CONCEPTUAL REVIEW ARTICLE

Verbal Self-Monitoring in the Second Language

Wouter P. J. Broos, Wouter Duyck, and Robert J. Hartsuiker

Department of Experimental Psychology, Ghent University

Speakers monitor their own speech for errors. To do so, they may rely on perception of their own speech (external monitoring) but also on an internal speech representation (internal monitoring). While there are detailed accounts of monitoring in first language (L1) processing, it is not clear if and how monitoring is different in a second language (L2). Here, we ask whether L1 and L2 monitoring differ and, if so, where the differences lie. L1 and L2 might differ in the speed with which monitoring is performed but also in their monitoring foci. We discuss studies on bilingual language control and suggest that self-monitoring might function as a last-resort control process. We conclude with speculation on the role self-monitoring might play in L2 learning and suggestions for future research.

Keywords L2 monitoring; bilingualism; speech production; speech comprehension

Introduction

When people produce language, they monitor their own speech for errors. An error can, for instance, be made in the grammatical structure of a sentence or in a certain word in a phrase that is pronounced in the wrong manner. For example, a speaker pronounces the sentence "The ban, the man got very angry." (Poulisse, 1999). In this phrase, the word "ban" is initially produced with the wrong consonant and is corrected shortly after. The system that is responsible for monitoring speech, the self-monitoring system, was not able to intercept the error "ban" in time. Yet, it did perform a repair when the error was noticed by

producing the correct word "man." Another possible error that could be realized is, for instance, "v-horizontal" (Levelt, 1989). Given the context (a task in which people often mentioned spatial attributes), it seems that the speaker wanted to produce the word "vertical" but quickly changed it to "horizontal." Notice that only a small part of the word "vertical" is realized, indicating that the self-monitoring system intercepted the error before it was produced in its entirety. Hence, the monitoring system can either monitor speech that has already been produced via external monitoring but also speech that has not been (entirely) realized yet by means of internal monitoring. Bilingual speakers can monitor speech in any of their languages. In this article we ask whether monitoring in a second language (L2) works the same as in a first language (L1).

Before going more deeply into the system that monitors for speech errors, we first ask whether speech errors themselves differ in the L1 and L2. Poulisse (1999) wrote an elaborate review on slips of the tongue in L1-speaking children and found that the nature of these slips is very similar to those found in L2 adults; the spoonerism (i.e., exchanges like "mad bunny" instead of "bad money") made by both groups were affected by the factors context, lexical stress, and number of syllables (MacKay, 1970). Additionally, Aitchison and Straf (1981) examined malapropisms (i.e., speech errors that sound similar to the intended word) made by L1 children and L2 adults and concluded that the same phonological features of words were used by both groups during word retrieval. But in a later paper, Poulisse (2000) discussed some differences between L1 and L2 and the underlying processes that might be responsible for the discrepancies. Importantly, Poulisse (1999) observed precisely 2000 slips in L2 while only 137 slips were made in L1. She explained this dissimilarity by arguing that speech production is less automatic in L2 learners than L1 speakers. In fact, highly proficient speakers made fewer errors than low-proficient ones (Poulisse, 1999). Additionally, L1 speakers mostly made phonological slips in content words (e.g., "flan" for "plan"). In contrast, L2 speakers also frequently produced phonological slips in function words. Finally, L1 intrusions are sometimes seen in L2 production (e.g., the nonword "luisten," which is a blend of the English word "listen" and its Dutch translation "luisteren"). So, based on speech error evidence, more and different errors are made in L2 speech production than in L1 speech.

In addition to differences in error distributions, there are also dissimilarities in other aspects of speech production in L1 and L2, including longer naming latencies (Kroll & Stewart, 1994), more Tip of the Tongue states (Gollan & Silverberg, 2001), and more disfluencies (de Bot & Schreuder, 1993; Flege & Frieda, 1995) in L2. Additionally, language processing in L2 is slower and more error prone in general, for instance in sentence parsing (Papadopoulou

& Clahsen, 2003), auditory word recognition (Lagrou, Hartsuiker, & Duyck, 2011), and reading (Van Assche, Duyck, Hartsuiker, & Diependaele, 2009).

Because it is conceivable that self-monitoring uses some of the same general mechanisms involved in language comprehension or production (see below), these processing differences between L1 and L2 suggest that self-monitoring may also differ between L1 and L2 speakers. Additionally, if speaking in L2 is less automatic than speaking in L1, this may have consequences for the amount of attentional resources that can be spent on self-monitoring. This may lead to a reduction of monitoring accuracy or an attentional shift towards one of several monitoring mechanisms (Oomen, Postma, & Kolk, 2005; see below). Finally, bilinguals may use the same self-monitoring mechanism that they use for detection of "regular" errors, for the function of language control (i.e., detecting nontarget language intrusions). Of course, monolingual speakers also need to determine how to phrase their speech and need to take context into consideration (e.g., to choose an appropriate speech register) but so do bilinguals (in both their languages). Hence, language control of bilinguals might be an additional monitoring task in the speech of bilinguals. The remainder of this article will review the literature with respect to several aspects of the self-monitoring system. First, we briefly discuss theories of self-monitoring in L1. Second, we discuss possible differences between L1 and L2 monitoring mechanisms, focusing on monitoring speed and resources. Third, monitoring foci of both L1 and L2 will be elaborated upon. Fourth, we turn to the role of self-monitoring on language learning. We end with a discussion and suggestions for future research.

Theories of Self-Monitoring in L1

There are two main approaches in the literature on self-monitoring. On the one hand, there is the perception-based approach, which argues that monitoring depends on the comprehension system. On the other hand, there are production-based approaches, which argue that monitoring is based on production processes. The Perceptual Loop Theory is a perception-based approach that claims that the comprehension system is used to monitor both one's own speech and someone else's (Levelt, 1983, 1989). This particular approach assumes three distinct loops that transfer information to a central monitor: the conceptual loop, the inner loop, and the auditory loop. The conceptual loop decides whether the words and sentences that are used are appropriate in a specific context. The inner loop monitors the speech plan before it is articulated, and the auditory loop monitors speech that is already produced. Evidence for an inner monitor was provided by Motley, Camden, and Baars (1982) who asked

participants to perform the spoonerisms of laboratory induced predisposition (SLIP) technique (in which participants were asked to pronounce word pairs [e.g., "mad-back"] after being primed with word pairs of a different structure [e.g., "big-men"]). Consonant exchanges that led to taboo words occurred significantly less often than if these did not form taboo words.

The production-based approach differs from the perceptual-based approach in that it assumes several independent monitors, which are situated at different levels of the speech production system (De Smedt & Kempen, 1987; Laver, 1973; Schlenk, Huber, & Willmes, 1987). A recent example of this approach is the interactive two-step model of Nozari, Schwartz, and Dell (2011), which states that error detection is based on competing representations. If all goes well, only the representation of the correct response (i.e., word or phoneme) will be activated (low conflict); but if there is an error, both the representation of the correct and an incorrect response will tend to be activated (high conflict). Consistent with work in the domain of action monitoring more generally (e.g., Botvinick, Braver, Barch, Carter, & Cohen, 2001) it assumes that conflict between alternative representations at a given layer of representation (words or phonemes) is indicative of an error (where conflict can be defined, for instance, in terms of the activation difference between the two nodes with the highest activation: the smaller this difference, the larger the conflict). Nozari et al. (2011) found that aphasic patients showed no significant correlation between comprehension measures and error detection but there was a significant correlation between production skills and error monitoring.

Finally, a mixture of production and perception monitoring is the forward modeling account of Pickering and Garrod (2014). There is much evidence for forward modeling in the domain of motor control in general and speech motor control in particular (Tian & Poeppel, 2014). The forward modeling account assumes that speakers monitor by means of predictions. They first create a "production command," which denotes the intention that people create before linguistic encoding takes place. This command is used to start two parallel processes. First, the command feeds into the production implementer, which in turn creates an utterance that contains information on semantics, syntax, and phonology. The utterance is processed in order to create the utterance percept, which also includes semantic, syntactic, and phonological information. Second, an efference copy of the production command is sent into a forward production model. This model creates a predicted utterance, which is fed into the forward comprehension model. The comprehension model creates a predicted utterance percept. Finally, the monitor is able to compare the utterance percept and the predicted utterance percept at several linguistic levels.

Differences in Monitoring Mechanisms Between L1 and L2

In this section, we review studies that asked whether L1 and L2 monitoring mechanisms (only) differ in aspects such as processing speed and capacity demands or whether there are fundamental differences (e.g., in terms of types of monitoring channels used).

Monitoring Speed

The speed of monitoring in language production has received some attention (Blackmer & Mitton, 1991; Declerk & Hartsuiker, 2016; Levelt, 1983; Oomen & Postma, 2002; Seyfeddinipur, Kita, & Indefrey, 2008; Van Hest, 1996) and there is an ongoing debate whether monitoring speed is the only difference between L1 versus L2 or whether other elements of the monitoring system differ as well. Two intervals that are often distinguished in discussions about the speed of monitoring are the error-to-cutoff and cutoff-to-repair intervals. The former represents the time between the moments when the error is made and when the speaker stops speaking and the latter denotes the amount of time between the interruption and repair onset (Levelt, 1983). Van Hest (1996) argued that the difference between L1 and L2 monitoring is quantitative because she only found a difference in monitoring speed. In particular, cutoff-to-repair intervals were longer in L2 for appropriateness repairs (e.g., the dot—the red dot) and covert repairs[1] (repairing an error before it is made) than in L1 and error-to-cutoff intervals were only longer for appropriateness repairs in L2.

Oomen and Postma (2001) manipulated these time intervals by means of a visual network task in which participants were asked to describe the trajectory of a red dot that was displayed on the screen. The red dot either moved according to a normal rate or a fast rate. Both the error-to-cutoff and cutoff-to-repair interval were shorter in the fast rate condition than the normal rate condition. The number of corrected errors did not differ significantly between conditions. A more recent study, Declerck and Hartsuiker (2016) used the same timing manipulations as Oomen and Postma in order to simulate speech and monitoring speed in L1 versus L2. Timing was manipulated in such a way that normal speech rate in L1 was similar to fast speech rate in L2. The relationship between the error-to-cutoff and cutoff-to-repair intervals was tested, as well as the effects of speech rate on the length of the two intervals. In both L1 and L2, both the cutoff-to-repair times and error-to-cutoff times were descriptively shorter in fast versus normal speech indicating that intervals vary as a function of objective speech rate (although the speech rate effect was only significant for the error to cutoff intervals). There was also a positive correlation between both time intervals. Importantly, it was also observed that the cutoff-to-repair

interval was significantly longer in L2 than in L1. Hartsuiker and Kolk (2001) implemented a computational model based on Levelt's perceptual loop theory and added the assumption of a global speech rate parameter that influences the speed of every part of language comprehension and production. Faster comprehension in this model leads to earlier error detection and hence a shorter error-to-cutoff interval. Faster production results in faster repair planning and therefore a shorter cutoff-to-repair interval. This model simulated Oomen and Postma's speech rate effect on both intervals, and it is compatible with the results of Declerck and Hartsuiker (2016) finding that objective speech rate determined the intervals.

Reduced Resources in Monitoring

Besides the interruption and repair times, monitoring speed can also be influenced by the amount of attentional resources that are available during monitoring. Oomen and Postma (2002) performed a study in which they focused on the effects of reduced processing resources during speech monitoring. Specifically, they wanted to observe whether the speed of monitoring was affected when fewer attentional resources were available. This was done by observing speech monitoring in a dual-task paradigm. It was found that fewer errors were repaired in the dual-task than in the single-task condition in both speech production and speech perception. Furthermore, the error-to-cutoff time and cutoff-to-repair time were shorter in the dual-task than the single-task condition. Their explanation was that speakers shift attention toward the prearticulatory channel when resources are scarce and that error detection is faster when focusing on the internal monitoring system than the external one.

Monitoring speed in dual task conditions has also been examined in L2. Using the same paradigm as Oomen and Postma (2002), Declerck and Kormos (2012) tested whether the efficiency and accuracy of speech monitoring in L2 are affected by single and dual task conditions. Significantly more lexical errors were made and fewer errors were corrected in the dual-task condition. Moreover, more proficient L2 speakers had a higher speech rate than less proficient speakers and made significantly fewer errors. However, there was no dual-task effect on the time course of monitoring. The authors argued that considerable attention and conscious processing is needed in L2 speech production. Adding another task would therefore not affect speech production because conscious attention is already required to perform this task.

Summarizing, both studies show that more errors are made and fewer errors are corrected in the dual-task condition. However, results differ with respect to the time course: Oomen and Postma (2002) observed shorter interval times

in the dual-task condition in L1, but Declerck and Kormos (2012) found no dual-task effect on the time intervals in L2.

Differences in Monitoring Foci Between L1 and L2

One other possible difference between monolinguals and bilinguals with regard to monitoring is that bilinguals need to exert language control in order to ensure that they will speak in the proper language. Such control might involve monitoring, in addition to other mechanisms. A possible mechanism of language control is proposed by La Heij (2005) who suggested that the language in which a bilingual intends to speak is part of the preverbal message. Thus, the decision to use a certain language is made early on in the speech production process. Moreover, he claimed that monolingual speakers perform a similar action in that they need to decide which register to use (e.g., formal speech when talking to a professor or informal when speaking to a family member). Therefore, language control is part of the choice of register in case of bilinguals. A similar notion is proposed in the Inhibitory Control (IC) model by Green (1998) and Poulisse and Bongaerts (1994) who argue that a language tag (Albert & Obler, 1978; Green, 1986, 1993; Monsell, Matthews, & Miller, 1992) is already attached to the conceptual representation. Other models assume that language control is specified postlexically. In comprehension, the BIA+ model (Dijkstra & Van Heuven, 2002) claims that information from the phonological and orthographic word identification processes is used to help select a language. Models of lexical-syntactic representations in production (Hartsuiker, Pickering, & Veltkamp, 2004) assume connections between lemmas and language nodes. Whatever the precise locus and mechanism of language selection, we argue that there are two ways in which self-monitoring can help this process, namely an early process of monitoring the context to decide upon which language to use and a late process of checking whether speech adheres to the initial language choice.

Prearticulatory Control for Language

Can prearticulatory monitoring prevent language errors? In order to answer this question, we must know which monitoring criteria are used and whether different criteria are used in L2 than L1. A possible monitoring criterion, that has generated much debate in the literature, is lexical status (i.e., *is an upcoming utterance a word or not?*). The lexical bias effect, the phenomenon that phonological errors result in words more often than predicted by chance, has been taken as evidence that the monitor uses this criterion. Specifically, discrete models argue that the lexical bias effect is a result of prearticulatory monitoring

(Levelt, 1989; Levelt, Roelofs, & Meyer, 1999; Nooteboom, 2005). Interactive models, however, claim that this effect represents feedback between the lexical and the phonological level as claimed by interactive models (Dell, 1986; Harley, 1993; Rapp & Goldrick, 2000). The self-monitoring account states that the self-monitoring system filters out more word errors than nonwords errors before speech production. The feedback account states that phoneme representations can only prime representations of existing words, not of nonexisting ones, increasing the chance of a real-word error. Finally, Hartsuiker, Corley, and Martensen (2005) showed that the lexical bias effect in L1 was modulated by context (also see Baars, Motley, & MacKay, 1975). Based on their pattern of results, they argued that both feedback and self-monitoring created the lexical bias effect.

Importantly, this same issue was also investigated in second language processing; Costa, Roelstraete, and Hartsuiker (2006) examined whether there is also feedback between the phonological and lexical level in second language production and whether phonological activation can spread from one language to another in Catalan-Spanish bilinguals. Results revealed a lexical bias effect when the SLIP task was performed in the L2 of the participants. During the SLIP task, participants are presented with certain constructions of (non)word pairs (e.g., coag–roan) in order to elicit a speech error when a certain word pair has to be pronounced (e.g., road–coat instead of coad–roat). A lexical bias effect was also seen when the resulting error was a word that existed in the nonresponse language (Catalan). Thus, the lexical bias effect (arguably resulting from self-monitoring and feedback) can spread across languages in bilinguals. Importantly, these results do not argue for language as a monitoring criterion. If language had been a criterion, then all Catalan words would have been considered errors and the lexical bias effect in Catalan would have been eliminated. However, a study that elicited language intrusions (Hartsuiker & Declerck, 2009) did find evidence for language being a criterion of the monitoring system as half of the language intrusions were repaired.

Summarizing, the lexical bias effect occurs both in L1 and L2, suggesting similar effects of feedback and internal self-monitoring in both languages. However, it is not clear whether target language is a monitoring criterion: the finding of the lexical bias effect in a nontarget language argues against monitoring for language, but the many self-corrections of language intrusions argue in favor of it. Further research is needed to determine whether language monitoring can be viewed as a last resort mechanism to prevent language intrusions or whether only (external) monitoring repairs language intrusions after they have become overt.

External Cues in Language Selection

Several studies have asked whether bilinguals use external cues to help determine what language should be used or expected (Duyck, Van Assche, Drieghe, & Hartsuiker, 2007; Elston-Güttler & Friederici, 2005; Elston-Güttler, Gunter, & Kotz, 2005; Gollan & Ferreira, 2009; Lagrou, Hartsuiker, & Duyck, 2012; Paulmann, Elston-Güttler, Gunter, & Kotz, 2006; Van Assche et al., 2009). This will be discussed in the following sections.

Comprehension

In comprehension, external cues can be visually present, as in written sentence context or be part of auditory input of speech. Elston-Güttler et al.(2005) focused on the visual aspect of comprehension and performed a semantic priming study in which they looked at processing of German-English homographs (e.g., "Gift" meaning "poison" in German and "present" in English) in sentence contexts. German-English bilinguals were first presented with a silent movie containing either German or English subtitles after which they were asked to perform a lexical decision task. Semantic priming (on reaction times and Event-related potential ERP components) was only observed for speakers who saw the German version of the movie in the first half of the experiment. In a further sentence context study, Van Assche et al. (2009) showed that the cognate effect (faster recognition of words like "ring" in Dutch and English than words with dissimilar translations) survived even in an L1 sentence context, indicating that representations of the L2 are sufficiently activated to affect word recognition. Hence, even though the language of a sentence could be used as a strong cue to facilitate lexical search by eliminating almost half of all available lexical candidates, a unilingual sentence context is apparently not used as a cue for language selection in bilinguals.

A study that focused on the *auditory* modality of comprehension was carried out by Lagrou et al. (2012). They investigated whether knowledge of the first language is influential when listening to a L2 and vice versa. Moreover, they observed whether the L1 of the participant affects the selectivity of lexical access of the listener. Dutch-English bilinguals were slower when listening to cross-lingual homophones (e.g., Dutch "bos" [forest] and English "boss") in their L1 or L2 than when listening to control words. Moreover, the homophone effect was independent of the native language of the speaker, indicating that speaker accent was not used as a cue to narrow down language selection. The effect was found when listening to Dutch and to English, suggesting that sentence context is not used by the listener to fully attend to one single language. If this were the case, then this effect should not occur at all. It must be noted

that Duyck et al. (2007) found that sentence context may nullify L2 effects of nonidentical cognates (perhaps because activation spreading across language is weaker in nonidentical cognates). There is also some evidence that factors like predictability can reduce cognate effects (Schwartz & Kroll, 2006). So, even though sentence context might affect the amplitude of the cognate effects, external cues in speech comprehension are not used by bilinguals in order to monitor what language should be used or is expected.

Production

In research on languages cues in production, two types of cues have been considered: the language used in a specific context and properties of the speaker (e.g., faces). Even though a considerable amount of research has been done on language switching (Costa & Santesteban, 2004; Hernandez & Kohnert, 1999; Meuter & Allport, 1999), not many studies have focused on switching in a dialogue setting. In such a situation, the language of an interlocutor might act as a language cue and affect the production of the other speaker. Gambi and Hartsuiker (2016) asked whether bilingual speakers are slower when switching to the other language than the language that was just used by an interlocutor. This was examined by means of a joint language switching task in which a pair of Dutch-English bilinguals was asked to name pictures. Nonswitching participants were slower in naming pictures in their L1 after the interlocutor named pictures in L2 than in L1. This effect was even stronger for highly proficient bilinguals. Obtained results suggest that the process of choosing languages is shared between production and comprehension as speech production is slower after hearing a language switch (see also Peeters, Runnqvist, Bertrand, & Grainger, 2014).

Next to the language or speech of the interlocutor, faces can also be used as an external cue for bilinguals to tune into a certain language (Li, Yang, Scherf, & Li, 2013; Molnar, Ibáñez-Molina, & Carreiras, 2015; Woumans et al., 2007; Zhang, Morris, Cheng, & Yap, 2013). Woumans et al. (2015) performed experiments in which bilinguals were asked to answer questions from their interlocutor after being familiarized with their faces. The question whether bilinguals use the face of an interlocutor in order to decide what language to speak was of particular interest. It was found that congruent trials were reacted to significantly faster than incongruent trials and, more importantly, the effect disappeared after some incongruent trials. Hence, evidence suggests that a face is not used as a cue for language selection anymore from the moment that participants know that the interlocutor may speak more than one language.

The above study has shown that faces are indeed used as cues for bilinguals when deciding what language to speak (see also Gollan, Schotter, Gomez,

Murillo, and Rayner [2014] for language intrusions in reading aloud mixed-language paragraphs). Hartsuiker and Declerck (2009) asked whether face cues can also lead speakers astray during language production. In particular, they wanted to see whether function word intrusions would occur in a second language if inconsistent cues are presented. This was investigated by a "famous faces" paradigm in both Dutch and English in which three pictures of famous faces (Dutch or English) were presented, some of which move up or down the screen. Participants were asked to tell which pictures went in what direction. When the task was performed in Dutch, the Dutch function word "en" (and) was often replaced with its English translation "and" while in the English task, the word "and" was more often substituted with "en." Yet, the effect was much stronger when the task was performed in the L2 suggesting that words in L1 might be stronger competitors. Hence, the association between the faces of famous people and the language they speak yields more language intrusions, indicating that faces activate a certain language even when this is not beneficial to the speaker.

To summarize, recent studies on external language cues have shown that language context is not used as a strong external cue for language selection, neither in visual nor in auditory perception. In language production, however, faces can be used as an external cue to zoom into a certain language up until the point that speakers know that an interlocutor is bilingual. The question that remains is to which extent external cues are used and why this differs between production and comprehension. In order to answer this, studies have to be performed in which the type of cues are kept constant across modalities.

The Nature of Monitoring Foci in L2

As speaking in the L2 is more difficult than in the L1, other foci might be part of the L2 monitoring system. The L2 language system in L2 learners is not fully developed yet and their production skills are less than optimal. Some speakers struggle with creating grammatical sentences while this is easier for other speakers, but all L2 speakers (more so than L1 speakers) are concerned with conveying their intentions in their L2 in an appropriate manner. In general, more syntactic errors are made by L2 speakers than L1 speakers and low frequency words yield a higher number of lexical and phonological errors (Kovač, 2011). Different types of repair are also observed in L2 speakers depending on their proficiency level: low-proficient L2 speakers make more lexical and phonological error repairs while highly proficient speakers use more appropriateness repairs for lexical items (Van Hest, 2000). This suggests that the focus of monitoring for less proficient L2 speakers is more on the

content of the message while more proficient speakers can pay more attention to appropriateness.

Another monitoring focus that may be emphasized more in L2 than in L1 is the effect that speech production has on the interlocutor (a monitoring loop that Postma [2000] called "knowledge of results"). By observing the reactions of the interlocutor (either explicitly or implicitly), L2 speakers will know whether their communicative efforts were successful or not. This is the first part of the process where there is an emphasis on the perception system in that the L2 learner interprets the reaction of the interlocutor. The second part is concerned with incorporating the information in the language production system. If explicit feedback is received (e.g., when the interlocutor says that a certain word is used in the wrong way), adjustments of internal representations can be performed. If positive feedback is received, then this is a confirmation that representations are already set in the right manner. Less proficient speakers will presumably rely more on this monitor than more proficient speakers as they have less confidence in their ability to communicate in their L2. Overall, the amount of emphasis on feedback of the interlocutor will depend on proficiency level and the nature of this feedback. Summarizing, L2 speakers are likely to be more concerned with the content of their speech than the form and might focus more on feedback of their interlocutor.

The Role of Self-Monitoring on Language Learning
L2 Pronunciation

When L2 speakers converse with native speakers, they adjust their speech to that of their conversation partner (Hwang, Brennan, & Huffman, 2015; Kim, Horton, & Bradlow, 2011; Kim, 2012). L2 speakers hear nativelike pronunciation of phonemes that do not exist in their L1 that might make them create new phonemic categories depending on the proficiency of the speaker and the similarity of the new phoneme with other similar phonemes in the L1 inventory of the speaker (Best & Tyler, 2007; Flege, 1995). L2 speakers will use self-monitoring to compare the pronunciation of the L1 speaker (using it as a standard) and one's own attempt to pronounce the nonnative phoneme. By monitoring speech output of the native speaker, they might be able to learn new phonemes, which in turn help speakers determine whether they sound like a native speaker.

Speech alignment is not only seen when one's own feedback is perceived but also when a speaker has a conversation with another speaker. Hwang et al. (2015) focused on phonetic alignment of Korean-English bilinguals. It was found that participants pronounce nonnative phonemes in a more

native-English manner after having spoken with a native speaker as opposed to a nonnative speaker. Hence, L1 production of the native English confederate has greater influence on L2 production of L2 learners than speech production of a nonnative confederate, indicating that the monitoring system not only monitors speech production from an interlocutor but is also able to regulate the amount of alignment depending on the nature of speech production of the interlocutor.

Next to learning from speech output from native speakers, Linebaugh and Roche (2015) have shown that L2 speakers can also adjust phonemic boundaries of nonnative phonemes more accurately after pronouncing them more nativelike. First-year Arabic learners of English learned to pronounce nonnative phonemes more nativelike after having received articulation training. During articulatory training, participants were first asked to listen and repeat the phonemes after which detailed instructions on the exact positioning of the tongue and jaw were given. At the end of the training, participants produced the contrastive phonemes in rapid succession. After training, L2 speakers were able to distinguish nonnative contrasts more accurately than before training indicating that perception is positively influenced by a more nativelike production. Hence, L2 speakers benefit from increased self-monitoring during production of nonnative phonemes in that their perception of these phonemes improves as well.

In short, the above studies have shown that the self-monitoring system is able to adapt phonemic boundaries and can positively affect L2 pronunciation. It does so by monitoring one's own speech as well as someone else's and can in fact determine how much speech adaptation is needed to optimize L2 speech production. Because the system only adapts language production when a more nativelike realization is perceived, it can be considered an effective learning mechanism. Thus, L2 speakers can use the self-monitoring system in order to validate whether their speech is nativelike, even though subsequent speech production might not always improve as a result.

L2 Learning on the Lexical Level

There is more to learning a second language than just pronunciation; mastering the lexicon of a particular L2 is of vital importance, too. Costa, Pickering, and Sorace (2008) argue that some degree of lexical alignment is seen in any conversation whenever possible in which case representations of the interlocutors become more similar (Pickering & Garrod, 2006). As in pronunciation, the self-monitoring system is used to compare the native realization and one's own (as exemplified in (1) below).

(1) L2 speaker: I am not able to call with my mobile phone anymore
     L1 speaker: Then you should buy a new cell phone
     L2 speaker: But the cell phone I want is too expensive

The L2 speaker imitates the word "cell phone" since s/he realizes that it sounds more natural. The monitoring system detects that "cell phone" sounds more natural and incorporates it in the speech plan of the L2 speaker. This information is subsequently used when pronouncing the next utterance. In this case, the L2 speaker learns from the interlocutor.

Yet, additional factors might affect the amount of lexical alignment and therefore the ability to monitor by comparing speech output (Costa et al., 2008). Lack of knowledge of the L2 can prevent alignment, for instance, when the L2 speaker is not sure what a certain word means (2):

(2) L2 speaker: The top of the trees in that forest is always green
     L1 speaker: It is known for its beautiful canopy.
     L2 Speaker: Since it is autumn, it surprises me that the top of the trees
     does not turn brown

The L1 speaker uses the word "canopy" for the description of the L2 speaker (i.e., the top of the trees). However, sentential context is not enough to extract its meaning. As the L2 speaker does not know what the word means (or that the L1 just used a word that covers his/her description), the word can also not be applied. Hence, lexical alignment is not realized due to lack of knowledge of the second language, an aspect that was not influential during phonemic alignment. Finally, the L1 of the speaker can also influence the amount of lexical alignment. In particular, the amount of lexical alignment is sometimes correlated with the phonological similarity between the L1 and L2 (Costa et al., 2008). For instance, an L1 speaker of English might use the word "skinny" after which an L1 speaker of Dutch uses the English word "thin." This word is phonologically similar to the Dutch equivalent "dun" (thin), which might cause a lack of lexical alignment. Still, it is clear that the monitoring system plays an important role in the extraction of nonnative sounds and words while an increased amount of self-monitoring helps to subsequently apply this knowledge during L2 speech production.

Discussion

The current article provided a brief overview of the different self-monitoring theories and examined potential differences between the L1 and L2 monitoring mechanisms. It also considered the role of self-monitoring on L2 learning

with regard to pronunciation and lexical learning. We end here with some speculation on possible differences of the L1 and L2 monitoring system by discussing speech error data, forward models, and conflict monitoring. Finally, we offer suggestions for future research on the use of register in L1 and L2, the effect of reduced resources in L2 monitoring, and how L2 studies can distinguish between the current self-monitoring accounts.

Speech Errors

Findings from speech error studies revealed that certain errors (spoonerisms and malapropisms) are formed in the same manner in L1 as in L2. This suggests that the monitoring system uses identical phonological and prosodic criteria in both the L1 and L2 monitoring. Still, different error patterns of L1 and L2 speakers indicate that monitoring in L2 is not identical to monitoring in L1. L2 speakers make significantly more slips during speech production, especially in function words. Native speakers, however, make fewer slips, meaning that the monitoring system detected more covert errors. Additionally, blends between L1 and L2 translation equivalents are produced by L2 speakers meaning that their L1 influences the types of errors that are made. Hence, the L2 monitoring system seems to have a different focus because it prioritizes content over appropriateness and form. Whether language is an additional criterion is still up for debate.

The Quality of Prediction

Another possible difference between monitoring in L1 and L2 concerns the quality of predictions (forward models) of how L2 speech will sound. That is, if an L2 speaker has difficulty producing and perceiving a certain phoneme (e.g., one that does not occur in L1), it stands to reason that it is also difficult to create a nativelike forward model of that nonnative phoneme. Imagine that an L1 speaker of Dutch is confronted with the nonnative phoneme /ʌ/ as in "monkey." Native speakers of Dutch tend to substitute this phoneme with the vowel /ʉ/ (Collins & Mees, 2003). When considering forward modeling, the phonology of the predicted utterance (that follows from the forward production model) and the predicted utterance percept will not be as optimal as that of a native speaker. The semantic and syntactic information will most likely be well defined because "monkey" is a relatively simple word. Yet, the production representation is still not ideal. When the comparison is made between the two percepts, subsequent speech production is not nativelike; the realization of the vowel /ʌ/ is more similar to /ʉ/. Note that even if the L2 learner produces a vowel that is identical to his/her forward model, it does not mean that the pronunciation

is nativelike. Hence, the nature of the forward models in L2 is different as they are not optimal when compared to those of the L1. Consequently, alignment in pronunciation will not be observed because the L2 speaker is unaware of the less than optimal representation. Awareness of the nonnative pronunciation can be gained by recording one's own speech and playing it back. This recording can then be compared to that of native speech after which the L2 speaker's percept can become more nativelike.

Conflict Monitoring

As mentioned, the interactive two-step model of Nozari et al. (2011) is a model that uses conflict as a basis for error detection. When considering participants that speak in their L1, activation patterns are strong and abnormal patterns act as a cue and encourage the monitoring system to increase monitoring. Contrary to representations of L2 speakers, the lexical and phonemic representations in the minds of the L1 speaker are well established (see also Gollan, Montoya, Cera, & Sandoval [2008] the weaker links hypothesis). Weaker representations in L2 speakers lead to weaker connections between layers of representations, making conflict less useful for the detection of errors (as is shown in the case of aphasics by Nozari et al.).

A main difference with representations in L2 when compared to L1 is that these representations on several layers can be influenced by the native language. Translation equivalents might cause more conflict at the word level as the word forms are different. Words that have a similar meaning and a similar form in L1 and L2 (i.e., cognates) would positively affect the stored lexical representation of that word. This particular lexical representation of the L2 word is then much better established than words that are not identical or similar in this respect. Note, however, that there will most likely be more conflict on the phonemic level because the pronunciation of these words is different. This holds for translations in which the dissimilar phonemes both exist in the L1 (e.g., English "ten" vs. Dutch "tien" where both vowels exist in Dutch). Importantly, conflict will be greater if the dissimilar phoneme does not exist in the L1. Thus, the representation of the phonemic representations of the nonnative vowel of the L2 speaker is not as accurate as that of the L1 speaker. This in turn leads to weaker activation patterns and reduces monitoring success. Consequently, the lexical representation will be less accurate as well. In short, monitoring success might be correlated with the characteristics of the production weights, which is where the difference between L1 and L2 monitoring lies in this case.

Challenges for Future Research

Let us revisit the matter of language control and register. As discussed above, both monolinguals and bilinguals need to decide what register they will use during language production. Monolinguals must choose to use formal or informal language (depending on context) while bilinguals must also decide in what language to speak. We argue that appropriateness monitoring not only decides whether a certain word or grammatical construction is used in the correct context but that it can also be used to select the appropriate language. One major challenge is to decide whether language monitoring can be seen as a last resort in order to prevent language intrusions or whether external monitoring only repairs intrusions after they have become overt.

Another question that is yet to be answered is why L1 speakers make more errors and detect fewer errors when resources are reduced, whereas this difference is not seen in L2 speakers. The effect in L1 is explained by arguing that their attention shifts toward the preverbal message (the internal loop) when having fewer resources available because it is faster, which in turn indicates that it is more automatic. The lack of such an effect in the L2 suggests that attention does not shift toward the internal loop and further supports the notion that the monitoring process is less automatic than in L1. Additionally, it can be argued that the monitoring system is already more active in a L2 than in the first language but in what way is it more occupied and how exactly is this applied? By performing dual task studies where different monitoring loops (external vs. internal) are involved, more insights into the relation between resources and the workings of the monitoring system in L2 might be obtained.

Finally, studies on L2 monitoring might support one of the current self-monitoring theories that exist in L1. By focussing on different modalities such as production and comprehension during specific monitoring tasks in both the L1 and L2, the question by which modality monitoring is driven might be answered. If it turns out to be a combination of multiple modalities, then the forward modeling account is supported. If only one modality drives monitoring, then this strengthens the claim of either the production- or perception-based approaches.

Conclusion

This article provides an overview of the differences between verbal self-monitoring in the L1 and L2 of speakers. In particular, it evaluated the mechanisms of monitoring in both L1 and L2 and considered potential differences in monitoring foci. The main difference in monitoring mechanisms between L1

and L2 is the length of the time intervals, especially the cutoff to repair interval. We identified several major issues that have remained unaddressed, such as differences in monitoring foci between L1 and L2 in which we argue that monitoring acts as a last resort in preventing language intrusions. Moreover, insights into the nature of L1 and L2 monitoring foci were provided. Finally, we interpreted the role of self-monitoring on different levels of L2 learning and speculated on further differences in monitoring by describing the workings of self-monitoring accounts in the L2 while suggesting topics for future research.

<div style="text-align: right">Final revised version accepted 4 May 2016</div>

Note
1 It must be noted that Van Hest (2000) assumed that disfluencies were interpreted as covert repairs.

References

Aitchison, J., & Straf, M. (1981). Lexical storage and retrieval: A developing skill? *Linguistics*, *19*, 751–796. doi: 10.1515/ling.1981.19.7-8.751

Albert, M. L., & Obler, L. K. (1978). *The bilingual brain: Neuropsychological and neurolinguistic aspects of bilingualism*. New York: Academic Press. doi: 10.2307/3586321

Baars, B. J., Motley, M. T., & MacKay, D. G. (1975). Output editing for lexical status in artificially elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior*, *14*, 382–391. doi: 10.1016/s0022-5371(75)80017-x

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In B. Ocke-Schwen & M. J. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 13–34). John Benjamins Publishing Company, Amsterdam/Philadelphia. doi: 10.1075/lllt.17.07bes.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Evaluating the demand for control: Anterior cingulate cortex and conflict monitoring. *Psychological Review*, *108*, 624–652. doi: 10.1038/46035

Blackmer, E. R., & Mitton, J. L. (1991). Theories of monitoring and the timing of repairs in spontaneous speech. *Cognition*, *39*, 173–194. doi: 10.1016/0010-0277(91)90052-6

Collins, B., & Mees, I. M. (2003). *The phonetics of English and Dutch*. Leiden: Brill Academic Publishers. doi: 10.1017/s0025100305212264.

Costa, A., Roelstraete, B., & Hartsuiker, R. J. (2006). The lexical bias effect in bilingual speech production: Evidence for feedback between lexical and sublexical levels across languages. *Psychonomic Bulletin & Review*, *13*, 972–977. doi: 10.3758/bf03213911

Costa, A., Pickering, M. J., & Sorace, A. (2008). Alignment in second language dialogue. *Language and Cognitive Processes*, *23*, 528–556. doi: 10.1080/01690960801920545

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, *50*, 491–511. doi: 10.1016/j.jml.2004.02.002

de Bot, K., & Schreuder, R. (1993). Word production and the bilingual lexicon. *The Bilingual Lexicon*, *191*, 214. doi: 10.1075/sibil.6.10bot

Declerck, M., & Hartsuiker, R. J. (2016). The timing of speech interruptions during error repairs. Manuscript submitted for publication.

Declerck, M., & Kormos, J. (2012). The effect of dual task demands and proficiency on second language speech production. *Bilingualism: Language and Cognition*, *15*, 782–796. doi: 10.1017/s1366728911000629

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological review*, *93*, 283. doi: 10.1037/0033-295x.93.3.283

De Smedt, K., & Kempen, G. (1987). Incremental sentence production, self-correction, and coordination. In G. Kempen (Ed.), *Natural language generation: Recent advances in artificial intelligence, psychology, and linguistics* (pp. 365–376), Dordrecht, Netherlands: Martinus Nijhoff Publishers. doi: 10.1007/978-94-009-3645-4_23

Dijkstra, T., & Van Heuven, W. J. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*, 175–197. doi: 10.1017/s1366728902003012

Duyck, W., Van Assche, E., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 663–679. doi: 10.1037/0278-7393.33.4.663

Elston-Güttler, K. E., & Friederici, A. D. (2005). Native and L2 processing of homonyms in sentential context. *Journal of Memory and Language*, *52*, 256–283. doi: 10.1016/j.jml.2004.11.002

Elston-Güttler, K. E., Gunter, T. C., & Kotz, S. A. (2005). Zooming into L2: Global language context and adjustment affect processing of interlingual homographs in sentences. *Cognitive Brain Research*, *25*, 57–70. doi: 10.1016/j.cogbrainres.2005.04.007

Flege, J. (1995). Second language speech learning: Theory, finding and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research* (pp. 233–277). Timonium, MD: York Press. doi: 10.1075/lllt.17.08str

Flege, J., & Frieda, E. (1995). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics*, *25*, 169–186. doi: 10.1006/jpho.1996.0040

Gambi, C., & Hartsuiker, R. J. (2016). If you stay, it might be easier: Switch costs from comprehension to production in a joint switching task. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *42*, 608–626. doi: 10.1037/xlm0000190

Gollan, T. H., & Ferreira, V. S. (2009). Should I stay or should I switch? A cost–benefit analysis of voluntary language switching in young and aging bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 640–665. doi: 10.1037/a0014981

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, *58*, 787–814. doi: 10.1016/j.jml.2007.07.001

Gollan, T. H., Schotter, E. R., Gomez, J., Murillo, M., & Rayner, K. (2014). Multiple levels of bilingual language control evidence from language intrusions in reading aloud. *Psychological science*, *25*, 585–595. doi: 10.1177/0956797613512661

Gollan, T. H., & Silverberg, N. B. (2001). Tip-of-the-tongue states in Hebrew–English bilinguals. *Bilingualism: Language and Cognition*, *4*, 63–83. doi: 10.1017/s136672890100013x

Green, D. W. (1986). Control, activation, and resource. *Brain and Language*, *27*, 210–223. doi: 10.1016/0093-934x(86)90016-7

Green, D. W. (1993). Towards a model of L2 comprehension and production. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 249–277). Amsterdam/Philadelphia: John Benjamins. doi: 10.1075/sibil.6.12gre

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*, 67–81. doi: 10.1017/s1366728998000133

Harley, T. A. (1993). Phonological activation of semantic competitors during lexical access in speech production. *Language and Cognitive Processes*, *8*, 291–309. doi: 10.1080/01690969308406957

Hartsuiker, R. J., Corley, M., & Martensen, H. (2005). The lexical bias effect is modulated by context, but the standard monitoring account doesn't fly: Related reply to Baars et al. *(1975)*. *Journal of Memory and Language*, *52*, 58–70. doi: 10.1016/j.jml.2004.07.006

Hartsuiker, R. J., & Declerck, M. (2009). *Albert Costa y Julio Iglesias move up, but Fidel Castro stays put: Language attraction in bilingual language production*. Paper presented at the AMLaP 2009 Conference, Barcelona, Spain.

Hartsuiker, R. J., & Kolk, H. H. J. (2001). Error monitoring in speech production: A computational test of the perceptual loop theory. *Cognitive Psychology*, *42*, 113–157. doi: 10.1006/cogp.2000.0744

Hartsuiker, R. J., Pickering, M. J., & Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English

bilinguals. *Psychological Science*, *15*, 409–414. doi:
10.1111/j.0956-7976.2004.00693.x

Hernandez, A. E., & Kohnert, K. J. (1999). Aging and language switching in bilinguals.
Aging, Neuropsychology, and Cognition, *6*, 69–83. doi: 10.1076/anec.6.2.69.783

Hwang, J., Brennan, S. E., & Huffman, M. K. (2015). Phonetic adaptation in
non-native spoken dialogue: Effects of priming and audience design. *Journal of
Memory and Language*, *81*, 72–90. doi: 10.1016/j.jml.2015.01.001

Kim, M. (2012). *Phonetic accommodation after auditory exposure to native and
nonnative speech*. Northwestern University. Retrieved from
http://www.linguistics.northwestern.edu/documents/dissertations/linguistics-
research-graduate-dissertations-kimdissertation2012.pdf

Kim, M., Horton, W. S., & Bradlow, A. R. (2011). Phonetic convergence in
spontaneous conversations as a function of interlocutor language distance.
Laboratory Phonology, *2*(1). doi: 10.1515/labphon.2011.004

Kovac, M. (2011). Speech errors in English as foreign language: A case study of
engineering students in Croatia. *English Language and Literature Studies*, *1*, 20.
doi: 10.5539/ells.v1n1p20

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture
naming: Evidence for asymmetric connections between bilingual memory
representations. *Journal of Memory and Language*, *33*, 149–174. doi:
10.1006/jmla.1994.1008

La Heij, W. (2005). Selection processes in monolingual and bilingual lexical access. In
Kroll, J. F., & De Groot, A. M. (Eds.), *Handbook of bilingualism* (pp. 289–307).
Oxford University Press: Oxford.

Lagrou, E., Hartsuiker, R. J., & Duyck, W. (2011). Knowledge of a second language
influences auditory word recognition in the native language. *Journal of
Experimental Psychology: Learning, Memory, and Cognition*, *37*, 952. doi:
10.1037/a0023217

Lagrou, E., Hartsuiker, R. J., & Duyck, W. (2012). The influence of sentence context
and accented speech on lexical access in second-language auditory word
recognition. *Bilingualism: Language and Cognition*, *16*, 508–517. doi:
10.1017/s1366728912000508

Laver, J. D. M. (1973). The detection and correction of slips of tongue. In V. A.
Fromkin (Ed.), *Speech errors as linguistic evidence* (pp. 132–143). The Hague,
Netherlands: Mouton. doi: 10.1515/9783110888423.132

Levelt, W. (1983). Monitoring and self-repair in speech. *Cognition*, *14*, 41–104. doi:
10.1016/0010-0277(83)90026-4

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

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech
production. *Behavioral and Brain Sciences*, *22*, 1–38. doi:
10.1017/s0140525x99001776

Li, Y., Yang, J., Scherf, K. S., & Li, P. (2013). Two faces, two languages: An fMRI study of bilingual picture naming. *Brain and Language*, *127*, 452–462. doi: 10.1016/j.bandl.2013.09.005

Linebaugh, G., & Roche, T. B. (2015). Evidence that L2 production training can enhance perception. *Journal of Academic Language and Learning*, *9*, A1–A17.

MacKay, D. G. (1970). Spoonerisms: The structure of errors in the serial order of speech. *Neuropsychologia*, *8*, 323–350. doi: 10.1016/0028-3932(70)90078-3

Meuter, R. F., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, *40*, 25–40. doi: 10.1006/jmla.1998.2602

Molnar, M., Ibáñez-Molina, A., & Carreiras, M. (2015). Interlocutor identity affects language activation in bilinguals. *Journal of Memory and Language*, *81*, 91–104. doi: 10.1016/j.jml.2015.01.002

Monsell, S., Matthews, G. H., & Miller, D. C. (1992). Repetition of lexicalization across languages: A further test of the locus of priming. *Quarterly Journal of Experimental Psychology, 44A*, 763–783. doi: 10.1080/14640749208401308

Motley, M. T., Camden, C. T., & Baars, B. J. (1982). Covert formulation and editing of anomalies in speech production: Evidence from experimentally elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior*, *21*, 578–594. doi: 10.1016/s0022-5371(82)90791-5

Nooteboom, S. G. (2005). Lexical bias revisited: Detecting, rejecting and repairing speech errors in inner speech. *Speech Communication*, *47*, 43–58. doi: 10.1016/j.specom.2005.02.003

Nozari, N., Dell, G. S., & Schwartz, M. F. (2011). Is comprehension necessary for error detection? A conflict-based account of monitoring in speech production. *Cognitive Psychology*, *63*, 1–33. doi: 10.1016/j.cogpsych.2011.05.001

Oomen, C. C., & Postma, A. (2001). Effects of time pressure on mechanisms of speech production and self-monitoring. *Journal of Psycholinguistic Research*, *30*, 163–184.

Oomen, C. C. E., & Postma, A. (2002). Limitations in processing resources and speech monitoring. *Language and Cognitive Processes*, *17*, 163–184. doi: 10.1080/01690960143000010

Oomen, C. C., Postma, A., & Kolk, H. H. (2005). 12 Speech monitoring in aphasia: Error detection and repair behaviour in a patient with Broca's aphasia. *Phonological encoding and monitoring in normal and pathological speech*, 209.

Papadopoulou, D., & Clahsen, H. (2003). Parsing strategies in L1 and L2 sentence processing. *Studies in Second Language Acquisition*, *25*, 501–528. doi: 10.1017/s0272263103000214

Paulmann, S., Elston-Güttler, K. E., Gunter, T. C., & Kotz, S. A. (2006). Is bilingual lexical access influenced by language context? *NeuroReport*, *17*, 727–731. doi: 10.1097/01.wnr.0000214400.88845.fa

Peeters, D., Runnqvist, E., Bertrand, D., & Grainger, J. (2014). Asymmetrical switch costs in bilingual language production induced by reading words. *Journal of*

Experimental Psychology: Learning, Memory, and Cognition, *40*, 284. doi: 10.1037/a0034060

Pickering, M. J., & Garrod, S. (2006). Alignment as the basis for successful communication. *Research on Language and Computation*, *4*, 203–228.

Pickering, M. J., & Garrod, S. (2014). Self-, other-, and joint monitoring using forward models. *Frontiers in Human Neuroscience*, *8*. doi: 10.1007/s11168-006-9004-0

Poulisse, N. (1999). *Slips of the tongue: Speech errors in first and second language production* (Vol. 20). Amsterdam: John Benjamins.

Poulisse, N. (2000). Slips of the tongue in first and second language production. *Studia Linguistica*, *54*, 136–149. doi: 10.1017/s002222670100888x

Poulisse, N., & Bongaerts, T. (1994). First language use in second language production. *Applied Linguistics*, *15*, 36–57. doi: 10.1093/applin/15.1.36

Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological review*, *107*, 460. doi: 10.1037//0033-295x.107.3.460

Schlenck, K. J., Huber, W., & Willmes, K. (1987). "Prepairs" and repairs: Different monitoring functions in aphasic language production. *Brain and Language*, *30*, 226–244. doi: 10.1016/0093-934x(87)90100-3

Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *Journal of Memory and Language*, *55*, 197–212. doi: 10.1016/j.jml.2006.03.004

Seyfeddinipur, M., Kita, S., & Indefrey, P. (2008). How speakers interrupt themselves in managing problems in speaking: Evidence from self-repairs. *Cognition*, *108*, 837–842. doi: 10.1016/j.cognition.2008.05.004

Tian, X., & Poeppel, D. (2014). Dynamics of self-monitoring and error detection in speech production: Evidence from mental imagery and MEG. *Journal of Cognitive Neuroscience*, *27*, 352–364. doi: 10.1162/jocn_a_00692

Van Assche, E., Duyck, W., Hartsuiker, R. J., & Diependaele, K. (2009). Does bilingualism change native-language reading? Cognate effects in a sentence context. *Psychological Science*, *20*, 923–927. doi: 10.1111/j.1467-9280.2009.02389.x

Van Hest, E. (1996). *Self-repair in L1 and L2 production*. Tilburg, Netherlands: Tilburg University Press.

Van Hest, E. (2000). Analyzing self-repair: An alternative way of language assessment. In Ekbatani, G. & DeLeon Pierson, H. (Eds.), *Learner-Directed Assessment in ESL* (pp. 75–88). Routledge: New York/London. doi: 10.4324/9781410605283.

Woumans, E., Martin, C., Vanden Bulcke, C., Van Assche, E., Costa, A., Hartsuiker, R., et al. (2015). Can faces prime a language? *Psychological Science*, *26*, 1343–1352. doi: 10.1177/0956797615589330

Zhang, S., Morris, M. W., Cheng, C.-Y., & Yap, A. J. (2013). Heritage-culture images disrupt immigrants' second-language processing through triggering first-language interference. *Proceedings of the National Academy of Sciences*, *110*, 11272–11277. doi: 10.1073/pnas.1304435110