



Two tongues, one brain: imaging bilingual speech production

Anna J. Simmonds^{1*}, Richard J. S. Wise^{1,2} and Robert Leech²

¹ Medical Research Council Clinical Sciences Centre, Imperial College London, UK

² Division of Experimental Medicine, Imperial College London, UK

Edited by:

Guillaume Thierry, Bangor University, UK

Reviewed by:

Jubin Abutalebi, Vita-Salute San Raffaele University, Italy
Narly Golestani, Université de Genève, Switzerland

*Correspondence:

Anna J. Simmonds, Medical Research Council Clinical Sciences Centre, Imperial College London, Hammersmith Hospital Campus, Du Cane Road, London W12 0NN, UK.
e-mail: anna.simmonds08@imperial.ac.uk

This review considers speaking in a second language from the perspective of motor–sensory control. Previous studies relating brain function to the prior acquisition of two or more languages (neurobilingualism) have investigated the differential demands made on linguistic representations and processes, and the role of domain-general cognitive control systems when speakers switch between languages. In contrast to the detailed discussions on these higher functions, typically articulation is considered only as an underspecified stage of simple motor output. The present review considers speaking in a second language in terms of the accompanying foreign accent, which places demands on the integration of motor and sensory discharges not encountered when articulating in the most fluent language. We consider why there has been so little emphasis on this aspect of bilingualism to date, before turning to the motor and sensory complexities involved in learning to speak a second language as an adult. This must involve retuning the neural circuits involved in the motor control of articulation, to enable rapid unfamiliar sequences of movements to be performed with the goal of approximating, as closely as possible, the speech of a native speaker. Accompanying changes in motor networks is experience-dependent plasticity in auditory and somatosensory cortices to integrate auditory memories of the target sounds, copies of feedforward commands from premotor and primary motor cortex and post-articulatory auditory and somatosensory feedback. Finally, we consider the implications of taking a motor–sensory perspective on speaking a second language, both pedagogical regarding non-native learners and clinical regarding speakers with neurological conditions such as dysarthria.

Keywords: bilingualism, fMRI, language, motor–sensory, speech

INTRODUCTION

Speech production is a complex motor act, involving rapid sequential motor movements that often extend over many seconds before a pause. It depends on the integration of feedforward motor and feedback sensory signals, with online self-monitoring guiding rapid modification of motor commands to the larynx, pharynx, and articulators. This allows the maintenance of intelligible speech, even under adverse speaking conditions, for example, when speaking (less than politely) with one's mouth full at mealtimes. It depends on motor (frontal), auditory (temporal), and somatosensory (parietal) cortex, as well as the insulae, cerebellum, and subcortical nuclei (Guenther et al., 2006; Ventura et al., 2009; Golfopoulos et al., 2010).

A range of linguistic stages are involved in speech production. Speech is the final expression of concepts and emotions, translated through linguistic pathways that involve lexical, syntactic, phonological, and phonetic stages (Levelt, 1989), as well as prosody. Although these stages have been defined and refined over many decades using behavioral measures in normal subjects (Levelt, et al., 1999) and lesion-deficit analyses of patients with focal lesions (Shallice, 1988), more recently their relationship to brain anatomy and function have become intensively studied in normal subjects, first with positron emission tomography (PET) and then with functional magnetic resonance imaging (fMRI). The large meta-analyses that have been made possible because of numerous individual functional neuroimaging studies have

considered native speech comprehension in terms of semantics, syntax, and phonology (Vigneau et al., 2006; Binder et al., 2009), and native single word speech production in more fine-grained terms of lemma retrieval and selection, phonological code retrieval, syllabification, and final motor output (Indefrey and Levelt, 2004). These studies have been followed by a more detailed analysis of the motor and sensory representations and processes that govern speech production, with studies designed to investigate speech-related breath control (Loucks et al., 2007), laryngeal function (Simonyan et al., 2009), and articulatory movements (Sörös et al., 2010). To allow ease of interpretation of the results, study designs have often been restricted to the study of syllable or single word production. One problem, particularly with fMRI, is speech-related movement artifact, and a number of studies have discussed motor–sensory integration during speech production while requiring the participants to speak covertly; that is, the results do not include activity associated with motor output and sensory feedback (Hickok et al., 2009). There are now well-established techniques (“sparse” scanning) to acquire functional data with fMRI while minimizing movement- and respiratory-related artifact during overt speech at the level of both single words and sentences (Hall et al., 1999; Gracco et al., 2005). In addition, “sparse” fMRI can be used to minimize auditory masking, as the overt speech is produced during silent periods during which functional images are not being acquired. This is possible because the signal in fMRI, the hemodynamic response function

(HRF), which relies on changes in blood flow in response to net regional synaptic activity (the blood oxygen-level dependent – BOLD – signal), extends over many seconds.

Despite the recent interest in imaging the neural systems supporting the execution and sensory monitoring of speech, this has only recently been extended into bilingualism and the effects of speaking a second language with an accent. As well as the interest in understanding the differences between native speech production and speaking a foreign language with an accent, these studies offer the potential for insights into the compensatory mechanisms that may be engaged in patients with abnormal speech, such as dysarthria and stuttering (Curio et al., 2000).

This review focuses on successive bilingualism, in which a second language (L2) was learnt after the first (L1) was already established; despite high levels of linguistic proficiency in L2, it is clearly a non-native language because of the persistence of a foreign accent. Simultaneous bilinguals, who acquired L1 and L2 concurrently in childhood, may not demonstrate the motor and sensory differences that we discuss in this article; but this has yet to be studied. Young learners who acquire a second language in childhood and go on to reach native proficiency sometimes stop using their L1 and only use L2. This requires different processing skills than those bilingual speakers who master a second language while maintaining their first (Snow, 2002). Within the scope of this review, we are focusing on bilingual speakers who maintain L1 whilst acquiring L2. The first section briefly reviews previous research on bilingualism, to place the motor–sensory aspects of speaking in L2 in context. We then consider how learning to speak L2 requires the retuning of the neural circuits involved in motor control of articulation, the effects on auditory and somatosensory feedback systems, and speculate on why the system rarely becomes so finely and accurately tuned that L2 can be spoken without an accent. Clearly, what applies to L2 also applies to dialects in L1, and there is the interesting and specialized area of voice training undergone by actors, who are perhaps most adept at mimicking dialects and accents. Finally, we consider the pedagogical and clinical implications of taking a motor–sensory perspective on speech production.

PREVIOUS RESEARCH ON BILINGUALISM: LINGUISTIC AND COGNITIVE ASPECTS

As articulation interacts with syntactic, lexical, phonological, and phonetic constraints, we review the previous functional imaging studies of bilingualism, which has largely concentrated on these aspects, and consider why control of articulation during L2 production has received so little interest in the functional neuroimaging literature in the past.

Learning a new language is normally understood as adaptations of the neural systems involved in linguistic processing; that is, those involved in acquiring a new syntax and lexicon (Francis, 1999; Kovelman et al., 2008; Kovacs and Mehler, 2009), and those components of domain-general cognitive control systems involved in translation and switching between languages (Price et al., 1999; Hernandez et al., 2000; Rodriguez-Fornells et al., 2002; Costa and Santesteban, 2004; Crinion et al., 2006; Abutalebi et al., 2007). Age and order of acquisition and/or proficiency in the different languages, as well as the way in which new languages are learnt, and the modality (signed/spoken) are all considered to play a role in

how L2, L3, etc., are represented in the brain relative to L1, and the control processes that operate on these representations (Vaid and Hull, 2002). This review focuses on the motor–sensory aspects of bilingual speech production, rather than linguistic or cognitive processing, both of which have been reviewed elsewhere. Bilingual fMRI studies of both production and comprehension at the word-, sentence-, and discourse-level, as well as inflectional morphology are reviewed by Indefrey (2006). Control mechanisms in bilingual language production are reviewed in the article by Abutalebi and Green (2008). Abutalebi and Green (2007) specify a model that integrates separable neural systems responsible for distinct aspects of cognitive control involved in bilingual speech production. These systems include the prefrontal cortex (up-dating the language, inhibition of the language not in use and error correction), the anterior cingulate cortex (attention, conflict monitoring, and error detection), the basal ganglia (language selection), and the inferior parietal lobule (maintenance of representations and working memory). Abutalebi et al. (2009) have also shown how this model can be applied to bilingual aphasic research, using dynamic causal modeling, in combination with behavioral and imaging data. This cognitive control model is the predominant neurocognitive model that exists in the bilingual literature, and the model of speech production we discuss in this review complements this cognitive control model.

DIFFERENCES BETWEEN FIRST AND SECOND LANGUAGE ACQUISITION

From a theoretical perspective, the stages by which first and second languages are acquired differ, especially when the latter begins once the former has already been established. In this vein, acquisition and learning can be defined as two separate processes (Krashen and Terrell, 1983). *Acquisition* can be thought of as an implicit process that enables the speaker to develop functional skills without theoretical knowledge. In contrast, *learning* can be thought of as knowing about the language, a more explicit and conscious process. In the late 1970s, a distinction between first and second language acquisition was proposed by Lamendella (1977), who defined primary language acquisition (PLA) as a child's acquisition of one or more languages, before the age of 5 years. In contrast, experience in a second language after the age of 5 years (SLA) was defined as both explicit formal foreign language learning and the natural acquisition of another language (acquired without formal instruction, but through frequent exposure and use). Lamendella suggested that PLA is largely dependent on innate neural systems, whose plasticity declines after a critical period, after which, different neural systems are recruited when learning a language. Thus, the hypothesis is that there are two “hierarchies” of language processing: communicative and cognitive. PLA and natural SLA use the communicative hierarchy, and formal SLA uses the cognitive hierarchy (with the implication that this is more dependent on top-down frontal executive control systems). Functional imaging research into bilingualism has investigated both these hierarchies, even if not explicitly. These hierarchies cope more or less well with different stages of linguistic processing; for example, the vocabulary of L2 can be thoroughly acquired by adults, but its phonological and morphological aspects are less problematic for children.

For a child, language acquisition begins with speech perception. Initially, the speech-perception skills of infants are language-general, which offers the potential to acquire any human language

to which the child is exposed. The phonetic repertoire of a language is based on both the individual consonant and vowel sounds and the permissible combinations of these sounds in the creation of words and phrases (Jacquemot et al., 2003). Before an infant is 6 months old, their perceptual system has become tuned to the phonetic repertoire of their native language, and making distinctions between non-native phonemes may be more difficult (the classic example being the lack of distinction between /r/ and /l/ in Japanese infants; Kuhl, 2004; see **Figure 1**). This early stage is apparent as a “silent period,” during which the infant listens to language without attempts to produce speech sounds. There then follows the babbling phase, with imitation of simplified syllables. By the time the child is 9 months old, its babbling is becoming language-specific, and a skilled child developmental psychologist can tell the difference between the babblings of children from different language cultures. This is the earliest evidence of language-specific motor-sensory processing. By 1 year, babbling turns to speech, beginning with single words, followed by short phrases and then sentences. Although linguistic errors are frequent at this stage, the accent of the child is clearly that of a native speaker. In contrast, those who learn a second language as an older child or an adult speak with an accent that clearly marks them as a non-native speaker.

These observations indicate some form of “critical periods” in language development, and after these periods the cortical plasticity of the motor and auditory systems become more limited. The formation of representations of foreign speech sounds may be less accurate later in development, so that the learner may not be able to distinguish perception of certain vowels and consonants. Sounds in L2 that do not exist or are “phonologically ungrammatical” in L1 are often assimilated into the closest acceptable form in the native language (Dupoux et al., 1999). The acquisition of L2 production is quite different from that for L1. There is no babbling stage, and the learners acquire new words with the explicit knowledge about their meaning. There is also a strong tendency to translate a word in L2 into its corresponding word in L1 when listening (Thierry and

Wu, 2007; Wu and Thierry, 2010), and both L2 and L1 phonological representations are retrieved during covert word production in L2 (Wu and Thierry, 2011). Second language learners may also “fossilize” at a relatively early stage, when the subject feels that they have developed L2 sufficient for their purpose, reflecting the more explicit nature of L2 acquisition and an acceptance of what is enough to “get by.” This means that they may not strive to tune their auditory cortex to generate accurate, long-term representations of L2 syllables, words, and phrases, as spoken by a native speaker. They may also be unable to accurately perceive the non-native phonemes or sequences of phonemes. L2 consonant or vowel strings that are not permissible in the L1 may be corrected through insertion of an “epenthetic” vowel, or through substitution of a sound (Jacquemot et al., 2003). Therefore, the online perceptual monitoring of what they are producing is never sufficiently fine-tuned to drive improvements in the long-term motor representations of articulation. The same factors that are used to investigate linguistic processing differences between two languages, such as the age of acquisition, the level of proficiency reached, the amount of exposure to the second language and the degree of language use in everyday life (Abutalebi et al., 2001), also apply to the differences in motor and sensory representations for the two languages.

AGE OF ACQUISITION AND PROFICIENCY

Second language acquisition research has undergone many changes since it was first established. In the 1950s there was a shift from considering language acquisition as learning parrot-fashion to viewing it as a cognitive process, by which a limited resource, i.e., vocabulary, can be combined and used to express an unlimited range of concepts. The two most common claims for late bilingual learners relate to the actual age of acquisition and the level of proficiency attained. The debate about age of acquisition began in 1967 when Lenneberg first put forward his “critical period hypothesis,” which proposed a specific time period in which language acquisition can occur, with only a language learnt before puberty being

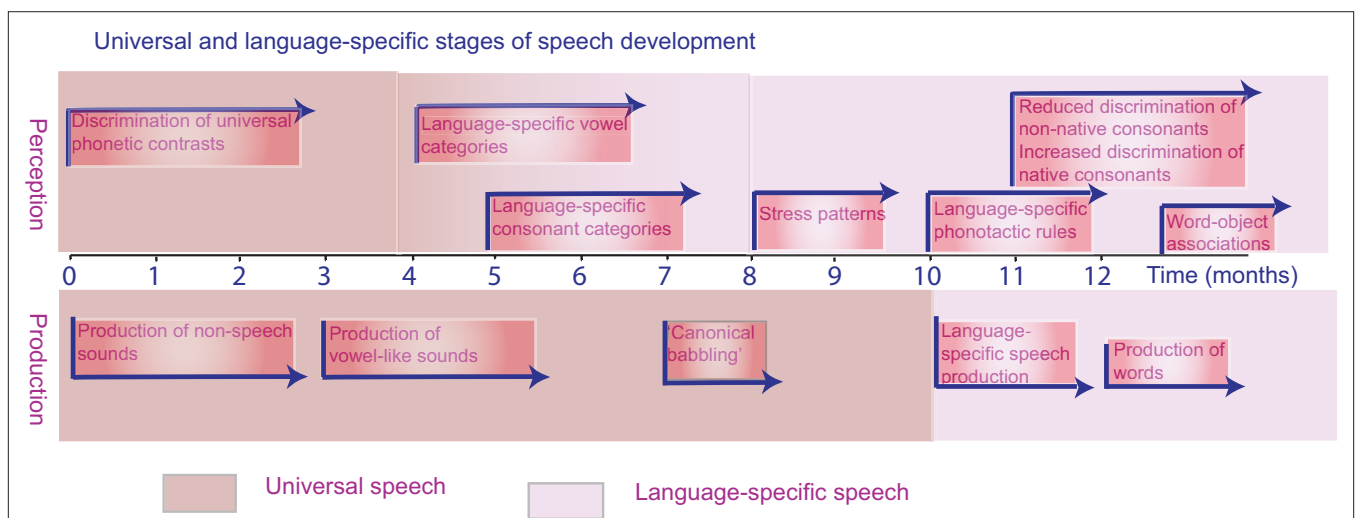


Figure adapted from Kuhl (2004) *Nat. Rev. Neurosci.* 5, 831-843

FIGURE 1 | Universal and language-specific stages of development for speech perception and speech production in typically developing human infants from birth to 1 year.

mastered to native proficiency (Lenneberg, 1967). The concept of this “critical period” was based on the idea that certain unspecified “electro-chemical” changes in the brain had reached maturity by the age of 10–12 years, after which implicit language acquisition can no longer occur (Lenneberg, 1967). Lateralization of cognitive functions has also been proposed as support for critical periods, as hemispheric specialization is established by early puberty. These proposals carry the inherent suggestion of loss of neocortical plasticity with maturation, meaning that with development language acquisition becomes increasingly difficult.

The critical period hypothesis was considered in more detail in the 1970s when Krashen reexamined Lenneberg’s data and concluded that whilst the critical period exists, it ends much sooner than Lenneberg suggested. Krashen proposed that the process of first language acquisition is complete by the age of 5 years and a second language learnt after that period would not reach native-like proficiency (in Danesi, 1994).

The motor and sensory aspects of bilingual speech production are clearly susceptible to some sort of “critical period,” in that it is much easier for younger children to learn a second language. We have already discussed the tuning of auditory cortex to language-specific speech sounds during the first year of life, and this must impact on the phonological competence accompanying the acquisition of L2, both in its perception and production (Scovel, 1969). Therefore, certain linguistic skills (vocabulary and grammar) are not as susceptible to age of acquisition limitations. However, it has been suggested that certain aspects of grammar acquisition are also less plastic than vocabulary acquisition; while semantic development uses associative learning mechanisms that are adaptable to later L2 learners, syntactic development uses a computational mechanism that is less plastic (Neville and Bavelier, 2001, in Scherag et al., 2004). Scovel (1988) redefined the critical period hypothesis to apply to a specific age (before puberty), a specific neurobiological change (lateralization), and a specific linguistic skill (the ability to sound like a native speaker). He claimed that native-like pronunciation is not possible for adult second language learners. This forms the major theme of this review. There are, of course, individual differences in levels of proficiency reached by later L2 learners, even with similar amounts and types of language training. It has previously been shown that there are brain structural differences between individuals that partly predict their ability at perceiving non-native speech sounds (Golestani et al., 2007).

The notion of critical periods for motor control (as involved in control of the articulators) can apply to fields outside language. It has been suggested that for experts in fields requiring great muscular dexterity (music, dance, skating, etc.), the acquisition process began in childhood, and the same could be said of language experts (Archibald, 1988). Penfield and Roberts (1959) also suggested that speech is not the only skill that is better acquired in childhood; it is also true for talents as diverse as piano and violin playing and skiing. The critical period hypothesis, especially as applied to speech production, might have less to do with cognition, and more to do with fine motor control. The articulation of human speech uses the many muscles that control breathing, the larynx, and the articulators, and so it could be that the reason adults have difficulty in mastering L2 speech is due to declining motor dexterity. This would explain why the mastery of the cognitive aspects of language, such as syntax and

vocabulary, remains possible for adult learners, whereas control of pronunciation reaches a level of proficiency below that of native speakers. Pronunciation is the only “physical” part of language with complex neuromuscular demands (gestures and handwriting use simple movements compared with speech production), and correct pronunciation is strongly dependent on sensory feedback of how and where the articulators are moving, with specific timings and sequences (Scovel, 1988). Other language aspects are “cognitive” or “perceptual,” rather than “physical” (Scovel, 1988). Scovel suggested that it is self evident that the motor expression of language would be most affected by the loss of neural plasticity that is hypothesized to occur with age. Even those who dispute the “critical period” hypothesis in general often accept that pronunciation may be one aspect in which the hypothesis could be valid (Walsh and Diller, 1978). This view is also supported by Long (1990), who has proposed that the age of a learner affects phonological attainment, with supra-segmental phonology being possible up the age of 6 years but the ability to acquire segmental phonology ending soon after that. However, Flege (1981) argues against this view, and suggests instead that accents arise not as a result of loss of plasticity but rather incorrect use of acoustic models of L2, due to interference from L1 (in Scovel, 1988). Under this psychomotor view, it could be argued that accents can be overcome if learners adapt their phonetic model of L2 phonemes to be less affected by L1 phonemes. However, Flege does point out that even if an L2 speaker could adapt their pronunciation to be more closely matched to the L2 phonology, there would still be differences between their pronunciation and that of a speaker for whom the L2 is a native language, even if these differences are only detectible using highly detailed acoustic analyses. He also claims that such phonetic learning of the L2 would affect pronunciation in the learner’s L1. Other arguments against Scovel’s biological constraints include the suggestion that accents can arise as a result of an adult L2 learner’s attitude and lack of motivation or discipline (Taylor, 1974), sociocultural expectations of language learning (Hill, 1970), and “cognitive maturation,” which causes adults to learn L2 differently from children (Dulay et al., 1982).

In contrast to a “critical period” as a hard biological constraint to later acquisition of a second language, there are occasional exceptions. A few adult learners of a second language manage to do so without an accent; and, in contrast, there are young learners who have slight accents (Flynn and Manuel, 1991). Abilities in foreign speech sound learning and articulation vary according to the individual, and previous work has shown that these individual differences correlate with brain structural differences in left insula/prefrontal cortex, left temporal cortex, and bilateral parietal cortices (Golestani et al., 2007, 2006). Salisbury (1962) and Sorensen (1967) also report adult learners who reach native-like proficiency in several languages, in contexts where multilingualism is necessary, such as in New Guinea and the Northwest Amazon. The social context of the language learner affects the level of proficiency attained. In contexts where it is necessary to speak as a native, the speaker will continue to progress, rather than fossilizing at the level of adequate communication, albeit with a non-native accent. To understand the neural basis of these exceptions requires research into the motor and sensory control of bilingual speech production.

In addition to age of acquisition, it has been suggested that language networks are affected by proficiency level. Abutalebi et al. (2007) demonstrated that activity in prefrontal regions reduces as the level of proficiency increases. Using a picture-naming task with German–French bilinguals, they found that when using the less proficient language, activation in the left caudate and anterior cingulate cortex was more extended. One interpretation is that the processes required to produce language become more automatic, requiring less domain-general executive control, as the language becomes more familiar. A number of studies have shown greater activation for lower proficiency (less proficient bilinguals compared with more proficient bilinguals, or the less proficiently spoken language of a bilingual compared with their more fluent language; Chee et al., 2001; Golestani et al., 2006; Stein et al., 2009). Several studies have also shown greater activation in prefrontal regions for tasks that require greater “top-down” processing (Frith et al., 1991; Raichle et al., 1994). These studies have also shown that as a task becomes more automated and the processing is more “bottom-up,” frontal activity decreases and activity in more posterior regions increases. The networks involved in bilingual speech production are more extensive than those concerned with linguistic processing; the increased activation for less proficient language production likely relates to cognitive processing as well. In line with this, processing a less fluent language can be considered more effortful and top-down, whereas a more fluent language is more automatic and bottom-up. This fits with Green’s (2003) “convergence” hypothesis, which states that convergence is possible because networks adapt; as proficiency in L2 increases, the representation of L2 and its processing profile converge with those of native speakers of that language. Qualitative differences between native and L2 speakers disappear as proficiency increases.

COGNITIVE DIFFERENCES IN ADULT L2 LEARNERS

It has also been argued, however, that not only do adults have the ability to acquire a second language as proficiently as young learners, they can be even more successful language learners. In some respects adults can be considered to have improved language-learning capabilities (Walsh and Diller, 1978). They can do better at certain cognitive levels, such as those involving grammatical and semantic complexity, as the neural systems responsible for these processes develop with age (Schlepppegrell, 1987). Older learners also have the advantage of a well-established first language, and they have the ability to integrate L2 with what they already know about L1. Their cognitive systems are more highly developed than those of young learners, enabling them to make higher-order associations and generalizations (Schlepppegrell, 1987). It has been shown that the skill of phonetic learning is stable within individuals and has structural correlates (Golestani et al., 2007). These cognitive differences in bilingualism are still being widely researched.

WHY HAS PREVIOUS BILINGUALISM RESEARCH LARGELY IGNORED MOTOR–SENSORY ASPECTS OF LEARNING?

It is usually obvious when a late bilingual is speaking either their native L1 or their later-acquired L2, because of their accent. Research into second language acquisition has considered the degree of accent, but not in terms of motor–sensory control. Within

the scope of this review, we suggest why functional imaging has largely ignored motor–sensory aspects of second language learning and how it can build upon data provided by linguistics research.

INTERLANGUAGE PHONOLOGY

A common analysis technique for L2 acquisition research is Contrastive Analysis, comparing two languages by investigating L1 and describing how it is different or similar to L2. The 1960s gave rise to the new notion of interlanguage phonology, in which comparisons are not only made between the target L2 and the L1, but also between the target L2 and variations of L2 that the speaker develops as proficiency increases. Learners have their own interlanguage phonologies, based on temporary rules that they develop throughout the learning process. Although research into phonology has been largely neglected, particularly by the functional imaging community, Alario et al. (2010) have recently investigated contrastive phonology in L1 and L2. They found that late bilinguals, but not early bilinguals, were sensitive to non-target syllable frequency. They interpret this by suggesting that syllable representations differ for the two groups of bilinguals. Early bilinguals are proposed to have independent syllable representations, whereas late bilinguals use the same representation for the two languages. For late bilinguals the syllable representation for L2 is based on their earlier L1 experience and consequently their L2 representations included non-native L1-like patterns. It has been suggested that the failure of non-native speakers to accurately produce L2 speech sounds may be a problem of phonetic implementation (articulation), rather than one of phonological encoding (auditory discrimination of speech sounds). It is apparent that a learner’s interlanguage phonology results from adapting the motor–sensory system during the course of acquiring L2 and these aspects of non-native speech production could be investigated with functional neuroimaging techniques, especially using more sensitive analyses, such as multivariate pattern analysis (see Raizada et al., 2010).

The acquisition of phonetic and articulatory skills must depend as much on sensory as motor processing. The study of these lower-level processes may not be best served by investigating highly constrained speech production tasks. For example, single word tasks, such as naming, reading aloud or repeating, do not reflect well what occurs during self-generated propositional speech. This may explain why earlier studies of motor differences in the production of L1 and L2 only demonstrated altered activity in the basal ganglia (Klein et al., 1994; Frenck-Mestre et al., 2005). Understanding the motor–sensory aspects of speech production, either in L1 or L2, is more sensitively studied with subjects producing whole phrases and sentences, rather than stimulus-led single word production, such as naming or repeating.

DIFFICULTY ASSESSING SPEECH

Another reason why the motor–sensory aspects of bilingual speech production are overlooked could relate to the difficulty in comparing these lower-level processes. Tasks such as picture naming, direct translation, or grammaticality judgments (all of which are frequently used in bilingual imaging studies) are easily assessed in an objective manner. Answers given by the subjects are either correct or incorrect, enabling comparison across conditions (i.e., native or non-native languages) only for trials in which the task was

successfully completed. In the case of studies investigating cognitive control, responses with longer reaction times or containing errors can be investigated. These trial-by-trial analyses are made possible by the development of event-related fMRI. However, with studies of overt propositional speech involving sentences, assessment is more subjective. The way in which one judges speaking proficiency is difficult to define. Whilst a native speaker may be able to listen and deem someone to be a “good” speaker, they often find it more difficult to explain why. The notion of fluency can be used, which is being able to communicate a message effectively, in real time, without undue hesitation or delay, as speech is a “real-time” phenomenon (Bygate, 1987). Therefore, although the correct use of vocabulary and syntax, spoken at a normal conversational rate, is central to language proficiency, the same sentence produced by two different proficient bilinguals may vary widely in accent. The problem is to come up with a reliable measure of “accent.” Perhaps the most sensitive indicator is a rater scale (Southwood and Flege, 1999). Other approaches, such as analysis of speech spectrograms (Arslan and Hansen, 1997) or possibly electropalatograms, may trade improved accuracy in assessing specific motor movements and consequences, possibly with more limited generalizability to all the dimensions that characterize human articulation.

SCAN ARTIFACTS OF SPEECH PRODUCTION

Scan artifacts arise in fMRI studies of overt speech production, encouraging many researchers to investigate covert speech production instead. These motion-induced artifacts can be categorized as two types: direct (resulting from head or jaw movements) and indirect signal changes (variations in the magnetic field; Gracco et al., 2005). Motion artifacts can both mask genuine signal changes due to neural activity and generate apparent “activity” time-locked to speech production.

Thus, a number of functional neuroimaging studies have reported activity within left temporo-parietal cortex, in a region termed Spt, during the *covert* production of both speech and non-speech vocalizations (Hickok and Poeppel, 2000; Pa and Hickok, 2008; Hickok et al., 2009), and related this to a motor–sensory interface for speech production. However, the investigation of imagined movements is only a partial substitute for the investigation of actual movements, and the results obtained with covert speech need to be directly contrasted with overt speech to fully understand speech-related activity in auditory and somatosensory association cortex. Some fMRI studies have compared covert and overt speech (Yetkin et al., 1995; Barch et al., 1999; Birn et al., 1999; Lurito et al., 2000; Rosen et al., 2000; Huang et al., 2001; Palmer et al., 2001; Shuster and Lemieux, 2005), but these have only included 10 or fewer participants. Larger studies will be required to allow closer investigation of covert and overt speech-related activity in auditory and somatosensory association cortex.

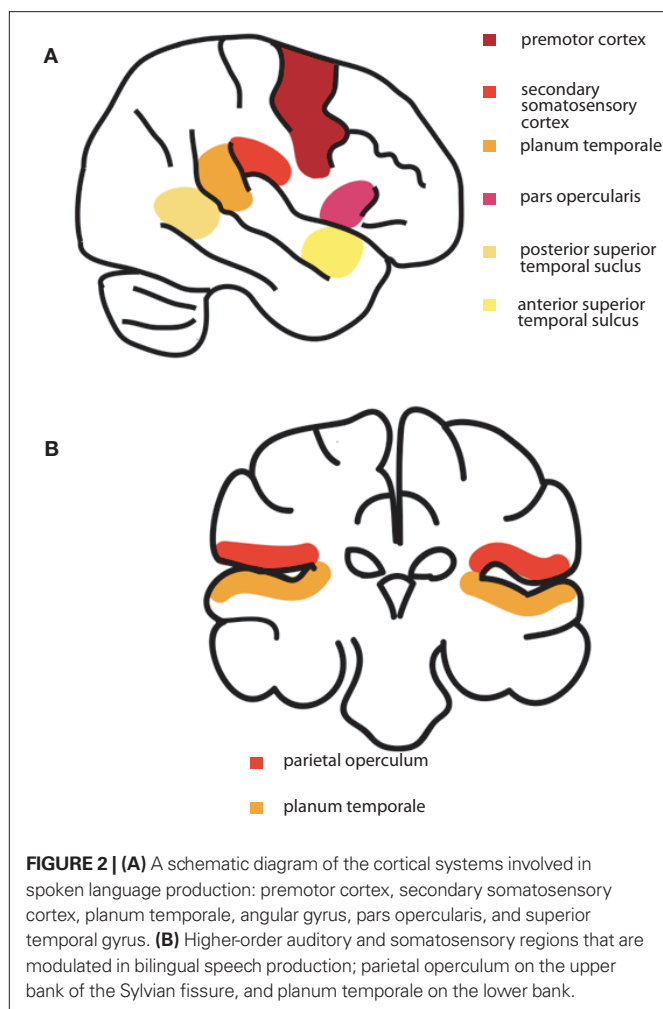
EVIDENCE FOR MOTOR–SENSORY CONTRIBUTIONS IN BILINGUALISM

In speech production models the cognitive planning stages are often described in great detail, whereas articulation is listed as a simple motor output (Indefrey and Levelt, 2004). Models that do provide detail of phonological encoding and articulation, and can encompass developmental changes in size of the articulators, have

been developed for monolingual speech production (Guenther et al., 2006; Golfinoopoulos et al., 2010). It has been proposed that the bilingual speaker is more than the sum of two monolingual speakers (Abutalebi et al., 2007) and consequently, monolingual speech production models are not necessarily sufficient to explain bilingual speech production.

Speech production requires the coordination of up to 100 muscles (Ackermann and Riecker, 2010). Feedforward signals from premotor and motor cortex, as well as contributions from other regions, including the insula, control the passage of air through the larynx and the movement of the articulators (Gracco and Lofqvist, 1994), and the shaping of the vocal tract results in the production of the intended acoustic signal (Nieto-Castanon et al., 2005). In parallel, feedback sensory systems, in superior temporal auditory and parietal somatosensory regions, provide online monitoring of the sensory consequences of the utterance, so that rapid motor adjustments can be made (Guenther et al., 2006). Other regions may also play a role in feedback monitoring. For example multi-modal processing, particularly integrating auditory and visual inputs, is likely carried out by the insula, demonstrated by its involvement in the McGurk effect (Bamiou et al., 2003). The final neural pathways integrate feedforward motor discharges with auditory and somatosensory feedback, to match expectation with outcome and monitor online for articulatory errors (Guenther et al., 2006; Golfinoopoulos et al., 2010) and speech adaptation involves both input and output processes simultaneously (Shiller et al., 2009). This comparison between predicted and actual movement and their sensory consequences, as well as any mismatch signals, are likely sent via climbing fibers that project from the inferior olive to the cerebellum (Blakemore et al., 2001). **Figure 2** shows a schematic diagram of the cortical systems involved in spoken language production. Both somatosensory and auditory feedback systems show some suppression of activation during self-initiated actions such as speech production (Paus et al., 1996; Aliu et al., 2008). This reduction in activity reflects cortical modulation by a parallel copy of the motor command, which could originate from premotor cortex, primary motor cortex, or both, sent to sensory cortical regions processing articulatory feedback signals. This allows more efficient feedback-appropriate sensory processing (Eliades and Wang, 2008), permitting the rapid detection and correction of articulatory errors (Guenther et al., 2006).

There is convincing evidence that the motor system generates internal representations of speech sounds (Wilson and Iacoboni, 2006). In a native language, these representations match auditory input received. The oral motor movements necessary for producing native speech sounds are highly over-learned and automatic, integrating feedforward motor and feedback auditory, and somatosensory information. However, in a foreign language, auditory, and somatosensory input does not match the internal representations. Ludlow and Cikaaja (1998) propose that, for a fluent speaker, the internal representation of speech produced by oneself is highly similar to the perception of speech produced by other native speakers. However, when learning a foreign language, perception, and internal representations likely match less well. The oral motor movements necessary for non-native phonemes are unfamiliar and require greater engagement of motor–sensory neural feedback systems. Speakers need to become aware of the



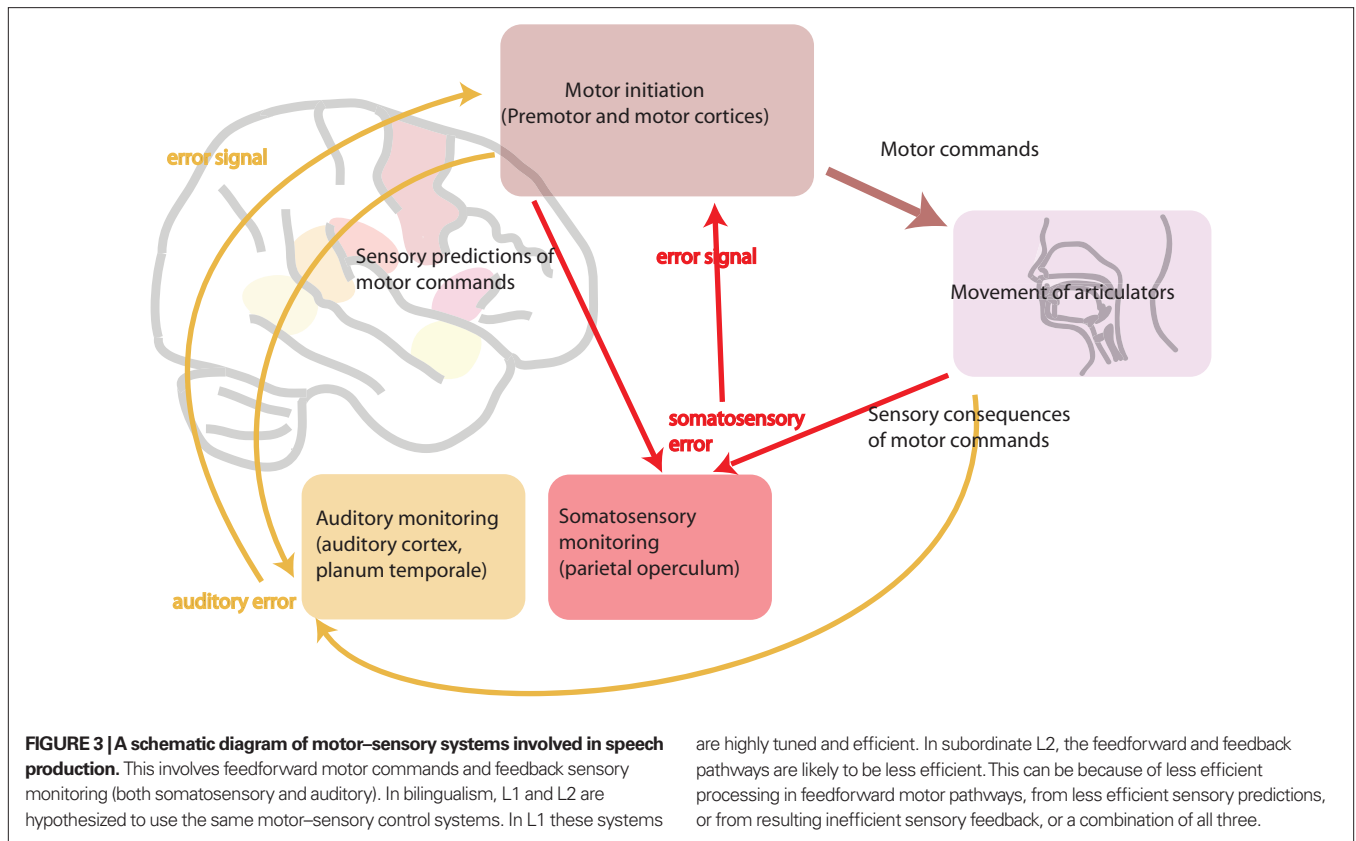
differences between their internal representation of speech and the perception of speakers of that foreign language. It is then necessary to map that new perception onto their own internal representation, in order to be able to produce accurate speech sounds in the foreign language. We suggest that auditory feedback is crucial to this process of modifying the speech motor-control system. When learning and speaking a foreign language, it is also hypothesized that online modifications to existing articulatory-acoustic relationships are necessary to produce accurate speech sounds. The persistent accent in late learners of a second language is likely to be the result of a failure to achieve the same proficiency in integrating the motor feedforward and sensory feedback control of articulation that we achieve when speaking in our native tongue. **Figure 3** shows a schematic diagram of motor-sensory systems involved in speech production.

It has been shown, using diffusion tensor imaging (DTI) to reveal white matter tracts *in vivo*, that the component of the putamen that forms part of the “motor loop” is connected to both primary sensory and motor areas and to medial premotor cortex in the posterior part of the SMA (Lehericy et al., 2004), areas active during overt speech production. In addition, Booth et al. (2007) suggested that the cerebellum and basal ganglia may be recruited in the modulation of articulatory or phonological output

representations, as demonstrated by the use of a rhyming task with fMRI. With L2 production, consideration has to be given to the effects on the recorded BOLD signal resulting from the production of novel or partially trained sequences of motor commands, and their effects on feedback signals that are only partially attuned to the sound and somatosensations of L2. These signals will be further confounded by an increased number of speech errors, whether or not the subject is consciously aware of these errors, and if so whether attempts at self-correction are or are not initiated. This contrasts with the highly automatic processing of native speech. More recently, it has been shown that processes such as articulation and post-articulatory monitoring result in greater activation for bilinguals than monolinguals (Parker Jones et al., 2011).

Moser et al. (2009) demonstrated that the production of unfamiliar speech sounds resulted in greater extent and intensity in the BOLD signal, compared to the production of familiar speech sounds. Increased activity in motor speech networks may directly reflect the lack of familiarity with the motor commands necessary to produce the target. Moser and colleagues discuss their results with particular reference to the anterior insula and adjacent frontal operculum and their roles in the formation and sequencing of articulatory gestures for novel native and non-native speech sounds embedded in non-words. It has been suggested that the insula is involved in allocating auditory attention and is activated more strongly for processing unfamiliar rather than familiar auditory stimuli (Bamiou et al., 2003). However, insula contributions may be modality-independent, reflecting salience of the stimuli (Seeley et al., 2007).

Klein et al. (1994) found activation in the putamen when subjects repeated words in L2, which they attribute to the role of left basal ganglia in articulation, particularly the precise motor timing of speech output, which is less automatic and more “effortful” in a language acquired later in life. This study was followed up in 1995 with three lexical search tasks (rhyme generation, synonym generation and translation), with word repetition as a control task (Klein et al., 1995). They found similar areas of activation for both within- and across-language searches, i.e., there were no significant differences related to whether the task used phonological or semantic cues or whether it used L1 or L2. Contrasting L1 translation with L1 repetition resulted in increased activation in the left putamen, as did the contrasts of L1 translation – L1 rhyming, L2 synonyms – L1 synonyms and L2 translation – L1 translation. Klein et al. (1999) extended their studies from English–French bilinguals to English and Mandarin Chinese bilinguals, using highly proficient speakers who had learnt L2 during adolescence. Using a noun–verb single word generation task, they demonstrated that common cortical areas were activated in fluent bilinguals who use both languages in daily life. Therefore, it appeared that similar brain regions are active even when the languages are typologically distant, such as English and Mandarin Chinese, and when L2 is acquired later in life. This is in line with their previous finding of similar brain regions for word repetition in L1 and in L2 (Klein et al., 1994, 1995). However, only the earlier studies (1994; 1995) demonstrated activity that could be attributed to motor control, and no increased activity in the left putamen associated with L2 was observed in their more recent study (1999), despite the fact that the L2 Mandarin Chinese was heavily accented. One suggestion put forward by the authors



to explain this disparity between studies was that the latter study required mono- or bi-syllabic production, whereas in the earlier study responses were mostly bi- or multi-syllabic. Further studies are required in order to investigate the effect of syllable counts on brain regions involved in articulation, particularly the basal ganglia.

Frenck-Mestre et al. (2005) also found increased activity in the left putamen for bilinguals who had learnt L2 after the age of 12, compared to early bilinguals, in a reading aloud task, both at word- and sentence-level. Otherwise activity in cortical, subcortical and cerebellar regions was identical for both L1 and L2 in both groups. The authors suggest that learning to produce new articulatory patterns necessary for speaking an L2 requires adaptation of existing neural networks, rather than recruitment of new networks. Of course, relating activity in the basal ganglia to particular processes is problematic, as different regions of the striatum are connected to widely distributed cortical areas that subserve very different functions, motor, cognitive, and emotional.

Aglioti et al. (1996) discuss the first neurolinguistically assessed case of bilingual subcortical aphasia and found that, due to a left capsulo-putaminal lesion, the patient had a speech deficit in their L1, with the much less practiced L2 being relatively spared. The patient spontaneously spoke only L2, and when L1 speech was elicited, it was non-fluent, slow, and characterized by a low voice. They also report that the patient spoke L2 with a foreign accent, which they attributed to left basal ganglia pathology. The left basal ganglia is involved in implicit memory and lesions here tend to affect the most automatic language of a bilingual. Despite the rarity of this dissociation, it has also been reported by Gelb (1937, in

Paradis, 1983). He described a patient who could no longer speak L1 (German) but could speak L2 (Latin), which he had studied formally as an adult. Speedie et al. (1993) reported a bilingual patient with a basal ganglia lesion, but on the right, and while propositional speech was unaffected, automatic (non-propositional) speech was impaired, in both L1 and in L2 (in Aglioti et al., 1996).

ADAPTATIONS TO AUDITORY AND SOMATOSENSORY FEEDBACK DURING L1 PRODUCTION

Speech sounds are defined acoustically and therefore rely on auditory feedback. However, congenitally deaf speakers have demonstrated that somatosensory feedback plays an important role and it has previously been shown that the somatosensory information constitutes a principal component of the speech target, independently of the acoustic information (Tremblay et al., 2003). In that study, subjects received altered somatosensory feedback due to mechanical jaw perturbations, although auditory feedback remained unaltered. Even when speech acoustics were unaltered by the jaw perturbation, subjects adapted their jaw movements when producing speech, indicating that the somatosensory target in speech is monitored independently of the acoustics. Nasir and Ostry (2006) provided evidence for the central role of somatosensory feedback in speech production by using a robotic device to alter jaw movements during speech. Speech acoustics were unaffected, demonstrating a dissociation between the influences of auditory and somatosensory feedback on speech production. Houde and Jordan (1998) investigated how altering auditory feedback affects speech motor control, relating their study to previous work in which

limb motor-control systems adapt to changes in visual feedback. Using native-English speakers, they demonstrated that the control system involved in vowel production adapts to altered auditory feedback. Similarly, using native Mandarin speakers, Jones and Munhall (2005) found that in response to altered auditory frequency feedback, subjects automatically adjusted the pitch of their speech. This suggests that the motor control of vocal pitch requires continuous mapping between the laryngeal motor system and the vocal output and that this mapping relies on auditory feedback. A further study, by Tourville et al. (2008), who altered auditory feedback during single word reading in L1, demonstrated an increase in activity in posterior auditory association cortex (including planum temporale) and in the parietal operculum (second-order somatosensory cortex). The subjects had no awareness of this alteration, and yet they automatically altered speech production as a motor compensation to counter the auditory perturbation.

INCREASED ACTIVITY IN RESPONSE TO SENSORY FEEDBACK DURING L2 PRODUCTION

When speakers use L1, it has been shown that there is a paradoxical suppression of neural activity (“sensory gating”) in second-order somatosensory association cortex in the parietal operculum (Dhanjal et al., 2008). Although auditory association cortex is active during overt speech, this activity is less than when listening to the voice of another (Wise et al., 1999; Ventura et al., 2009), providing further evidence of the importance of internal sensory feedback. This suppression of activity may indicate the efficiency of online sensory–motor monitoring during L1 speech production, which is highly automatic. In contrast, L2 speech production is less automatic and may result in increased activity in response to sensory feedback.

Previous work from our group has shown that, by considering self-generated, fluent, sentential speech, we see a more extensive picture of the distributed neural systems involved in non-native speech production (Simmonds et al., 2011). We used fMRI to compare native and non-native propositional speech production with the specific aim of revealing the changes in motor–sensory control when switching from native speech production to speech in a second language. Subjects were instructed to give definitions of visually presented pictures, in either their native language (L1) or English (L2 for all subjects). Rest was included as the baseline condition. We predicted and observed that altered feedback processing in the non-native language resulted in increased activity in sensory regions, particularly in second-order somatosensory cortex. This network involves both auditory and somatosensory areas not previously revealed by previous functional imaging studies of bilingualism. Our results demonstrate that learning to speak a second language after the normal period of “innate” language acquisition (i.e., an L2 that is acquired after L1 has already been established) has functional consequences on cortex involved in auditory and somatosensory feedback control of articulation (Simmonds et al., 2011). Demonstrating sensory feedback in bilingualism was made possible by using a task dependent on propositional speech production, rather than single word production. A prospective training study on novel speech sounds is required to understand the relative motor–sensory contributions from feedforward motor systems, sensory predictions, and sensory feedback.

POTENTIAL LINKS BETWEEN BILINGUAL ARTICULATION AND COMPREHENSION

In this section we examine whether perception and production of phonological features can be dissociated or whether production abilities depend on accurate perception of the target phonological distinctions. The criteria by which you can judge “being good at speaking” are vast, and even in the target language there are regional variations, yet it is possible to use functional imaging data to investigate language proficiency in a deeper way than behavioral measures alone. We discuss how neuroimaging data of native and non-native speech production can be used to inform current theoretical models of bilingual language processing.

LANGUAGE INPUT AND ITS IMPACT ON SPEECH OUTPUT

Flege and Hillenbrand (1984) investigated limits on phonetic accuracy of adult L2 learners. They suggest that a non-native accent is in part due to phonological and phonetic differences between the speaker’s L1 and L2. They cite Weinreich’s (1953) hypothesis, that a non-native learner makes substitutions for phones or phonemes when there are similarities between them in L1 and L2. For example, the phone/s/sounds similar in French and English but the place of articulation is different in the two languages (dental in French and alveolar in English). A native-English adult learner of French perceives the acoustic similarity between the native and non-native/s/and does not adapt production, pronouncing/s/in French as an alveolar phone (Flege and Hillenbrand, 1984). Using the/r-/l/contrast varying in frequency of the second and third formants in a study with English, Japanese, and German speakers, Iverson et al. (2003) showed that speakers of different languages attend to different dimensions of the perceptual input, even when the same stimulus is used. There are many other examples of perceptual biases with respect to how speech sounds are heard, particularly as a function of the L1 phonetic repertoire, which likely influence L2 production. See Kuhl et al. (2008) for a review of phonetic perception development models and Hickok et al. (2011) for a review of auditory input affecting language output.

SPEECH OUTPUT AND ITS IMPACT ON LANGUAGE INPUT

Lenneberg (1962, in Krashen, 1982) presented the case of a boy with severe congenital dysarthria. However, he was able to understand spoken English perfectly. Lenneberg claimed that the boy had acquired “competence” in the language, without ever speaking it himself. In Krashen’s (1982) view, producing speech output does play a role in language acquisition, but only indirectly. The benefit of speaking is not that it improves language acquisition itself, but rather that the acquisition and use of fluent speech encourages dialog with others. Thus, speaking increases being spoken at, the quantity of which is matched by the quality, as native speakers use more natural language to learners they deem to be at a higher level. Native speech is often modified to include “foreigner talk,” i.e., simplifying the language to make it accessible to a non-native speaker. A second language learner who does not attempt to speak much, who makes lots of mistakes, has a strong non-native accent and speaks hesitantly, will often be spoken to in a more simple version of the target language than a speaker who appears fluent.

IMPLICATIONS OF TAKING A MOTOR–SENSORY APPROACH TO BILINGUAL SPEECH PRODUCTION

PEDAGOGIC IMPLICATIONS: ADULTS ACQUIRING NATIVE-LIKE LEVELS OF PRODUCTION

Taking a more motor–sensory approach to understanding bilingualism implies that, in line with L1 acquisition, adult second language learners might benefit from a mute period – a period of intense auditory exposure to L2 before attempting to produce the sounds. This may prove beneficial in enabling the learner to hear (and thus produce) subtly different phonetic features, new phoneme distinctions and unfamiliar sequences of stress patterns. One possibility is that an artificially induced mute period may protect the learner from using first language phonological categories to represent the L2 system, thus enabling higher levels of production performance and avoiding L1 transfer or interference.

Neufeld (1979) showed that, with the right method of instruction, adults can acquire native-like pronunciation (in Archibald, 1988). Students were trained on pronunciation of certain sounds from Inuktitut, a language to which they had not been exposed previously. The learning process involved a lot of time listening to the language, with no attempt at producing the sounds. They were later instructed to produce the sounds and their attempts was rated by native speakers and much of the speech was deemed to be native-like. Neufeld claimed that the silent period at the beginning helped the students to accurately produce the language later. Removing students' own attempts allowed perception to remain more plastic, such that the L2 acoustic template is heard accurately before erroneous phonetic utterances in L2 become crystallized. Producing the sound too early, and therefore incorrectly, would have influenced this acoustic template and thus hindered their production.

BABBLING ADULTS?

As well as the benefits adult second language learners might gain from a mute period, there is the possibility that a babbling phase might also improve non-native speech learnt later in life. By imitating the target speech sounds in isolation, before attempting to produce them in word form, adult learners might develop more accurate efference copies of the motor commands required for the production of the sounds, allowing more efficient feedback for the monitoring and correction of articulatory errors.

THE ROLE OF LITERACY IN L1 AND IN L2

Skilled readers are very familiar with the written form of their native language, and automatically decode the grapheme by producing a phonological representation of the sound (Snow, 2002; although when reading text skilled readers progress to recognize the whole word form and read at an appropriate speed). For example, the letter “p” is highly familiar to a literate adult native speaker of English and each time they read that letter, they associate it with the typical English phonological representation, which is aspirated and has a relatively long voice-onset time. If a native-English speaker begins learning Spanish, it is likely that they would transfer the typical phonological representation used in spoken English and therefore pronounce the Spanish “p” with an English accent, rather than the non-aspirated bilabial stop with a short voice-onset time that a native Spanish speaker would produce. It would seem possible,

therefore, that learning a language without being exposed to the written form would lead to more accurate pronunciation. In line with this, it could also be expected that learning a second language with a different orthography to one's native language would also result in more accurate foreign language pronunciation. Learning orally has similar advantages to listening before speaking, discussed in an earlier section of this review. Representations of L2 speech sounds are more plastic and less influenced by automated activation of native representations that might be triggered by reading a letter.

CLINICAL IMPLICATIONS

Comparing speech production in a foreign language with that in the native language provides insights into how feedforward and feedback systems operate, which could help explain what goes wrong in the case of motor speech disorders, such as dysarthria. Dysarthria is characterized by problems with pronunciation, making speech sound slurred. Focal lesions resulting in dysarthria can be very differently located: the cerebral cortex, white matter, basal ganglia, thalamus, cerebellum, and brain stem, anywhere that might compromise the motor–sensory pathways that control the final outflow from motor neurons to the articulatory muscles. Patients show slow and uncoordinated movements of the muscles required for speech (Sellars et al., 2005), possibly with impaired prosody (Ackermann and Hertrich, 2000). Although communication disorders after stroke are often the result of aphasia and associated with problems with word understanding and word retrieval, in dysarthric patients, the predominant disorder is with the motor aspects of speech. Although natural recovery can occur, so that dysarthria, although still present, is so mild as not to impair intelligibility, in a proportion of patients intelligibility is so compromised that alternative communication systems need to be used.

Greater understanding of how brain processes underlying speech production adapt to non-native speech may help develop more effective therapy techniques to be developed in order to improve dysarthric speech. Similar to previous research on bilingualism, most research on intervention in communication disorders following stroke has focused on impairment of linguistic and semantic representations and processes. In contrast, the motor–sensory control of speech production has received much less attention.

BENEFITS OF RESEARCHING NEURAL BASES OF BILINGUALISM

Theories of second language acquisition can inform our understanding of how the brain works. Conversely, understanding how the brain can adapt to multiple languages could provide an empirical basis to constrain cognitive and linguistic theories of second language acquisition (Danesi, 1994). For the learning of second languages to become more effective, the teaching of such languages needs to become more “brain-compatible.” This review has discussed the largely overlooked motor–sensory aspects of bilingual speech production and argued that investigating these aspects will lead to a more comprehensive understanding of bilingual speech and ways in which the teaching of this can become more effective.

ACKNOWLEDGMENTS

This work was supported by the Medical Research Council and the Research Councils UK.

REFERENCES

- Abutalebi, J., Annoni, J.-M., Zimine, I., Pegna, A. J., Seghier, M. L., Lee-Jahnke, H., Lazeyras, F., Cappa, S. F., and Khateb, A. (2007). Language control and lexical competition in bilinguals: an event-related fMRI study. *Cereb. Cortex* 18, 1496–1505.
- Abutalebi, J., Cappa, S. F., and Perani, D. (2001). The bilingual brain as revealed by functional neuroimaging. *Biling. Lang. Cogn.* 4, 179–190.
- Abutalebi, J., and Green, D. (2007). Bilingual language production: the neurocognition of language representation and control. *J. Neurolinguist.* 20, 242–275.
- Abutalebi, J., and Green, D. (2008). Control mechanisms in bilingual language production: neural evidence from language switching studies. *Lang. Cogn. Process.* 23, 557–582.
- Abutalebi, J., Rosa, P. A. D., Tettamanti, M., Green, D. W., and Cappa, S. F. (2009). Bilingual aphasia and language control: A follow-up fMRI and intrinsic connectivity study. *Brain Lang.* 109, 141–156.
- Ackermann, H., and Hertrich, I. (2000). The contribution of the cerebellum to speech processing. *J. Neurolinguist.* 13, 95–116.
- Ackermann, H., and Riecker, A. (2010). The contribution(s) of the insula to speech communication: a review of the clinical and functional imaging literature. *Brain Struct. Funct.* 214, 5–6.
- Aglioti, S., Beltramello, A., Giradi, F., and Fabbro, F. (1996). Neurolinguistic and follow-up study of an unusual pattern of recovery from bilingual subcortical aphasia. *Brain* 119, 1551–1564.
- Alario, F. X., Goslin, J., Michel, V., and Laganaro, M. (2010). The functional origin of the foreign accent. *Psychol. Sci.* 21, 15–20.
- Aliu, S. O., Houde, J. F., and Nagarajan, S. S. (2008). Motor-induced suppression of the auditory cortex. *J. Cogn. Neurosci.* 21, 791–802.
- Archibald, J. (1988). *Second Language Phonology*. Amsterdam/Philadelphia: John Benjamins Publishing Co.
- Arslan, L. M., and Hansen, J. H. L. (1997). A study of temporal features and frequency characteristics in American English foreign accent. *J. Acoust. Soc. Am.* 102, 28–40.
- Bamiou, D.-E., Musiek, F. E., and Luxon, L. M. (2003). The insula (Island of Reil) and its role in auditory processing. *Brain Res. Rev.* 42, 143–154.
- Barch, D. M., Sabb, F. W., Carter, C. S., Braver, T. S., Noll, D. C., and Cohen, J. D. (1999). Overt verbal responding during fMRI scanning: empirical investigations of problems and potential solutions. *Neuroimage* 10, 642–657.
- Binder, J. R., Desai, R. H., Graves, W. W., and Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb. Cortex* 19, 2767–2796.
- Birn, R. M., Bandettini, P. A., Cox, R. W., and Shaker, R. (1999). Event-related fMRI of tasks involving brief motion. *Hum. Brain Mapp.* 7, 106–114.
- Blakemore, S.-J., Frith, C. D., and Wolpert, D. M. (2001). The cerebellum is involved in predicting the sensory consequences of action. *Neuroreport* 12, 1879–1884.
- Booth, J. R., Wood, L., Lu, D., Houk, J. C., and Bitan, T. (2007). The role of the basal ganglia and the cerebellum in language processing. *Brain Res.* 1133, 136–144.
- Bygate, M. (1987). *Speaking*. Oxford: Oxford University Press.
- Chee, M. W. L., Hon, N., Ling Lee, H., and Siong Soon, C. (2001). Relative language proficiency modulates BOLD signal change when bilinguals perform semantic judgments. *Neuroimage* 13, 1155–1163.
- Costa, A., and Santesteban, M. (2004). Lexical access in bilingual speech production: evidence from language switching in highly proficient bilinguals and L2 learners. *J. Mem. Lang.* 50, 491–511.
- Crinion, J., Turner, R., Grogan, A., Hanakawa, T., Noppeney, U., Devlin, J. T., Aso, T., Urayama, S., Fukuyama, H., Stockton, K., Usui, K., Green, D. W., and Price, C. J. (2006). Language control in the bilingual brain. *Science* 312, 1537–1540.
- Curio, G., Neuloh, G., Numminen, J., Jousmaki, V., and Hari, R. (2000). Speaking modifies voice-evoked activity in the human auditory cortex. *Hum. Brain Mapp.* 9, 183–191.
- Danesi, M. (1994). The neuroscientific perspective in second language acquisition research: a critical synopsis. *Int. Rev. Appl. Ling.* 32, 201–229.
- Dhanjal, N. S., Handunnetthi, L., Patel, M. C., and Wise, R. J. S. (2008). Perceptual systems controlling speech production. *J. Neurosci.* 28, 9969–9975.
- Dulay, H., Burt, M., and Krashen, S. D. (1982). *Language Two*. New York: Oxford University Press.
- Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., and Mehler, J. (1999). Epenthetic vowels in Japanese: a perceptual illusion? *J. Exp. Psychol. Hum. Percept. Perform.* 25, 1568–1578.
- Eliades, S. J., and Wang, X. (2008). Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature* 453, 1102–1107.
- Flege, J. E. (1981). The phonological basis of foreign accent. *TESOL Quart.* 15, 443–455.
- Flege, J. E., and Hillenbrand, J. (1984). Differential use of temporal cues to the [s-z] contrast by native and non-native speakers of English. *J. Acoust. Soc. Am.* 79, 508–517.
- Flynn, S., and Manuel, S. (1991). “Age-dependent effects in language acquisition: an evaluation of critical period hypotheses,” in *Point Counterpoint: Universal Grammar in the Second Language*, ed. L. Eubank (Amsterdam: Benjamins), 117–145.
- Francis, W. S. (1999). Cognitive integration of language and memory in bilinguals: semantic representation. *Psychol. Bull.* 125, 193–222.
- Frenck-Mestre, C., Anton, J. L., Roth, M., Vaid, J., and Viallet, F. (2005). Articulation in early and late bilinguals’ two languages: evidence from functional magnetic resonance imaging. *Neuroreport* 16, 761–765.
- Frith, C., Friston, K. J., Liddle, P. F., and Frackowiak, R. S. J. (1991). A PET study of word finding. *Neuropsychologia* 29, 1137–1148.
- Gelb, A. (1937). Zur medizinischen Psychologie und philosophischen Anthropologie. *Acta Psychol. (Amst.)* 3, 193–271. [Excerpt translated in: Paradis, M. (ed.). (1983). *Readings on Aphasia in Bilinguals and Polyglots*. Montreal: Marcel Didier, 383–4].
- Golestani, N., Alario, F.-X., Meriaux, S., Le Bihan, D., Dehaene, S., and Pallier, C. (2006). Syntax production in bilinguals. *Neuropsychologia* 44, 1029–1040.
- Golestani, N., Molko, N., Dehaene, S., LeBihan, D., and Pallier, C. (2007). Brain structure predicts the learning of foreign speech sounds. *Cereb. Cortex* 17, 575–582.
- Golfinopoulos, E., Tourville, J. A., and Guenther, F. H. (2010). The integration of large-scale neural network modeling and functional brain imaging in speech motor control. *Neuroimage* 52, 862–874.
- Gracco, V. L., and Lofqvist, A. (1994). Speech motor coordination and control: evidence from lip, jaw and laryngeal movements. *J. Neurosci.* 14, 6585–6597.
- Gracco, V. L., Tremblay, P., and Pike, B. (2005). Imaging speech production using fMRI. *Neuroimage* 26, 294–301.
- Green, D. W. (2003). “Neural basis of lexicon and grammar in L2 acquisition,” in *The Lexicon-Syntax Interface in Second Language Acquisition*, eds R. van Hout, A. Hulk, F. Kuiken, and R. Towell (Amsterdam: John Benjamins), 197–218.
- Guenther, F. H., Ghosh, S. S., and Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain Lang.* 96, 280–301.
- Hall, D. A., Haggard, M. P., Akeroyd, M. A., Palmer, A. R., Summerfield, A. Q., Elliott, M. R., Gurney, E. M., and Bowtell, R. W. (1999). “Sparse” temporal sampling in auditory fMRI. *Hum. Brain Mapp.* 7, 213–223.
- Hernandez, A. E., Martinez, A., and Kohnert, K. (2000). In search of the language switch: an fMRI study of picture naming in Spanish-English bilinguals. *Brain Lang.* 73, 421–431.
- Hickok, G., Houde, J., and Rong, F. (2011). Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron* 69, 407–422.
- Hickok, G., Okada, K., and Serences, J. T. (2009). Area Spt in the human planum temporale supports sensory-motor integration for speech processing. *J. Neurophysiol.* 101, 2725–2732.
- Hickok, G., and Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends Cogn. Sci.* 4, 1463–1467.
- Hill, J. (1970). Foreign accents, language acquisition, and cerebral dominance revisited. *Lang. Learn.* 20, 237–248.
- Houde, J. F., and Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Science* 279, 1213–1216.
- Huang, J., Carr, T. H., and Cao, Y. (2001). Comparing cortical activations for silent and overt speech using event-related fMRI. *Hum. Brain Mapp.* 15, 39–53.
- Indefrey, P. (2006). A meta-analysis of hemodynamic studies on first and second language processing: which suggested differences can we trust and what do they mean? *Lang. Learn.* 56, 279–304.
- Indefrey, P., and Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition* 92, 101–144.
- Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., and Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition* 87, B47–B57.
- Jacquemot, C., Pallier, C., LeBihan, D., Dehaene, S., and Dupoux, E. (2003). Phonological grammar shapes the auditory cortex: a functional magnetic resonance imaging study. *J. Neurosci.* 23, 9541–9546.
- Jones, J. A., and Munhall, K. G. (2005). Remapping auditory-motor representations in voice production. *Curr. Biol.* 15, 1768–1772.
- Klein, D., Milner, B., Zatorre, R. J., Meyer, E., and Evans, A. C. (1995). The neural substrates underlying word generation: a bilingual functional-imaging study. *Proc. Natl. Acad. Sci. U.S.A.* 92, 2899–2903.
- Klein, D., Milner, B., Zatorre, R. J., Zhao, V., and Nikelski, J. (1999). Cerebral

- organization in bilinguals: a PET study of Chinese-English verb generation. *Neuroreport* 10, 2841–2846.
- Klein, D., Zatorre, R. J., Milner, B., Meyer, E., and Evans, A. C. (1994). Left putaminal activation when speaking a second language: evidence from PET. *Neuroreport* 5, 2295–2297.
- Kovacs, A. M., and Mehler, J. (2009). Flexible learning of multiple speech structures in bilingual infants. *Science* 325, 611–612.
- Kovelman, I., Baker, S. A., and Petitto, L.-A. (2008). Bilingual and monolingual brains compared: a functional magnetic resonance imaging investigation of syntactic processing and a possible “neural signature of bilingualism”. *J. Cogn. Neurosci.* 20, 153–169.
- Krashen, S. D. (1982). *Principles and Practice in Second Language Acquisition*. Oxford: Pergamon Press.
- Krashen, S. D., and Terrell, T. D. (1983). *The Natural Approach: Language Acquisition in the Classroom*. London: Prentice Hall Europe.
- Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nat. Rev. Neurosci.* 5, 831–843.
- Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., and Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 979–1000.
- Lamendella, J. T. (1977). General principles of neurofunctional organization and their manifestations in primary and non-primary language acquisition. *Lang. Learn.* 27, 155–196.
- Lehericy, S., Ducros, M., Van de Moortele, P.-F., Francois, C., Thivard, L., Poupon, C., Swindale, N., Ugurbil, K., and Kim, D.-S. (2004). Diffusion tensor fiber tracking shows distinct corticostriatal circuits in humans. *Ann. Neurol.* 55, 522–529.
- Lenneberg, E. (1962). Understanding language without ability to speak: a case report. *J. Abnorm. Soc. Psych.* 65, 117–122.
- Lenneberg, E. (1967). *The Biological Foundations of Language*. New York: Wiley.
- Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.
- Levelt, W. J. M., Roelofs, A., and Meyer, A. S. (1999). A theory of lexical access in speech production. *Behav. Brain Sci.* 22, 1–75.
- Long, M. H. (1990). Maturation constraints on language development. *Stud. Sec. Lang. Acquisit.* 12, 251–285.
- Loucks, T. M. J., Poletto, C. J., Simonyan, K., Reynolds, C. L., and Ludlow, C. L. (2007). Human brain activation during phonation and exhalation: common volitional control for two upper airway functions. *Neuroimage* 36, 131–143.
- Ludlow, C. L., and Cikaoja, D. B. (1998). Is there a self-monitoring speech perception system? *J. Commun. Disord.* 31, 505–510.
- Lurito, J. T., Karekan, D. A., Lowe, M., Chen, S. H. A., and Mathews, V. P. (2000). Comparison of rhyming and word generation with fMRI. *Hum. Brain Mapp.* 10, 99–106.
- Moser, D., Fridriksson, J., Bonhilha, L., Healy, E. W., Baylis, G., Baker, J. M., and Rorden, C. (2009). Neural recruitment for the production of native and novel speech sounds. *Neuroimage* 46, 549–557.
- Nasir, S. M., and Ostry, D. J. (2006). Somatosensory precision in speech production. *Curr. Biol.* 16, 1918–1923.
- Neufeld, G. (1979). Towards a theory of language learning ability. *Lang. Learn.* 29, 227–241.
- Neville, H. J., and Bavelier, D. (2001). “Specificity of developmental neuroplasticity in humans: evidence from sensory deprivation and altered language experience,” in *Toward a Theory of Neuroplasticity*, eds C. A. Shaw and J. C. McEachern (Hove: Taylor & Francis), 261–274.
- Nieto-Castanon, A., Guenther, F. H., Perkell, J. S., and Curtin, H. D. (2005). A modeling investigation of articulatory variability and acoustic stability during American English/r/ production. *J. Acoust. Soc. Am.* 117, 3196–3212.
- Pa, J., and Hickok, G. (2008). A parietal-temporal sensory-motor integration area for the human vocal tract: evidence from an fMRI study of skilled musicians. *Neuropsychologia* 46, 362–368.
- Palmer, E. D., Rosen, H. J., Ojemann, J. G., Buckner, R. L., Kelley, W. M., and Petersen, S. E. (2001). An event-related fMRI study of overt and covert word stem completion. *Neuroimage* 14, 182–193.
- Paradis, M. (1983). *Readings on Aphasia in Bilinguals and Polyglots*. Montreal: Marcel Didier.
- Parker Jones, O., Green, D. W., Grogan, A., Pliatsikas, C., Filippopolitis, K., Ali, N., Lee, H. L., Ramsden, S., Gazarian, K., Prejawa, P., Seghier, M. L., and Price, C. J. (2011). Where, when and why brain activation differs for bilinguals and monolinguals during picture naming and reading aloud. *Cereb. Cortex*. doi: 10.1093/cercor/bhr161. [Epub ahead of print].
- Paus, T., Perry, D. W., Zatorre, R. J., Worsley, K. J., and Evans, A. C. (1996). Modulation of cerebral blood flow in the human auditory cortex during speech: role of motor-to-sensory discharges. *Eur. J. Neurosci.* 8, 2236–2246.
- Penfield, W., and Roberts, L. (1959). *Speech and Brain Mechanisms*. Princeton, NJ: Princeton University Press.
- Price, C. J., Green, D. W., and von Studnitz, R. (1999). A functional imaging study of translation and language switching. *Brain* 122, 2221–2235.
- Raichle, M. E., Fiez, J. A., Videen, T. O., Macleod, A.-M. K., Pardo, J. V., Fox, P. T., and Petersen, S. E. (1994). Practice-related changes in human brain functional anatomy during nonmotor learning. *Cereb. Cortex* 4, 8–26.
- Raizada, R. D. S., Tsao, F.-M., Liu, H.-M., and Kuhl, P. K. (2010). Quantifying the adequacy of neural representations for a cross-language phonetic discrimination task: prediction of individual differences. *Cereb. Cortex* 20, 1–12.
- Rodriguez-Fornells, A., Rotte, M., Heinze, H.-J., Nosselt, T., and Munte, T. F. (2002). Brain potential and functional MRI evidence for how to handle two languages with one brain. *Nature* 415, 1026–1029.
- Rosen, H. J., Ojemann, J. G., Ollinger, J. M., and Petersen, S. E. (2000). Comparison of brain activation during word retrieval done silently and aloud using fMRI. *Brain Cogn.* 42, 201–217.
- Salisbury, R. F. (1962). Institute of social research in New Guinea. *Anthropol. News* 3, 1–2.
- Scherag, A., Demuth, L., Rosler, F., Neville, H. J., and Roder, B. (2004). The effects of late acquisition of L2 and the consequences of immigration on L1 for semantic and morpho-syntactic language aspects. *Cognition* 93, B97–B108.
- Schleppegrell, M. (1987). “The older language learner,” in *ERIC Digest* (Washington, DC: Center for Applied Linguistics), ERIC identifier ED287313.
- Scovel, T. (1969). Foreign accents, language acquisition and cerebral dominance. *Lang. Learn.* 19, 245–253.
- Scovel, T. (1988). *A Time to Speak: a Psycholinguistic Inquiry into the Critical Period for Human Speech*. New York: Newbury House Publishers.
- Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., Reiss, A. L., and Greicius, M. D. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *J. Neurosci.* 27, 2349–2356.
- Sellers, C., Hughes, T., and Langhorne, P. (2005). Speech and language therapy for dysarthria due to non-progressive brain damage. *Cochrane Database Syst. Rev.* doi:10.1002/14651858.CD002088.pub2
- Shallice, T. (1988). *From Neuropsychology to Mental Structure*. Cambridge, UK: Cambridge University Press.
- Shiller, D. M., Sato, M., Gracco, V. L., and Baum, S. R. (2009). Perceptual recalibration of speech sounds following speech motor learning. *J. Acoust. Soc. Am.* 125, 1103–1113.
- Shuster, L. I., and Lemieux, S. K. (2005). An fMRI investigation of covertly and overtly produced mono- and multisyllabic words. *Brain Lang.* 93, 20–31.
- Simmonds, A. J., Wise, R. J. S., Dhanjal, N. S., and Leech, R. (2011). A comparison of sensory-motor activity during speech in first and second languages. *J. Neurophysiol.* 106, 470–478.
- Simonyan, K., Ostuni, J., Ludlow, C. L., and Horwitz, B. (2009). Functional but not structural networks of the human laryngeal motor cortex show left hemispheric lateralization during syllable but not breathing production. *J. Neurosci.* 29, 14912–14923.
- Snow, C. E. (2002). “Second language learners and understanding the brain,” in *The languages of the brain*, eds A. M. Galaburda, S. M. Kosslyn, and Y. Christen (Cambridge, MA: Harvard University Press), 151–165.
- Sorensen, A. P. (1967). Multilingualism in the Northwest Amazon. *Am. Anthropol.* 69, 670–684.
- Sörös, P., MacIntosh, B. J., Tam, F., and Graham, S. J. (2010). fMRI-compatible registration of jaw movements using a fiber-optic bend sensory. *Front. Hum. Neurosci.* 4:24. doi: 10.3389/fnhum.2010.00024
- Southwood, M. H., and Flege, J. E. (1999). Scaling foreign accent: direct magnitude estimation versus interval scaling. *Clin. Linguist. Phon.* 13, 335–349.
- Speedie, L. J., Wertman, E., Tair, J., and Heilman, K. M. (1993). Disruption of automatic speech following a right basal ganglia lesion. *Neurology* 43, 1768–1774.
- Stein, M., Federspiel, A., Koenig, T., Wirth, M., Lehmann, C., Wiest, R., Strik, W., Brandeis, D., and Dierks, T. (2009). Reduced frontal activation with increasing 2nd language proficiency. *Neuropsychologia* 47, 2712–2720.
- Taylor, B. P. (1974). Toward a theory of language acquisition. *Lang. Learn.* 24, 23–35.
- Thierry, G., and Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proc. Natl. Acad. Sci. U.S.A.* 104, 12539–12535.
- Tourville, J. A., Reilly, K. J., and Guenther, F. H. (2008). Neural mechanisms

- underlying auditory feedback control of speech. *Neuroimage* 39, 1429–1443.
- Tremblay, S., Shiller, D. M., and Ostry, D. J. (2003). Somatosensory basis of speech production. *Nature* 423, 866–869.
- Vaid, J., and Hull, R. (2002). “Re-envisioning the bilingual brain using function neuroimaging,” in *Advances in the Neurolinguistics of Bilingualism*, ed. F. Fabbro (Udine: Forum), 315–355.
- Ventura, M. I., Nagarajan, S. S., and Houde, J. F. (2009). Speech target modulates speaking induced suppression in auditory cortex. *BMC Neurosci.* 10, 58–69. doi: 10.1186/1471-2202-10-58
- Vigneau, M., Beaucousin, V., Herve, P. Y., Duffau, H., Crivello, F., Houde, O., Mazoyer, B., and Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *Neuroimage* 30, 1414–1432.
- Walsh, T. M., and Diller, K. C. (1978). Neurolinguistic foundations to methods of teaching a second language. *Int. Rev. Appl. Ling.* 16, 1–14.
- Weinreich, U. (1953). *Languages in Contact*. Mouton: The Hague.
- Wilson, S. M., and Iacoboni, M. (2006). Neural responses to non-native phonemes varying in producibility: evidence for the sensorimotor nature of speech production. *Neuroimage* 33, 316–325.
- Wise, R. J. S., Greene, J., Buchel, C., and Scott, S. K. (1999). Brain regions involved in articulation. *Lancet* 353, 1057–1061.
- Wu, Y. J., and Thierry, G. (2010). Chinese-English bilinguals reading English hear Chinese. *J. Neurosci.* 30, 7646–7651.
- Wu, Y. J., and Thierry, G. (2011). Event-related brain potential investigation of preparation for speech production in late bilinguals. *Front. Psychol.* 2:114. doi: 10.3389/fpsyg.2011.00114
- Yetkin, F. Z., Hammeke, T. A., Swanson, S. J., Morris, G. L., Mueller, W. M., McAuliffe, T. L., and Haughton, V. M. (1995). A comparison of functional MR activation patterns during silent and audible language tasks. *AJNR Am. J. Neuroradiol.* 16, 1087–1092.
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 20 February 2011; accepted: 02 July 2011; published online: 15 July 2011.

Citation: Simmonds AJ, Wise RJS and Leech R (2011) Two tongues, one brain: imaging bilingual speech production. *Front. Psychology* 2:166. doi: 10.3389/fpsyg.2011.00166

This article was submitted to *Frontiers in Language Sciences*, a specialty of *Frontiers in Psychology*.

Copyright © 2011 Simmonds, Wise and Leech. This is an open-access article subject to a non-exclusive license between the authors and Frontiers Media SA, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and other Frontiers conditions are complied with.