /ga/, resulting in subjects reporting perception of the syllable /da/. Marian (2009) discusses how this integration affects bilingual language processing, suggesting that "as a word unfolds, incoming auditory input is combined online with incoming visual input and the two sources mutually interact to exclude options that are not plausible in at least one modality."

It could be argued that bimodal bilinguals are even better attuned than monolinguals or unimodal bilinguals at integrating information from separate modalities, because their processing systems are trained to do just that. In addition, research suggests that users of sign languages show greater visual attention in the periphery, and increased peripheral motion processing (Bavelier, Dye, and Hauser, 2006; Bavelier, Tomann, Hutton, Mitchell, Corina, Liu and Neville, 2000). If bimodal bilinguals are more attentive to the visual field due to their experience with a signed language, we might expect to see a greater influence of visual information in language processing.

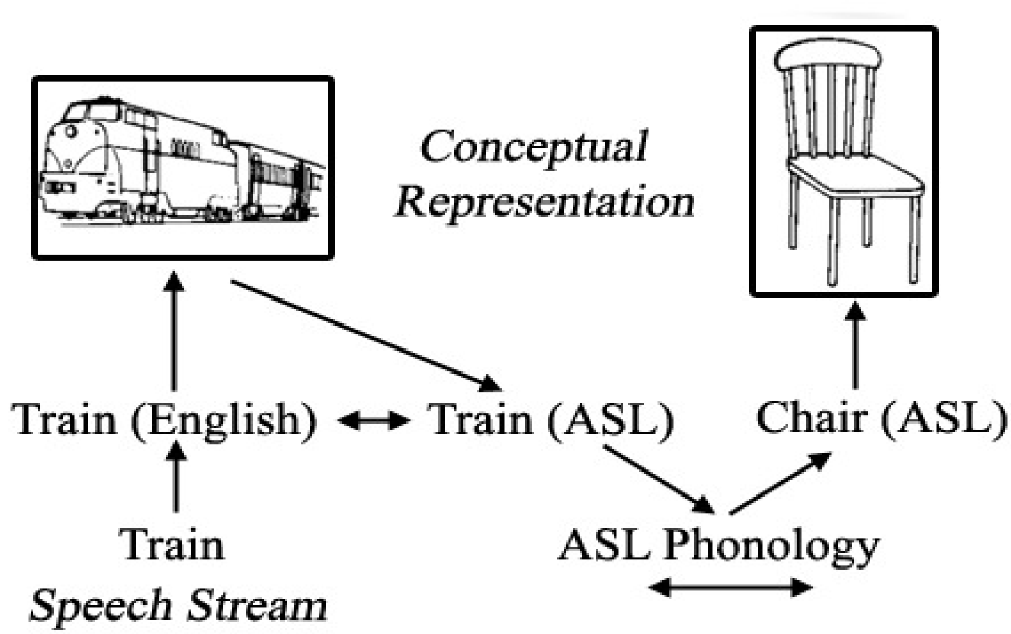

Figure 3. A graphical representation of top-down activation patterns in bimodal bilinguals activating both English and American Sign Language during a comprehension task.

A key difference between bimodal and unimodal audio-visual integration arises when one considers the mechanism by which lexical items are activated. In the unimodal case, the net activation of a lexical item is the sum of activation garnered by the phonological (or auditory information) as well as the extra, top-down activation provided by visual information. Conversely, bimodal bilinguals viewing congruous code-blends will activate lexical items in both languages that correspond to the same referent in the semantic system. Activation from both lexical items feeds to the semantic level and results in high activation of the semantic node, which can feed back to both lexical items, causing their activation levels to increase in turn. Note that this implies a significant top-down contribution to increased lexical activation. Not only that, it also implies that simultaneous parallel activation of lexical items between languages that do not share modality in bimodal bilinguals, when considering code-blend contexts, is due to bottom-up information.

Conversely, the audio-visual integration concept could work in one of two ways for bimodal bilinguals using only one of their languages. As previously mentioned, eye-tracking studies with unimodal bilinguals have shown that bilingual subjects performing an experiment in their L1 are more likely to look at competitor objects in a visual scene when the L2 lexical item for the competitor overlaps phonologically with the L1 target (e.g., Marian and Spivey, 2003). This is explained via bottom-up accounts where phonological information activates items from both lexicons, which is similar to the mechanism that drives parallel activation in bimodal bilinguals viewing code-blends. However, if bimodal bilinguals using only one language (e.g., spoken English) in these eye-tracking tasks also show parallel activation patterns, it must be due to a top-down mechanism. This can occur in two possible ways. First, as in the case of code-blends, activation from the phonological level feeds upward to the semantic level, which then feeds back down to corresponding lexical entries for both languages. Second, it could occur in the same fashion that Marian (2009) suggests audio-visual integration occurs. While the auditory signal is sending information up the processing chain, the visual information has supplied the conceptual information, which once again feeds back down to the corresponding lexical items in both languages, and to the cross-modal phonological information for both entries. This dichotomy is testable by virtue of the fact that each account makes different predictions. Should the information need to be fed up to the semantic level before feeding back to activate lexical entries in the non-target language, we should expect subjects to take more time to disambiguate between the target and the competitor relative to unimodal bilinguals, since the competition occurs later in the processing stream. If the second account is true, and the visual information is integrated at the conceptual level, which then feeds back down, we might expect the rate of disambiguation for bimodals and unimodals to be about equal.

With the increase of research on bimodal bilingualism in recent years, we are beginning to see comparisons between unimodal bilinguals and their bimodal peers. While there seem to be similarities in function between the two groups, there are also some differences in the mechanisms that underlie their language processing. The research outlined in the previous sections has established certain similarities between the processing patterns of bimodal and unimodal bilinguals, enabling us to consider how phenomena found in studies with unimodal bilinguals might manifest in bimodal bilinguals. In other words, how might the architecture of the language processing system vary? Do bimodal and unimodal bilinguals utilize the same information (e.g., bottom-up segmental information or top-down linguistic context) to the same degree? To begin to answer these questions, it may be useful to consider current models of bilingual language processing in light of bimodal bilingual research. In doing so, we may gain a deeper understanding of how and why bimodal bilinguals process language the way that they do, as well as how bilinguals in general utilize linguistic input.

3. Modeling Language Processing in Bimodal Bilinguals

3a. Modeling Language Production in Bimodal Bilinguals

One of the most immediate challenges of developing a model of language production in bimodal bilinguals is accommodating the ability to produce both languages simultaneously. This phenomenon greatly increases the complexity of the processing system. Consider the notion of lexical access. Rather than designing a system in which one target is chosen from many activated lexical items, the Emmorey et al. (2008) results suggest the need for a mechanism that is able to choose separate (but semantically related) lexical items from languages of variant modality for simultaneous production. Furthermore, the notion that the majority of the simultaneously-produced items are semantically related implies that prior to the lexical selection process, a shared conceptual store provides information to the different languages. Also, one needs to consider the way in which syntactic information interacts between the two languages, given how morpho-syntactically meaningful ASL facial expressions occur with English speech (Pyers and Emmorey, 2008) and how bimodal bilinguals' English constructions sometimes show influence from ASL syntax (Bishop and Hicks, 2005).

To our knowledge, only one model of bilingual language processing has been specifically proposed for bimodal bilinguals (see Figure 4). Emmorey et al. (2008) adapted Levelt's model of Speech Production (1989), and integrated it with a model of speech and gesture production proposed by Kita and Özyrük (2003). In the model proposed by Emmorey et al., the grammatical, phonological and lexical aspects of production of both ASL and English are separate but connected, and are all activated by a Message Generator, which relays conceptual information to the two languages. In other words, propositional or semantic information is sent to both languages simultaneously, and that conceptual unit is determined before lexical selection occurs. Each formulator then encodes the input based on the matrix language (which is the language that submits the syntactic frame). The input is encoded with English grammatical constructs when English is the matrix language and with ASL grammatical constructs when ASL is the matrix language. The result is that code-switched or code-blended productions which are of the non-matrix language are produced in accordance with the matrix language's grammar.

Kita and Özyrük's (2003) addition is the Action Generator, which is a general mechanism used for creating an "action plan." In this sense, it is not inherently linguistically biased. Rather, it generates movements in both real and imagined space, guided by knowledge of spatial features and information. Though independent of the Message Generator, the two interact; this can result in the gestures generated by the Action Generator being influenced by linguistic information. Casey and Emmorey (2008) suggested that bimodal bilinguals produced more iconic gestures and gestures from a character viewpoint than non-signers.

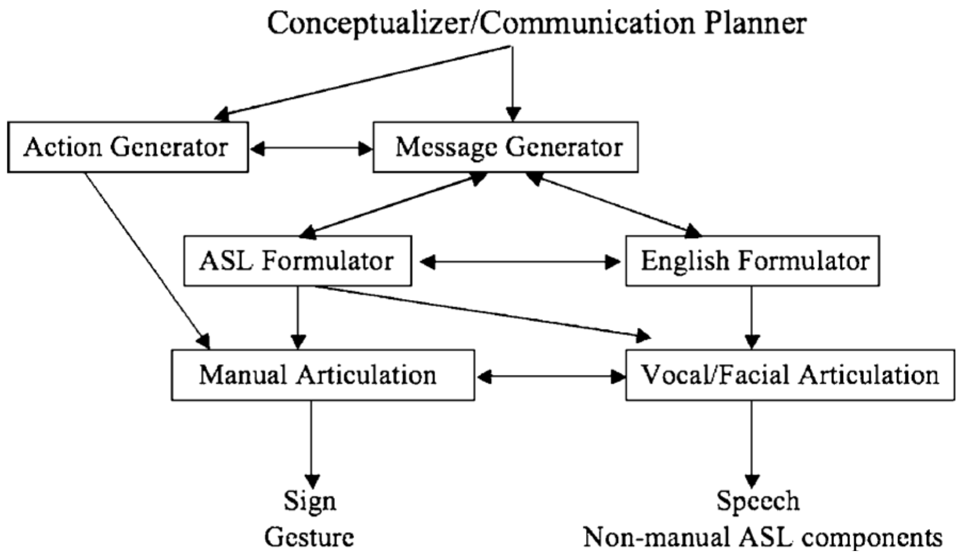

Figure 4. Model of Code-blend production in bimodal bilinguals as proposed in Emmorey, Borinstein, Thompson, and Gollan (2008). This model integrates Levelt's (1989) model of speech production with Kita and Özyrük's (2003) model of co-speech gesture.

They hypothesized that this was due to the interaction between the Action Generator and Message Generator. Feedback from the Formulator levels of the model (e.g., the ASL Formulator, which includes phonological and lexical information) to the Message Generator likely imbues the latter with information about how to encode the spatial properties of ASL. Furthermore, the authors suggested that this can prime the Action Generator to produce more iconic gestures even when the speaker is simply using English.

As a whole, the ASL-English code-blend model proposed by Emmorey et al. (2008) sufficiently explains the results shown in the code-blend literature. It also provides an insight into the timing of bilingual language production systems. While there is some debate about the level at which lexical selection is made in bilingual production (e.g., Finkbeiner, Gollan, and Caramazza, 2006), Emmorey et al.'s (2008) results suggest a late-selection mechanism. The majority of code-blended signs produced by bilingual participants in their study were semantic equivalents to the simultaneously produced English word (approximately 81%). If lexical selection occurred at an earlier stage, say with concept formation, one would expect either no code-blends (rather, code-switches), or more code-blends that produced unrelated, or non time-locked, signs (also see Casey and Emmorey, 2008).

However, it is unclear exactly how well this model can generalize to unimodal bilingual data, or conversely, whether that function is necessary. On the one hand, we can simply view the model as specific to bimodal production. In this scenario, we assume variant architectures for unimodal and bimodal bilinguals. This has the negative side-effect of rendering the support of late lexical selection via bimodal data useless when talking about unimodal bilinguals. However, it seems that bilingual production models for unimodal bilinguals are not dissimilar from the basic architecture of the ASL-English code-blend model, and as Emmorey et al. claim, differences seen between the outputs of unimodal versus bimodal productions are not due to systematic or architectural differences, but rather to the biological constraint of not being able to articulate two words at the same time that bimodal bilinguals are not subject to. In light of this notion of biological constraint perhaps masking the architectural or systematic nature of the bilingual language production system, one should examine models created to accommodate unimodal bilingual production patterns, and consider how well they are able to account for bimodal data. There are at least four such models of bilingual production, each of which provides a slightly different explanation for how bilinguals access their lexicons, specifically in regards to how they correctly choose items from the intended target language.

One such model was proposed by Costa, Miozzo and Caramazza (1999) and contends that lexical selection during bilingual speech production is language specific. In other words, when choosing a lexical item to produce, the production system has inherent knowledge of both that lexical item's language category, as well as the intended target language of the speaker, and the lexical selection mechanism will only choose members of the intended language. While this account suggests a language-specific lexical-level selection mechanism, the bimodal data seem to suggest otherwise. Selection likely occurs later in the stream, given that speakers produce simultaneous ASL and English fairly often (35.71% of all utterances in Emmorey et al., 1998, occurred with a code-blended sign). Also, if the lexical selection mechanism were language specific, we would not expect to see code-blends at all.

Another model of bilingual language production is Green's Inhibitory Control (IC) model (1998). Green proposed that rather than having a language specific selection mechanism that activates only one language at a time, conceptual information actually activates all candidates, regardless of language, and the non-target language is then suppressed through some inhibition mechanism (support for this notion often comes from studies showing increased inhibitory control in bilinguals even in non-linguistic tasks, see Bialystok, 1999; and Bialystok, Craik, Klein, and Viswanathan, 2004). In the IC model, since both lexicons are activated, the notion of language non-specific selection suggested by the bimodal data is supported. Furthermore, the IC model posits specific asymmetries in language activation in that a speaker's L1 is likely to be more strongly activated by the semantic system than a speaker's L2. Also, since the amount of suppression put forth by the inhibitory mechanism is directly proportional to the amount of activation, L1 words should be more strongly suppressed during L2 use than L2 words during L1 use. Emmorey et al. (2008) found that while producing English, bimodal bilinguals were highly likely to produce concurrent single-word ASL signs. However, when signing ASL, no examples of English single-word intrusions were found. In Emmorey et al.'s subjects, ASL was the L1, and English was the L2. If ASL is more suppressed during production of English, and English less suppressed during ASL production, we would expect to see the opposite result. Specifically, if L1 (ASL) is more strongly suppressed during L2 (English) use, and L2 less strongly suppressed during L1 use, then more English intrusions should be found during ASL productions than ASL intrusions during English production, which was not the case in Emmorey's results.

A third model of lexical selection in bilingual language production comes from La Heij (2005) who suggests that *concept* selection is more important than lexical selection. In this model, preverbal messages carry information about the speaker's intended language, as well as information about things like register, and the conceptual notion itself. La Heij's model presupposes language-specific selection (like Costa et al., 1999), while as previously mentioned, the bimodal data suggest language-non-specific selection. However, La Heij's model also supposes that the preverbal message at the semantic level is capable of activating related semantic concepts at the lexical level. One possible point of support comes from differences found in the rates of code-blending between Emmorey et al. (2008) and Casey and Emmorey (2008). In the former, bimodal bilingual subjects relayed stories to other bimodal bilingual subjects. In the latter, bilinguals relayed stories to monolinguals, and showed fewer code-blends than in the Emmorey et al. study. According to La Heij, this would be due to the fact that the preverbal message was more strongly balanced towards English in the Casey and Emmorey study, based on the subjects' knowledge that English was the only shared language.

However, this model also seems to predict symmetrical interference across languages. The preverbal message contains the conceptual idea and the intended target-language, which then activates the target language lexical representation, as well as the semantic translation-equivalent, to a lesser degree. It's intuitive to claim, then, the activation of the non-target language translation equivalent should be the same, regardless of which of the two languages is the intended language (at least in equally proficient bilinguals). This prediction should result in symmetrical interference between languages, which was not seen in the Emmorey et al. study.

More recently, Finkbeiner et al. (2006) proposed an account in which lexical selection occurs at the verbal output stage, rather than at the lexical stage, called the response-selection account. According to this account, well-formed phonological information about activated lexical items is held within an output buffer, and from those low-level phonological entries, some selection mechanism chooses the correct target. Within each entry in the output buffer is an item's phonological and lexical information (e.g., language identity and grammatical class). Intuitively, this is appealing in that it allows for language-non-specific activation, and posits a late-stage selection mechanism that can explain the parallel activation of ASL and English during production as shown in the studies by Emmorey and colleagues (which suggest that lexical items are chosen late during the language production process). The proposal presumes that items in the output buffer are examined serially, beginning with the response that comes first. So, for translation equivalents, semantic priming causes the target and the translation equivalent to be highly active and occur early within the list – the non-target language item is then quickly rejected based on the target-language mismatch.

If one considers only semantic priming, the mechanism by which bimodal bilinguals might code-blend is fairly straightforward. If the two items are both highly active, and the biological "one thing at a time" constraint is not present, both items may be produced. However, the response-output account suggests that the fact that the two items occur in different languages means that the non-target item (in this case, the ASL sign) should be rejected quickly, and not have time to maintain enough activation to reach the production stage. The response-output account does not explain what happens to an item when it is rejected. If activation decays slowly, high levels of activation may still cause the language system to produce a rejected item. If non-target lexical items are actively suppressed, then the account fails to explain the bimodal data. Even if we accept that the selection-mechanism can choose two possible referents for simultaneous production, we still run into the issue of how to deal with non-translation equivalent, non-target language distractors. Empirical findings suggest that these types of distractors cause slower response times in unimodal bilinguals - in other words, they take longer to be rejected. If they take longer to be rejected, then they should continue to gain activation over time, and be more likely to be produced simultaneously in bimodal bilinguals. This does not seem to be the case however, as the vast majority of code-blend productions are translation-equivalents (Emmorey et al. 2008).

In support of the serial-nature of the response-output model, there is some recent research to suggest that bimodal bilinguals may actually produce code-blends serially. Emmorey, Petrich and Gollan (2008) performed a picture-naming study where bimodal bilinguals were asked to produce an English word, an ASL sign, or both simultaneously. They found that for bimodal bilinguals, the time required to produce an English word *during a code-blend* was significantly longer than during the production of an English word or an ASL sign alone. This suggests that during a code-blend, the ASL structure is being constructed first, followed by the English structure. This seems to contrast with Emmorey et al.'s (2008) results that showed a time-lock between production of speech and signs during code-blends. One possible explanation is that the motor planning phase for each language production is different, and the production system is able to coordinate the signal so that they both occur temporally locked. However, if this were the case, we'd expect to see differences in timing for sign-alone and speech-alone productions, which we do not. Another possibility is that the delay of English production during code-blends is simply due to an over-taxed motor coordination system. Further research is required to tease apart these possibilities. While the response-selection proposal seems to cover a sizable portion of empirical incongruities found in other bilingual production models, neither it, nor the other models discussed in this chapter, are currently equipped to explain the bimodal data.

Though the majority of bilingual models, in both production and perception, tend to focus on phonological, orthographic and lexical levels (so called lower-order levels), there is also some evidence to indicate syntactic transfer across languages. Research suggests that Spanish-English (Hartsuiker, Pickering, and Veltkamp, 2004) and Dutch-English (Schoonbaert, Pickering, and Hartsuiker, 2007) bilinguals show priming of sentence structures across their two languages, such that the use of a syntactic structure in L1 can prime the use of that same structure in L2. This implies a cross-linguistic, integrated syntactic system where sentence structures that overlap in both languages can utilize lexical items from both languages. Currently, no bilingual models have been developed beyond the level of the lexicon and though many include links to the semantic system, the morphosyntactic system is not usually incorporated in these models.

There is also evidence to suggest similar syntactic integration in bimodal bilinguals. Bishop and Hicks (2005) provided samples of written English by children of deaf adults (CODAs) and found that their English constructions showed a good deal of influence from ASL grammatical structure (e.g. missing determiners, dropped subjects, etc.). Emmorey et al. (2008) found similar English constructions in some of their subjects as well. In addition, Pyers and Emmorey (2008) showed that when producing English sentences, bimodal bilinguals tended to produce grammatically relevant ASL facial expressions. These studies may suggest the possibility that the syntactic system is somewhat overlapping even when the modalities of a bilingual's two languages do not match. Future work will need to determine the nature of syntactic interaction in both unimodal and bimodal bilingual populations, in order to capture a more comprehensive picture of bilingual language processing.

3b. Modeling Language Perception in Bimodal Bilinguals

Arguably the most well developed model of bilingual language processing is the Bilingual Interactive Activation+ model (BIA+, see Figure 5).

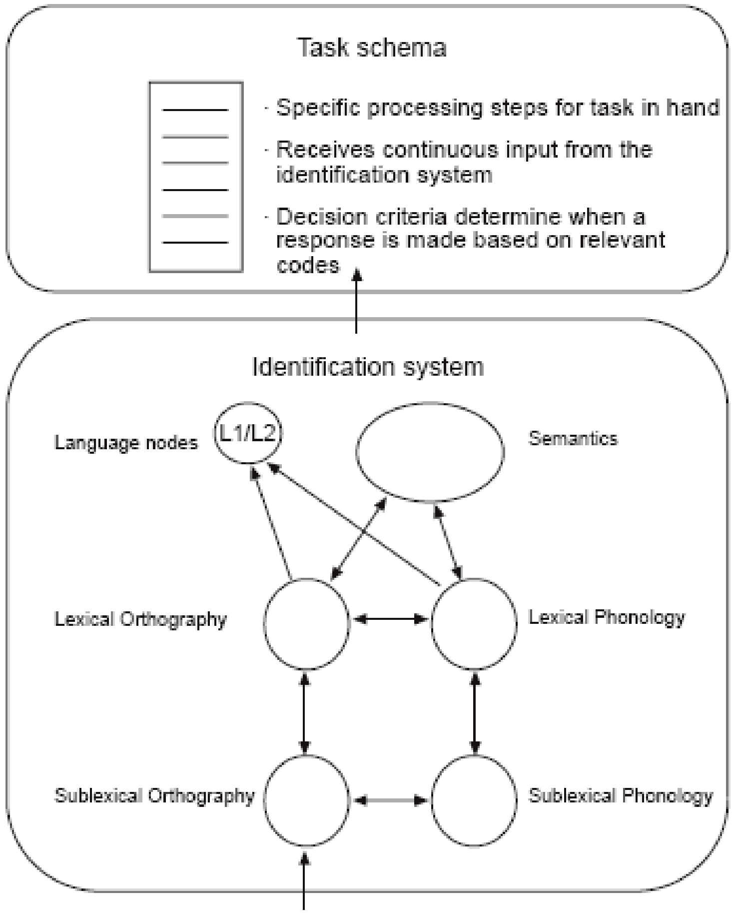

Figure 5. The Bilingual Interactive Activation+ model (BIA+) adapted from Dijkstra and Van Heuven (2002).

Initially a bilingual version of the Interactive Activation model (proposed by McClelland and Rumelhart, 1981), the BIA was created to explain how orthographic information is processed in bilinguals (Djikstra and van Heuven, 1998). The model consisted of an orthographic feature level that fed to a letter level. The letter level then provided weighted activation to lexical entries that shared orthographic information with the activated features, according to position (so, graphemes in initial position activated all words with that grapheme in initial position, and inhibited all entries that did not). At the word level, all words, regardless of language, were stored in one lexicon, and had lateral inhibition. This supported the notion of non-selective access, where words in both languages could be equally activated by bottom-up information. The final level consisted of language nodes, which represented each language in the bilingual. These nodes had two functions – first, they categorized words in the single lexicon as belonging to one language or the other. Secondly, summed activation levels from words of one lexicon could activate the language node they corresponded to, which they could further facilitate, and inhibit the words of the other language.

One major issue with the BIA, in its initial implementation, was its focus on purely orthographic stimuli. Dijkstra and Van Heuven (2002) proposed an extension to the BIA, labeled the BIA+, which included extensions for phonological and semantic representations via a separate model called SOPHIA (the Semantic, Orthographic, PHonological Interactive Activation model; see Figure 6). Language nodes are still present but no longer inhibit the non-target language - They are simply used to supply lexical items with a category. However, though SOPHIA adds the ability for BIA+ to process phonological information, the model still uses orthographic information as input. While it is capable of modeling the interactions between orthographic and phonological information, it is not very good at making predictions regarding solely phonological input. Due to this reliance on written language, BIA+ is limited in its ability to explain results found in bimodal bilingual studies using hearing signers.

At first glance, the BIA+ *may* be well equipped to explain the results found within the sign-text bilingual subgroup of bimodal bilinguals (see Hanson, 1982). Since one of a sign-text bilingual's languages relies entirely on orthographic information, the BIA+ model may be able to capture the interactions between the written language and the signed language.

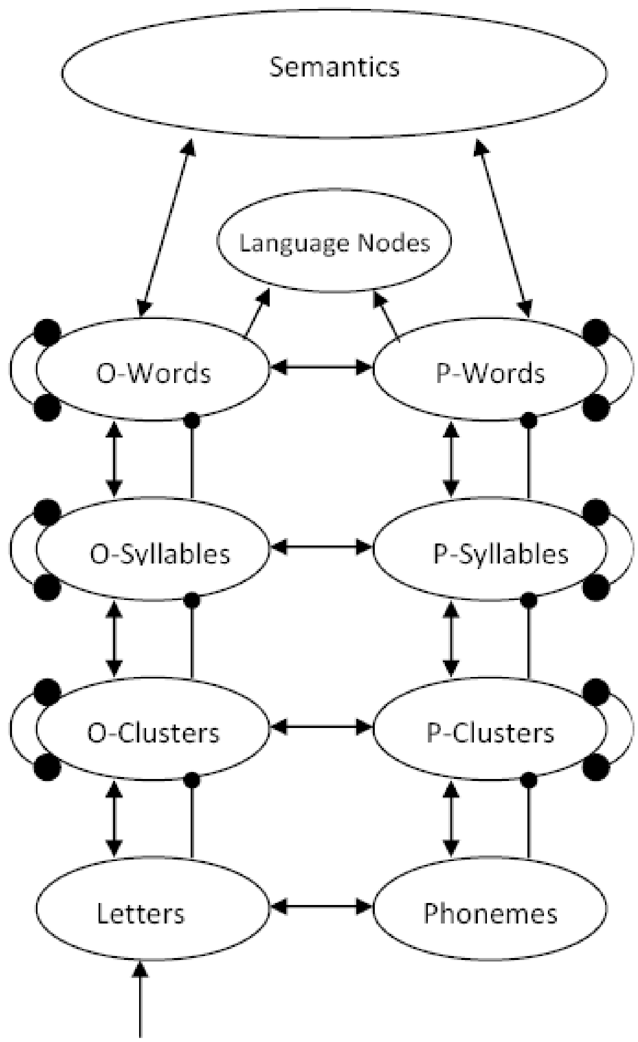

Figure 6. The Semantic, Orthographic, Phonological Interactive Activation model (SOPHIA) adapted from Dijkstra and Van Heuven (2002). Connections between orthographic and phonological pathways allow for interaction between the two domains.

However, further examination of results from sign-text bilingual studies suggests a problem for BIA+, as the data imply that the two lexicons of English written form versus ASL are separate, but capable of being encoded simultaneously when accessed via L1. The BIA+/SOPHIA combination actually posits a temporal delay between the processing of orthographic information and phonological information within L2 items relative to L1 items. In certain tasks, where understanding is guided by L1 orthographic codes, the L2 semantic or phonological information may not have time to influence the perception of L1 – in other words, we should see no L1-L2 interference when the task demands deem orthographic L1 input as most significant. Furthermore, the data from sign-text studies seem to advocate the notion of separate lexicons, which is in direct contrast to the integrated lexicon of the BIA+.

Yet, we must consider whether to view sign-text bilingualism in the same category as unimodal (speech-speech) bilinguals, or bimodal (sign-speech) bilinguals. The two languages in a sign-text bilingual are uniquely separated, in that one is housed entirely within the phonological (as it applies to the phonological properties of sign languages) realm while the other is exclusively orthographic. They also make use of entirely different grammars. This could result in weakened links between the orthographic and phonological pathways, affecting their interaction. In unimodal bilinguals, the systems and mechanisms utilized by the two languages are shared. In bimodal sign-speech bilinguals (e.g., ASL-English), lexical items in English could map to translation-equivalent lexical items in ASL via associative learning. This mapping could be generated initially by semantic links, but could also result in the development of lateral links between the two lexical systems (or modalities within the same lexical system). Since the orthography of language in the SOPHIA model is connected laterally to the phonology of the same language at every level of structure, one could then design a system where English orthography is able to connect to ASL phonology through an exclusively lateral chain. This suggests a greater similarity between unimodal and bimodal sign-speech bilingual processing systems than either unimodal or bimodal sign-speech systems with sign-text.

Perhaps the simplest way to circumvent the issues of the BIA+ is to change the input structure. Most models of language processing are based on single-modality input – often this is done to simplify the enormous task of modeling the entire language system. In order to accurately capture the pattern of bimodal bilingual processing, however, models of bilingual language processing require multiple input structures and need to re-envision the connections between them. Also, the input structures themselves will have to be altered. The BIA+ utilizes a positional system –orthographic input at the letter level activates lexical items whose orthography matches not only in character, but also in position. It isn't enough for two lexical items to share a "d" in their orthography, but that "d" has to occur in the same position in order for two words to coactivate. For sign, it is much more difficult to recognize a serialized structure, and so the sign input structure must likely do away with the positional form.

Still, the bimodal sign-speech bilingual system consists of distinctive language pathways. It would be difficult to imagine, due to the cross-modal nature of the bimodal bilingual, that the phonologic, lexical and feature levels represented in the BIA+ could be appropriately shared in a bimodal bilingual. The total lack of overlap at any level seems to preclude this notion. Instead, were we to attempt to adapt the BIA+ to bimodal results, one way to design it could be the inclusion of a third pathway that includes lateral links – phonology of L1, phonology of L2 and orthography of L2. The lateral links, as previously mentioned, could be borne from associative links via a semantic pathway (like the Hebbian notion of cells that fire together, wire together; Hebb, 1949). In fact, Emmorey et al. (2008) showed that bimodal bilinguals who produced code-blends (simultaneous production of sign and speech) did not produce signs that were propositionally different from the speech output – the signs were often closely related semantically, suggesting that the same conceptual notion from a shared semantic level activates items in both lexicons.

In order to better account for results from studies of sign-speech bimodal bilinguals, it may be more appropriate to look at models that consider auditory information as the primary source of input. One such model is the Bilingual Model of Lexical Access (BIMOLA, Grosjean, 2008; see Figure 7), a model of bilingual *speech* perception based on the TRACE model of speech perception (McClelland and Elman, 1986). The BIMOLA posits separate clusters of phonemes and words for each language, though they are housed within a single set at each level. This means that L1 words do not compete with L2 words during auditory recognition at the lexical level. This does not mean that the two languages cannot be activated in parallel; rather, the separation of language sets acts as a categorization tool like the language nodes found in the BIA+. The BIMOLA also has a "global language information" node that informs the system of contextual information.

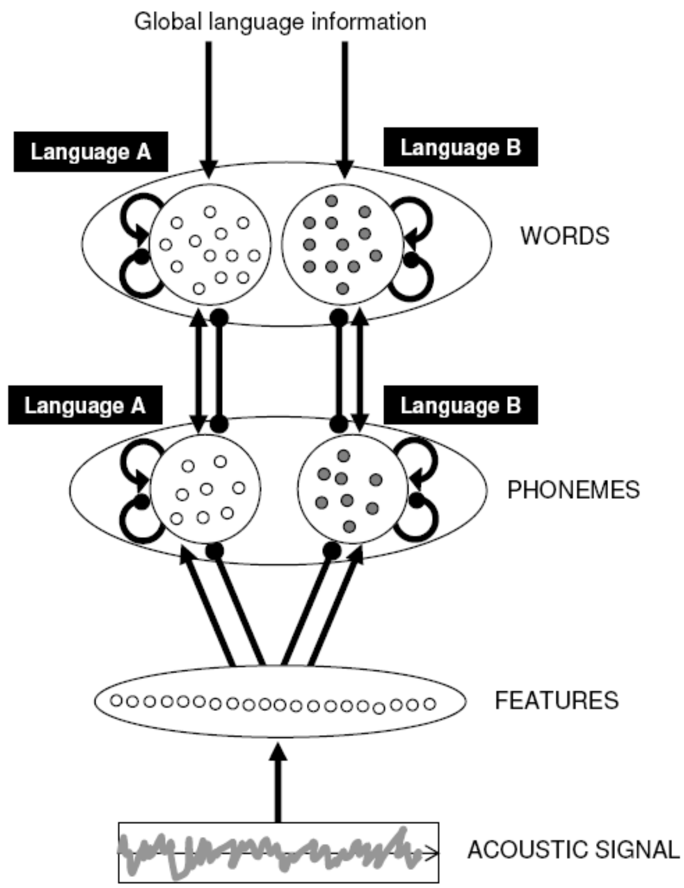

Figure 7. The Bilingual Model of Lexical Access (BIMOLA) as proposed in Thomas and Van Heuven (2005).

The most salient issue with the BIMOLA regarding its ability to cope with bimodal bilingual data is that its input is restricted to the auditory modality. However, it should be possible to change the modality of input features. Let us suppose then, that the BIMOLA were equipped with featural information for signed languages as well as spoken languages. Since the features of signed phonology and spoken phonology do not overlap, then we must immediately change the overlapping feature level to two separate feature levels that feed independently upward to the phonological level.

The fact that BIMOLA already posits separate lexicons and phonological levels is congruent with bimodal processing; however, there may be an issue in that the language-specific lexical and phonological sets are housed within the same larger set. BIMOLA does not clearly define how the set is categorized. If the set is based on shared featural information that guides the larger set of phonological entities (which would allow for shared phonemes across languages) then separate language clusters aren't enough – the model would require separate *sets* at each level.

This restructuring of the model has important implications for bimodal bilingual processing patterns during language comprehension. Since previous studies show parallel activation in unimodal bilinguals (Blumenfeld and Marian, 2007; Canseco-Gonzalez, et al., 2005; Marian and Spivey, 2003; Weber and Cutler, 2004;), it is possible that bimodal bilinguals also activate their two languages in parallel during language comprehension (e.g., Shook and Marian, in progress).

If this is the case, then the featural level, which is the presumed primary locus of parallel activation in unimodal bilinguals, cannot be the cause of parallel activation in bimodal bilinguals. Instead, it must be due to feedback from the semantic system, or connections between languages at a lexical level (see Figure 3).

BIMOLA does not include a semantic level in its architecture, nor does it allow for lateral activation or inhibition. Notably, the BIA+ also does not specifically include the semantic system in its architecture either, but both models posit a shared conceptual system (for support, see Finkbeiner, Nicol, Nakamura and Greth, 2002 and Li and Gleitman, 2002). This is, however, a fairly straightforward fix.

If we assume a shared conceptual store, then it is simply another level above the lexicon which, importantly, must include both feed-forward and feed-back connections in order to explain parallel processing in bimodal bilinguals.

Note that one model that was not included in the present discussion is the Self-Organizing Model of Bilingual Processing (SOMBIP, Li and Farkas, 2002). While the SOMBIP is uniquely qualified to look at the influence of developmental patterns on the structure of the bilingual lexicon, it currently makes no predictions regarding how a bilingual's two languages might interact (regardless of modality) or concerning the mechanisms that underlie language processing in a fully developed user. In summary, no perceptual model of bilingualism currently explains bimodal bilingual data. In order to account for these data, the models must be altered to include visual-linguistic input structures beyond orthography, and they require feedback systems to allow for top-down influence on the activation of lexical items, primarily for the unused language. Further research needs to be done, however, before we understand more fully the level of interaction between the two languages in a bimodal bilingual.

CONCLUSIONS

This chapter summarizes existing knowledge about language processing in bimodal bilinguals and compares phenomena found in bimodal research to those found in unimodal research in order to create a more complete model of bilingual language processing. The question of whether languages that vary in modality, such as ASL and English, are represented the same way in the brain is not a trivial one. It requires that the two languages be placed on an even playing field. In other words, one could argue that similar representation requires similar function. While on the surface, ASL seems vastly different from English, there are structural similarities. For example, the phonological systems in both languages are based on the combination of finite, meaningful parts. Perhaps more importantly, sign languages display many of the same features of spoken languages, such as phonological priming, an articulatory loop in working memory, and categorical perception. It is intuitive to believe that the functional similarities between the two languages may lead to representational similarities as well, where signed and spoken languages exist in the same language processing system. Furthermore, this overlap of two languages in one system should look much like that found in a unimodal system.

However, the processing patterns of bimodal and unimodal bilingual groups are not always congruous. For example, bimodal bilinguals code-blend (produce a sign and speech simultaneously) while unimodal bilinguals code-switch (produce lexical items from both languages at different times in the same sentence). It is possible that this difference between groups is due solely to the unimodal bilinguals' biological constraint of simply being unable to produce two spoken languages at once. If so, it is unclear whether the relationship between two languages in unimodal and bimodal bilinguals is similar. Results by Emmorey et al. (2008) suggest a semantic overlap between lexical items in code-blends, which implies that a single concept from the semantic system activates items from two lexicons. Bimodal bilinguals also seem to show increased occurrence of interfering morphosyntactic markers across languages (Pyers and Emmorey, 2008). These studies indicate that languages of different modalities seem to strongly interact in bimodal bilinguals.

The question of *how* exactly two languages of different modalities interact remains. One possibility is that language, regardless of modality, is represented similarly in the brain. Work by Emmorey and colleagues seems to suggest that although the modalities aren't shared, there is overlap between languages in bimodal bilinguals, much like we might expect to see in unimodal bilinguals. While spoken and signed languages vary in surface and structural aspects, research suggests that they are processed very similarly. In light of this, one could argue that the distinction between spoken and signed languages on a representational level is seamless – they utilize the same processes within the same architecture. In this account, the differences seen between the two groups may be based on two things – non-linguistic differences or constraints, and *degree* of processing. The first refers to surface level differences based on the nature of the language itself. For example, as we have discussed, bimodal bilinguals' code-blending is not necessarily based on processing differences, but the lack of a biological constraint found in unimodal bilinguals. The second refers to the degree to which certain mechanisms are utilized. As previously discussed, users of signed languages have been shown to have greater visual field perception (Bavelier et al. 2000), which could result in greater attention to visual detail during language processing. This suggests the possibility that visual and linguistic information are more strongly linked in users of signed languages, and subsequently, bimodal bilinguals. In this scenario, the structure of the processing system is not inherently different between unimodal and bimodal groups, but the degree to which certain mechanisms influence speech is distinct.

However, we must still entertain two different possibilities for *modeling* the bilingual language system. The first is the notion that perhaps bimodal bilinguals utilize an entirely different system than unimodal bilinguals. The differences found between bimodal bilinguals and unimodal bilinguals can suggest that the reason current models of bilingual language processing are unable to accommodate bimodal data is due to differences in the two systems. Indeed, arguments that explain differences between groups by positing separate processing patterns or architectures are used to compare modality differences within unimodal bilingual models as well. Thomas and Van Heuven (2005) explain the differences between the BIA+ and the BIMOLA by saying "the modelers implicitly assume that the different demands of recognition in each modality have led to different functional architectures." This is an immediately appealing argument, in that it allows researchers to parse up the different domains of language processing. Note that Thomas and Van Heuven are simply talking about the functional differences between the visual and auditory domain in *perception*. This says nothing about production, and the way in which the production system in bilinguals (as well as monolinguals) maps onto the receptive system.

Note that while this path of separate domain/separate task modeling is tempting, there are issues with implementation. It would be difficult to argue that separate domains within language processing do not interact in some way, whether it be through shared semantic representations or influence at lower levels of processing. Consider the process of audio-visual integration – even in *monolinguals*, there is an obvious combination of the auditory and visual modalities that influences speech perception. What this means is that, even if we model each domain separately, the different domains still need to be integrated in some way. For the sake of parsimony, it seems simpler to assume shared architecture from the first stages of development, and build upward. While creating a fully implemented model is an enormous task, it seems like the most cost-effective method of developing models that can handle data from bimodal bilingual studies, considering that these require a system that can integrate information from different domains or modalities.

This idea leads directly to the second possibility, that unimodal and bimodal bilinguals utilize the same architecture to process language. Indeed, studies of sign language suggest similar functional capacity between signed languages and spoken languages at a cognitive level. Furthermore, the bimodal studies suggest that bimodal bilinguals show similar language function as unimodal bilinguals, such as parallel processing of their two languages. While Emmorey et al. (2008) point out *differences* in the production patterns of bimodal bilinguals compared to unimodal bilinguals, the authors also provide evidence to suggest the differences are due to output ability (bimodal bilinguals can produce two lexical items simultaneously) rather than the structure of the system itself (since code-blends tended to be semantically linked).

With the similarities in processing between unimodal and bimodal groups in mind, rather than creating separate models for unimodal and bimodal bilingual groups, it may be most useful to develop models that are capable of adjudicating unimodal and bimodal bilingual data without forcing variant architectures. Perhaps the best method would be to consider localist models in which weights between nodes can be readily shifted not just between phonology and orthography, but also between separate phonological pathways for languages of different modalities. Furthermore, bilinguals show an influence of audio-visual integration, and models of language processing should be capable of implementing that integration. With language processing architecture that includes connections between modalities, both in terms of input (visual or auditory) and language (L1 or L2), one may be able to make systematic predictions about the way bilinguals process their two languages, regardless of modality, simply by varying the interactions between domains instead of the structure of the system itself.

By looking at bimodal bilinguals, it becomes clear that computational models require the ability to accommodate and integrate information from a variety of modalities and domains. This necessitates the development of more harmonious models which take into account both the functional convergence and structural divergence found in bimodal bilingual processing patterns, and consider language processing not to be domain-specific, but experience-based, in the sense that the underlying mechanisms that govern language are the same, but how they are used varies dependent on the type of input the system receives. Furthermore, by studying bimodal bilinguals, we may begin to gain a clearer picture of how language processing is affected by cross-modal input. If we are to expand our understanding of bilingual language processing, we must take into account those groups (like bimodal bilinguals) who are on the periphery of current research, in order to widen the boundaries of our knowledge about what the bilingual language system is capable of. In doing so, we may more fully understand bilingual language processing, and be better equipped to describe the mechanisms that govern language.

AUTHOR NOTE

The authors would like to thank James Bartolotti, Caroline Engstler, Jenna Luque and Scott Schroeder for comments and feedback on earlier drafts of this manuscript.
Correspondence should be directed to Anthony Shook at a-shook@northwestern.edu or at the Department of Communication Sciences and Disorders, Northwestern University, 2240 Campus Drive, Evanston IL 60208.

REFERENCES

Baker, S.A., Idsardi, W.J., Golinkoff, R.M., and Pettito, L. (2005). The perception of handshapes in American Sign Language. *Memory and Cognition, 33(5),* 887-904.

Baker, S.A., Golinkoff, R.M., and Pettito, L. (2006) New insights into old puzzles from infants' categorical discrimination of soundless phonetic units. *Language Learning and Development, 2(3),* 147-162.

Bavelier, D., Tomann, A., Hutton, C., Mitchell, T., Corina, D., Liu, G., and Neville, H. (2000). Visual attention to the periphery is enhanced in congenitally deaf individuals. *Journal of Neuroscience, 20,* 1-6.

Bavelier, D., Dye, M.W.G., and Hauser, P.C. (2006). Do deaf individuals see better? *Trends in Cognitive Sciences, 10(11),* 512-518.

Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Development, 70,* 636-644.

Bialystok, E., Craik, F.I.M., Klein, R., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging, 19,* 290-303.

Bishop, M., and Hicks, S. (2005). Orange eyes: Bimodal bilingualism in hearing adults from deaf families. *Sign Language Studies, 5(2),* 188-230.

Blumenfeld, H., and Marian, V. (2007). Constraints on parallel activation in bilingual spoken language processing: Examining proficiency and lexical status using eye-tracking. *Language and Cognitive Processes, 22(5),* 633-660.

Bonvillian, J.D., and Folven, R.J. (1993). Sign Language Acquisition: Developmental Aspects. In M. Marschark and M.D. Clark. (Eds.), *Psychological Perspectives on Deafness,* pp. 229-265. Hillsdale, NJ: Lawrence Erlbaum Associates.

Brentari, D. (1998). *A Prosodic Model of Sign Language Phonology*. MIT Press.

Canseco-Gonzalez, E. (2005). "Carpet or Carcel" effects of speaker type, fluency, and language mode on bilingual lexical access. *Abstracts of the 5th International Symposium on Bilingualism*, pp. 156-157.

Casey, S., and Emmorey, K. (2008). Co-speech gesture in bimodal bilinguals. *Language and Cognitive Processes*, 1-23.

Chomsky, N. (2000). *New Horizons in the Study of Language and Mind*. Cambridge University Press.

Colomé, A. (2001). Lexical activation in bilinguals' speech production: Language-specific or language-independent? *Journal of Memory and Language, 45(4),* 721-736.

Costa, A. and Caramazza, A. (1999). Is lexical selection in bilingual speech production language-specific? Further evidence from Spanish-English and English-Spanish bilinguals. *Bilingualism: Language and Cognition, 2,* 231-244.

Costa, A., Miozza, M. and Caramazza, A. (1999). Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection? *Journal of Memory and Language, 41,* 365-397.

Dijkstra, T. and van Heuven, W. (1998). The BIA model and bilingual word recognition. In J. Grainger and A.M. Jacobs (Eds.), *Localist Connectionist Approaches to Human Cognition* (pp. 189-225). Mahwah, NJ: Erlbaum.

Dijkstra, A., and Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5,*175-197.

Dufour, R. (1997). Sign language and bilingualism: Modality implications for bilingual language representation. In J.F. Kroll and A. M. B de Groot (eds.), *Tutorials in Bilingualism: Psycholinguistic Perspectives.* pp. 301-330. Mahwah, NJ: Lawrence Erlbaum Associates.

Dye, M. W. G., and Shih, Shui-I. (2006). Phonological priming in British Sign Language. In L. M. Goldstein, D. H. Whalen and C. T. Best (Eds.), *Laboratory Phonology 8: Varieties of Phonological Competence*. Berlin: Mouton de Gruyter.

Emmorey, K., McCullough, S., and Brentari, D. (2003). Categorical Perception in American Sign Language. *Language and Cognitive Processes, 18(1),* 21-45.

Emmorey, K., Borinstein, H.B., Thompson, R., and Gollan, T.H. (2008). Bimodal bilingualism. *Bilingualism: Language and Cognition, 11(1),* 43-61.

Emmorey, K., Petrich, J., and Gollan, T.H. (2008). Simultaneous production of ASL and English implies serial access in production. Presented at the *49th Annual Meeting of the Psychonomic Society*, November 16th, 2008, Chicago IL.

Finkbeiner, M., Nicol, J., Nakamura, K., and Greth, D. (2002). The role of language in memory for actions. *Journal of Psycholinguistic Research, 31(5),* 447-457.

Finkbeiner, M., Gollan, T.H., and Caramazza, A. (2006). Lexical access in bilingual speakers: What's the (hard) problem? *Bilingualism: Language and Cognition, 9(2),* 153-166.

Gathercole, S. E., and Baddeley, A. D. (1993). *Working Memory and Language* . Hillsdale, NJ: Erlbaum.

Goldinger, S.D., Luce, P.A., Pisoni, D.B., and Marcario, J.K. (1992). Form-based priming in spoken word recognition: The roles of competition and bias. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *18*, 1210–1238.

Grainger, J. (1993). Visual word recognition in bilinguals. In R. Schreuder and B. Weltens (Eds.), *The Bilingual Lexicon* (pp. 11–26). Amsterdam: John Benjamins.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1*, 67-82.

Grosjean, F. (1997). Processing mixed languages: Issues, findings and models. In A.M.B. de Groot and J.F. Kroll (Eds.), *Tutorials in Bilingualism: Psycholinguistic Perspectives* (pp. 225-254). Mahwah, NJ: Erlbaum.

Grosjean, F. (2008). *Studying bilinguals.* Oxford: Oxford University Press.

Hartsuiker, R.J., Pickering, M.J., and Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals. *Psychological Science, 15(6),* 409-414.

Hanson, V. (1982). Short-term recall by deaf signers of American Sign Language: Implications of encoding strategy for order recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 572-583.

Hebb, D.O. (1949) *The Organization of Behavior*. New York: Wiley.

Hildebrandt, U., and Corina, D. (2002). Phonological similarity in American Sign Language. *Language and Cognitive Processes, 17(6),* 593-612.

Kita, S. and Özyrük, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language, 48,* 16-32.

Krakow, R., and Hanson, V. (1985). Deaf signers and serial recall in the visual modality: Memory for signs, fingerspelling and print. *Memory and Cognition, 13,* 265-272.

La Heij, W. (2005). Selection processes in monolingual and bilingual lexical access. In J.F. Kroll and A.M.B de Groot (eds.), *Handbook of Bilingualism: Psycholinguistics Approaches,* pp. 289-307. New York: Oxford University Press.

Levelt, W.J.M. (1989). *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.

Leybaert, J. (2000). Phonology acquired through the eyes and spelling of deaf children. *Journal of Experimental Child Psychology, 75(4),* 291-318.

Li, P. and Farkas, I. (2002). A self-organizing connectionist model of bilingual processing. In R. Heredia and J. Altarriba (Eds.), *Bilingual Sentence Processing,* pp. 59-85. North Holland: Elsevier.

Li, P., and Gleitman, L. (2002). Turning the tables: Language and spatial reasoning. *Cognition, 83,* 265-294.

Liddell, S. (1980). *American Sign Language Syntax.* The Hague: Mouton.
Marian, V. (2009). Audio-visual integration during bilingual language processing. In A. Pavlenko (Ed.), *The Bilingual Mental Lexicon*. Clevedon, UK: Multilingual Matters.
Marian, V., and Spivey, M. (2003). Bilingual and monolingual processing of competing lexical items. *Applied Psycholinguistics*, *24*, 173-193.
Marslen-Wilson, W.D. (1990). Activation, competition, and frequency in lexical access. In G. T. M. Altmann (Ed.), *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives.* Cambridge, MA: MIT Press.
Marslen-Wilson, W. D., and Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology, 10*, 29-63.
McClelland, J. L., and Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of Basic Findings. *Psychological Review, 88,* 375-407.
McClelland, J. L., and Elman, J. L. (1986). The TRACE Model of Speech Perception. *Cognitive Psychology, 18*, 1-86.
McGurk, H. and MacDonald, J. (1976). Hearing lips and seeing voices: A new illusion. *Nature, 264*, 746-748.
McMurray, B., Tanenhaus, M.K., and Aslin, R.N. (2002). Gradient effects of within-category phonetic variation on lexical access. *Cognition, 86(2),* 33-42.
McQuarrie, L., and Parrila, R. K. (2008). Deaf children's awareness of phonological structure: Rethinking the 'functional-equivalence' hypothesis. *Journal of Deaf Studies and Deaf Education,* doi:10.1093/deafed/enn025.
Meuter, R.F.I, and Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs in language selection. *Journal of Memory and Language, 40*, 25-40.
Miller, P. (2004). Processing of written words by individuals with prelingual deafness. *Journal of Speech, Language, and Hearing Research, 47,* 979-989.
Petitto, L.A., and Marentette, P.F. (1991). Babbling in the manuel mode: Evidence for the ontogeny of language. *Science, 251,* 1493-1496.
Praamsta, P., Meyer, A.S., and Levelt, W.J.M. (1994). Neuropsychological manifestations of phonological processing: Latency variation of a negative ERP component timelocked to phonological mismatch. *Journal of Cognitive Neuroscience, 6(3),* 204-219.
Pyers, J.E., and Emmorey, K. (2008). The face of bimodal bilingualism: Grammatical markers in American Sign Language are produced when bilinguals speak English to monolinguals. *Psychological Science, 19(6),* 531-536.
Schoonbaert, S., Hartsuiker, R.J., and Pickering, M.J. (2007). The representation of lexical and syntactic information in bilinguals: Evidence from syntactic priming. *Journal of Memory and Language, 56(2),* 153-171.
Shook, A., and Marian, V. (in progress). Parallel processing in bimodal bilinguals. Manuscript in Preparation.
Slowiaczek, L.M., and Pisoni, D.B. (1986). Effects of phonological similarity on priming in auditory lexical decision. *Memory and Cognition, 14(3),* 230-237.
Spivey, M., and Marian, V. (1999). Crosstalk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science, 10,* 281–284.
Stokoe, W. (1960). Sign language structure: An outline of the visual communication systems of the American Deaf, *Studies in Linguistics, Occasional Papers 8*. Available from Silver Spring, MD: Linstok Press.

Supalla, T., and Webb, R. (1995). The grammar of International sign: A new look at pidgin languages. In K. Emmorey and J. Reilly (eds.), *Sign, Language and Space,* pp. 333-352. Hillsdale, NJ: Lawrence Erlbaum Associates.

Sutton-Spence, R. and Woll, B. (1999). *The Linguistics of British Sign Language.* Cambridge, U.K.: Cambridge University Press.

Thomas, M.S.C., and van Heuven, W.J.B. (2005). Computational models of bilingual comprehension. In J. F. Kroll and A. M. B. de Groot (Eds.), *Handbook of Bilingualism: Psycholinguistic Approaches.* Oxford: Oxford University Press.

Weber, A., and Cutler, A. (2004). Lexical competition in non-native spoken- word recognition. *Journal of Memory and Language, 50*, 1-25.

Werker, J.F., and Tees, R.C. (1984). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America, 75(6),* 1866-1877.

Wilson, M., and Emmorey, K. (1997). A visuospatial "phonological loop" in working memory: evidence from American Sign Language. *Memory and Cognition, 25,* 313-320.

Wilson, M., and Emmorey, K. (1998). A "word length effect" for sign language: Further evidence on the role of language in structuring working memory. *Memory and Cognition, 26,* 584-590.