BILINGUAL EDUCATION & BILINGUALISM
Series Editors: Nancy H. Hornberger, *University of Pennsylvania, USA* and Colin Baker, *Bangor University, Wales, UK*

The Bilingual Mental Lexicon
Interdisciplinary Approaches

Edited by
Aneta Pavlenko

MULTILINGUAL MATTERS
Bristol • Buffalo • Toronto

Library of Congress Cataloging in Publication Data
A catalog record for this book is available from the Library of Congress.

British Library Cataloguing in Publication Data
A catalogue entry for this book is available from the British Library.

ISBN-13: 978-1-84769-125-5 (hbk)
ISBN-13: 978-1-84769-124-8 (pbk)

Multilingual Matters
UK: St Nicholas House, 31-34 High Street, Bristol BS1 2AW, UK.
USA: UTP, 2250 Military Road, Tonawanda, NY 14150, USA.
Canada: UTP, 5201 Dufferin Street, North York, Ontario M3H 5T8, Canada.

The policy of Multilingual Matters/Channel View Publications is to use papers that are natural, renewable and recyclable products, made from wood grown in sustainable forests. In the manufacturing process of our books, and to further support our policy, preference is given to printers that have FSC and PEFC Chain of Custody certification. The FSC and/or PEFC logos will appear on those books where full certification has been granted to the printer concerned.

Typeset by Datapage International Ltd.
Printed and bound in Great Britain by the Cromwell Press Ltd.

Contents

Chapter 3

Audio-visual Integration During Bilingual Language Processing

VIORICA MARIAN

Introduction

Despite the fact that bilingual language processing generally takes place in complex natural environments with multiple simultaneous modes of input, laboratory studies of bilingualism rarely consider the multimodal nature of bilingual language comprehension. For bilingual as well as monolingual listeners, multimodal input includes both linguistic input (such as the phonological and orthographic forms of a word), as well as nonlinguistic input (such as seeing the face of the speaker or the objects in the listener's immediate surrounding environment). While the interaction between phonology and orthography has received consideration in psycholinguistic studies of bilingual language processing, research on cross-modal audio-visual sensory integration in bilinguals is more limited. In general, audio-visual integration in bilinguals has yet to receive careful consideration in the literature. The objective of this chapter is to contribute to the understanding of how bilingual language processing is impacted by multimodal integration and audio-visual interaction. To accomplish this goal, two bodies of literature are considered. The first incorporates paradigms that consider multimodal integration of auditory and visual input during language comprehension and discusses them within the context of bilingualism. The second focuses on evidence suggesting that orthographic and phonological information interact during bilingual spoken and written language comprehension. The chapter also reviews existing models of bilingual language processing that may be able to account for audio-visual interaction, and draws implications for representation in the bilingual mental lexicon.

1. Audio-visual Integration During Language Processing

One of the most striking phenomena illustrating the interaction between auditory and visual modalities during spoken language comprehension is the McGurk effect (McGurk & MacDonald, 1976). The McGurk effect refers to the finding that if listeners are played the sound /ba/ auditorily, while presented visually with a face that is pronouncing /ga/, they will report hearing the sound /da/. As counterintuitive as this seems, this effect is very robust and consistent (e.g. Van Wassenhove *et al.*, 2007) and, in fact, many lecturers now include it as an easy-to-replicate demo in their undergraduate courses. The McGurk effect demonstrates that listeners constantly combine input from the two modalities when perceiving language. In general, in healthy individuals, language processing is cross-modal, with input from the auditory and visual modalities interacting as one hears, reads, writes or pronounces words.

The cross-modal nature of language processing can be traced to the cortical level. For example, cortical areas typically associated with speech processing have been found to be active during observation of visual articulatory movement (Calvert *et al.*, 1997, 2000; Campbell *et al.*, 2001; Nishitani & Hari, 2002). Similarly, neural responses to speech sounds in a modality-specific region of the auditory cortex appear to be modified by the simultaneous visual presentation of letters (e.g. Van Atteveldt *et al.*, 2004). Integration of auditory and visual information appears to happen so early in the processing stream that it can be detected even at the level of the brainstem (Musacchia *et al.*, 2006). Musacchia and associates (2006) found that seeing facial movement (lip-reading) changed the amplitude of the brainstem response to acoustic speech and suggested that the brain showed enhanced responses to auditory-visual stimuli relative to the sum of unimodal responses.

The role of visual input during spoken language processing is clear when one considers findings that visual input alone, in the absence of any auditory information, can be sufficient for listeners to determine what language a speaker is using. For instance, Ronquest and Hernandez (2005) asked participants to watch visual-only speech stimuli and decide if the language spoken was Spanish or English. They found that listeners were able to identify the language correctly well above chance, simply by watching the speaker's face. Moreover, speech perception is improved not only by viewing the mouth/lip movements, but also by having access to a view of the top of the head only (e.g. Davis & Kim, 2006) and to head movements (e.g. Munhall *et al.*, 2004). Although these latter effects are smaller than those observed for watching lip movements, it is striking that such visual information significantly enhances language comprehension. This suggests that listeners are adept at perceiving visual input during language processing, and integrate it with auditorily perceived input.

Interestingly, while both monolinguals and bilinguals can distinguish a language by visual correlates of speech alone, bilinguals seem to do so with greater accuracy than monolinguals, but only for languages they know (Soto-Faraco *et al.*, 2007). Soto-Faraco and associates (2007) found that Spanish–Catalan bilinguals were very successful at discriminating Spanish and Catalan based on visual information alone (watching the speaker's face), Spanish monolinguals were able to do so less successfully, while Italian–English bilinguals were unable to differentiate Spanish and Catalan. This suggests that previous experience is instrumental in shaping the way listeners recognize and rely on visual information during spoken language comprehension. This experience is likely largely automatic and not consciously monitored by the speaker or listener. Because bilinguals' experiences are more varied and extensive, they may be more adept at interpreting input from the visual modality during language processing and may be more skilled at relying on both modalities during language comprehension.

This suggestion is consistent with findings from the speech perception literature. Consider, for instance, a study by Sumby and Pollack (1954) who found that reliance on visual properties during speech processing increases as the signal-to-noise ratio (SNR) decreases (i.e. as the speech signal becomes weaker in relation to the noise around it). Viewing a speaker's articulatory movements is known to particularly benefit a listener's ability to understand speech in noisy environments. The gain from viewing visual articulation relative to an auditory-alone condition can be up to threefold under certain SNRs (Ross *et al.*, 2007). Specifically, Ross and associates (2007) found that multisensory integration was maximally beneficial at intermediate SNR levels. Where exactly bilingual language comprehension falls on the SNR continuum is likely influenced by many factors, such as language proficiency, degree of accented speech, similarity between the two languages, etc. However, overall, the SNR is likely to be lower in bilinguals (especially in their weaker language) because their decreased ability to perceive sound contrasts in a non-native language may yield a larger 'noise' category (e.g. Bradlow & Bent, 2002; Bradlow & Pisoni, 1999). As a result, bilingual and multilingual listeners, especially those with lower language proficiency, may be more likely to rely on the speaker's face when listening to speech, in order to boot-strap information that they cannot access auditorily. This suggestion is consistent with the 'inverse effectiveness' principle of multisensory integration, according to which the gain from viewing articulatory movements is most pronounced when auditory input is decreased relative to noise (Ross *et al.*, 2007).

Taking into account SNRs, spoken language comprehension in bilinguals is likely to rely even more on cross-modal integration than in monolinguals, with visual information supplementing auditory input. This greater reliance on multimodal input when processing auditory information may explain why non-native speakers sometimes experience greater difficulties comprehending their less-proficient language over the telephone and why native speakers have greater difficulty comprehending accented speech over the telephone. In fact, the argument that over-the-telephone conversations are negatively impacted by the speaker's linguistic background have been used to deny employment to non-native speakers and have been upheld in court rulings (e.g. *Clau v. Uniglobe* and *Guillen v. Dufour,* in Munro, 1998), with the tribunal commissioner asserting that 'he knew from personal experience that accented speech was hard to understand on the phone' (Munro, 1998: 141). Although much of the evidence on intelligibility over the telephone is anecdotal, it is consistent with evidence that intelligibility of foreign speech is negatively impacted by noise (e.g. Munro, 1998).

Consequently, it might be argued that speakers of multiple languages may be especially likely to rely on and be impacted by cross-modal integration when processing linguistic input. Previous research shows that visual speech information can even be used to help distinguish sounds in a second language that cannot be perceived using auditory input alone. For instance, Navarra and Soto-Faraco (2007) have found that adding visual information about lip movements could enhance Spanish speakers' ability to distinguish between the Catalan sounds /ɛ/ and /e/. Specifically, while Spanish-dominant bilinguals were unable to distinguish between the two Catalan sounds based on auditory-only information, they were able to perceive the contrast when visual information about lip movements was added (Catalan-dominant bilinguals were able to differentiate between the two sounds in both conditions). The authors concluded that visual speech information enhances second language perception by way of multisensory integration. This idea that nonproficient speakers rely more on cross-modal input when processing language is consistent with evidence from the second language acquisition literature showing that bimodal input benefits implicit and explicit memory (Bird & Williams, 2002), speaking performance (Borras & Lafayette, 1994) and language learning (Danan, 1992; Vanderplank, 1988).

Another line of convincing evidence supporting the cross-modal nature of bilingual spoken language processing comes from psycholinguistic studies using eye tracking. Initially, the eye-tracking methodology (e.g. Tanenhaus *et al.*, 1995) was adapted for use with bilinguals to test whether bilinguals process their two languages sequentially or in parallel (e.g. Marian, 2000; Marian & Spivey, 2003a, 2003b; Spivey & Marian, 1999). In a typical experimental set-up, bilinguals were presented with a display of objects, including one whose name (e.g. *chair*) overlapped across languages with the name of another object in the display (e.g. *cherepaha*, Russian for 'turtle') (Figure 3.1). When instructed to pick up the *chair*, Russian–English bilinguals made eye movements to the *cherepaha* significantly more often than to control objects. Such findings suggest that, during early stages of processing, unfolding auditory input activates multiple word alternatives within and between languages. The influence of sublexical phonological overlap on word processing remains

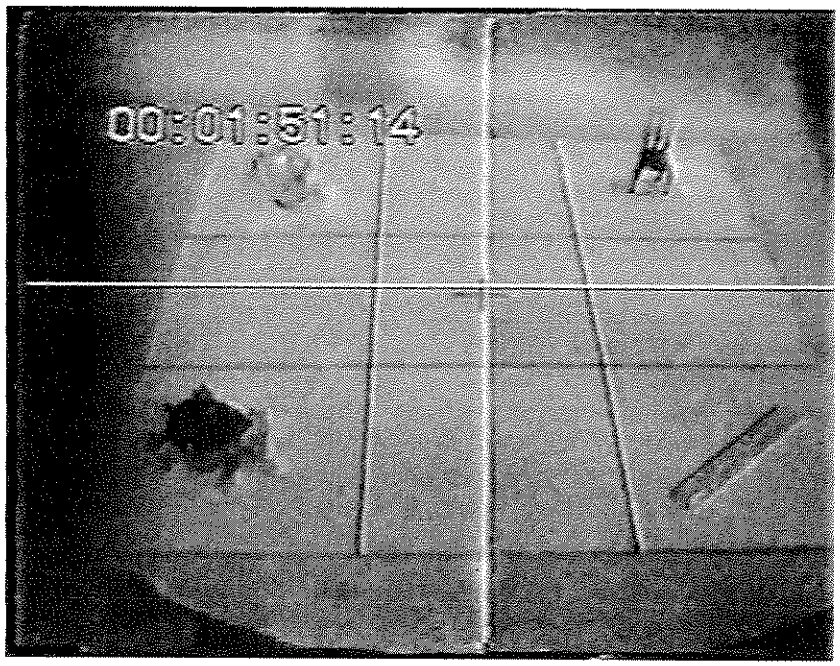

Figure 3.1 Sample display from an eye-tracking study of spoken word recognition in bilinguals. The display shows a light bulb, a chair, a ruler and a turtle. The Russian word for turtle is *cherepaha* (the 'che' overlaps with 'chair'). Studies show that when instructed in English to pick up the chair, Russian–English bilinguals frequently look at the turtle. Similarly, when instructed in Russian to pick up the turtle, Russian–English bilinguals frequently look at the chair. Monolinguals do not show these effects.

apparent until the lexical decision stage. Such gradual activation, leading up to word selection, is consistent with auditory word recognition models (e.g. Luce & Pisoni, 1998; Marslen-Wilson, 1987; McClelland & Elman, 1986) and indicates that a bilingual's two languages remain active in parallel and that bilinguals simultaneously map phonemic input onto both of their lexicons (with it cascading to higher levels of representation, e.g. Blumenfeld & Marian, 2007; Marian *et al.*, 2008) as a word unfolds in real time.

These findings have since been replicated and extended in eye-tracking studies with Dutch–English bilinguals (Weber & Cutler, 2004), Japanese–English bilinguals (Cutler *et al.*, 2006), French–German bilinguals (Weber & Paris, 2004) and Spanish–English bilinguals (Canseco-Gonzalez, 2005; Ju & Luce, 2004). They have shown that the visual environment comes into play both on microscales, such as computer screens (e.g. Blumenfeld & Marian, 2007; Weber & Cutler, 2004) and on macroscales, such as actual objects in the surrounding environment (e.g. Marian & Spivey, 2003a, 2003b; Spivey & Marian, 1999). Across languages and experimental designs, these eye-tracking studies not only support parallel processing of both languages during spoken word comprehension, but also suggest that the visual array surrounding a bilingual listener in real-world settings interacts with auditory input to influence language comprehension.

Specifically, 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, thus making the recognition process faster and more efficient than it would be unimodally. In other words, the two modalities work together to facilitate comprehension, with the spoken word recognition component incrementally decoding the speech signal, mapping it onto multiple plausible lexical items as the word unfolds, and finally zeroing in on one specific lexical item with partial matches no longer activated. At the same time, the visual modality speeds up the process in a top-down fashion by limiting the options that the auditory input can map onto. Alternatively, the visual modality may increase activation of the target via the additional pathway, so that the cumulative effect of both modalities makes activation of the target more robust and faster. This highly interactive, highly dynamic process happens online in a matter of milliseconds and is a testimony to the astounding human linguistic capacity. Just as monolingual language comprehension is a 'hungry' process (Spivey, 2007), where the system continuously and insatiably seeks and integrates new information, so is the bilingual system one in which expectations about the upcoming input continuously influence the computation of probabilities for what is likely to come next. This 'hunger' is a well-fitting description for the opportunistic cognitive processes that constantly integrate signals across modalities and information sources and, in the case of bilinguals, also across languages.

In fact, bilingualism may be one of a number of experiences that can influence how auditory and visual information is integrated cross-modally. Another interesting example comes from music. Kraus (2007) suggested that a hallmark of the musician's brain is enhanced multisensory processing and that musicians neurally combine audio-visual information more precisely than nonmusicians in the auditory brainstem early in the processing stream. This work suggests that previous experience (in this case, with music) can influence neural patterns of multisensory integration and has important implications for bilinguals, a group whose previous experience with other languages may yield similar outcomes. For instance, both groups (musicians and bilinguals) are 'auditory experts', in the sense that they receive a lot of rich and varied auditory input that they often have to integrate with other modalities, visual or sensorimotor, very quickly online (see also Mechelli *et al.*, 2004).

In future research, it would be interesting to examine empirically whether bilingual and multilingual speakers who have extensive practice with more than one language show the same type of neurological changes in the auditory brainstem and similar performance advantages as those exhibited in research with musicians (see Wong *et al.*, 2007, for an example of advantages in musicians). These advantages may be more pronounced for bilinguals whose experience with audio-visual integration is more diverse due to speaking languages that vary more, for example, when one language is alphabetic and the other logographic, or when one language is tonal and the other is not (for studies of neural changes as a result of acquiring a tonal language, see Krishnan *et al.*, 2005; Wang *et al.*, 2003). Another question for future empirical research with bilinguals is whether this greater experience with integrating sensory input across multiple modalities translates to other advantages in the cognitive system. It is possible that a greater experience with cross-modal integration constitutes one of the sources on which a bilingual advantage builds, and works alongside or cumulatively with other sources examined by bilingualism scholars, such as cognitive control (e.g. Ben-Zeev, 1977) and inhibition skills (e.g. Bialystok, 2005, 2006, for a review, see Cook, 1997). Future research exploring the role of cross-modal integration alongside and vis-à-vis other sources of bilingual advantage may be fruitful.

In sum, previous research has shown that language comprehension relies on cross-modal integration from the auditory and visual modalities, and that this is especially true for bilinguals, whose SNR may be lower. Sources of visual input include visual information about the speaker and visual information about the surrounding environment. Eye-tracking studies suggest that, during language comprehension, auditory input is continuously coupled with and augmented by visual input. Bilinguals rely on cross-modal integration as a matter of course to enhance intelligibility (sometimes even to perceive sounds that cannot be discriminated unimodally) and to speed-up comprehension, and they do so continuously and automatically during language use.

2. Interaction Between Phonology and Orthography During Language Processing

In addition to cross-modal integration, some evidence for audio-visual interaction comes from studies that focus on the interplay between phonology and orthography during language processing. Although visual word recognition is often conceived of as processing of written input, it is in fact the case that the auditory shape of the word (i.e. its phonetic and phonological form) also becomes active. That is, when reading a word, the reader automatically coactivates the auditory form of the visual input, regardless of whether the latter is alphabetic or logographic. Some evidence of phonological involvement in reading comes from monolingual studies showing that regular letter-to-phoneme mappings are read faster than irregular letter-to-phoneme mappings. For example, the word *mint* has regular letter-to-phoneme mappings for *-int* that are consistent with the majority of words including this letter sequence (e.g. *hint*, *stint*, *interesting*, etc). In contrast, the word *pint* has irregular letter-to-phoneme mappings for *–i- in -int* (pronounced /aI/ as in *mile* and *kind*) and therefore takes longer to process (e.g. Baron & Strawson, 1976). Other evidence comes from studies of cross-modal priming, which consistently find that written primes facilitate performance on auditory tasks, i.e. previous exposure to the visual form of a stimulus facilitates recognition of that stimulus when presented auditorily (e.g. Berry *et al.*, 1997; Lovemann *et al.*, 2002; McClelland & Pring, 1991). Moreover, empirical evidence suggests that a word's phonological similarity to other words influences its recognition in the visual modality (e.g. Dijkstra *et al.*, 1999; Ferrand & Grainger, 1994; Perfetti & Bell, 1991; Van Orden, 1987; Van Orden *et al.*, 1988).

In the bilingual literature, a number of studies have examined the role of phonology in visual word recognition (e.g. Doctor & Klein, 1992; Lam *et al.*, 1991; Nas, 1983). Increased phonological similarity has been found to influence bilingual language processing. For example, studies of masked phonological priming (i.e. the prime was presented too briefly for the subject to be consciously aware of it) revealed facilitative interlingual homophone priming from both the native to the non-native language, and from the non-native to the native language (e.g. Brysbaert *et al.*, 1999; Van Wijnendaele & Brysbaert, 2002). In contrast, cross-linguistic form primes (i.e. the participant was consciously aware of perceiving the phonological prime) have been found to inhibit target words in the native language (e.g. Silverberg & Samuel, 2004) and in the non-native language (e.g. Dijkstra *et al.*, 1999; Nas, 1983). These differences in results likely emerged because the masked priming tasks that yielded facilitation had activated sublexical phonological representations only, while lexical decision tasks that yielded cross-linguistic inhibition had activated both lexical and sublexical representations. Together, both facilitatory and inhibitory types of evidence suggest that phonology and phonological overlap play a notable role in orthographic decoding and impact both lexical and sublexical processing in bilingual visual word recognition.

Moreover, studies of bilingual visual word recognition have found that the phonological form of a word is activated not only for the target language, but also for the nontarget language (e.g. Dijkstra *et al.*, 1999; Van Wijnendaele & Brysbaert, 2002). Furthermore, the phonology of the nontarget language is activated not only when the two languages share orthography (for a review, see Doctor & Klein, 1992), but also when the two orthographic systems are distinct, further confirming the interactive nature of the bilingual mental lexicon. For instance, English-Hebrew bilinguals (Tzelgov *et al.*, 1990) and Mandarin-English bilinguals (Chen & Ho, 1986), tested with a cross-linguistic Stroop task, experienced interference from the nontarget language. Because all stimuli were presented only visually and because the two languages did not share orthography, any cross-linguistic interference observed in these bilingual Stroop studies (e.g. Chen & Ho, 1986; Tzelgov *et al.*, 1990) indicated that this interference was driven by nontarget language phonology, thus suggesting that the orthography-to-phonology mappings activated the phonological system of the nontarget language.

With the increased use of eye-movement monitoring to study bilingual language processing during reading (e.g. Altarriba *et al.*, 1996, 2001; McDonald & Thompson, 2006), eye-tracking technology also prompted research on bilingual visual word recognition. For example, Kaushanskaya and Marian (2007) provided evidence for activation of both the phonological and orthographic forms of a word during bilingual language processing in a study that tracked eye movements of Russian–English bilinguals during a Picture-Word Interference task. Specifically, bilinguals were asked to name pictures presented on a computer-screen in the presence of written words or pseudowords that appeared elsewhere on the same screen, while their eye movements were recorded. The picture label and the words/pseudowords overlapped across languages in orthography, phonology or both. Eye movements to written words that competed cross-linguistically with a target were taken as indicative of parallel activation during the comprehension component of the task. During picture naming, reaction time differences for words and pseudowords that shared orthography/phonology compared to words and pseudowords that did *not* share orthography/phonology were taken as indicative of parallel activation during the production component of the task. Results showed that both the auditory form and the written form of the target and the nontarget language were activated during this combined word recognition and naming task. These findings were interpreted as supporting a nonselective view of bilingual language processing in which both languages are active in parallel, but they also provide support for automatic online audio-visual integration during bilingual language comprehension and production.

In addition to visual word recognition, coactivation of both written and auditory forms of a word has been found during auditory language comprehension. Although less intuitive, it is clearly the case that the orthographic shape of a word is frequently activated during spoken language processing. For instance, auditory primes have been found to influence performance on written word processing tasks (e.g. Lovemann *et al.*, 2002). Moreover, bimodal (auditory and written) presentation of information has been found to impact performance during auditory language processing tasks and to improve word recognition in monolingual speakers, as well as in bilingual speakers (Dijkstra *et al.*, 1993; Erdener & Burnham, 2005; Frost *et al.*, 1988; Massaro *et al.*, 1990). Other support for the rôle of orthography in spoken word recognition comes from monolingual studies that find that auditory word recognition is influenced not only by phonological neighborhood size, but also by orthographic neighborhood size (e.g. Ziegler *et al.*, 2003). The term *orthographic neighborhood* refers to all words that differ by one letter from the target word, whereas the term *phonological neighborhood* refers to all words that differ by one phoneme from the target word. Studies that have manipulated the size of the orthographic neighborhood of a word during an auditory word recognition task have found that neighborhood size influences recognition rates.

In sum, psycholinguistic studies have demonstrated that phonological information influences written word processing and orthographic information influences auditory word processing (e.g. Van Orden & Goldinger, 1994; Van Orden *et al.*, 1990). The finding that bilinguals process phonological and orthographic information in parallel during visual and auditory word recognition lends further support to a bilingual language processing system that is highly interactive across languages and modalities.

3. Modeling Audio-visual Integration During Bilingual Language Processing

Computational models of bilingual language processing currently include the Bilingual Interactive Activation model (BIA and its modified version BIA+) (Dijkstra *et al.*, 1998; Dijkstra & Van Heuven, 2002; Van Heuven *et al.*, 1998), the Bilingual Activation Verification model (BAV) (Grainger, 1993), the Bilingual Interactive Model of Lexical Access (BIMOLA) (Grosjean, 1997), the Semantic Orthographic and Phonological Interactive Activation model (SOPHIA) (Van Heuven, 2000) and the Self-Organizing Model of Bilingual Processing (SOMBIP) (Li & Farkas, 2002). New models are continuously being developed.

An example of a typical early model of bilingual language processing is the BAV model, shown in Figure 3.2 (Grainger, 1993). The BAV is a model of bilingual visual word recognition modeled after the monolingual Activation Verification model (Paap *et al.*, 1982). According to this model, a given letter string activates all lexical representations that share letters in the same position as the stimulus, and the greater the number of shared letters, the higher the activation level of the corresponding lexical representation. In the BAV, incoming orthographic information initially activates lexical representations in both languages independently of language context. The most strongly activated of these representations form two separate candidate sets in each lexical system. Language context information then comes into play during selection/verification processes. This information guides the verification process to the appropriate lexical system and thus diminishes the number of possible candidates by half.

Although the BAV allows initial parallel access to lexical representations in both languages, it does posit two distinct lexical systems in

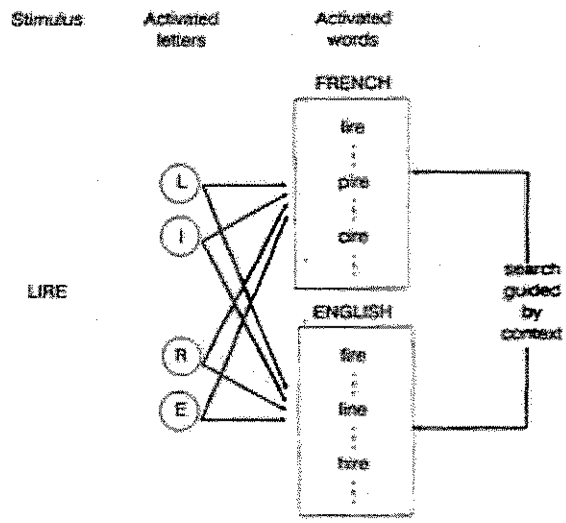

Figure 3.2 A graphic representation of the Bilingual Activation Verification (BAV) model (based on Grainger, 1993).

bilinguals. The utility of separate lexical systems for the two languages remains a matter of debate and is inconsistent with some of the more recent empirical evidence suggesting an integrated lexical system in bilinguals (for instance, the BAV cannot account for cross-language neighborhood effects, e.g. Grainger & Dijkstra, 1992; Grainger *et al.*, 1992). As it is, the model works at the letter level, but not at the feature level (the feature level stores separate visual letter features and precedes letter recognition), and can be used to account for results from languages that share orthography, but not from languages with different orthographies. In order for the BAV to account for audio-visual interaction and for languages that do not share orthographies, it would have to be revised to include a phonological component and a feature component, both of which would interact bidirectionally with the letter component. Without such revisions, the BAV is representative of what early models of bilingual language processing looked like: none of them incorporated the interplay between phonology and orthography and between audio-visual sensory input. However, efforts to do so have been undertaken in recent years and are considered next.

As far as accounting for the interaction between phonology and orthography during word recognition, current monolingual models typically incorporate a phonological processing component (e.g. Coltheart *et al.*, 2001; Seidenberg & McClelland, 1989). Moreover, some monolingual accounts of auditory speech perception also incorporate an orthographic component (e.g. Dijkstra *et al.*, 1995; Frauenfelder *et al.*, 1990; Ziegler *et al.*, 2003). Although models of bilingual *spoken* word recognition have not yet included the orthographic form of the input, bilingual models of *visual* word recognition have started to incorporate the phonological form of the word. For instance, the BIA+ model (Dijkstra & Van Heuven, 2002) shown in Figure 3.3 is one model of bilingual visual word recognition that incorporates the role of phonology during bilingual reading and has been implemented computationally.

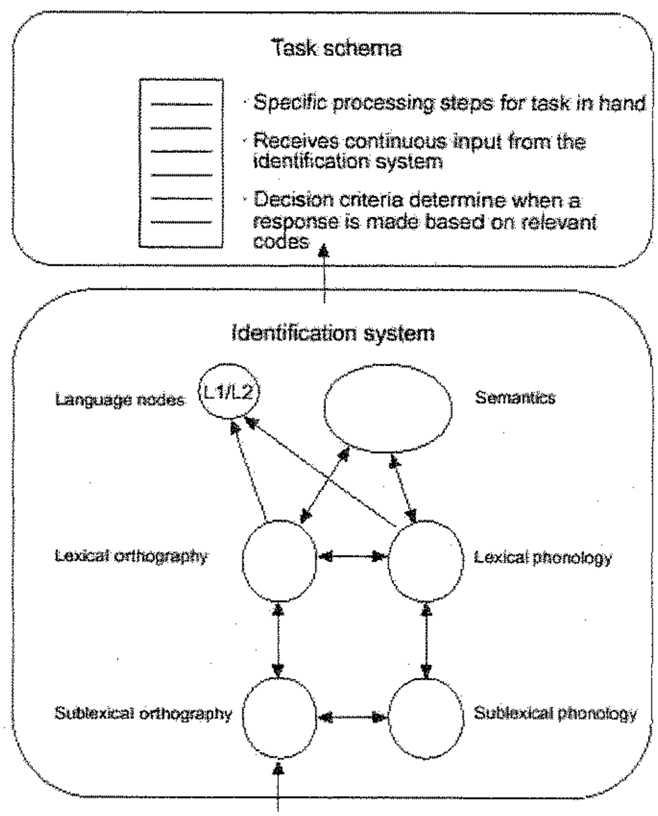

Figure 3.3 A graphic representation of the Bilingual Interactive Activation Plus (BIA+) model (based on Dijkstra & Van Heuven, 2002).

The BIA+ constitutes a revision and enhancement of the original BIA model (Dijkstra & Van Heuven, 1998), which in turn was modeled after the Interactive Activation model of visual word recognition in monolinguals proposed by McClelland and Rumelhart (1981). The BIA+ model consists of four representational levels: a feature level, a letter level, a word level and a supralexical 'language node' (Dijkstra *et al.*, 1998). Activation in the BIA+ model takes place bidirectionally both within and between levels. The model is interactive in the sense that higher-level nodes can now send input to lower-level nodes, in addition to previously postulated bottom-up input.

In the BIA+, sensory input from orthographic stimuli activates feature representations in memory. The feature representations send activation to letter representations, which in turn send activation to lexical representations in both languages. These lexical representations then send activation to the supralexical language nodes. An activated language node sends back excitatory feedback to all the lexical nodes in that language, and those lexical nodes in turn send excitatory feedback to their component letters. Activation of the language node in the BIA+ is influenced by previous lexical recognition. A previously recognized word in L1 increases the activation level of the L1 language node and decreases the activation level of the L2 language node. As a result, the activation level of all words in the L1 increases, while the activation level of all words in the L2 decreases. It is through this feedback process from language nodes to word nodes that the BIA+ explains how bilingual subjects limit interference from the nontarget language.

The BIA+ model suggests that almost immediately following the initial nonselective activation phase, language context information allows the suppression of context-incompatible lexical representations via top-down inhibitory connections from language nodes to word representations. In this way, the BIA+ can account for the occurrence of cross-language interference and for the fact that bilinguals are able to keep such interference to a minimum by using language context. In addition, the BIA+ model (Dijkstra & Van Heuven, 2002) postulates separate nodes for lexical and sublexical *phonological* information, as well as for lexical and sublexical *orthographic* information. This ability to incorporate the interplay between phonology and orthography during visual language processing is a major strength of the BIA+ model, yielding an elegant account of audio-visual integration in terms of phonological and orthographic linguistic input. The BIA+ model was not, however, designed to integrate other forms of linguistically relevant visual information, such as items in the surrounding visual environment.

In fact, none of the existing models of bilingual language processing currently accommodate cross-modal audio-visual interaction such as that described in the first section of this chapter. Thus, beyond the integration of phonology and orthography, there remains a need to model the interplay of auditory and visual input during bilingual spoken language comprehension in complex visual environments. One model of monolingual language processing that can be adopted particularly well to reflect bilingual processing is Marslen-Wilson's COHORT model. The original COHORT model (Marslen-Wilson & Welsh, 1978) is based on the assumption that word recognition takes place on the basis of analyzing all words that match the onset of a target word (see Figure 3.4, panel A). For example, as a spoken word unfolds over time, listeners may initially coactivate the words *marker, marble,* and *mop* (among others) upon hearing *m-*, rule out *mop* after hearing *mar-* and identify the target word as *marker* after hearing *mark-*.

According to the COHORT model, recognition takes place in two stages. During the first stage, all words that exactly match the onset of a

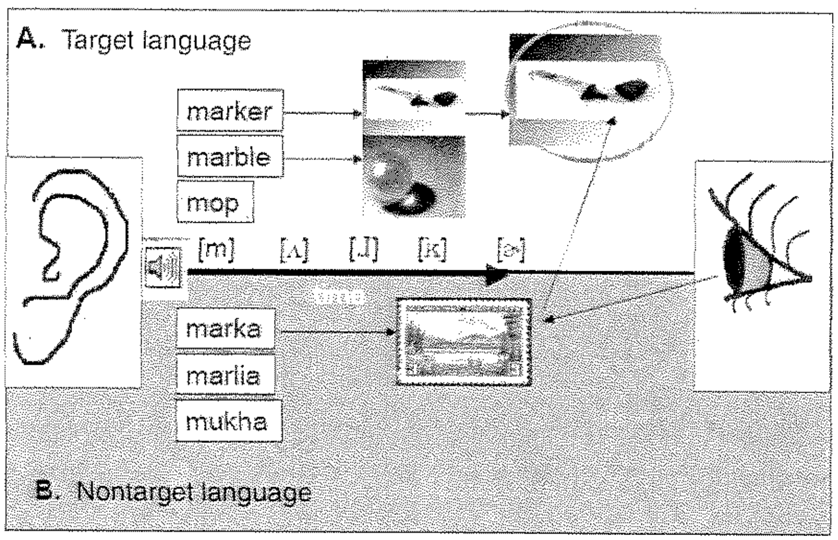

Figure 3.4 Unfolding of the acoustic signal, and cohort activation within and across languages, as indexed by eye movements. Panel A illustrates Marslen-Wilson's COHORT model (1987) in a monolingual scenario (Tanenhaus *et al.*, 1995). Panel B extends the model to a bilingual scenario, as proposed by Marian (2000) and Blumenfeld and Marian (2007).

target word are activated, generating the word-initial cohort. Words that do not match the onset do not enter the cohort, are not activated and do not compete for recognition. The second stage is a deactivation stage during which the cohort members that do not match subsequent sensory input are eliminated from the cohort as more input is received. The cohort members do not have any effect upon each other; therefore, neither the number nor the frequency of the cohort members affects the time course of word recognition in this early model. Only the final cohort members that match the target word the longest determine word recognition. Thus, the COHORT model predicts recognition of a given word on the basis of analyzing its cohort members. The status of words is binary, i.e. the words are either in or out of the cohort. All processing takes place online and in parallel. The activated lexical competitors at any given time are the words that match, and are aligned with, the target word. The recognition point of a presented word corresponds to the moment that the word becomes unique with respect to the other words in the lexicon. The model can make precise and testable predictions about the time course of word recognition. One shortcoming of the COHORT model is the binary status of the cohort words, where the matching between acoustic input and each member of the cohort is an all-or-none process. Another major problem with the COHORT model seems to be its failure to take into account that the number and the frequency of the cohort members affect the time course of word recognition. Both number and frequency effects have been observed repeatedly in monolingual language processing (e.g. Luce *et al.*, 1990), as well as in bilingual language processing (e.g. Beauvillain & Grainger, 1987). To address these shortcomings, the initial model was modified to the COHORT II model.

The COHORT II model (Marslen-Wilson, 1987), like the COHORT, is a strictly bottom-up model, with no top-down effects involved in access or selection. It retains the fundamental characteristics of a cohort-based word recognition model, but differs from the original model in a few important ways. First, the cohort membership is increased to include words that mismatch the sensory input to some degree. The advantage of this modification is that the proposed system is more tolerant of minor deviations in the input; the disadvantage is that the cohort set is now not clearly specified. Second, the model allows for different levels of activation for the cohort members, depending upon their fit with the input and their frequency. The advantage of this modification is that it expresses the varying degree of match between the input and the different competitors; the disadvantage is that it does not specify how word frequency and the degree of match with the input determine activation. Finally, the model is more explicit on the issue of sublexical representations. The advantage of this modification is that now recognition is based on feature matrices, mapping features directly onto lexical units, with no intermediate phoneme decision. Consequently, cohorts can be activated whose initial phoneme differs from that of the input but has partial feature overlap. Thus, while the COHORT model makes clear and testable predictions, it makes several simplifying assumptions that are not always accurate with respect to lexical processing. COHORT II fits better with what we know about lexical processing, but is not clear about the competitor set and cannot always predict the time course of word recognition (Frauenfelder, 1996).

To account for bilingual language processing, the COHORT II model needs to be modified to postulate activation of initial cohorts in both languages (see Figure 3.4, panel B). A Bilingual COHORT model would assume that as the listener perceives auditory input, lexical items from both languages are activated in parallel, resulting in a combined bilingual cohort that extends across the two languages. In the end, as the auditory input unfolds, only the appropriate target is selected, as it is in a monolingual context. A Bilingual COHORT model can successfully accommodate findings from the bilingual empirical literature showing that both languages are activated in parallel in a bilingual, while at the same time incorporating the interplay between the auditory and visual modalities during spoken word recognition in naturalistic environments. That is, in everyday language comprehension, the auditory pathway is rarely the only source of input; rather, the listener is typically situated in complex visual environments with the visual pathway providing additional input that may or may not be related to the auditory input (for example, when words have multiple meanings, as is the case of homophones within and across languages, processing of auditory input is augmented by visual input that is congruent with the relevant word meaning).

The proposed Bilingual COHORT model could account for bilinguals relying more on visual information during auditory processing relative to monolinguals. For example, parallel visual access could be posited and modeled after Seidenberg and McClelland's (1989) triangle model of reading, where all words are processed by a single system containing distributed orthographic, phonological and semantic codes. In the Bilingual COHORT model, integration of the auditory and visual information would likely take place at an intermediate stage, perhaps with amplification of auditory information that is consistent with the visual input. The same could be true for orthography, as in the case of subtitles, where one experiences amplification of auditory input that is consistent with the subtitles. Such multimodal parallel access would be active in both monolinguals and bilinguals, but in bilinguals it might result in greater reliance on the other modality when input from one modality is incomplete or noisy.

Despite its advantages, the proposed Bilingual COHORT model still has the drawback of an entirely bottom-up, feed-forward architecture. Although the model seems to account well for the initial stages of cross-linguistic parallel activation of lexical items in bilinguals, its strictly feed-forward architecture may not allow for any contextual effects on processing. For example, in the case of exact homonyms, the lexical items in both languages are activated, but only one of the two is chosen as the target word, depending upon linguistic context. One way to accommodate these kinds of data is to incorporate 'language tags' in the model to specify the language in use, similar to the language tags postulated by the BIA model (Dijkstra & Van Heuven, 1998) or the Inhibitory Control (IC) model (Green, 1998). Another way would be to postulate the existence of a language node. Or, if the model is a connectionist one, the selectivity could be further explained by higher initial probabilities of the vector of lexical nodes in the active language (e.g. by residual activation from prior use). Mathematically, that could be expressed by adding a positive constant to the nodes in the active language, increasing their baseline activation. The connectionist model may also incorporate variations in the strength of the weights between items by language, to account for higher activation of the target language than activation of the nontarget language. In this case, stronger connection weights between sublexical and lexical levels in the native language could account for auditory input mapping more strongly onto the native system than onto the non-native one. In the same vein, a language node account may also influence selectivity by increasing pathway strength or increased activation from the language node associated with the stronger language.

Each of these solutions may work to different degrees, and efforts are currently underway to develop a computational, interactive model of bilingual spoken language processing that would adequately account for audio-visual interaction (laterally within levels, as well as bottom-up and top-down between levels). A Bilingual COHORT model, like any other bilingual model, should also be able to accommodate mixed language input. In any case, one problem that models of bilingual spoken language processing are guaranteed to encounter is the difficulty of accounting for identification when the phonemes and features do not carry much overlap in the two languages, as is often the case (even when the perceptual categories of the bilingual are not sensitive to these differences, their existence in the input must be accounted for computationally). This problem, however, is not unique to bilingual processing. Models of monolingual language processing are also plagued by the difficulty of accounting for differences among speakers (accents, voice properties) and variability in the rate of speech, ambient noise, etc.

Even with these changes and caveats, the Bilingual COHORT model remains very limited. Its scope is much narrower than that of, for instance, BIA+. In an ideal world, a global model that would combine the different foci and strengths of existing bilingual models would be particularly desirable. This global model of bilingualism would combine BIA+ ability to allow for interaction between phonology and orthography during written comprehension, SOMBIP and BIMOLA relevance to spoken comprehension, Bilingual COHORT model ability to integrate nonlinguistic visual information as the word unfolds, the Revised Hierarchical Model's (RHM) (Kroll & Stewart, 1994) focus on links between forms and concepts, and IC focus on control of the two languages. This, of course, is a tall order, and one that cannot currently be implemented. However, with further development of computational tools and additional knowledge about the bilingual cognitive architecture, it is only a matter of time until a comprehensive model of bilingual language processing can be developed. After all, the capacity of bilingual models has notably increased in the 10 years between BAV and BIA+, and it is not unreasonable to expect that a few more decades of modeling would produce further strides. Until then, the task of modeling bilingualism is broken down into multiple smaller components (e.g. written comprehension, spoken comprehension, production, nonlinguistic visual input, control, concepts). (In addition to the original texts, concise summaries of the models discussed here, along with their graphic representations, can be found in a number of recent review papers, including a thorough overview of computational models by Thomas and Van Heuven (2005) and a historical review of bilingual cognitive models by Kroll and Tokowicz (2005).)

Conclusions: Audio-visual Integration and Representation in the Bilingual Mental Lexicon

In sum, it is clear that audio-visual integration is an inherent part of bilingual language processing, be it integration of auditory and visual modalities or integration of phonological and orthographic word-forms. Bilingual models of language processing are beginning to account for the role of phonology during visual word recognition, but they have yet to address the role of orthography during auditory word recognition. The COHORT model (Marslen-Wilson, 1987; Marslen-Wilson & Welsh, 1978) of monolingual language processing is one model that can be adapted to allow for bimodal integration during bilingual spoken language processing. However, a model of audio-visual integration during bilingual language comprehension would have to allow for both – the cross-modal interaction between auditory and visual input and the interplay between phonology and orthography. Such a model will also need to be able to account for the fact that, by its very nature, bilingualism is not a static phenomenon, but a dynamic, developmental process, which undergoes changes continuously. As the level of proficiency or the manner of acquisition change, so do language processing and representation. The concept of a dynamically evolving bilingual system has previously been incorporated in the RHM (Kroll & Stewart, 1994) and the SOMBIP (Li & Farkas, 2002). Both of these models have posited changes in the representation of the second language, and in the relationship between the first and second languages, as proficiency levels change. Although neither the RHM nor the SOMBIP address the issue of audio-visual integration in bilinguals, they provide valuable insights into developmental factors (such as changes in proficiency levels) that have to be incorporated into a multimodal integration model of bilingual language processing.

The recognition that language comprehension is a cross-modal, interactive process carries implications for representation in the bilingual mental lexicon. It suggests that, as linguistic input unfolds, it becomes integrated across modalities incrementally and thus, its representation in the bilingual mental lexicon is constantly changing and evolving. This 'in-flux' state is inevitable, as language comprehension is a continuous, never static, process. In his book, *The Continuity of Mind*, Spivey (2007: 206) describes the mental representations of sentences, words and phonemes as *a process* and not *a thing*, and writes the following about language comprehension:

> ... there is no point in time when the mental trajectory through state space, which is propelled by a combination of environmental sensory input and goal-oriented expectations, stops and stands still. It is always in motion. The patterns of neural activation in the brain are in perpetual flux. An important consequence of this temporal continuity of mind is that there can be no mediating states (e.g., Dietrich & Markman, 2003), because states require stasis. ...these dynamic

> patterns, or continuous processes... are not states, and therefore cannot be static symbols that are discretely separable from one another in time or in representational space. What this means is that language is not a string of symbols whose grammatical relationships are encoded by discrete hierarchical structures in an encapsulated linguistic module. Language, like the rest of perception and cognition, is a continuous trajectory through a high-dimensional state space...

Though these statements are both sensible and intuitive, they in fact carry revolutionary implications for the way we think about representations in the bilingual mental lexicon. Current models of bilingual language organization and processing do not allow for this continuity of mental states and instead idealize representations in the bilingual mental lexicon as discrete states. The need to describe mental representations in the bilingual lexicon as dynamic, continuous processes is the likely next step in the psycholinguistics of bilingualism. In general, as bilingualism scholars, we need to move away from unimodal perspectives of bilingual language processing and recognize that language comprehension in the real world is a combination of multiple sources of input that interact to form a final composite interpreted by the brain. In essence, we hear with our brain and not with our ears, as the McGurk effect (described at the beginning of this chapter) so elegantly shows. As research on bilingual language processing moves from the laboratory to the real world, it becomes increasingly clear that studies and models of bilingualism need to reflect and to be rooted in the understanding that language is a multimodal and dynamic process.

Acknowledgements

Preparation of this chapter was supported in part by grants NSF-BCS-0418495 and NICHD-1R03HD046952-01A1. I thank Henrike Blumenfeld, Margarita Kaushanskaya, Melissa Baese, Caroline Engstler, Aneta Pavlenko and the anonymous reviewers for helpful comments and suggestions.

References

Altarriba, J., Kambe, G., Pollatsek, A. and Rayner, K. (2001) Semantic codes are not used in integrating information across eye fixations in reading: Evidence from fluent Spanish-English bilinguals. *Perception & Psychophysics* 63, 875–890.

Altarriba, J., Kroll, J., Sholl, A. and Rayner, K. (1996) The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory and Cognition* 24, 477–492.

Baron, J. and Strawson, C. (1976) Use of orthographic and word-specific knowledge in reading words aloud. *Journal of Experimental Psychology: Human Perception and Performance* 2, 386–393.

Beauvillain, C. and Grainger, J. (1987) Accessing interlexical homographs: Some limitations of a language-selective access. *Journal of Memory and Language* 26, 658–672.

Ben-Zeev, S. (1977) The influence of bilingualism on cognitive strategy and cognitive development. *Child Development* 48, 1009–1018.

Berry, D., Banbury, S. and Henry, L. (1997) Transfer across form and modality in implicit and explicit memory. *The Quarterly Journal of Experimental Psychology* 50A (1), 1–24.

Bialystok, E. (2005) Consequences of bilingualism for cognitive development. In J. Kroll and A.M.B. De Groot (eds) *Handbook of Bilingualism: Psycholinguistic Approaches* (pp. 417–432). Oxford: Oxford University Press.

Bialystok, E. (2006) Effect of bilingualism and computer video game experience on the Simon task. *Canadian Journal of Experimental Psychology* 60 (1), 68–79.

Bird, S. and Williams, J. (2002) The effect of bimodal input on implicit and explicit memory: An investigation into the benefits of within-language subtitling. *Applied Psycholinguistics* 23, 509–533.

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), 1–28.

Borras, I. and Lafayette, R. (1994) Effects of multimedia courseware subtitling on the speaking performance of college students of French. *The Modern Language Journal* 78, 61–65.

Bradlow, A. and Bent, T. (2002) The clear speech effect for non-native listeners. *Journal of the Acoustical Society of America* 112 (1), 272–284.

Bradlow, A. and Pisoni, D. (1999) Recognition of spoken words by native and non-native listeners: Talker-, listener- and item-related factors. *Journal of the Acoustical Society of America* 106 (4), 2074–2085.

Brysbaert, M., Van Dyck, G. and Van de Poel, M. (1999) Visual word recognition in bilinguals: Evidence from masked phonological priming. *Journal of Experimental Psychology: Human Perception and Performance* 25, 137–148.

Calvert, G., Bullmore, E., Brammer, M., Campbell, R., Williams, S., McGuire, P., *et al.* (1997) Activation of auditory cortex during silent lipreading. *Science* 276, 593–596.

Calvert, G., Campbell, R. and Brammer, M. (2000) Evidence from functional magnetic resonance imaging of cross-modal binding in the human heteromodal cortex. *Current Biology* 10, 649–657.

Campbell, R., MacSweeney, M., Surguladze, S., Calvert, G., McGuire, P., Suckling, J., *et al.* (2001) Cortical substrates for the perception of face actions: An fMRI study of the specificity of activation for seen speech and for meaningless lower-face acts. *Cognitive Brain Research* 12, 233–243.

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). Barcelona, Spain: Universidad Autonoma de Barcelona.
Chen, H.C. and Ho, C. (1986) Development of Stroop interference in Chinese-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 12, 397–401.
Clau v. Uniglobe Pacific Travel (1995) *Canadian Human Rights Reporter* 23, D/515–D/518.
Coltheart, M., Rastle, K., Perry, C., Langdon, R. and Ziegel, J. (2001) DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review* 108, 204–256.
Cook, V. (1997) The consequences of bilingualism for cognitive processing. In A.M.B. De Groot and J. Kroll (eds) *Tutorials in Bilingualism: Psycholinguistic Perspectives* (pp. 279–300). Hillsdale, NJ: Lawrence Erlbaum.
Cutler, A., Weber, A. and Otake, T. (2006) Asymmetric mapping from phonetic to lexical representations in second-language listening. *Journal of Phonetics* 34, 269–284.
Danan, M. (1992) Reversed subtitling and dual coding theory: New directions for foreign language instruction. *Language Learning* 42 (4), 497–527.
Davis, C. and Kim, J. (2006) Audio-visual speech perception off the top of the head. *Cognition* 100, B21–B31.
Dietrich, E. and Markman, A. (2003) Discrete thoughts: Why cognition must use discrete representations. *Mind and Language* 18, 95–119.
Dijkstra, A. Frauenfelder, U. and Schreuder, R. (1993) Grapheme-phoneme interaction in a bimodal detection task. *Journal of Experimental Psychology: Human Perception and Performance* 19 (5), 1–20.
Dijkstra, A., Grainger, J. and Van Heuven, W. (1999) Recognition of cognates and interlingual homographs: The neglected rôle of phonology. *Journal of Memory and Language* 41, 496–518.
Dijkstra, A., Roelofs, A. and Fiews, S. (1995) Orthographic effects on phoneme monitoring. *Canadian Journal of Experimental Psychology* 49 (2), 264–271.
Dijkstra, A. and Van Heuven, W. (1998) The BIA-model and bilingual word recognition. In J. Grainger and A. Jacobs (eds) *Localist Connectionist Approaches to Human Cognition* (pp. 189–225). Mahwah, NJ: Lawrence Erlbaum.
Dijkstra, A. and Van Heuven, W. (2002) The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition* 5, 175–197.
Dijkstra, A., Van Heuven, W. and Grainger, J. (1998) Simulating cross-language competition with the Bilingual Interactive Activation model. *Psychologica Belgica* 38, 177–196.
Doctor, E. and Klein, D. (1992) Phonological processing in bilingual word recognition. In R. Harris (ed.) *Cognitive Processing in Bilinguals* (pp. 237–252). Amsterdam: Elsevier.
Erdener, V. and Burnham, D. (2005) The role of audiovisual speech and orthographic information in non-native speech production. *Language Learning* 55 (2), 191–228.
Ferrand, L. and Grainger, J. (1994) Effects of orthography are independent of phonology in masked form priming. *The Quarterly Journal of Experimental Psychology* 47A, 365–382.

Frauenfelder, U. (1996) Computational models of spoken word recognition. In T. Dijkstra and K. de Smedt (eds) *Computational Psycholinguistics: AI and Connectionist Models of Human Language Processing* (pp. 114–138). London: Taylor & Francis.

Frauenfelder, U., Segui, J. and Dijkstra, A. (1990) Lexical effects in phoneme processing: Facilitatory or inhibitory? *Journal of Experimental Psychology: Human Perception and Performance* 16 (1), 77–91.

Frost, R., Repp, B. and Katz, L. (1988) Can speech perception be influenced by simultaneous presentation of print? *Journal of Memory and Language* 27, 741–755.

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

Grainger, J. and Dijkstra, T. (1992) On the representation and use of language information in bilinguals. In R. Harris (ed.) *Cognitive Processing in Bilinguals* (pp. 207–220). Amsterdam: Elsevier.

Grainger, J., O'Regan, J., Jacobs, A. and Segui, J. (1992) Neighborhood frequency effects and letter visibility in visual word recognition. *Perception & Psychophysics* 51, 49–56.

Green, D. (1998) Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition* 1, 67–81.

Grosjean, F. (1997) Processing mixed language: Issues, findings, and models. In A.M.B. De Groot and J. Kroll (eds) *Tutorials in Bilingualism: Psycholinguistic Perspectives* (pp. 225–254). Hillsdale, NJ: Lawrence Erlbaum.

Guillen v. R. Dufour Enterprises Ltd. (1995) Unpublished decision. British Columbia Council of Human Rights.

Ju, M. and Luce, P. (2004) Falling on sensitive ears: Constraints on bilingual lexical activation. *Psychological Science* 15, 314–318.

Kaushanskaya, M. and Marian, V. (2007) Non-target language recognition and interference in bilinguals: Evidence from eye tracking and picture naming. *Language Learning* 57, 119–163.

Kraus, N. (2007) Music experience shapes subcortical sensory circuitry. Paper presented at the Music Cognition Colloquium, Northwestern University School of Music.

Krishnan, A., Xu, Y., Gandour, J. and Cariani, P. (2005) Encoding of pitch in the human brainstem is sensitive to language experience. *Cognitive Brain Research* 25, 161–168.

Kroll, J. and Stewart, E. (1994) Category interferences in translation and picture naming: Evidence for asymmetric connection between bilingual memory representation. *Journal of Memory and Language* 33, 149–174.

Kroll. J. and Tokowicz, N. (2005) Models of bilingual representation and processing. In J. Kroll and A.M.B. De Groot (eds) *Handbook of Bilingualism: Psycholinguistics Approaches* (pp. 531–553). Oxford: Oxford University Press.

Lam, A., Perfetti, C. and Bell, L. (1991) Automatic phonetic transfer in bidialectal reading. *Applied Psycholinguistics* 12, 299–311.

Li, P. and Farkas, I. (2002) A self-organized connectionist model of bilingual processing. In R. Heredia and J. Altarriba (eds) *Bilingual Sentence Processing* (pp. 59–88). North Holland: Elsevier Science.

Lovemann, E., van Hoff, J.C. and Gale, A. (2002) A systematic investigation of same and cross modality priming using written and spoken responses. *Memory* 10 (4), 267–276.

Luce, P. and Pisoni, D. (1998) Recognizing spoken words: The neighborhood activation model. *Ear & Hearing* 19, 1–36.

Luce, P., Pisoni, D. and Goldinger, S. (1990) Similarity neighborhoods of spoken words. In G. Altmann (ed.) *Cognitive Models of Speech Perception: Psycholinguistic and Computational Perspectives* (pp. 122–147). Cambridge, MA: MIT Press.

Marian, V. (2000) Bilingual language processing: Evidence from eye-tracking and functional neuroimaging. Unpublished doctoral dissertation, Cornell University.

Marian, V. and Spivey, M. (2003a) Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition* 6, 97–115.

Marian, V. and Spivey, M. (2003b) Bilingual and monolingual processing of competing lexical items. *Applied Psycholinguistics* 24, 173–193.

Marian, V., Blumenfeld, H. and Boukrina, O. (2008) Sensitivity to phonological similarity within and across languages: A native/non-native asymmetry in bilinguals. *Journal of Psycholinguistic Research* 37, 141–170.

Marslen-Wilson, W. (1987) Functional parallelism in spoken word recognition. *Cognition* 25, 71–102.

Marslen-Wilson, W. and Welsh, A. (1978) Processing interactions and lexical access during word-recognition in continuous speech. *Cognitive Psychology* 10, 29–63.

Massaro, D., Cohen, M. and Thompson, L. (1990) Visible language in speech perception: Lipreading and reading. *Visible Language* 22, 8–31.

McClelland, A. and Pring, L. (1991) An investigation of cross-modality effects in implicit and explicit memory. *Quarterly Journal of Experimental Psychology* 43A, 19–33.

McClelland, J. and Elman, J. (1986) The TRACE model of speech perception. *Cognitive Psychology* 18, 1–86.

McClelland, J. and Rumelhart, D. (1981) An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review* 88, 375–407.

McDonald, S. and Thompson, J. (2006) Using eye-tracking to contrast lexical-level and conceptual-level connections in the bilingual lexicon. Unpublished manuscript. School of Philosophy, Psychology, and Language Sciences, University of Edinburgh.

McGurk, H. and MacDonald, J. (1976) Hearing lips and seeing voices: A new illusion. *Nature* 264, 746–748.

Mechelli, A., Crinion, J., Noppeny, U., O'Doherty, J., Ashburner, J., Frackowiak, R., *et al.* (2004) Structural plasticity in the bilingual brain: Proficiency in a second language and age at acquisition affect grey-matter density. *Nature* 431 (7010), 757.

Munhall, K., Jones, J., Callan, D., Kuratate, T. and Vatikiotis-Bateson, E. (2004) Visual prosody and speech intelligibility – Head movement improves auditory speech perception. *Psychological Science* 15, 133–137.

Munro, M. (1998) The effects of noise on the intelligibility of foreign-accented speech. *Studies in Second Language Acquisition* 20, 139–154.

Musacchia G., Sams, M., Nicol, T. and Kraus, N. (2006) Seeing speech affects acoustic information processing in the human brainstem. *Experimental Brain Research* 168, 1–10.

Nas, G. (1983) Visual word recognition in bilinguals: Evidence for a cooperation between visual and sound based codes during access to a common lexical store. *Journal of Verbal Learning and Verbal Behavior* 22, 526–534.

Navarra, J. and Soto-Faraco, S. (2007) Hearing lips in a second language: Visual articulatory information enables the perception of secondary language sounds. *Psychological Research-Psychologische Forschung* 71, 4–12.

Nishitani, N. and Hari, R. (2002) Viewing lip forms: Cortical dynamics. *Neuron* 19, 1211–1220.

Paap, K., Newsome, S., McDonald, J. and Schvaneveldt, R. (1982) An activation-verification model for letter and word recognition: The word-superiority effect. *Psychological Review* 89, 573–594.

Perfetti, C. and Bell, L. (1991) Phonemic activation during the first 40 ms of word identification: Evidence from backward masking and priming. *Journal of Memory & Language* 30 (4), 473–485.

Ronquest, R. and Hernandez, L. (2005) Lip-reading skills in bilinguals: Some effects of L1 on visual-only language identification. *Research on Spoken Language Processing, Progress Report* 27. Indiana University, Bloomington, IN.

Ross, L., Saint-Amour, D., Leavitt, V., Javitt, D. and Foxe, J. (2007) Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environment. *Cerebral Cortex* 17, 1147–1153.

Seidenberg, M. and McClelland, J. (1989) A distributed, developmental model of word recognition and naming. *Psychological Review* 96, 523–568.

Silverberg, S. and Samuel, A. (2004) The effect of age of second language acquisition on the representation and processing of second language words. *Journal of Memory and Language* 51, 381–398.

Soto-Faraco, S., Navarra, J., Weikum, W., Vouloumanos, A., Sebastian-Galles, N. and Wereker, J. (2007) Discriminating language by speech-reading. *Perception & Psychophysics* 69, 218–231.

Spivey, M. (2007) *The Continuity of Mind*. Oxford/New York: Oxford University Press.

Spivey, M. and Marian, V. (1999) Cross-talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science* 10, 281–284.

Sumby, W. and Pollack, I. (1954) Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America* 26, 212–215.

Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K. and Sedivy, J. (1995) Integration of visual and linguistic information in spoken language comprehension. *Science* 268, 632–634.

Thomas, M. and Van Heuven, W. (2005) Computational models of bilingual comprehension. In J. Kroll and A.M.B. De Groot (eds) *Handbook of Bilingualism: Psycholinguistics Approaches* (pp. 202–225). Oxford, NY: Oxford University Press.

Tzelgov, J., Henik, A. and Leiser, D. (1990) Controlling Stroop interference: Evidence from a bilingual task. *Journal of Experimental Psychology* 16, 760–771.

Van Atteveldt, N., Formisano, E., Goebel, R. and Blomert, L. (2004) Integration of letters and speech sounds in the human brain. *Neuron* 43, 271–282.

Van Heuven, W. (2000) Visual word recognition in monolingual and bilingual readers: Experiments and computational modeling. PhD thesis, University of Nijmegen.

Van Heuven, W., Dijkstra, T. and Grainger, J. (1998) Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language* 39, 458–483.

Van Orden, G. (1987) A rows is a rose: Spelling, sound and reading. *Memory and Cognition* 15, 181–198.

Van Orden, G. and Goldinger, S. (1994) Interdependence of form and function in cognitive systems explains perception of printed words. *Journal of Experimental Psychology: Human Perception and Performance* 20, 1269–1291.

Van Orden, G., Johnston, J. and Hale, B. (1988) Word identification in reading proceeds from spelling to sound to meaning. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 14, 371–386.

Van Orden, G., Pennington, B. and Stone, G. (1990) What do double dissociations prove? *Cognitive Science* 25, 111–172.

Van Wassenhove, V., Grant, K. and Poeppel, D. (2007) Temporal window of integration in auditory-visual speech perception. *Neuropsychologia* 45, 598–607.

Van Wijnendaele, I. and Brysbaert, M. (2002) Visual word recognition in bilinguals: Phonological priming from the second to the first Language. *Journal of Experimental Psychology: Human Perception and Performance* 28, 616–627.

Vanderplank, R. (1988) The value of teletext subtitles in language learning. *ELT Journal* 42 (4), 272–281.

Wang, Y., Sereno, J., Jongman, A. and Hirsch, J. (2003) fMRI evidence for cortical modification during learning of Mandarin lexical tones. *Journal of Cognitive Neuroscience* 15, 1–9.

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

Weber, A. and Paris, G. (2004) The origin of the linguistic gender effect in spoken-word recognition: Evidence from non-native listening. Poster presented at the 26th Annual Meeting of the Cognitive Science Society, Chicago, IL.

Wong, P., Skoe, E., Russo, N., Dees, S. and Kraus N. (2007) Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience* 10, 420–422.

Ziegler, J., Muneaux, M. and Grainger, J. (2003) Neighborhood effects in auditory word recognition: Phonological competition and orthographic facilitation. *Journal of Memory and Language* 43, 779–793.