Bilingualism: Language and Cognition 9 (2), 2006, 167–176 © 2006 Cambridge University Press doi:10.1017/S1366728906002513

Modeling the control of phonological encoding in bilingual speakers*

ARDI ROELOFS
Max Planck Institute for Psycholinguistics; F. C. Donders Centre for Cognitive Neuroimaging, Nijmegen Institute for Cognition and Information, Nijmegen
KIM VERHOEF
Nijmegen Institute for Cognition and Information, Nijmegen

Phonological encoding is the process by which speakers retrieve phonemic segments for morphemes from memory and use the segments to assemble phonological representations of words to be spoken. When conversing in one language, bilingual speakers have to resist the temptation of encoding word forms using the phonological rules and representations of the other language. We argue that the activation of phonological representations is not restricted to the target language and that the phonological representations of languages are not separate. We advance a view of bilingual control in which condition-action rules determine what is done with the activated phonological information depending on the target language. This view is computationally implemented in the WEAVER++ model. We present WEAVER++ simulations of the cognate facilitation effect (Costa, Caramazza and Sebastián-Gallés, 2000) and the between-language phonological facilitation effect of spoken distractor words in object naming (Hermans, Bongaerts, de Bot and Schreuder, 1998).

People are generally good but not perfect at dealing with distraction. "I can resist everything except temptation", a character in an Oscar Wilde play once said. When talking in a foreign language, bilingual speakers have to resist the temptation of their mother tongue. The temptation to use the native language should be strong especially for bilingually unbalanced speakers, who are more proficient in the native than in a foreign language. This difference in proficiency presumably holds for selecting the appropriate words as well as for encoding their forms. The present article deals with the ability of bilingual speakers to encode the phonological forms of words in one language rather than another.

Encoding the phonological forms of words in one language would be unproblematic if the phonological representations of the two languages were completely separate and if phonological representations were selectively activated in the target language only. In this article, we review evidence that the activation of phonological representations is not restricted to the target language and that the phonological representations of the languages of a bilingual individual are not separate. The evidence for language-nonspecific phonological activation includes the cognate facilitation effect, which is the finding that object naming is facilitated when the object names in the two languages have similar phonological forms, like Spanish *gato* and Catalan *gat* for "cat" (Costa, Caramazza and Sebastián-Gallés, 2000; for an extensive discussion, see Costa, Santesteban and Caño, 2005). Evidence that phonological representations are shared between languages comes from the finding that phonemic segments of a picture name may be prepared without knowing whether one or the other language has to be used (Roelofs, 2003a). The language-nonspecific activation and the sharing of phonological representations raise the issue of the control of phonological encoding in bilingual individuals.

We briefly discuss a number of theoretical views on cognitive control and advance a view of bilingual control in which condition-action rules determine what is done with the activated phonological information depending on the target language. This view is implemented in the WEAVER++ model of spoken word production (e.g. Roelofs, 1992, 1997, 2003a). In order to provide some form of theoretical test of our claims (i.e. a proof of principle), we present WEAVER++ simulations. The simulations concerned the cognate facilitation effect (Costa et al., 2000) and the time course of the between-language phonological facilitation effect of spoken distractor words in object naming (Hermans, Bongaerts, de Bot and Schreuder, 1998). The temptation to encode forms in the nonintended language should be especially strong when bilingually unbalanced individuals have to produce words in their weaker second language while hearing words in their stronger first language, especially when the words are similar in sound (e.g. for Dutch–English bilingual speakers: the English picture name *mountain* and Dutch distractor word MOUW, *sleeve*). Bilingual speakers appear to be able to resist the

* We thank Albert Costa and three anonymous reviewers for helpful comments. The preparation of this article was supported by a VICI grant awarded to the first author by the Netherlands Organisation for Scientific Research (NWO).

Address for correspondence: Ardi Roelofs, Max Planck Institute for Psycholinguistics; Wundtlaan 1, 6525 XD, Nijmegen, The Netherlands
E-mail: *ardi@mpi.nl*

https://doi.org/10.1017/S1366728906002513 Published online by Cambridge University Press

temptation without apparent effort – the overall interference is similar when hearing first- and second-language words (Hermans et al., 1998). The ability to encode phonological forms in a second language while hearing words in the first language provides a strong test for models of bilingual control. We demonstrate that WEAVER++ passes the test, just like real bilingual speakers.

Bilingual phonological encoding

In this article, we restrict ourselves to phonological encoding in the production of single words, which is an essential component of the production of larger utterances. As a theoretical framework, we use the WEAVER++ model of word planning, which implements theoretical claims about phonological encoding in a single language (e.g. Roelofs, 1997) and about phonological encoding in bilingual speakers (Roelofs, 2003a). The WEAVER++ model (e.g. Roelofs, 1992; 1997; Levelt, Roelofs and Meyer, 1999; Roelofs, 2003a) holds that after conceptually driven retrieval of lemmas, which specify the syntactic properties of words, word forms are encoded in three steps: morphological encoding, phonological encoding, and phonetic encoding. Morphological encoding is the process by which the morphemes of a word are activated and selected. In phonological encoding, the phonemic segments of the morphemes are spelled out in parallel and they are prosodified in a sequential fashion from the beginning of a word to its end. Prosodification involves syllabification and stress assignment, yielding phonological word representations. These representations make explicit the syllables of a word and, for polysyllabic words, the stress pattern across syllables. Phonetic encoding translates the abstract phonological word representation into a context-dependent phonetic representation that can guide articulation (i.e. it makes explicit articulatory tasks such as lip protrusion and lowering of the jaw). WEAVER++ assumes that phonetic encoding includes accessing a syllabary, which is a store of ready-made motor programs for syllables. Figure 1 illustrates the processes assumed by WEAVER++ for phonological encoding and its relation to hearing words (relevant for the simulations of the experiments of Hermans et al., 1998).

Whereas morphemes differ between languages, some phonemic segments are common. For example, whereas the English language has the phonemic segment /θ/ (e.g. the last segment of *mouth*) and the Dutch language does not (Booij, 1995), both languages have phonemic segments such as /m/ (i.e. a voiced nasal labial that is contrastive within the languages). Phonemic segments such as /m/, /t/, /p/ and /k/ occur with only slight differences in most of the world's languages (Ladefoged and Maddieson, 1996). WEAVER++ implements the claim that bilingual speakers have a single memory token for each phonemic segment (e.g. /m/), which is shared between languages (Roelofs, 2003a). During segmental spellout, these segments are made available for the subsequent prosodification process.

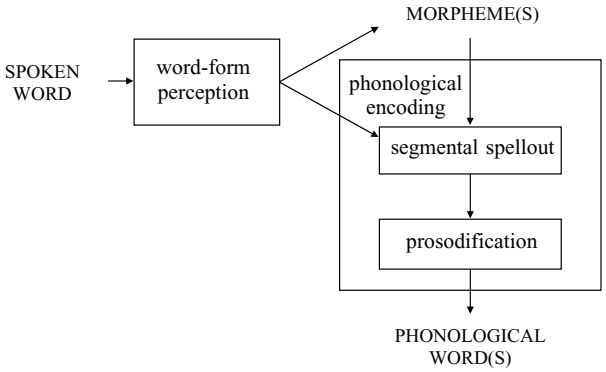


Figure 1. Phonological encoding in the WEAVER++ model. During phonological encoding, the phonemic segments of the morphemes of a word are spelled out in parallel and they are prosodified in a serial fashion from the beginning of a word to its end. Hearing words activates morphemes and phonemic segments in parallel (Roelofs, 1997).

Other aspects of phonological encoding, however, cannot be shared between languages. Languages often differ in the rules for syllabification and stress assignment that are employed during prosodification. For example, word stress is typically on the first syllable in English, but on the last syllable in French. Thus, a French–English bilingual speaker should have separate stress-assignment rules for English and French. In speaking English, a French–English bilingual individual should prevent the application of the French stress rule in constructing phonological words in English. In addition, there are language-specific rules as to how many consonants can surround a vowel and as to the internal organization of the sequence of phonemic segments within a syllable. Language specific rule-application is also required at the phonemic segment level. For example, Dutch and German have a phonological rule that devoices syllable-final obstruents (Booij, 1995), whereas English has no such rule. For example, the devoicing rule turns the underlying phonemic segment /d/ of Dutch *rood* "*red*" into a /t/ (the underlying /d/ surfaces as /d/ in the onset of the second syllable of the disyllabic inflected form *rode*). In talking in English, a Dutch–English bilingual speaker has to prevent the application of the devoicing rule to English phonemic segments. Some, but certainly not all, aspects of prosodification have been implemented in WEAVER++ (Roelofs, 1997; Levelt et al., 1999). Future developments of the model may include detailed aspects of prosodification in bilingual speakers. The WEAVER++ simulations that are reported below, however, do not depend on these detailed aspects.

https://doi.org/10.1017/S1366728906002513 Published online by Cambridge University Press

Languages may have common phonemic segments that differ in phonetic detail. For example, both Dutch and English have the phonemic segment /w/, which is pronounced differently in the two languages (e.g. Booij, 1995; Ladefoged, 2001). WEAVER++ captures this distinction between phonemic segments and their language-specific phonetic implementation by assuming that the phonemic segments may be shared between the two languages of bilingual individuals, whereas the syllable motor programs that phonetically implement the segments differ between languages. Thus, whereas phonological encoding (i.e. segmental spellout and prosodification) may happen in terms of shared segments, the phonetic encoding of the segments will differ between languages.

WEAVER++ makes a distinction between activation of information through spreading activation in a lexical network, on the one hand, and selection of activated information and construction of phonological word representations, on the other hand. In the model, selection and construction are achieved by condition-action rules that make explicit reference to behavioral goals. In particular, the condition-action rules determine what is done with the activated lexical information depending on the task. When a goal (e.g. to speak English rather than Dutch) is placed in working memory, the processing in the system is focussed on those condition-action rules that include the goal among their conditions. For example, in naming a pictured mountain in English (say "mountain") rather than in Dutch (say "berg"), a lexical concept is selected for the pictured object (i.e. MOUNTAIN(X)) and flagged to indicate the target language, here English. Lemmas are connected to language nodes that indicate the language to which the word belongs. The lexical concept MOUNTAIN(X) activates both the Dutch lemma *berg* and the English lemma *mountain*. However, only the condition-action rule of English *mountain* will fire, because only the language of the lemma *mountain* matches the target language, English. Subsequently, the corresponding morpheme <mountain> is selected for the selected lemma *mountain*, the phonemic segments /m/, /au/, /n/, /t/, and /I/ are selected for the selected morpheme <mountain> and used to generate a phonological word representation, and the syllable motor programs [maun] and [tIn] are accessed in the syllabary for the constructed phonological word.

For word production in a single language, WEAVER++ implements the assumption that only the phonological representations corresponding to a selected lemma are activated, called the discreteness assumption (Levelt et al., 1999). The discreteness assumption implies that activation at the lemma level does not cascade to the word forms of translation equivalents. However, discreteness does not exclude phonological activation of the nonintended language. In particular, phonological representations of the nonintended language are activated to the extent that representations are shared between languages in memory. Moreover, phonological rules of the other language, such as the Dutch devoicing rule in speaking English, will be triggered through these shared representations. And syllable motor programs will be activated in the other language via the shared phonemic segment representations. For example, in planning the form of the English word *mountain*, the phonemic segment /m/ will not only activate English syllable motor programs including [maun], but also Dutch syllable motor programs that phonetically realize the /m/. Thus, form activation in WEAVER++ is not restricted to the target language because of shared phonological representations.

Moreover, it is possible that there are links between the morphemes of translation equivalents in the form network. For example, Kroll and colleagues (Kroll and Stewart, 1994; Kroll and De Groot, 1997) argued that there exist both conceptual and lexical links in bilingual memory. The lexical links are established during early stages of second language learning. As individuals become more proficient in their second language, conceptual links are established and increasingly determine processing, but the lexical links do not disappear. If links between the morphemes of translation equivalents exist, the nonintended language will also inadvertently be activated via these links.

Language-nonspecific activation and shared phonological representations

Encoding the phonological forms of words in one language rather than another would be unproblematic if the phonological representations of the two languages were completely separate, unlike what is assumed by WEAVER++. Moreover, encoding the phonological forms of words in one language rather than the other would be unproblematic if the phonological representations of the two languages were selectively activated, unlike what is assumed by WEAVER++. We first review evidence suggesting that phonological activation is not restricted to the target language (for a more extensive review, see Costa, 2005). Next, we discuss evidence that suggests that phonological representations are shared between languages.

Is phonological activation restricted to the intended language?

The types of speech errors made by bilingual speakers cast doubt on the assumption that phonological activation is restricted to the target language. Sometimes bilingual speakers erroneously use phonemic segments or apply phonological rules of their native language when speaking a foreign language. In speaking English, for instance, Dutch–English bilingual speakers sometimes use the English pronunciation of phonemic segments

https://doi.org/10.1017/S1366728906002513 Published online by Cambridge University Press

in accidentally accessed Dutch words (Poulisse, 1999). For example, while intending to produce the English word *piece*, a speaker erroneously selected the Dutch translation equivalent *stuk* (/stUk/, where the vowel has the phonological feature [−back]) and pronounced it as the English word *stuck* (/stUk/, where the vowel has the feature [+back], in particular, the English /U/ is a lax high back rounded vowel). Such between-language interference suggests that phonological activation is not restricted to the target language.

Other evidence has been taken to suggest, however, that in naming objects only the form of the word in the intended language becomes activated. Hermans and colleagues (Hermans et al., 1998; Hermans, 2000) asked Dutch-English bilingual speakers to name pictures in their second language English, while trying to ignore spoken first-language (i.e. Dutch) distractors. For example, participants had to name a pictured mountain (say "mountain"), while they heard Dutch distractor words that were phonologically related to the English picture name (e.g. Dutch MOUW, *sleeve*), Dutch distractor words that were phonologically related to the Dutch translation equivalent ("berg") of the English picture name (e.g. Dutch BERM, *verge*), or phonologically unrelated Dutch words (e.g. Dutch KAARS, *candle*). The onset of spoken distractor presentation was before picture onset (i.e. −300 or −150 msec, henceforth referred to as distractor-first SOAs, indicated by a minus sign), the onsets of distractor and picture coincided (SOA = 0 msec), or the onset of distractor presentation was after picture onset (i.e. 150 msec, henceforth distractor-second SOAs). It was observed that the phonologically related Dutch distractors speeded up picture naming in English compared to phonologically unrelated Dutch distractors. The facilitation was obtained at the distractor-second SOA of 150 msec but not at the earlier SOAs. The Dutch distractors that were phonologically related to the Dutch translation equivalent of the English picture name (e.g. Dutch BERM, phonologically related to Dutch "berg", *mountain*) yielded interference compared to the unrelated distractors. The interference was obtained at the early SOAs (i.e. the SOAs of −300, −150, and 0 msec) but not at the distractor-second SOA of 150 msec. Thus, the phonological facilitation effect (of Dutch MOUW, *sleeve*) and the phonological interference effect (of Dutch BERM, *verge*) occurred at different SOAs. According to Hermans et al. (1998), the absence of overlap in the timing of effects suggests that the Dutch translation equivalents of the English picture names are not activated at the phonological level.

However, Costa, Colomé, Gómez and Sebastián-Gallés (2003) observed that the phonological facilitation effect and the phonological interference effect occur at the same distractor-second SOA of 150 msec in bilingually balanced Spanish-Catalan speakers. Moreover, the outcomes of other studies provided converging evidence that translation equivalents do become phonologically activated in object naming. Costa et al. (2000) asked Catalan-Spanish bilingual speakers to name pictured objects whose translation equivalents were phonologically related (the cognates) and pictures whose translation equivalents were phonologically unrelated (the noncognates). For example, the Catalan and Spanish words for cat, *gat* and *gato*, are phonologically related. It was observed that the pictures were named faster when their translation equivalents were phonologically related than when they were unrelated. This phonological facilitation effect was not observed when Spanish monolingual speakers named the pictures. Recently, we replicated the cognate facilitation effect in our own lab with bilingually unbalanced Dutch-English speakers. The cognate facilitation effect suggests that the phonological forms of the picture name in the target language and its translation equivalent in the other language are activated in naming objects.

This conclusion was confirmed in two experiments by Gollan and Acenas (2004), who observed that tip-of-the-tongue states (TOTs) in English-Spanish and English-Tagalog speakers are less likely to occur for cognates than for noncognates. For monolingual individuals, this difference was not obtained. In a TOT, a speaker is able to access the form of the word only partly. Often, the initial phonemic segments of the word are accessible, but access to the remainder of the word fails. The finding that fewer TOTs happen for cognates than for noncognates in bilingual individuals suggests that translation equivalents are activated phonologically, which leads to an increase in activation of the target in case there is phonological overlap (cognates) compared to when there is no overlap (noncognates). This makes a TOT less likely for cognates than for noncognates in bilingual speakers.[1]

The cognate effects suggest that translation equivalents are phonologically activated. This is also supported by other evidence. Colomé (2001) observed that the speed of internal monitoring for target phonemic segments in picture names by Catalan–Spanish bilingual speakers is facilitated by the presence of the phonemic segment in the translation equivalent. The facilitation was not observed for monolingual speakers. This suggests that

[1] Gollan and Acenas (2004) claimed that the WEAVER++ model cannot account for the cognate facilitation effect on TOTs, because the effect reflects facilitation of phonemic segment selection. This claim was based on the assumption that TOTs in the model reflect failures of lexeme (i.e. morpheme) node selection. However, according to Levelt et al. (1999) and Roelofs, Meyer and Levelt (1998), "a speaker in a tip-of-the-tongue state has the feeling of knowing a particular word, but can only access part of the information about the word form" (Roelofs et al., 1998, p. 224). Failures in accessing word forms in the model do not need to be restricted to trouble in morpheme retrieval, but may also involve problems in phonemic segment node selection.

https://doi.org/10.1017/S1366728906002513 Published online by Cambridge University Press

the translation equivalents of picture names become phonologically activated. Similarly, Rodriguez-Fornells, Van der Lugt, Rotte, Britti, Heinze and Münte (2005) observed that deciding whether picture names begin with a vowel or consonant by German-Spanish bilingual speakers is affected by whether the translation equivalent of the picture name also begins with a vowel or consonant.

The cognate effects in picture naming and tip-of-the-tongue experiments and the findings from internal phoneme monitoring suggest that phonological representations in the other language are activated. Costa et al. (2000) argued that the cognate facilitation effect indicates that the discreteness assumption (which says that only word forms corresponding to selected lemmas become activated) does not hold for the languages of bilingual speakers. The cognate effect suggests that shared phonological representations are activated by two sources of activation. The first source of activation is the lemma of the picture name in the target language. One possible second source of activation is the lemma of the translation equivalent of the picture name in the other language (as assumed by Costa et al., 2000). Another possible second source of activation is the morpheme of the translation equivalent if direct lexical links exist (this possibility was rejected by Gollan and Acenas, 2004). We put these two hypotheses about the cause of the cognate facilitation effect to a theoretical test by running WEAVER++ simulations.

In the first simulation, we dropped the discreteness assumption of WEAVER++ and allowed for cascading of activation of lemmas to word forms in the model. As argued by Roelofs (2003b), the amount of activation that cascades from lemmas to word forms needs to be restricted to be able to explain several other findings in the literature. The computer simulations showed that a weakly cascading version of WEAVER++ accounts for the cognate facilitation effect. The procedure in simulating the experiment of Costa et al. (2000) showing the cognate facilitation effect followed the simulations reported by Roelofs (1997). In simulating the bilingual group, the morphemes of the target word and its translation equivalent (e.g. Spanish <gato> and Catalan <gat>) were activated, whereas in simulating the monolingual group, only the target morpheme was activated. Figure 2 shows the cognate facilitation effect observed by Costa et al. (2000) and the results from the WEAVER++ simulations. Picture naming is faster for cognates than for noncognates in bilingual but not in monolingual individuals. WEAVER++ shows this pattern of results even when the activation input to the morpheme of the translation equivalent is only 5% of the activation input to the target morpheme. Thus, the cognate effect occurs in WEAVER++ even when the amount of activation that spreads from lemmas to word forms is limited.

In the second simulation, we tested whether links between morphemes of translation equivalents (and no cascading of activation between lemmas and forms) would also yield the cognate facilitation effect. This was indeed the case. When links between morphemes were present in the network (e.g. between Spanish <gato> and Catalan <gat>), a cognate facilitation effect of about 35 msec was obtained for the simulated bilingual group and no effect for the monolingual group. When the morphemes of translation equivalents are connected, the target phonemic segments receive activation from both the target morpheme and the morpheme of the translation equivalent in case of cognates but not in case of noncognates. For example, the segments of Spanish *gato* receive activation from both the Spanish morpheme node <gato> and the Catalan morpheme node <gat> if these morphemes are linked. However, in case of links between morphemes of Spanish and Catalan words that share no segments (noncognates), the target segments receive activation from only one source, namely the Spanish target morpheme. Consequently, encoding will be faster for cognates than for noncognates: the cognate facilitation effect.

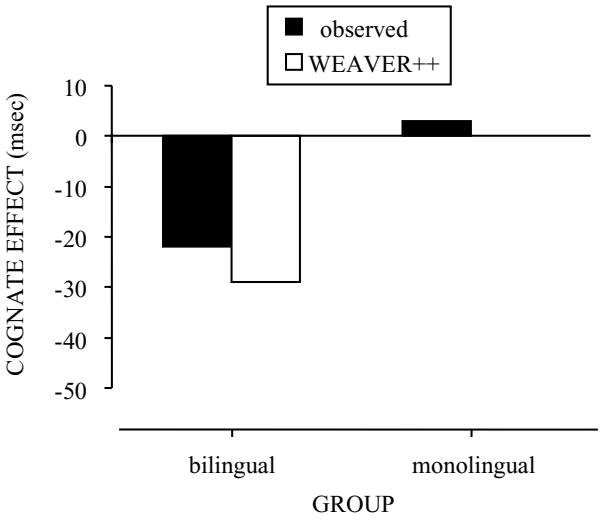


Figure 2. The cognate facilitation effect. Real data (Costa et al., 2000) and WEAVER++ simulation results.

To conclude, the WEAVER++ simulations suggest that the cognate facilitation effect in picture naming may arise because of cascading of activation from lemmas to word forms or because of links between morphemes of translation equivalents in the production system. Future experimental work may empirically test between these alternatives (we are conducting such tests currently in our lab).

Shared or separate phonological representations?

The cognate effects (Costa et al., 2000; Gollan and Acenas, 2004) and the findings from the internal phoneme

https://doi.org/10.1017/S1366728906002513 Published online by Cambridge University Press

monitoring (Colomé, 2001; Rodriguez-Fornells et al., 2005) suggest that phonological representations in the other language are activated. However, the evidence is inconclusive about whether representations are shared between languages or only closely linked. The cognate facilitation effect can happen if the phonemic segments of cognates are shared between languages in memory, but also if the segment tokens are separate but linked. For example, the cognate effect may occur if a single memory token for the first phonemic segment of the Spanish word *gato* and the Catalan word *gat* is shared between languages in memory, but also when there are different memory tokens for the segments that are linked. If different segment tokens are linked, the activation of the segments of the Catalan word *gat* can spread to those of the Spanish word *gato*, thereby yielding the cognate facilitation effect.

Roelofs (2003a) explicitly tested whether representations of common phonemic segments are shared between languages using a form-preparation paradigm. The participants were Dutch–English bilingual individuals. They were given words to produce that did or did not share onset phonemic segments in a block of trials, and that came or did not come from different languages. The words were not read aloud, but participants first learned prompt-response pairs (e.g. RIVER – "boat", WING – "bird", GIRL – "boy") and then produced the response words (e.g. "bird") when presented with the prompt words (e.g. WING), following seminal experiments by Meyer (1990, 1991). If representations of common phonemic segments are shared between languages, bilingual speakers should be able to plan initial phonemic segments that are common to both languages without knowing the language of the word. A preparation effect should be obtained when initial phonemic segments are shared even when the languages of the words in a block of trials differ. For example, it should be possible to plan /b/ when the block of trials includes English *boat* and *bird* and Dutch *boor* (*drill*). It was found that when onset phonemic segments were shared among the response words, those onsets were prepared, even when the words came from different languages. There was no preparation effect when phonological features rather than full phonemic segments were shared among the responses. For example, *boat*, *bird*, and *boy* yielded a preparation effect, but *boat*, *bird*, and *pain* did not. In the latter case, /b/ and /p/ share all phonological features except that /b/ is voiced and /p/ is voiceless. These results suggest that the representations of phonemic segments common to the languages are shared in memory.

Recently, Damian and Bowers (2003) questioned the phonological interpretation of the preparation effects by Roelofs (2003a) on the grounds that when initial phonemic segments are not identical but share phonological features only (e.g. /b/ vs. /p/), their spelling will also differ. In experiments conducted in English, using the prompt-response word generation task, Damian and Bowers (2003) observed that when the words in a set shared initial phonemic segments that differed in spelling, no preparation effect was obtained. For example, a preparation effect was obtained for *camel*, *coffee*, *cushion*, but not for *kennel*, *coffee*, *cushion*. However, in experiments in Dutch conducted in our own lab (Roelofs, in press), we failed to replicate the spelling disruption effect that Damian and Bowers (2003) obtained for English. We tested for spelling effects using word production tasks in which spelling is relevant (oral reading) or irrelevant (picture naming and prompt-response word generation). Response preparation was disrupted by spelling inconsistency only in reading. In picture naming and paired-associate word generation, a preparation effect was obtained for words that shared phonemic segments regardless of their spelling. The size of the preparation effect did not differ between phonemic segments that were spelled the same and phonemic segments that were spelled differently in a block of trials. These results suggest that the cross-language preparation effects observed by Roelofs (2003a) are due to shared phonemic segments.

To summarize, the available evidence suggests that the phonological representations of bilingual speakers are not separate and that the activation of phonological representations is not restricted to the target language. This raises the issue of the control of phonological encoding in bilingual individuals. How are bilingual speakers able to encode the phonological forms of words in one language rather than the other, given that the phonological forms of translation equivalents are co-activated with the target form? In the next section, we briefly discuss a number of theoretical views on cognitive control (for more extensive discussions, see Roelofs, 2003b, 2004). In the remainder, we test WEAVER++ on data on the time course of between-language phonological effects of spoken distractor words in bilingual object naming (Hermans et al., 1998). The temptation to encode forms in the nonintended language should be especially strong when bilingually unbalanced individuals have to produce words in their second language while hearing words in their first language. Bilingual speakers appear to be able to resist the temptation without apparent effort. The ability to encode phonological forms in the second language while hearing words in the first language provides a strong test for models of bilingual control. We demonstrate that WEAVER++ passes the test.

Views on control

According to a prominent view on cognitive control in the literature (e.g. Cohen, Dunbar and McClelland, 1990; Phaf, Van der Heijden and Hudson, 1990; Cohen and Huston, 1994; Botvinick, Braver, Barch, Carter and Cohen, 2001), the issue of control arises because selection

https://doi.org/10.1017/S1366728906002513 Published online by Cambridge University Press

processes are blind in that only the level of activation of memory information is taken into consideration. The highest activated information is always selected. This is also assumed by a prominent monolingual model of phonological encoding in the literature, the model of Dell (1986). Activation-based selection in bilingual phonological encoding requires that activation levels of phonological information are selectively enhanced for the intended language or that the phonological information of the nonintended language is selectively inhibited. Green (1998) proposed an inhibition model for the control of languages in bilingual individuals. According to this model, words in the nonintended language are reactively inhibited.

According to an alternative view, implemented in WEAVER++, selection in phonological encoding is guided in that not only the level of activation of information is taken into consideration but also symbolic information (Roelofs, 1997, 2003b, 2004). In bilingual performance, symbolic information about language membership is taken into account. Inhibition is therefore not required for selection in one language or another (Roelofs, 1998, 2003b). WEAVER++ combines a lexical network with condition-action rules that determine what is done with the activated lexical information depending on the task. When a goal symbol is placed in working memory (e.g. specifying the intended language), the processing in the system is focussed on those condition-action rules that include the goal among their conditions.

Control processes have been intensively studied by introducing distraction while participants perform a task (following Stroop, 1935). For example, in a picture-word interference experiment, participants have to name pictured objects while trying to ignore written or spoken distractor words (MacLeod, 1991). The picture-word interference paradigm has been a major workhorse in examining issues of selection and control in bilingual speakers (e.g. Hermans et al., 1998; Costa and Caramazza, 1999; Costa, Miozzo and Caramazza, 1999; Hermans, 2000; Costa, 2005).

Modeling the time course of between-language phonological facilitation

Seminal experiments by Schriefers, Meyer and Levelt (1990) and Meyer and Schriefers (1991) revealed that spoken distractors in picture naming delay responding compared to no distractor. In these experiments, participants named the pictures in their first language while hearing first-language distractor words. A phonological relation between picture name and spoken distractor word reduces the interference. For example, in naming a pictured mountain (say "mountain"), the naming latencies are shorter when the phonologically related spoken distractor word MOUTH is presented than when the phonologically unrelated spoken distractor word CAKE is presented (the example is in English, but the experiments of Schriefers et al., 1990; and Meyer and Schriefers, 1991 were conducted in Dutch). The time course of the phonological facilitation effect has been studied by manipulating the SOA. Phonological facilitation is typically first obtained at short distractor-first SOAs (i.e. −150 msec) and the magnitude of the facilitation increases when the SOA becomes more positive (i.e. from SOA = 0 msec to SOA = 150 msec). For example, the facilitation is typically larger at SOA = 0 than at SOA = −150 msec, and it is larger at SOA = 150 than at SOA = 0 msec.

In later experiments, Starreveld (2000) demonstrated that the onset of phonological facilitation by spoken distractors at the SOA of −150 msec is not sacrosanct. Again, participants named pictures in their first language while hearing first-language distractor words. When word-initial fragments are presented instead of whole words, phonological facilitation is already obtained at SOA = −300 msec. Moreover, whether facilitation is obtained at the early SOA (i.e. −300 msec) depends on the nature of the other distractors in an experiment. The facilitation effects at early SOAs seem to be influenced by strategic factors (see also Roelofs, 2003b). However, regardless of whether effects are obtained at the earliest SOAs, the facilitation increased when the SOA became more positive (i.e. from SOA = 0 msec to SOA = 150 msec) in the experiments of Starreveld (2000).

Elsewhere (Roelofs, 1997), it has been shown that WEAVER++ accounts for the time course of phonological facilitation of spoken distractor words in picture naming (first-language targets and distractors). Phonological facilitation effects are largest in the model when the activation of phonemic segment representations by the distractor is maximal at the moment in time that the phonemic segments are needed for constructing the phonological word representation for the picture name. Given that phonological encoding is a late process in planning a word (it follows conceptualizing, lemma retrieval, and morphological encoding), and that the spoken distractor develops over time, phonological distractors have the biggest impact in the model when they are presented at late (i.e. distractor-second) SOAs. This corresponds that what is empirically observed (e.g. Meyer and Schriefers, 1991; Starreveld, 2000).

Hermans et al. (1998) observed that the phonological facilitation of spoken distractor words in picture naming may occur across the languages of bilingual speakers. The between-language phonological facilitation suggests that any spoken word, regardless of the language to which it belongs, may activate the corresponding segments in the other language. This provides confirming evidence for the view that phonemic segments common to the languages of a bilingual speaker are shared (Roelofs, 2003a) or

https://doi.org/10.1017/S1366728906002513 Published online by Cambridge University Press

closely linked in memory. In the bilingual experiments of Hermans et al. (1998), the target picture names and the phonologically related spoken distractors shared the word-onset consonants and the first vowel. Although the Dutch and English vowels of the words were very similar, they were not identical. Following Hermans et al. (1998), it is therefore plausible to assume that because of phonetic mismatch, the input word forms of the second-language target words are activated less by first-language than by second-language spoken distractors during word-form perception (cf. Ladefoged and Maddieson, 1996). For example, the English target word *mountain* is presumably activated less by the Dutch word MOUW (where the vowel has the phonological feature [–back], Booij, 1995) than by the English word MOUTH (where the vowel has the phonological feature [+back], just as in the target *mountain*, Ladefoged, 2001). Consequently, the English morpheme <mountain> in the production network is activated less when hearing the phonologically related first-language distractor MOUW than when hearing the second-language distractor MOUTH. Because of these differences in activation, the phonological facilitation will be less for Dutch MOUW than for English MOUTH, as empirically observed by Hermans et al. (1998).

We tested WEAVER++ on the data that Hermans et al. (1998) obtained on the time course of the between-language phonological facilitation effect. The simulations involved picture naming in English (the second language) by Dutch–English bilingual speakers. The spoken distractors were phonologically related or unrelated Dutch words (the first language). The procedure in simulating the bilingual performance was identical to the simulations reported by Roelofs (1997) for experiments with first-language targets and distractors (the experiments of Meyer and Schriefers, 1991). All parameter values of WEAVER++ in the simulations were the same as those in the simulations reported by Roelofs (1997) except that the correction for the mental SOA was set at 125 msec (instead of 100 msec in Roelofs, 1997), capturing the fact that the second-language picture naming of Hermans et al. (1998) was slower than the first-language picture naming simulated in Roelofs (1997). Moreover, the between-language lexical activation was set at 20% of the within-language activation in the simulations reported in Roelofs (1997).

Figure 3 gives the time course of the between-language phonological effect obtained in the simulations together with the empirical data obtained by Hermans et al. (1998). In the WEAVER++ simulations, phonological facilitation was obtained at the distractor-second SOA of 150 msec but not at earlier SOAs. Phonological facilitation occurred because the Dutch distractor (e.g. MOUW) speeded up the selection of segments and syllable motor programs of the English target (e.g. *mountain*). Given that phonological encoding is a late process in planning a word and that

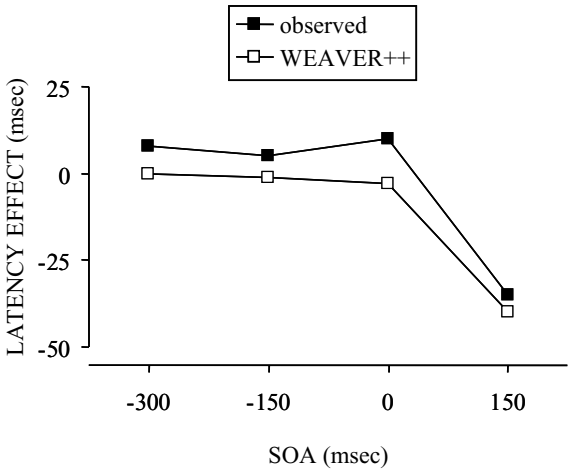


Figure 3. The time course of the between-language facilitation effect of phonologically related spoken distractors relative to phonologically unrelated distractors in picture naming by bilingual speakers. The target language was the speakers' second language English (L2), whereas the distractor language was their first language Dutch (L1). Real data (Hermans et al., 1998) and WEAVER++ simulation results.

the speech signal of the distractor unfolds over time, phonological distractors had the biggest impact in the model when they were presented at late SOAs. The simulations demonstrate that WEAVER++ encodes the right word form in the weaker, second language (English) despite the fact that a similar sounding word in the stronger, first language (Dutch) is presented as distractor. This was achieved in the model without inhibiting the nontarget language as a whole or inhibiting words of the nontarget language (cf. Green, 1998). When a goal symbol specifying the intended language English is placed in working memory, the processing in the model is focussed on those condition-action rules that include the target language among their conditions. Consequently, the morphemes, segments, and syllable motor programs for the English picture name are selected. The simulations demonstrate that WEAVER++ encoded the phonological form of the English picture name with a time course that is similar to that observed for real Dutch-English bilingual speakers.

Other models of the control of spoken word production predict a different time course of the distractor effect. For example, the models of Cohen et al. (1990), Phaf et al. (1990), Cohen and Huston (1994), and Botvinick et al. (2001) predict that distractors have the biggest impact at distractor-first SOAs (see Roelofs, 2003b for an extensive discussion). The model of Cohen et al. (1990) assumes that evidence for one spoken response or another accumulates in parallel in a word production network until a response threshold is reached. Consequently, more

https://doi.org/10.1017/S1366728906002513 Published online by Cambridge University Press

evidence will have accumulated for one response or another at distractor-first than at distractor-second SOAs. Thus, the biggest effects of distractors are predicted at long distractor-first SOAs and the smallest distractor effects are predicted at distractor-second SOAs, unlike the real data (see Figure 3). The models of Phaf et al. (1990), Cohen and Huston (1994), and Botvinick et al. (2001) assume that word planning occurs via a constraint-satisfaction network that settles onto one stable activation state or another depending on the inputs to the network. In perceiving a distractor word, the system quickly settles into a stable activation state corresponding to the distractor. The stable state does not vary with time. As a result, making the distractor-first SOA longer or shorter has no effect on the impact of distractors, except when an SOA is used that is too short for the distractor to induce a stable activation state. Consequently, the impact of distractors is maximal and constant at distractor-first SOAs and only reduces at distractor-second SOAs, unlike the real data (see Figure 3). To summarize, other models of control (i.e. Cohen et al., 1990; Phaf et al., 1990; Cohen and Huston, 1994; Botvinick et al., 2001) predict that distractors have the biggest impact at distractor-first SOAs and the smallest impact at distractor-second SOAs, unlike the real data. In contrast, WEAVER++ predicted maximal impact of the distractors at distractor-second SOAs, in agreement with the data of Hermans et al. (1998).

Conclusions

When talking in one language, bilingual speakers have to resist the temptation of encoding word forms using the phonological rules and representations of the other language. We argued that the phonological representations of the languages of bilingual individuals are not separate and that the activation of phonological representations is not restricted to the target language. We advanced a view of bilingual control in which condition-action rules determine what is done with the activated phonological information depending on the target language, as implemented in the WEAVER++ model of spoken word production. We presented the results of WEAVER++ simulations of bilingual phonological encoding, which showed that the model accounts for the cognate facilitation effect and for the time course of between-language phonological facilitation of spoken distractor words in object naming. The presented bilingual version of WEAVER++ is the first implemented model of bilingual phonological encoding and its control. In this article, we presented tests of the model against existing bilingual data. Future research may put the model to test on new sets of data, as we are currently doing by measuring response times and event-related brain potentials in bilingual spoken word production.

References

Booij, G. E. (1995). *The phonology of Dutch*. Oxford: Oxford University Press.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S. & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, 108, 624–652.

Cohen, J., Dunbar, K. & McClelland, J. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review*, 97, 332–361.

Cohen, J. D. & Huston, T. A. (1994). Progress in the use of interactive models for understanding attention and performance. In C. Umiltà & M. Moscovitch (eds.), *Conscious and nonconscious information processing: Attention and performance XV*, pp. 453–476. Cambridge, MA: MIT Press.

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

Costa, A. (2005). Lexical access in bilingual production. In J. F. Kroll & A. M. B. de Groot (eds.), *Handbook of bilingualism: Psycholinguistic approaches*, pp. 308–325. Oxford, UK: Oxford University Press.

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

Costa, A., Caramazza, A. & Sebastián-Gallés, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1283–1296.

Costa, A., Colomé, A., Gómez, O. & Sebastián-Gallés, N. (2003). Another look at cross-language competition in bilingual speech production: Lexical and phonological factors. *Bilingualism: Language and Cognition*, 6, 167–179.

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

Costa, A., Santesteban, M. & Caño, A. (2005). On the facilitatory effect of cognate words in bilingual speech production. *Brain and Language*, 94, 94–103.

Damian, M. F. & Bowers, J. S. (2003). Effects of orthography on speech production in a form-preparation paradigm. *Journal of Memory and Language*, 49, 119–132.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93, 283–321.

Gollan, T. H. & Acenas, L.-A. R. (2004). What is a TOT? Cognate and translation effects on tip-of-the-tongue states in Spanish–English and Tagalog–English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 246–269.

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

Hermans, D. (2000). Word production in a foreign language. Ph.D. dissertation. University of Nijmegen.

https://doi.org/10.1017/S1366728906002513 Published online by Cambridge University Press

Hermans, D., Bongaerts, Th., De Bot, K. & Schreuder, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language? *Bilingualism: Language and Cognition*, 1, 213–229.

Kroll, J. F. & De Groot, A. M. B. (1997). Lexical and conceptual memory in the bilingual: Mapping form to meaning in two languages. In A. M. B. de Groot & J. F. Kroll (eds.), *Tutorials in bilingualism: Psycholinguistic perspectives*, pp. 169–199. Mahwah, NJ: Lawrence Erlbaum Publishers.

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.

Ladefoged, P. (2001). *Vowels and consonants: An introduction to the sound of languages*. Oxford: Blackwell.

Ladefoged, P. & Maddieson, I. (1996). *The sounds of the world's languages*. Oxford: Blackwell.

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

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, 109, 163–203.

Meyer, A. S. (1990). The time course of phonological encoding in language production: The encoding of successive syllables of a word. *Journal of Memory and Language*, 29, 524–545.

Meyer, A. S. (1991). The time course of phonological encoding in language production: The phonological encoding inside a syllable. *Journal of Memory and Language*, 30, 69–89.

Meyer, A. S. & Schriefers, H. (1991). Phonological facilitation in picture–word interference experiments: Effects of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 1146–1160.

Phaf, R. H., Van der Heijden, A. H. C. & Hudson, P. T. W. (1990). SLAM: A connectionist model for attention in visual selection tasks. *Cognitive Psychology*, 22, 273–341.

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

Rodriguez-Fornells, A., Van der Lugt, A., Rotte, M., Britti, B., Heinze, H.-J. & Münte, T. F. (2005). Second language interferes with word production in fluent bilinguals: Brain potential and functional imaging evidence. *Journal of Cognitive Neuroscience*, 14, 422–433.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, 42, 107–142.

Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, 64, 249–284.

Roelofs, A. (1998). Lemma selection without inhibition of languages in bilingual speakers. *Bilingualism: Language and Cognition*, 1, 94–95.

Roelofs, A. (2003a). Shared phonological encoding processes and representations of languages in bilingual speakers. *Language and Cognitive Processes*, 18, 175–204.

Roelofs, A. (2003b). Goal-referenced selection of verbal action: Modeling attentional control in the Stroop task. *Psychological Review*, 110, 88–125.

Roelofs, A. (2004). The seduced speaker: Modeling of cognitive control. In A. Belz, R. Evans, & P. Piwek (eds.), *Natural language generation*, pp. 1–10. Berlin: Springer-Verlag.

Roelofs, A. (in press). The influence of spelling on phonological encoding in word reading, object naming, and word generation. *Psychonomic Bulletin & Review*.

Roelofs, A., Meyer, A. S. & Levelt, W. J. M. (1998). A case for the lemma-lexeme distinction in models of speaking: Comment on Caramazza and Miozzo (1997). *Cognition*, 69, 219–230.

Schriefers, H., Meyer, A. & Levelt, W. J. M (1990). Exploring the time-course of lexical access in language production: Picture–word interference studies. *Journal of Memory and Language*, 29, 86–102.

Starreveld, P. A. (2000). On the interpretation of auditory context effects in word production. *Journal of Memory and Language*, 42, 497–525.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643–662.

Received 2 November 2004
Revision received 14 September 2005
Accepted 18 October 2005

https://doi.org/10.1017/S1366728906002513 Published online by Cambridge University Press