



Access provided by Australian National University (14 Mar 2018 23:07 GMT)

LANGUAGE MONITORING IN BILINGUALS AS A MECHANISM FOR RAPID LEXICAL DIVERGENCE

T. MARK ELLISON

Australian National University

LUISA MICELI

University of Western Australia

Recent studies have highlighted divergent change as a more common outcome of language contact than previously thought. While convergent change is often attributed to bilingual cognitive pressures, divergent change has usually been explained by appealing to sociocultural factors. We argue that the effects of social pressures on linguistic systems must nevertheless be realized in how language is processed in the individual bilingual speaker and, therefore, that divergent change is also ultimately rooted in bilingual cognition. Since lexical forms are most susceptible to contact-induced divergent change we focus on their production. We begin by developing a cognitive model that combines Grosjean's language mode with a later output-monitoring stage. The parameters to the model are then fit to the results of an experiment in which bilinguals are seen to avoid shared lexical items. These best-fit parameters form the basis of a series of multi-agent simulations that show rapid divergence in the lexica of languages with large proportions of bilinguals. We consider the implications of these findings for the psycholinguistic study of bilingual lexical selection, the construction of phylogenies, and the reconstruction of language family histories.

Keywords: bilingual lexicon, cognitive bias, contact-induced change, divergence, language contact, phylogeny

1. INTRODUCTION. The default assumption in linguistics is that language contact leads to convergent change. However, as recent works highlight—for example, Muysken 2013, Braunmüller et al. 2014, and Evans's recent plenary address at the International Conference for Historical Linguistics (Evans 2015)—this is not the only possible outcome. Increasingly, empirical studies of contact situations identify stability and even divergence as potential outcomes, and also find that convergent and divergent change occur at the same time (Höder 2014). For example, there are now several well-documented cases of scant lexical borrowing coupled to extensive structural convergence (e.g. Ross 2001, Aikhenvald 2002, Epps 2007), while François (2011) describes a situation of ongoing contact that has given rise to divergence of word forms and convergence in structure (semantics and morphosyntax). François argues that the languages of Northern Vanuatu can now be effectively described as having one grammar but seventeen distinct lexicons (François 2011:223–24). Arnal (2011) proposes a similar trajectory for Catalan in recent decades: convergence with Spanish in structure, along with emerging divergence in the lexicon (see also Kasstan 2013 on lexical 'distanciation'). O'Shannessy (2015a,b) has likewise observed ongoing differentiation of lexical forms between Warlpiri and Light Warlpiri in the community of Lajamanu (northern Western Australia). Harvey (2011:365–66) suggests that certain contact situations in precolonial Australia would have favored the creation of neologisms and derivation, rather than convergence due to borrowing, consequently giving rise to divergence in the repertoire of lexical forms.

Thus, it can be argued that divergence is more common as an outcome of language contact than the traditional emphasis on convergence in the literature suggests. The

* The authors' names are listed alphabetically, and the authors take equal responsibility for the work presented here. The article has benefited from discussions and/or feedback from a number of colleagues, the editors, and two anonymous referees. In particular, we wish to thank Bethwyn Evans, Nicholas Evans, Nicolas Fay, Simon Greenhill, Shelly Harrison, John Henderson, and Erich Round. All flaws remain our own. Data and code used in the preparation of this article are available at <http://bld.markellison.net/> and <http://muse.jhu.edu/resolve/13>. T. Mark Ellison was in part supported in this work by Australian Research Council grants DP120104237 and FL130100111.

overrepresentation of convergence in the contact literature may therefore be partly due to the fact that it is easier to identify and describe, since there are limited possibilities as to how languages may converge, whereas divergence can be achieved in many different ways. It may also stem from linguists' greater interest in grammatical structure than the lexicon, since the impact of divergence appears to be most obvious in word forms. The privileged status of convergence may furthermore be linked to the assumption that it is a more natural outcome of bilingual cognitive processing—associated with cognitive overload of the bilingual speaker (see e.g. Matras 1998, 2007)—while divergence, when discussed, is always attributed to specific, idiosyncratic social circumstances. For example, in the northern Vanuatu case, François's explanation for lexical differentiation is that speakers' greater awareness of lexical forms renders them potential emblems of locales, while the relatively unconscious use of language structures makes them unsuitable for this role (François 2011:224). The key motivator, he contends, is a Melanesian ideological bias that fosters cultural and linguistic differentiation—namely, a predilection to ascribe things and people to specific places (François 2011:228). Harvey (2011), likewise appealing to speakers' awareness of word forms, analyzes the low rates of shared lexical items in certain contiguous related languages across northern Australia as the by-product of emblematic code-switching. Using a distinctive lexeme at the point of code-switching highlights the switch, while using shared vocabulary may result in the switch passing unnoticed. Both of these proposed explanations link the divergence of the lexicon to the conscious expression of language identity—whether emblematic of group and/or place.

Although we do not deny the role of social factors, in this article we argue that the effects of social pressures on linguistic systems must nevertheless be realized in how language is processed in the individual bilingual speaker—and therefore that lexical divergence, like convergence of structure, is also ultimately rooted in bilingual cognition. In order to explain the impact of social pressure on individual linguistic behavior, we develop a probabilistic model of bilingual language control in lexical production. Our model combines Grosjean's LANGUAGE MODE (Grosjean 1985a,b, 2001) with a production output-monitor (as in e.g. Levelt 1989). Both of these components find ample support in the psycholinguistic literature, but we argue that since neither, on its own, adequately accounts for the full range of published experimental results, they should be integrated into a single model. One of the predictions of the resulting model is that bilinguals who are at least partially in BILINGUAL MODE, motivated to respond in a particular language and hence to monitor their production, will avoid vocabulary common to their two languages. In the psycholinguistic literature, shared vocabulary items are referred to as cognates, regardless of whether this is by direct inheritance or borrowing. In historical linguistics the term is instead used exclusively to refer to word forms—however dissimilar—inherited directly from a common ancestor. To avoid confusion, we use the novel term DOPPEL for shared word forms in the general psycholinguistic sense, and reserve COGNATE for inherited items.¹

We present the results of an experiment designed to test our lexical production model and in particular the doppel-avoidance prediction. Dutch-English bilinguals completed a task specifically designed to trigger activation of both languages, while requiring a response in English. Response contexts were constrained semantically such that in each

¹ As examples, English *ten* [ten] and Polish *dziesięć* [dʲɛɕɛntɕ] are cognates, but since they are not similar to speakers (excluding perhaps Indo-Europeanists), they are not doppels. In contrast, the translation equivalents English *perfume* and Dutch *parfum* are doppels because of their similarity in meaning and form, but are not cognates since they do not derive via normal transmission from a common ancestor.

case either a doppel or an English-only word form could be used. In comparison to a monolingual English-speaking control group, bilinguals displayed a preference for distinctive vocabulary. We show that the observed degree of doppel avoidance supports our model of production where both output monitoring and language mode are involved, and that, in the long term, output monitoring is likely to result in bilinguals using doppels significantly less frequently than the doppels occur in their environment.

Using the levels of bilingual mode and monitoring inferred from the experimental data, we develop agent-based simulations of the longevity of shared vocabulary in the presence of variously sized communities of bilinguals. We show that even a small bias such as that identified in the experiment can potentially lead to quite rapid lexical replacement, and that the greater the overlap in speech communities the more rapid the loss of shared forms.

2. A BILINGUAL BIAS IN LEXICAL PRODUCTION. In this section, we present and justify a formal model of bilingual lexical production. We start with the simplest possible model of bilingualism—two monolingual systems in one head—and revise this progressively, in the light of insights gained into how language production works in the bilingual mind. The complete model is then tested against experimental evidence in §3.

2.1. MONOLINGUAL LEXICAL CHOICE. The modeling of bilingual production begins with a simple functional model of monolingual lexical choice. Uses of lexical items—and thus their functional meanings—can be inferred to a remarkable extent from their distribution (Baroni & Lenci 2010). In a basic account of production, the distribution of forms uttered for a given meaning is the same as the distribution of forms encountered by the user when that meaning has been inferred. Numerous psycholinguistic studies provide evidence of a frequency effect in lexical production, with more frequent words being faster to retrieve and more-often produced (Oldfield & Wingfield 1965, Rubenstein et al. 1970, Dell 1990, Griffin & Bock 1998, Garlock et al. 2001, Alario et al. 2002). So a model of lexical acquisition should ideally reflect not just the set of meaning-form pairings, but also their frequencies.

Using f for form, s for meaning, and l to denote the language, we express the relative frequency (R) with which a language learner hears a form-meaning combination, in the context of language l , as $R(f, s|l)$. In this minimal model, the probability (in monolingual mode: P_M) of using form f on a future occasion to express meaning s in language l is the ratio of the relative frequency of the combination to the marginal frequency of using that meaning in that language (1).

$$(1) P_M(f|s; l) = R(f, s|l) / \sum_f R(f, s|l)$$

This model does not take into account effects known to have an impact on production frequency, such as collocations and semantic contexts. However, these conditioning factors can be analyzed as components of the functional meaning of the form, by treating them as conditioning the semantic/pragmatic meanings.

2.2. TWO MONOLINGUALS IN ONE PERSON. The simplest model of bilingualism treats the bilingual as two monolinguals in one person. While rarely voiced explicitly, this model is often implicit in discussions of bilingual proficiency, in which monolingual norms are considered a gold standard. Under this account, we can model speech production in a bilingual as the combination of monolingual production engines plus a switch to select the language to be spoken.

In the TWO-MONOLINGUALS-IN-ONE-PERSON MODEL of bilingualism (hereafter the 2M MODEL for brevity), language selection occurs by manipulating the language parameter

to the conditional probability distribution. So, if the speaker attempts to use target language t , then they will produce forms f for meaning s with probability $P_{2M}(f|s; t)$ (the subscript $2M$ denoting the 2M MODEL). Another way of expressing this conditioning is to form a weighted sum of the distributions for each language by a selection coefficient δ_l^t , which is 1 if and only if the target language t and the actual language l from which the word form comes match, that is, $t = l$, and 0 otherwise.

$$(2) P_{2M}(f|s; t) = \sum_l \delta_l^t P_M(f|s; l)$$

Evidence has accumulated that this model provides an insufficient account of language processing in bilinguals. Its greatest shortcoming is that it does not admit any influence from the lexicon of the nontarget language in lexical production. This influence has been observed, however, in many naturalistic phenomena and in experiments (see e.g. Myers-Scotton 2006, de Groot 2011 for summaries). Furthermore, the model is inconsistent with the commonality of loanwords as an outcome of language contact. The two-monolinguals model would instead be more compatible with the unusually low levels of lexical borrowing exemplified by Athabaskan languages such as Montana Salish. In these languages, word forms for novel European concepts have typically been constructed by combining native components, rather than by borrowing (e.g. Montana Salish *p'ip'úyšn* 'automobile', literally 'wrinkled feet'; see Thomason 2001:80)—but such cases are infrequent and generally regarded as exceptional.

2.3. LANGUAGE MODE: THE SHIFTING WEIGHTS OF A BILINGUAL'S LANGUAGES. Due to these shortcomings, a number of studies by Grosjean (e.g. Grosjean 1985a, 1989) have made the case for treating bilinguals not as composite, bipartite language users, but rather as having a unique and specific linguistic configuration of their own. Grosjean's model of the bilingual elaborates the simple switching model presented in the previous section. There is still a switch that selects a target language (or 'base language', in his terminology) from the speaker's repertoire. However, this selection is not all or nothing. There is an additional parameter, LANGUAGE MODE (see e.g. Grosjean 2001), which determines how strictly language selection is applied: when the value of the parameter is MONOLINGUAL MODE, the speaker behaves similarly (as much as is possible) to a monolingual of the target language; when the value of the parameter is at the other extreme, full BILINGUAL MODE, the speaker freely mixes both of their languages in production.² A continuum of variation links these two extremes.³

There are a number of factors that influence selection of the target language and language mode. These include: the languages available to the interlocutor (if the interlocutor shares more than one language with the bilingual, the bilingual is more likely to be in bilingual mode, at least to some degree, for their interactions), languages recently encountered, the formality of the situation, the cultural appropriateness of languages in the situation, and topic.

² It may not be wholly clear whether language mode is about readiness to use a language or actual production of language forms, in other words, whether language mode is a cognitive parameter or a behavioral description. Grosjean (p.c.) confirms that his definition of language mode as '[t]he state of activation of the bilingual's languages and language processing mechanisms, at a given point in time' (2008:39) is intended to include readiness to process (hear, read) speech/written language and not necessarily produce it, and so that it is cognitive. This is indeed also de Groot's (2011:289) interpretation, who expresses the relationship between language mode and behavior in the following way: cognitive monolingual mode brings about monolingual behavior; cognitive bilingual mode brings about bilingual behaviors such as code-mixing or code-switching.

³ Green (1998) suggests an inhibition model of language control. This does not substantially differ from Grosjean's account, as a shift away from bilingual mode toward monolingual mode can be interpreted as increasing inhibition of language B (L_B) while simultaneously increasing activation of language A (L_A).

Formalizing Grosjean’s model requires a simple adjustment to the 2M model; see 3 below. Rather than the bilingual having two monolingual conditional distributions with an either/or switch, in Grosjean’s account, the two distributions are combined in a linear superposition. Once again, we use t for the TARGET LANGUAGE (Grosjean’s BASE LANGUAGE), s for the meaning, and f for the form. There is a new parameter, b , which represents how far the language mode is toward the BILINGUAL end of the continuum: if the speaker is in monolingual mode, then $b = 0$, and if fully in bilingual mode, then $b = 1$.

Monolingual mode is functionally equivalent to the 2M model developed in §3.1: one language is spoken at a time, with no intrusions from other languages. In contrast, full bilingual mode (represented with subscript BM below) means that each language is equally likely to be represented in the output. This full bilingual mode is like the double-monolingual model, but with uniform language coefficients: if a speaker has a set L of languages, then the probability of drawing from the lexicon of any one of them is $\frac{1}{|L|}$. These coefficients constitute a flat probability distribution over the speaker’s languages, and the derived conditional probability distribution over lexical forms, given this totally mixed lexicon, is the $P_{BM}(f|s)$ defined in 3.

$$(3) \ P_{BM}(f|s) = \sum_{l \in L} \frac{1}{|L|} P(f|s; l)$$

Grosjean’s account of bilingualism (G) places production on a spectrum between selective monolingualism (as in the P_{2M} model) on the one hand and free and balanced language mixing (as in the P_{BM} model above) on the other. The conditional distribution $P_G(f|s; t, b)$ in this account—expressing intermediate positions along the continuum—can be modeled as a linear combination of the two extreme models, weighted by the language-mode parameter b , as in 4.

$$(4) \ P_G(f|s; t, b) = (1 - b) P_{2M}(f|s; t) + bP_{BM}(f|s)$$

Here is how this model works in practice. A bilingual in Scottish Gaelic (SG) and English wishes to express the meaning ‘winter’ in SG, while in 70% bilingual mode. The English conditional distribution offers only a single form *winter* for the meaning, while Gaelic offers *geamhradh* and *gamh* (we ignore the synonyms *dùllaigh* and *samhnaich* as they seem marginal at best), with frequencies at likelihoods of 73% and 27%, respectively (based on a Google search). Table 1 shows the expected distribution of output forms according to our formalization of Grosjean’s model, the two-bilinguals model, and the fully bilingual model given above. Figure 1 shows how the frequency of use of each form varies with the language-mode parameter b and language target t .

LEXEME	P_G	P_{2M}	P_{BM}
geamhradh	0.47	0.73	0.365
gamh	0.17	0.27	0.135
winter	0.35	0.0	0.5

TABLE 1. Likelihoods (based on artificial frequency figures) in the three models for three candidates expressing the meaning ‘winter’: two from Scottish Gaelic (SG), one from English.
The Grosjean model values assumed a language mode of 0.7.

One key contrast between Grosjean’s model—when the speaker is in bilingual mode—and the 2M model is its treatment of shared vocabulary, that is, doppelts. In Grosjean’s model, if a form-meaning pair is the same or similar in both of a bilingual’s languages,⁴ then the likelihood of using that form for that meaning is greater than it would be for a monolingual. This is because the probability of using the form in either

⁴ In this article we do not address the issue of how similar two forms need to be to count as a doppel, though it is clear that doppelhood is relative and that what may count as a doppel for one bilingual speaker may not be so for another—for example, for a historical linguist who recognizes patterns of sound correspondence,

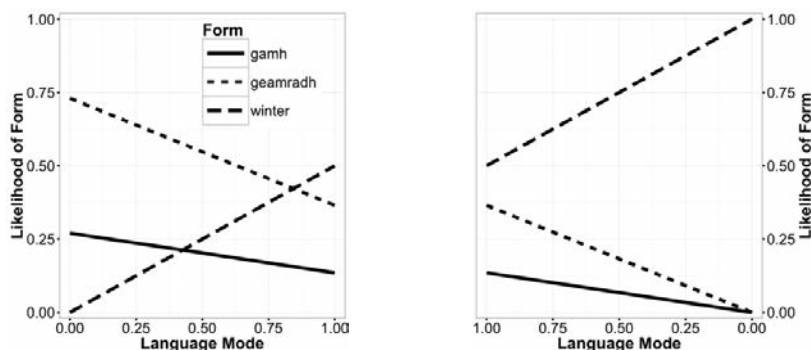


FIGURE 1. Plot of P_G by target language selection (SG left, English right) and language mode. Candidates are associated with only a single language each, and the two SG candidates (solid line and short-dashed line) have a web-derived distribution, while the English candidate (long-dashed line) is without synonyms.

Note that in the right-hand graph, bilingual mode increases right-to-left.

language contributes to the total probability of using the form; for unshared vocabulary, there is no boost for the target language form from the nontarget language.

For example, Dutch *zak* and French *sac* are doppel, sharing the meaning ‘bag’ and a similar pronunciation. Let us presume that in the respective language each form is 80% likely to represent that meaning.

$$(5) P_M(zak|BAG; l = Dutch) = 80\% = P_M(sac|BAG; l = French)$$

Due to their similarity in pronunciation and meaning, these words are categorized as doppel, so the forms are treated as identical in the model. Recall that at language mode 1.0 the bilingual’s languages will be evenly balanced, so the probability of using a Dutch word in bilingual mode b is $(1.0 - b) + b/2$; with $b = 0.6$, this number is 0.7. So in language mode 0.6, the speaker will draw on Dutch for 70% of their lexical selections and French for 30%. In that case, the probability of producing *zak* (or its doppel) is 0.8, through the calculation shown in 6.

$$(6) P_G(zak|s; Dutch, 0.6) = 0.7 P_M(zak|s; Dutch) + 0.3 P_M(sac|s; French) \\ = 0.7 * 0.8 + 0.3 * 0.8 = 0.8$$

So there is an 80% likelihood of saying something like *zak* for this meaning. However, no unshared Dutch vocabulary form can achieve a probability greater than 70% no matter what the meaning, because bilingual mode restricts the chance of drawing forms solely from the target language.

Figure 2 shows how the probability of using a doppel increases relative to the probability of using a nondoppel alternative as language mode becomes more bilingual.

This doppel advantage has been found in a number of psycholinguistic studies (which refer to it as the COGNATE ADVANTAGE or COGNATE FACILITATION). While we have characterized this advantage in terms of probability of realization, these studies look for the doppel advantage in naming latency and lexical-decision timing, among other metrics. For example, Costa and colleagues (2000:1296) show that Spanish-Catalan bilinguals respond more quickly in a picture-naming task when the name to be given is a doppel. Cristoffanini and colleagues (1986) show that Spanish-English bilinguals are able to

many more lexical items will be doppel than for the nonlinguist! Here we restrict our examples and stimuli to forms that are highly likely to count as doppel for the majority of bilingual speakers and leave further investigation of the issue for future research.

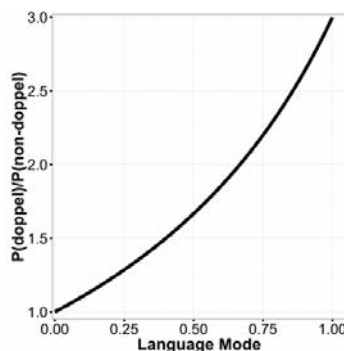


FIGURE 2. Ratio of the probability of choosing a doppel in L_A to that of choosing a nondoppel for meaning s .

L_B is assumed to also have a single form for this meaning—the doppel. The ratio increases from 1.0 as the bilingual shifts from pure monolingual mode (0.00) to full bilingual mode (1.00); in this latter case the bilingual is twice as likely to produce the doppel instead of the nondoppel.

make lexical judgments—decisions on whether a visually presented string of characters is an actual word in English—faster when the stimulus is a doppel with Spanish. Both of these types of facilitation seem likely to be reflected in higher production frequencies for doppels. There are many other studies confirming a special status for doppels (Roberts & Deslauriers 1999, Colomé 2001, Colomé & Miozzo 2010, Szubko-Sitarek 2011).

The existence of doppel facilitation proves that a bilingual's lexica interact. It does not of itself prove that the level of this interaction varies from situation to situation. Grosjean (1997) describes a study showing differences in lexical selection made by French-English bilinguals as they respond to imagined, nonpresent French native speakers who were described as falling into one of the following three categories: (i) a monolingual just arrived in the United States, (ii) a bilingual who has lived in the United States for seven years but uses French at work, at home, and socially (bilingual A), and (iii) a bilingual who has lived in the United States for seven years but uses English at work and a mix of English and French in social and home contexts (bilingual B). Grosjean found that the French words used (operationalized by syllable counts) decreased in speech directed at interlocutors from categories (i) to (iii), while English intrusions (by syllable counts) increased; see Figure 3a. He concludes that these differences reflect movement along the language-mode continuum from French-monolingual mode toward a more bilingual mode combining French with English. The graph in Figure 3b shows how the formalization in 8 below can account for these results by implementing increasing levels of bilingual mode.

2.4. LANGUAGE SELECTION BY MONITOR: MAKING SURE. De Groot offers an alternative explanation for Grosjean's results (de Groot 2011:293; see also Dewaele 2001, de Groot 2002). She argues that, rather than reflecting shifts in language mode, they could equally well be explained by differences in the level of output monitoring that is operating in the speaker. Psycholinguistic accounts of language production have long made use of a monitoring component to explain a number of language behaviors, such as error correction and taboo-word avoidance (Levelt 1989, Dell & O'Seaghdha 1992). The same mechanism has figured in some accounts of bilingual lexical production as well (de Bot 1992, de Bot & Schreuder 1993, La Heij 2005).

In de Groot's monitoring explanation, the French-English bilingual uses a high level of monitoring to catch and avoid English intrusions in order to deal with an interlocutor who is new to the speech community, and so is less likely to be fluent in English (ad-

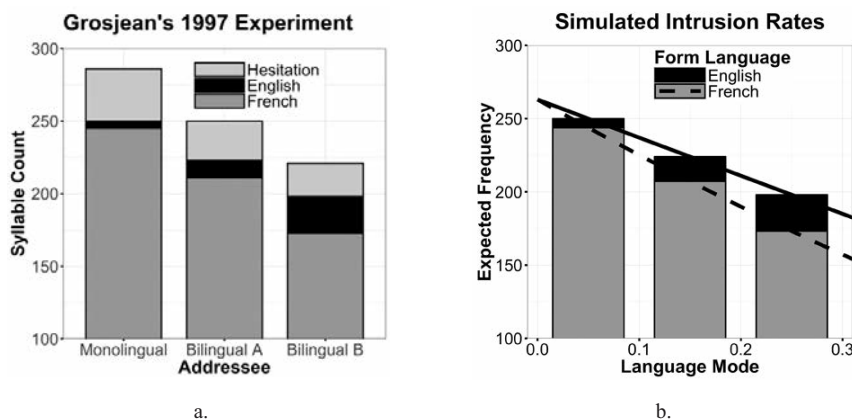


FIGURE 3. (a) Grosjean's experimental results showing differing degrees of intrusion of English (black) against French (gray) in response to interlocutor backstory (graph based on results reported in Grosjean 1997). Graph (b) shows the model predictions for level of intrusion as a function of bilingual mode. The bars show that levels of bilingual mode—0.05, 0.15, 0.25—result in intrusion frequencies matching Grosjean's experimental results. The number of items in the simulated data was scaled to match the syllable counts in the experimental results.

dressee (i) in the Grosjean experiment). In dealing with someone likely to be English-fluent as well as French-speaking, monitoring is relaxed and more intrusions are allowed (addressees (ii)–(iii)). After describing the various required components below, we provide a formalization of the model linked to this explanation.

MONITORING MODEL COMPONENTS. In the 2M model, on which the Grosjean (language mode) model is also based, there can be more than one way of expressing a meaning—for example, synonyms. We model this by having a probability distribution over possible forms associated with a meaning (for that language). In any single communicative moment, however, only one form can be used: that is, the speaker will only select one word from possible options. The distribution, even if never realized except via a single choice, has psychological reality. This psychological reality, underpinning the probability of usage, is known as **ACTIVATION**. Activation is subject to other causal influences, such as **PRIMING** (Meyer & Schvaneveldt 1971, Schvaneveldt et al. 1976, de Groot 1992), in which the activation of one percept or concept triggers the activation of another. For example, forms that are likely to be used in expressing a meaning are said to be **PRIMED** by that meaning.

Languages as a whole can also have a level of activation. In the Grosjean model, language-level activation is determined by the target language and the language mode of the speaker. Formally, this corresponds to the probability $P_T(l|t, b)$ in which the likelihood of use of a particular language l is conditioned by the intended target language t and the language mode b (which is responsive to topic, languages known to the interlocutor, and so on). The level of activation for each language, in turn, conditions which forms are used: the more activated a language, the more likely its forms are to be used to represent meanings. The expression $P_{Act}(f|s; t) = P_M(f|s; l) P_T(l|t, b)$ denotes the level of activation of forms primed by the pattern of activations of meanings and languages.

Activation lets us account for the disparity between context and experience, on the one hand, and linguistic production, on the other. For example, a simultaneous interpreter hearing input in Mandarin and rendering it into French has to deal with an environment awash with both languages—so both must be highly activated—which effectively places

the interpreter at the bilingual-mode end of the language-mode continuum, and yet these professionals manage to produce output in only a single language with scant intrusion. This mismatch between production and environment seemingly conflicts with Grosjean's account of language mode.

In accounts of production that include a monitor, the monitor narrows the choice of candidates from those activated by priming mechanisms. In the model suggested by de Groot, where there is no varying language mode, the forms of all languages will be equally activated or activated in proportion to their frequency in the experience of the speaker. Language selection happens when the monitor filters out all forms identified as not belonging to the target language. The equivalent of bilingual mode in this model is when the monitor is not applied, and so output forms are selected in a mix of languages. Crucially, therefore, this model replaces the continuum of language mode with a continuum of monitoring effort.

In general, lexical-production models involving a monitor must also include generation by a priming component. The priming component determines the probability of forms given a meaning and any other conditioning factors, such as language. In contrast, the monitor takes as input the form and meaning to be expressed, as well as any other parameters, such as monitoring effort. These determine a distribution describing how likely each language is to be the source of the form. For example, the monitor in a French-Dutch bilingual who is expressing the concept 'shoes' with *chaussures* should evaluate French as high probability and Dutch as low probability. But if the bilingual is expressing 'bag' with *sac*, French and Dutch are more closely matched in probability due the phonological similarity of *sac* and *zak*.

We envisage the monitor as not just a language-identification device, but also a filter. It acts to block production of forms whose identified language is not the target language, by modifying the activation of those forms (similar to Berg's 1986 account of active monitoring). In our model, the monitor modulates the probability of using a form for a particular meaning based on how likely it is to belong to the target language.

A number of neurological studies have localized monitoring processes to prefrontal areas of the cortex (Barch et al. 2000, van Veen & Carter 2002, Badre & Wagner 2004, Ridderinkhof et al. 2004, Christoffels et al. 2007, Brown 2013). For example, Petrides (1996), in discussing working memory, states that the mid-dorsolateral frontal region impacts free recall by its on-line monitoring of the output of long-term memory, in contrast to the mid-ventrolateral frontal cortex, which is more directly involved in active retrieval. The combination of free recall, from long-term memory, with monitoring is what we propose here as the mechanism for (mono- and) bilingual lexical production. Following Petrides (1996), then, we expect monitoring activity to be an integral part of production.

We must also consider that there are a variety of social settings for bilingualism. In some settings, there is little pressure to avoid using more than one language within an exchange (see e.g. the discussion of code-switching presented in Meeuwis & Blommaert 1998). In such contexts little effort needs to be allocated to monitoring. In other settings, using the 'wrong' language may be socially inappropriate, and so monitoring is likely to be strongly engaged (e.g. blocking the use of standard German in place of local Swiss German in everyday conversations in Switzerland).

We model code-switching by a conditional probability $P_L(l|f, s; p_m)$ over languages l given a form-meaning pair f, s and monitor parameter p_m . This term expresses the likelihood of a particular language being responsible for the candidate form-meaning pair f, s . Monitoring seeks to enforce the target language in output, so we require the identified language to match the target language ($l = t$), giving a monitoring term $P_L(l = t|f, s; p_m)$.

This term is then applied to the premonitoring probability $P_{PreL}(f|s; t, p_G)$, using Bayes's theorem to fix a posterior postmonitoring probability.

$$(7) P_{PostL}(f|s; t, p_m) = k_{PostL} P_L(l=t|f, s; p_m) P_{PreL}(f|s; t, p_G)$$

The level of monitoring can vary. We represent this variation by taking m to consist of a single monitoring-effort parameter varying between 0 (no monitoring effort) and 1 (maximal monitoring effort). At $m = 0$, the probability estimate P_L that the form is from the correct language has the fixed value of $\frac{1}{|L|}$. When $m = 1$, P_L takes the probability of producing the form from language l with each language as an equally likely source.

$$(8) P_L(l|f, s; t, p_G, m) = m k_L P_M(f|s; l) + \frac{1-m}{|L|}$$

When m is zero, no monitoring is applied, as $P_L(l=t|f, s; m)$ is uniformly unresponsive to the candidates for production. Consequently, the probability of realizing meaning s with particular form f is conditioned only by the effects of premonitoring priming. This models a scenario where speakers freely mix their languages, in response to priming—be it phonological, pragmatic, or other kinds of priming.

At the other extreme, we can imagine speakers exerting a maximum level of monitoring effort. In this case, the language-checking term $P_L(l=t|f, s; m)$ counterbalances differences in activation of the candidates due to priming, and selects for output a form most strongly identified with the target language.

FORMALIZATION OF THE MONITORING-ONLY MODEL. We are now in a position to offer a formalization of de Groot's (represented with subscript *DG* below) monitoring explanation of Grosjean's (1997) experimental results. In this model, bilinguals have a fixed language mode (fully bilingual) and implement language selection in their monitor. We modify equation 7 above to have fixed probability p_l associated with each language l .

$$(9) P_{Act}(f|s; t) = \sum_{l \in L} P_M(f|s; l) p_l, P_T(l|s; t) = p_l$$

This new value is then inserted in 7 and 8 to give a combined probability $P_{DG}(f|s; t, m)$ of expression of a form, given meaning and target language. There are no parameters p_{Act} , p_m to the activation component or the monitor (apart from monitoring effort m).

$$(10) P_{DG}(f|s; t, m) = k_{PostL} (m k_L P_M(f|s; t) + \frac{1-m}{|L|}) \sum_{l \in L} P_M(f|s; l) p_l$$

This formalization of the model successfully accounts for the data in Grosjean's experiment (Figure 4a).

One form of supplementary evidence for the monitoring account (also noticed by de Groot (2011:293–94)) is the number of hesitations encountered by Grosjean in the above-mentioned experiment. These decrease as production becomes more language-mixed. Grosjean interprets the difference in hesitation frequency as a reflection of the difficulty of finding the right word in French when in monolingual mode and of the need for finding an appropriate circumlocution, but de Groot suggests that these hesitations could just as well reflect the repression of non-French vocabulary by the monitor and the delay in expressing a low-activation candidate. Her conclusion is that this evidence does not favor either language-mode or language-monitoring accounts of language selection.

Figure 4b shows the predictions of this monitoring-only model for doppel and non-doppel forms, not investigated in Grosjean's study. At maximum bilingual mode, we can see that while the model predicts an increased probability of expression of the non-doppel candidate as monitoring increases, this probability never exceeds that of the doppel candidate even at maximum monitoring levels. Figure 4c provides comparison

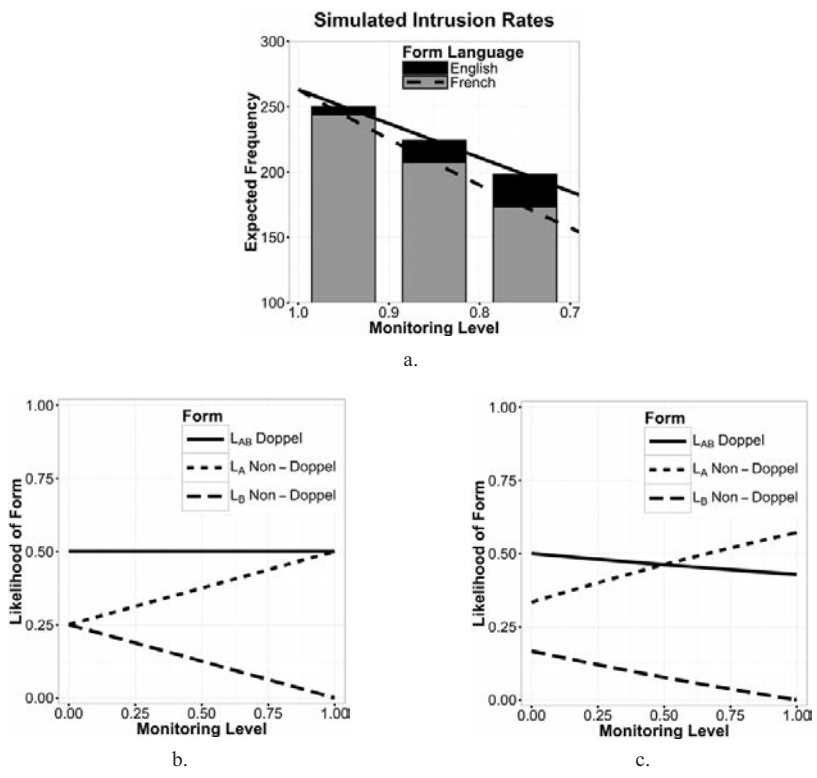


FIGURE 4. These graphs show the outcomes of de Groot’s monitor-only model. In (a), we see the simulation of the number of intrusions versus target-language forms for different levels of monitoring. Comparison with Grosjean’s results (reproduced in Fig. 3a) show that monitoring can account for his data. Panels (b) and (c), however, highlight a potential problem for the monitoring-only explanation. When monitoring is applied to equally activated languages (b), doppels are never deprecated in comparison to nondoppel alternative forms—no matter the degree of monitoring. At best, the doppel and the nondoppel achieve their monolingual relative frequency when monitoring is maximized. Panel (c) shows a different situation, one where monitoring is applied in a state of unequal language activation—67% activation of L_A, 33% of L_B, or a language mode of 67%. Doppel/nondoppel parity is achieved with approximately 50% monitoring, while further monitoring renders the doppel a disfavored option. Section 3 presents an experimental test of this situation.

with a situation where, rather than full bilingual mode, there is 67% activation of the target language and 33% activation of the other. In this case the nondoppel candidate DOES overtake the doppel candidate when monitoring effort exceeds the 50% level. This comparison demonstrates that combining differential language mode and differential monitoring into a single model of lexical production can potentially account for a wider range of speaker behaviors. This is the model that we discuss in the next section.

2.5. BOTH LANGUAGE MODE AND MONITORING. We agree with de Groot’s conclusion that published experimental evidence does not favor either the language-mode or the language-monitoring accounts of language selection. This may be because, in fact, BOTH processes are necessary to fully explain bilingual production. In this section, we argue that experimental evidence best supports a model that includes as essential parameters both language-mode variation and a monitor with variable degrees of strictness.

In the literature, the combined role of language mode and monitoring is most evident in those studies that control for language dominance of the speakers. Few bilinguals are

balanced, or equally proficient in each of their languages. This is a consequence of the fact that languages are often acquired under different conditions and are used for different purposes and extents at various stages of the speaker's life (Myers-Scotton 2006:38–45, Grosjean 2013a:7–11). The typical bilingual is therefore someone who has a dominant language. A study that takes into account participants' language dominance is that by Grosjean (2013b:58), who finds that greater language proficiency helps a bilingual speaker avoid intrusions in a production task. Since speakers are (usually) more proficient in their L1, it follows that more monolingual levels of language mode can be achieved in a speaker's L1 than in their L2. The more balanced the bilingual, the less true this will be—due to a smaller gap in proficiency between L1 and L2.

Given the discussion in §2.2, we should, on the one hand, expect a greater likelihood/activation of doppels while a speaker is using a weaker language (usually L2) rather than their dominant language (usually L1). On the other hand, we would expect that L2 speakers engage their monitors more forcefully to check for language appropriateness than do L1 speakers. One reason for this is that L1 speakers are just less likely to have intrusions from L2 than vice versa, so the need to check for them is decreased (Kormos 1999).

Doppel advantage varies in response to language mode: the more bilingual the language mode, the more doppel advantage we expect to see. However, this doppel advantage varies with the inverse of monitoring level: the less monitoring, the greater doppel advantage retained.

We summarize these expectations for L1 and L2 processing in Table 2.

	LANGUAGE MODE		MONITORING	
	MODE	DOPPEL BIAS	EFFORT	DOPPEL BIAS
L1	more monolingual	weak positive	weaker	weak negative
L2	more bilingual	strong positive	stronger	strong negative

TABLE 2. Bilingual language mode and monitoring are weaker when a bilingual uses their L1. This leads to conflicting predictions about doppel facilitation.

The expectations in Table 2 allow us to disentangle evidence for language mode from evidence for monitoring. Take, for example, Christoffels and colleagues' (2007) examination of doppel advantage in German-Dutch bilinguals. While doppel facilitation in reaction times is not identical with final production probability, it seems unlikely that reaction-time advantage would not vary positively with frequency of production. Christoffels and colleagues (p. 197) report a much higher facilitation of doppels in L2 than L1 in BLOCKED TRIALS in which all items are read in the same language. Similar L2 doppel-facilitation effects have been reported by others (e.g. Costa et al. 2000, Lalor & Kirsner 2000:393, Dijkstra & van Hell 2003, Szubko-Sitarek 2011:202, Poarch & van Hell 2012:426). This accords with speakers exhibiting a more bilingual language mode when using L2 than L1. In contrast, when participants were faced with language shifts between consecutive stimuli, the doppel advantage dropped in L2 production. We would argue that this resulted from the greater monitoring effort expended in order to ensure the correct language was used in this case.

This behavioral evidence for monitoring is supported by additional neural data. Christoffels and colleagues (2007:200–201) find a negative electrophysiological response signal associated with doppels in two time windows: 275–375 ms and 375–475 ms. These may reflect an N2 effect (a primarily frontal wave of negative potential in the 200–350 ms range), which has been identified as a mismatch detector and is also associated with conflict monitoring (Donkers & van Boxtel 2004). This signal is indicative of monitoring in the production of doppels.

Thus there is evidence supporting both a language-mode effect, in terms of the activation of lexical items from a bilingual’s languages, and a monitor which ensures that only candidate forms reflecting the target language are realized. Without both components in the model, it becomes difficult to fully explain the results from Christoffels et al. 2007. We therefore combine these two components into a single model giving an integrated account of bilingual lexical selection, summarized diagrammatically in Figure 5.

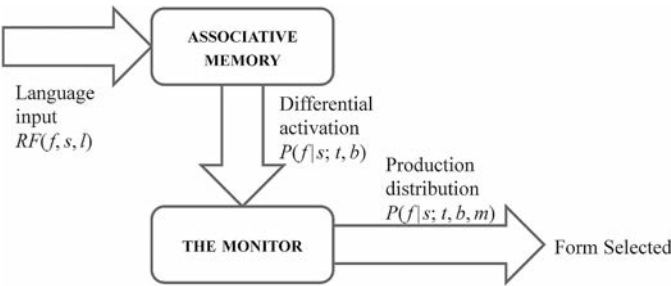


FIGURE 5. Diagram of proposed model structure including both language-mode activation level and a language-appropriateness monitor.

This model, expressed in 11, is a specialization of the general monitoring model (7). The activation component of the model is the one we developed on the basis of Grosjean’s language mode theory, expressed in equation 4. The monitor is the one in equation 8 in §2.4, with a specialized language-identification term (k as always is a normalizing term, varying with parameters s, t, m). The probability $P_C(f|s; t, b, m)$ (subscripted C for COMPLETE) gives the likelihood of producing form f after priming and monitoring.

(11) $P_C(f|s; t, b, m) = k P_L(l = t|f, s; t, b, m) P_G(f|s; t, b)$

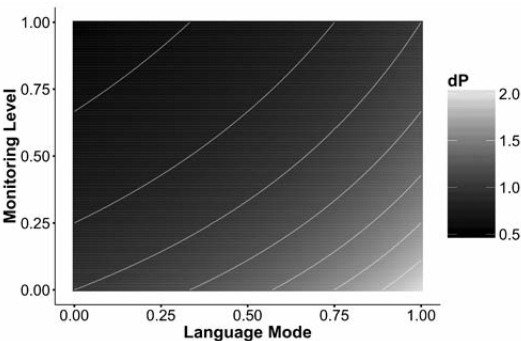


FIGURE 6. A graph of the impact of the language-mode and monitoring-effort parameters. The shading varies with dP —the ratio of the probabilities of doppel and nondoppel in L_A , when each language carries the doppel and one language-specific nondoppel of the same frequency, to express meaning s . Generally, more monitoring leads to a lower likelihood for doppels, while a higher language mode makes them more likely.

Since our model combines Grosjean’s language mode with monitoring, it can result in both facilitation and inhibition of doppels. This is seen in the likelihood ratio of doppels to nondoppels shown in Figure 6. Bilingual mode favors doppels, while monitoring discourages them.

Most previous experimental studies on doppel production test activation through production latency, lexical-decision times, and other similar measures, but did not examine production frequency. We know of no published picture-naming studies that report any

differential results based on the presence, or even possibility, of synonyms in the naming. In fact, some ignore productions that do not conform to a predetermined target form (e.g. Christoffels et al. 2007). The lack of synonym-production information is a major problem in understanding how doppel-related biases might bring about language change. In comparing the two categories of doppel and nondoppel words, one can be favored only if the other is disfavored for the same representational role, and vice versa. If a doppel has no nondoppel synonym, then it cannot be rendered either more likely or less likely by the action of a cognitive bias. This means that much of the psycholinguistic literature on doppels tells us little about whether they are more or less likely to be realized than their competitors for any given meaning.

One researcher has looked at whether doppels are used in translation tasks. Malkiel (2009a,b) argues that, given the ease of learning and using doppels, we should expect them to be the default choice in translation. However, examining the translations of two texts each from forty-five trainee translators, she finds that this is not the case. Instead, regardless of whether the translation is into a dominant or weaker language, the translators frequently preferred the nondoppel even though a translation with a doppel would have been as accurate and appropriate. We contend that the translators are exerting considerable monitoring effort, and both languages are quite highly activated. Consequently, doppels, while receiving the doppel activation boost, may be suffering at the hand of monitoring.

Unfortunately, Malkiel's studies do not carefully control for frequency in order to determine whether the doppels are used less frequently than they would be by a monolingual. It is possible that the doppels were lower frequency than the nondoppels even in the languages at large. To compensate for these shortfalls, we develop a frequency-controlled experiment to identify whether bilinguals whose languages are both activated to some degree avoid using doppels in comparison to monolinguals presented with the same stimuli. This experiment, described in the next section, therefore tests the prediction presented in Fig. 4c: that there will be a bias against doppels when there is some degree of bilingual mode and one language is the target for production.

3. TESTING FOR AN ANTI-DOPPEL BIAS. Any anti-doppel bias in bilinguals must be evaluated in comparison to a baseline of monolingual behavior. A differentiating bias will lower the frequency of doppels in bilingual productions in comparison to their frequency in corresponding contexts by monolinguals.

3.1. METHODS. The study aims to elicit cognitive biases in bilingual lexical-form selection, contrasting these with monolingual outputs. This section details the choices made in the experimental design and the reasons for them. The study received approval from the University of Western Australia Human Research Ethics Committee (RA/4/1/5639).

LANGUAGES. We chose Dutch and English as focus languages because they exhibit many doppels, both cognates and common borrowings. There were also a significant number of Dutch-English bilinguals available in the local community.

PARTICIPANTS. Twenty-four Dutch-English bilinguals and twenty-five English monolinguals volunteered to participate in the study. Both groups answered a questionnaire about their linguistic backgrounds. All of our bilinguals were living in an English-speaking country (Australia) and using English as their primary language of work or study.

STIMULI. Each stimulus consisted of two parts. The first part gave background to the target sentence, constraining the target meaning. In the bilingual condition, this part

was in Dutch; in the monolingual condition, the background text was in English. The second part of the stimulus consisted of a single English sentence frame with a single, noninitial gap.

The background and sentence-frame texts were designed to select a semantics that was as specific as possible, while allowing a choice between lexical forms. For each stimulus, one likely form was a doppel between Dutch and English; one or more likely alternatives were not doppels. Contexts that permit both doppels and nondoppels allow us to observe the differential frequency effects of the experimental conditioning.

DUTCH/ENGLISH BILINGUAL	ENGLISH MONOLINGUAL
Gisterenmiddag ben ik naar het strand geweest.	Yesterday afternoon I went to the beach.
I wanted to take a ____ of the sunset.	I wanted to take a ____ of the sunset.

TABLE 3. Example stimulus as it appears in the bilingual and monolingual conditions. Potential fillers include *photo* (doppel) and *picture* (nondoppel).

An example stimulus for bilingual and monolingual participants appears in Table 3. The two-part structure of the stimulus has a special role in the presentation to bilinguals. It serves to trigger a certain degree of bilingual mode in these participants, as it requires them to switch language systems before and after each background paragraph.⁵

We need to consider whether the switch costs involved in the bilingual condition—and not present in the monolingual condition—could be a confounding factor.⁶ For a survey of switch-cost phenomena, see Declerck & Philipp 2015. The concern is that switching seems to selectively disadvantage doppels and so may, at least in part, account for any results showing a bias against selecting doppels.

There are some crucial differences between our study and those showing greater switch costs in accessing doppels. The latter are isolated word experiments, such as picture naming or lexical decision. The current experiment is more like that of Gullifer et al. 2013, in which whole sentences were presented in a uniform language and participants were asked to name one word highlighted within the sentence. As with our experiment, language switching occurred only between—not within—sentences. Importantly, Gullifer and colleagues find no significant interaction between language switching and doppel status in the reaction time of participants. However, doppels show substantial facilitation in both switch and nonswitch items.

Given the close similarities between stimulus presentation in the current experiment and that in Gullifer et al. 2013, the experiment is likely to engender facilitation of doppel forms, but switch costs are unlikely to bias the results.

PROCEDURE. Participants first completed a questionnaire about their linguistic background. One reason for this was to ensure that they were suitable for the current study. For example, one bilingual participant’s answers left it unclear which language could be considered their L1 (see exclusions below).

A second purpose of the questionnaire was to identify any nonneutral attitudes toward either language. This was achieved by an open-ended request for comments. In

⁵ A referee expresses concern that the presentation-language differences between the monolingual and bilingual stimuli render the conditions incomparable. The target sentences and the responses are in English across both conditions, and so the forms of response are the same across conditions, and the distributions of these forms can be straightforwardly compared. Since our stimuli were solely written—as opposed to pictorial—it was necessary to provide relevant linguistic input in both of the bilinguals’ languages in order to activate their entire linguistic repertoire of pragmatically associated terms.

⁶ We thank a referee for bringing this to our attention.

the context of previous questions, this gave an opportunity for participants to express any motivated language choices that they make.

The study itself consisted of sequential presentation of the entire set of forty-one stimuli in the language-appropriate condition. The first two bilingual participants were administered the study in paper form. The remaining bilinguals and all monolinguals completed the study online. The order of the stimuli was randomized for all online participants.

3.2. RESULTS.

QUESTIONNAIRE. In the open question, no bilinguals revealed attitudinal preferences for either language. Comments focused solely on pragmatic reasons for language choice.

EXCLUSIONS. Where bilinguals self-reported only limited experience in either language, they were excluded from the pool (one subject). One English monolingual was also excluded, since they regularly gave atypical, whimsical answers, for example, *squiz* where others had *photo(graph)* or *picture*, or *fermentation* where others had *pain* or *ache*.

DESCRIPTIVE STATISTICS. The average number of distinct forms returned per stimulus was 5.78 for bilinguals and 5.39 for monolinguals. The use of more forms by bilinguals is backed up by differences in the amount of variation—measured by entropy—in the answers: on average 41.55 bits for bilinguals, and 40.24 bits for monolinguals.

DOPPEL USE. Bilinguals showed consistently less frequent use of doppels than monolinguals. Table 4 shows the distribution of responses for the stimulus given in Table 3 above.

BILINGUALS	MONOLINGUALS
picture (16)	<i>photo(graph)</i> (14)
<i>photo(graph)</i> (7)	picture (11)
view (1)	

TABLE 4. Distribution of answers to the example stimulus (Table 3). Note that the doppel *photo(graph)* (cf. Dutch *foto(graaf)*) is more frequently produced by monolinguals than bilinguals, who opt for the nondoppel alternative.

Figure 7 presents views of the different rates of doppel use among the two groups of participants. In both graphs, the use of doppels can be clearly seen to be higher among the monolingual participants. Of course, a difference in the frequency of doppels is only possible where nondoppel alternatives are available.

STATISTICS. The per-question rates of doppel use are significantly different between the two participant conditions ($t(40) = 4.77$, $p < 10^{-4}$). This significant disfavoring of doppel use by the bilingual participants confirms that doppels are less likely to be selected by a bilingual speaker when a nondoppel synonym is available.

3.3. FITTING THE MODEL TO THE EXPERIMENTAL RESULTS. In §2.5, we presented a probabilistic model of bilingual lexical selection. This model provided a distribution over possible forms for expressing meanings as a function of target language, language mode, and monitoring effort. We used the monolingual experimental results as the likelihood of English inputs to monolinguals and bilinguals. For one stimulus, all monolinguals used a nondoppel, so a frequency for the doppel could not be determined. This item was therefore excluded from parameter fitting. Lacking frequency data for Dutch words in our specific contexts, we assumed a flat distribution over one doppel and one nondoppel

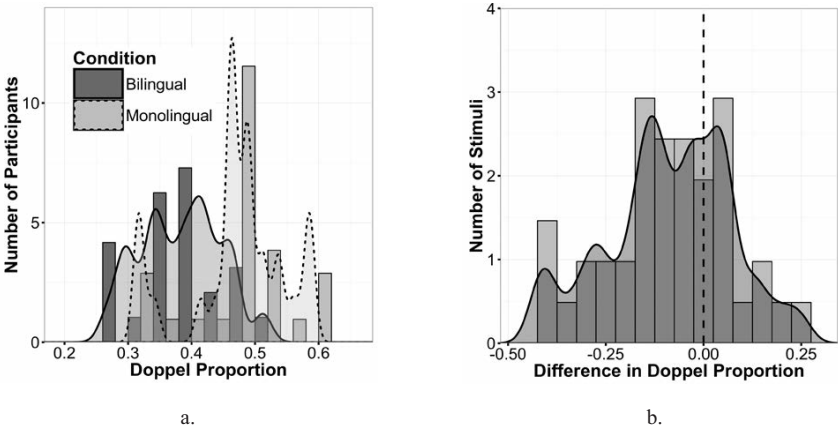


FIGURE 7. The contrasting distribution of doppel use between bilingual and monolingual participants. Histogram (a) shows the different per-participant doppel distributions. The size of the bars shows the number of participants with that doppel rate. The distribution is shifted up for monolinguals (light gray) and down for bilinguals (darker gray). Bar chart (b) shows the difference in frequency of doppel responses per stimulus in each condition. The gray bars show the number of questions with the given difference between bilingual and monolingual doppel percentages. The anti-doppel bias can be seen in the greater area under the curve to the left of the $x = 0$ axis, reflecting the fact that in the majority of questions, bilinguals used fewer doppels than monolinguals.

alternative. We then calculated the information in the results of the bilingual experiment as the negative $\log_{10}$ probability of the posterior likelihood of the data. Figure 8 shows the difference in likelihood of the data in the combined model compared to the two-monolinguals model. At the point of its greatest advantage, the combined model makes the data over a million times more likely than does the two-monolinguals model.

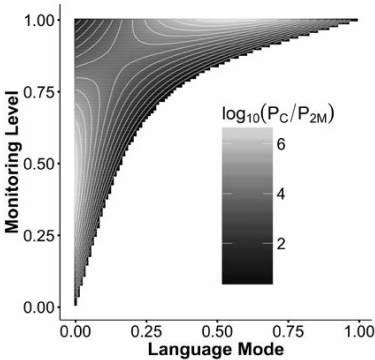


FIGURE 8. This graph shows the fit to the experimental results of the combined model compared to the 2M model, measured as log-likelihood ratios. The best-fit models surpass the 2M model by approximately a factor of 10^6 (light gray). Dark gray indicates parameter combinations where the likelihood of the combined model is closer to that of the 2M model. Where the combined model is less probable than the 2M model, no shading is given.

There were two parameter combinations in which the combined model accounted for the bilingual experimental data best. In the first, language mode was monolingual, and there was a moderate amount of monitoring ($b = 0.0, m = 0.46$). This fit scores highly because it results in few intrusions from Dutch (there were none in the experimental results) and will, to some extent, discourage doppels.

In the second parameter combination ($b = 0.54$, $m = 1.0$), the language mode was mixed (but with the target language favored), and monitoring was intense. Given that the experiment was designed to engender bilingual mode, we take this second parameter combination to be the more likely. The maximal level of monitoring reflects the anti-doppel bias in the data and the lack of intrusions from Dutch.

Note that neither a purely language-mode model (no monitoring, only language mode) nor a purely monitoring model (maximal bilingual mode, only variation in monitoring) produces competitive results.

In summary, the experimental results support the combined model of language production, over the three other models discussed in §2. Crucially, the results reflect the combined action of partial bilingual mode and strong monitoring in an anti-doppel bias.

4. FROM ANTI-DOPPEL BIAS TO LEXICAL REPLACEMENT. In the previous section, we presented evidence of a bias against doppels that arises as a result of cognitive processing in bilinguals. In this section, we show that this bias can both accelerate the rate of macroscopic lexical replacement and target particular language forms.

Agent-based modeling (see e.g. Boero & Squazzoni 2005, Smith & Conrey 2007) allows us to explore how biases modeled in individual speakers, multiplied over a population and projected through time, can account for diachronic language change. Some models are more useful than others in these extrapolations. Cognitive models can be safely assumed to be common cross-culturally and through time, and they therefore offer strong predictions.

The following section describes how an anti-doppel bias impacts the lexical choice between doppel and nondoppel alternatives in a single language.

4.1. THE SIGMOID SIMULATION. Sigmoids (also known as S-curves) have long been recognized as describing the process of language change (see e.g. Greenberg et al. 1954:155 and Weinreich et al. 1968:113). Blythe and Croft (2012) argue that sigmoidal curves can only arise in linguistic contexts if the replicators involved (grammatical or lexical forms) are subject to selection based on their properties. There is evidence from experimental semiotics that replicator-based selection does indeed take place with representational forms (see e.g. Fay & Ellison 2013 and Tamariz et al. 2014). Here, we show that the anti-doppel bias, modeled in the previous section, can give rise to sigmoidal change in the frequency of use of doppels.

In §3 we described the results of a study of lexical choice, comparing the frequencies of doppel use by bilinguals and monolinguals. We found that the best account of the data showed bilinguals in 54% language mode, with 100% monitoring, resulting in a consistent preference for bilinguals to choose the nondoppel. We investigate here what impact a constant preference against doppels can have on the distribution of doppels and synonymous nondoppels over an extended period of time.

This simulation pitched doppels against single nondoppel competitors. Simulated speakers produce new forms according to the parameters best fitting the experimental data. The productions of this simulated speaker then act as the inputs for the next generation of speakers. It is assumed that the nontarget language contains a single doppel-form for the meaning.

As can be seen in Figure 9, even if the nondoppel starts with a frequency one hundredth that of the doppel, over time this ratio can be reversed, effectively eliminating the doppel from use.

The next section follows the impact of the anti-doppel bias on the frequency of doppel use in two-language communities for different proportions of bilingual speakers.

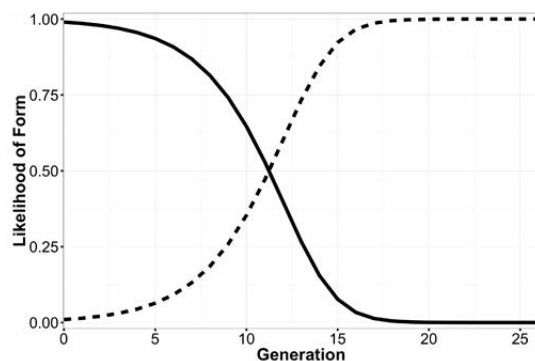


FIGURE 9. Competition between doppel (solid) and nondoppel (dashed) word forms for the same meaning under the action of the anti-doppel bias. The bias leads to progressive replacement of the doppel by the nondoppel alternative.

4.2. THE MULTI-AGENT MODEL. Here we have taken a distributional approach to language in order to unite two aspects of language that are often treated disparately: what individuals produce, and the properties of the language as a whole. In a distributional account, the individual acts of production by speakers together form the frequency profiles of the lexical items and structures that are the defining properties of the language. Language change is the process of modifying those distributions (Zuraw 2003). Understanding the key role of distributions has led to the adoption of new techniques in sociolinguistics for studying language change in progress, such as the use of spoken corpora, multivariate analysis, multidimensional scaling, and linear regression (Wälchli 2010).

This distributional linguistics is not in contradiction to traditional possibilistic linguistics. The frequency with which a form is used can diminish over time—this is a change in distribution, but not a change in the set of lexical forms—until it eventually reaches a frequency so low that new learners of the language do not encounter it, learn it, or use it. At this point, the change in distribution has become a categorical change in the forms available in the language.

In order to achieve a macroscopic change in the distribution of particular lexical forms or structures, the change at the level of individuals must be biased in some direction. If it were unbiased, then as some speakers increase their use of a form, others would likely decrease theirs, and the total frequency of the form would not necessarily vary (but might change due to random selection effects). Where there is a consistent bias in individuals, such that they use a particular form or structure more than they are exposed to it themselves, then this will necessarily lead to a progressive increase in the frequency of this form or structure. The bias need not be uniform, but needs to be sufficiently consistent that there is a net trend in the biased direction.

Our macrolinguistic model brings together a population of agents, each implementing our microlinguistic model (§2) as either monolinguals or bilinguals, arranged to interact in ways reflecting particular linguistic demographics. The model adds three further assumptions. The first assumption is the presumption that the fraction of speakers in each language who are bilingual is a constant proportion and does not change over time. This is not a necessary feature of the model, but it is true of all three cases that we look at here. There are two reasons for this: it is the most tractable assumption to make, and we are also interested specifically in the long-term effects of ongoing, stable bilingualism.

The second assumption is that only the most recent generation provides input to the next generation during their lexical acquisition phase. This assumption is made to simplify computation, but it also accords with the well-known iterated learning model of language transmission (see e.g. Smith et al. 2003).

The third assumption concerns lexical structure. We assume that there is only a single meaning to be communicated, but the lexica of both languages initially offer two lexical forms for this meaning. Of crucial importance is that one of the lexical forms is common to the two lexica. This means that initially, if selected at random, half of the lexical-form instances will be doppel. Multiple synonyms were chosen for simulation because in languages where we most expect to find this kind of long-term bilingualism and consequent divergence (e.g. Australian indigenous languages), it is not uncommon to have many synonyms (see Boretzky 1984).

To be a doppel means that a word form occurs in both languages. If a word form becomes progressively less frequent, to the point where it falls below a frequency of 1%, then it is regarded as unsustainable and presumed lost from the language.

The lexicon is transmitted from one generation to the next in the following way. The distribution over lexical forms that is associated with the current generation is used to randomly generate 100 lexical-form instances per agent. Collectively, these instances define the relative frequencies of items according to the distribution associated with the new generation.

Doppel forms have a reduced impact on the distribution associated with the new generation. For each doppel instance, the proportion of bilinguals in the population determines the likelihood that this instance will be recognized as a doppel, and so be subject to the anti-doppel bias. If it is recognized as a doppel, then the weighting assigned to this instance in the distribution over lexical items for the new generation will be devalued by the anti-doppel bias rate, and the difference assigned to a random nondoppel alternative.

We assume that the degree of bilingual mode and language monitoring remains constant.

4.3. SIMULATIONS AND RESULTS. Figure 10 shows the results for three distinct two-language simulations. The anti-doppel bias rate assumed in the simulations with bias is that derived from the experiment: bilinguals in 54% bilingual mode and using 100% monitoring, and so less likely to use a doppel than monolinguals. The first simulation (Fig. 10a) assumes no bilinguals in either language community. Consequently, the only change in lexical distribution results from drift in the sense in which it is used in evolution theory, that is, unbiased change due to the limited sample size and the associated error. Eventually, these languages can show a gradual loss of doppel, as chance alterations of the distribution increase the odds of selected instances not being shared lexical forms, but this takes many more generations than shown.

The second simulation (Fig. 10b) assumes that one language community is a proper subset of the other. This means that all speakers of language L_B are bilinguals, also speaking language L_A . However, many L_A speakers—in this simulation 50%—do not speak L_B . A real-world parallel is the situation in Spain. All persons in Catalonia, school age and above, are bilingual in Catalan and Spanish. However, only 17% of Spain's Spanish speakers are bilingual in Catalan. In this simulation, L_B 's speakers rapidly decrease their use of doppel. At a slower pace, L_A is also seen to lose doppel frequency. The slower pace is because the doppel available in L_A are dropping from the lexicon in L_B and so have less impact as doppel.

The final simulation (Fig. 10c) models balanced overlapping speech communities. Half of the speakers in each community are bilingual in the language of the other. This

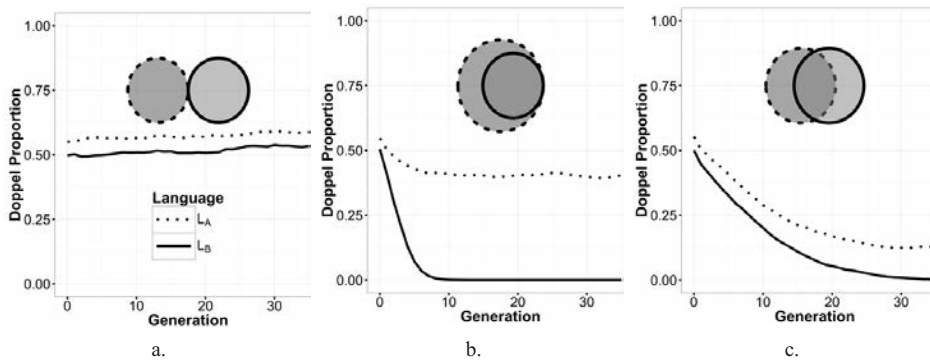


FIGURE 10. Graphs showing doppel ratios from simulations (thirty-five generations, 100 agents per generation, 100 samples per agent) of three scenarios: (a) distinct communities for L_A and L_B , with no bilinguals, (b) all speakers of L_B bilingual in L_A (forming half of its speech community), and (c) both languages having an equal proportion (50%) of bilinguals. In the initial state L_A has the doppel at 55%, while L_B has it at 50%, so that the lines are visible separately. With no bilinguals (a), there is only slow, random change in relative frequency. In the asymmetrical condition (b), the language without monolingual speakers loses its doppel rapidly. Once the doppel is lost from L_B , the anti-doppel bias no longer drives its counterpart's decline in the L_A , so that after an initial decline, its frequency levels off. With equal proportions of bilinguals (c), both languages show a consistent decline in doppel use, until it is lost in L_B .

kind of language relationship is likely to arise in the interaction of, and at the borders between, two small language communities. Real-world examples of this kind are difficult to find. This is partly because, on the one hand, colonization and nation building result in bilinguals combining a local language with a national or international language, leading to subset relationships between language communities, as examined in the second simulation. On the other hand, there may be many instances of balanced overlapping speech communities, but because they do not involve a major language, they are not well documented. In any case, this scenario describes a logically possible contact relationship between languages, one that may offer an explanatory account for historical phenomena occurring before the creation of national and world languages.

Figure 10c shows a rapid fall in doppel use in both languages. Simulated generations in both communities avoid doppel and so gradually decrease the frequency of these lexical items, leading eventually to their loss.

In this section we have explored the macrolinguistic consequences of the microlinguistic model developed on the basis of the experimental study. In particular, we see that a consistent bias can have a major effect on the historical development of a language when magnified by repeated application. The most important implication here is that an anti-doppel bias can lead not only to distributional change in lexical use, but also to an eventual loss of shared lexical forms.

5. DISCUSSION. In the preceding sections, we have examined a linguistic and cognitive model of lexical selection in bilinguals, and looked at the predictions of the model across time and populations. Here we explore some of the implications of an anti-doppel bias for accounts of language processing, language change, and phylogenetics.

5.1. COMPETITIVE VERSUS SINGLE-ANSWER BILINGUAL PRODUCTION EXPERIMENTS. Our experimental results are potentially at odds with the bulk of the psycholinguistic literature, which finds doppel facilitation rather than an anti-doppel bias. Nozari and colleagues (2011) argue that monitoring is not a function of error detection per se, but rather of competition between alternatives. This experimental result is matched in our

model by the role of monitoring as a multiplier on the probability of a candidate (equation 10). If there is only one candidate, normalization of the resulting probability distribution eliminates any impact that monitoring might have.

Many of the experiments that show doppel facilitation are picture-naming tasks. These tasks usually seek to elicit only a single target response, and alternative responses if they do occur are dropped from analysis as ‘errors’. To the extent that an experimental design achieves unambiguous priming of responses, to this same extent it eliminates conflict between alternative candidates and thus vitiates any role for monitoring.

With the experiment described in §3, we have demonstrated that monitoring leads to an anti-doppel bias. In order to confirm that the above explanation for doppel facilitation is correct, we plan to develop a picture-naming task where some of the stimuli trigger competing potential responses: one doppel and one or more nondoppels. Following Christoffels et al. 2007, we will also use electroencephalography (EEG) to relate our results to current thinking on the neurological and temporal localization of monitoring processes. Our hypothesis would be that in the absence of competitors, doppels would be facilitated, but in the presence of competitors, bilinguals would use doppels less frequently than monolinguals.

5.2. STRUCTURE VERSUS THE LEXICON. In this article, we have focused exclusively on lexical selection. We have shown an impetus in bilinguals for lexical differentiation. We would not, however, expect these results to be paralleled by divergence of structure. One argument for the difference comes from a lack of monitoring. François (2011:224) notes that lexical selection is more conscious than structural selection. This lack of ‘consciousness’ refers in part to speakers’ difficulty in monitoring for structure (Silverstein 1981 offers early commentary on the linguistic awareness of native speakers). As discussed in §2.3, an increase in crosslinguistic priming leads to higher probabilities for shared forms, while conflict-triggered monitoring leads to a bias against shared forms. The same outcomes would also apply to structures, but while structural priming, even crosslinguistic structural priming, is well known (Branigan et al. 1995, Pickering & Branigan 1998, Hartsuiker 1999, Salamoura & Williams 2006), as François suggests, monitoring for structure is difficult.

Further support for the apparent different trajectories of form versus structure in language contact, and how these may be linked to ease/difficulty in monitoring, can be found in the work of Matras (e.g. Matras 1998, 2007). His explanation for the high vulnerability of utterance modifiers in contact-induced change rests on their embedding in the ‘grammar of directing’. He argues that, due to their gesture-like, discourse-regulating function, their usage is highly automatized, making them more likely to escape the bilingual speaker’s control in instances of processing overload (Matras 1998:324–26). The more lexical the properties of an item and the less automatized the usage, the easier for the speaker to monitor its form. He concludes that it is therefore likely that lexical and grammatical borrowing ‘follow different rules’ (Matras 1998:326).

Consequently, we expect that lexical divergence will frequently cooccur with structural convergence in communities where there is long-term bilingualism and the languages, at least initially, share a lot of doppels—which is indeed what we see in the Northern Vanuatu case.

5.3. DIAGNOSING LONG-TERM BILINGUALISM. It is possible that this kind of mismatch between the evolution of structure and lexicon may constitute the kernel of a diagnostic for a history of long-term bilingualism between languages. As already noted, cases of extreme structural convergence, while lexicons are either held distinct or diverge, have

been observed in a number of contexts in different parts of the world (Weinreich et al. 1968, Gumperz & Wilson 1971, Nadkarni 1975, Thurston 1987, Enfield 2001, 2003, Ross 2001, 2007, Aikhenvald 2002, Epps 2007, François 2011). The observed structural similarity is extremely unlikely without either widespread bilingualism leading to convergence, or shared inheritance. In contrast, lexical dissimilarity suggests either lack of recent shared inheritance or rapid divergence due to widespread bilingualism. The presence of both structural similarity and lexical dissimilarity requires the intersection of potential causes: widespread stable bilingualism where monitoring is actively engaged (see Table 5 below for a summary of broader patterns and their likely causes).

To operationalize the diagnosis of these patterns, we need comparable measures for both lexical and structural similarity. We are exploring probabilistic methods—such as Markov models—as a way of gaining a common yardstick for both types of similarity. This will then enable us to identify and assess the degree to which there is a mismatch and the likelihood of long-running stable bilingualism in the past.

	FORM SIMILARITY	STRUCTURE SIMILARITY
PATTERN (a): widespread bilingualism & monitoring	less	more
PATTERN (b): contact but little bilingualism	more	less
PATTERN (c): little or no contact	balanced	balanced

TABLE 5. A summary of form/structure comparative patterns and their likely causes. Pattern (a) is the one we have focused on in this article. Pattern (b) may result from situations where one of the languages has higher prestige or power and there is lexical borrowing but little active bilingualism. Pattern (c) is expected when languages have undergone little contact-induced change. In situations of widespread bilingualism where monitoring is little engaged, there will be substantial language mixing and the outcome may be a mixed code.

5.4. SOCIAL PRESSURES AND THE ANTI-DOPPEL BIAS. According to our cognitive model (§2.5), the anti-doppel bias results from monitoring aimed at avoiding crosslinguistic intrusion in situations where both languages of a bilingual are activated to some degree. The intensity of this monitoring reflects the importance the speaker places on the avoidance of such intrusions. Two baseline levels of monitoring suggest themselves: no monitoring and pragmatic monitoring. In the first of these, the speaker does not mind which language s/he speaks. This option is functional only when the speaker’s bilingualism is matched by everyone, or at least most people, in the community (e.g. demographics simulated in Fig. 10c). In this case, there is no PRAGMATIC reason to ensure that productions are in one language or the other, and so there is no PRAGMATIC motivation to monitor.

In the second baseline, a sizable portion of the speaker’s broader community understands only one of their languages (e.g. demographics simulated in Fig. 10b). Consequently, it is pragmatically unhelpful to use the other language in many of the exchanges. In this situation, there is likely to be a strong pragmatic motivation to monitor, so as to avoid incomprehension on the part of interlocutors.

For the bilinguals taking part in our experiment, much of their interaction is with people who do not share both of their languages. Consequently, their motivation to monitor is at least as strong as the pragmatic baseline mentioned above. However, there is no obvious stigma or social approbation associated with speaking Dutch in Western Australia, and so it is unlikely that social pressures go very much beyond the pragmatic motivation to monitor.

Of course, this is not the case in all social situations. In some communities, people are bilingual in the same languages, so the no-monitoring option is a pragmatic possi-

bility. However, social mores about what language is appropriate to speak—when, where, and to whom—go beyond the pragmatic and motivate a strictness of language choice that is cultural. For example, Stanford (2009) reports on the social approbation that follows persons in the Sui community who do not maintain the dialectal features of their village of origin. Similarly, Epps (2007) describes a shared regional perspective in the Vaupés area of Amazonia linking language and identity that prompts negative reactions to obvious instances of language mixing, such as the use of lexical forms from the other language.

Social pressures of these kinds, we suggest, result in greater monitoring effort than would be required for pragmatic purposes alone. Consequently, we would expect to see even greater bias effects against doppels between related languages and against the borrowing of word forms.

5.5. CONTACT AND LANGUAGE RELATEDNESS. Recent literature has posed the question of whether contact between related languages is qualitatively different from contact between distantly related or unrelated languages (see the 2013 special issue of the *Journal of Language Contact*, edited by Patience Epps, John Huehnergard, and Na'ama Pat-El). Our results suggest that contact between closely related languages may indeed be different from contact between distantly related languages. Languages related closely enough to have many cognates as doppels will be subject to the anti-doppel bias in bilinguals and therefore will undergo rapid lexical divergence. If the languages in contact are unrelated, then monitoring and the anti-doppel bias cannot steer replacement since there are few shared words, but the same monitoring mechanism may account for slower-than-expected rates of borrowing, as observed, for example, in Amazonia (Aikhenvald 2002, Epps 2007).⁷ Thus, we would argue, the dynamics of lexical replacement depend crucially on the degree of relationship between the languages in contact—in one case monitoring may speed up replacement, and in the other it may slow it down.

It is worth noting that our approach of relating individual cognitive behavior to language change is presaged in discussions in the special issue of the *Journal of Language Contact* mentioned above. Bower (2013:413) suggests that the question of whether language relatedness matters in contact could be answered by looking at whether genetic relatedness matters in the mind of the bilingual. She does not pursue this approach, however, arguing that generalizing from the behavior of individuals to the impacts of contact seen in the linguistic record is difficult and that if one's primary goal is to understand the results of contact in the linguistic record, this approach would not be fruitful. We disagree.

5.6. PUNCTUATED LEXICAL CHANGE. If it is in contact with related languages that we are most likely to have a significant number of doppels, then it is in this situation that we are most likely to see an anti-doppel bias and, consequently, rapid lexical divergence. If there are no or few related languages available for contact, there will be little lexical differentiation, and so a slower rate of lexical replacement. If a language is in contact with a number of related languages, over a long period of time, then it will face more pressure to differentiate, and so end up with fewer reflexes of ancestral lexical items than a language that has had few relatives to interact with.

This implication of a connection between the presence of related languages and the rate of lexical replacement has been borne out in a recent study by Atkinson and col-

⁷ We thank a referee for highlighting this.

leagues (2008). They looked at the number of branchings intervening between proto-language and modern-day descendants in a number of language families, and at how many proto-language lexical items have been retained. They found that languages resulting from more branchings retained fewer lexical forms from the proto-language. We interpret this as evidence supporting our contention that having fewer related languages limits the opportunities of a language engaging in rapid lexical differentiation.

This interpretation is different from, but compatible with, that offered by Atkinson and colleagues (2008) themselves. They suggest that the correlation between lexical-replacement rate and number of branches results from punctuated equilibrium. When languages are in the process of splitting into two or more descendants, lexical replacement happens the fastest. This coheres with our model of differentiation if we make the not-unreasonable assumption that languages are most likely to be close in both space and social ties around the time that they first split, and most prone to change resulting from the anti-doppel bias as they share much of their lexicon. Whether rapid differentiation is due to the number of related neighbors or to a short, intense period of bilingualism and contact at the point of splitting, the correlation between the number of related languages and the loss of proto-language lexical forms fits what we would expect from lexical differentiation due to the anti-doppel bias. It is possible that the apparent conservatism of structure (as reported in, for example, Dediu & Levinson 2013) in contrast to the lexicon is at least partly due to in-family contact leading to reinforcement of shared structure, but promoting divergence in the lexicon.

5.7. PUNCTUATED LEXICAL CHANGE AND DIXON 1997. In their article, Atkinson and colleagues (2008) refer to Dixon's essay, published in 1997, which also proposes that languages do not evolve gradually throughout the course of their history, but rather that in their evolution there will be short periods of rapid change and long periods of equilibrium. A detailed review of Dixon's adaptation of PUNCTUATED EQUILIBRIUM is beyond the scope of this article (see e.g. Kuteva 1999, Bower 2006 for further discussion). We nevertheless wish to highlight that our model of lexical change, despite being in agreement with the general idea that there will be periods of more rapid change (periods of contact with related languages) and periods of gradual drift (when a language is not in contact with relatives), makes different predictions. Dixon (1997:15) writes: 'If two languages are in contact—some of the speakers of each having a degree of competence in the other—they are likely to borrow lexemes, grammatical categories and techniques, and some grammatical forms ... and gradually become more similar'. We contend, however, that there is an important distinction between what happens to lexical forms and to structure (since monitoring is more effective on the former), with periods of bilingualism giving rise to accelerated lexical divergence even if they lead to structural convergence. So while the languages may be converging in many respects, we predict that the average similarity of lexical forms will decrease as doppels are preferentially replaced—in one language or the other—by less similar forms.

5.8. DOPPELS AND ESTABLISHING LANGUAGE RELATEDNESS. As mentioned above, the typical case where we find an anti-doppel bias is between closely related languages with overlapping speech communities. This is because the preconditions include not only a sizeable proportion of bilinguals, but also of doppels. Closely related languages can be expected to share many cognates, having had little time for lexical replacement by drift. They can also be expected to have undergone few sound changes in the short time since separation. Thus many cognates will also be doppels. Consequently, closely related languages will have many potential targets for an anti-doppel bias. In contrast, if

there have been a lot of sound changes, cognates may not be doppels—that is, the word forms of the cognates may be quite dissimilar between languages, such as the English [æpl] and Polish [jabłko] cognates for ‘apple’, or the Greek [duo] and Armenian [jerku] cognates for ‘two’. If there are few doppels, the bias may occur but not affect a significant proportion of the vocabulary, and therefore leave little visible impact on the lexicon.

If there are many doppels, the effect on the lexicon will be proportionately large, and many lexical items may be affected and/or lost. In situations of common inheritance, these doppels are also cognates. Thus the loss of doppels reduces the evidence available to establish sound correspondences and build a case for language relatedness. This reduction does not eliminate the possibility of applying the comparative method, but reduces the certainty with which results can be stated, due to the smaller evidential basis.

Traditional lexicostatistics and modern phylogenetics frequently rely on the use of cognates as shareable characters to establish internal relationships within language families (see e.g. Gray et al. 2011). A bias against doppels is most likely to affect cognates, and so means that there will be fewer characters that can support particular phylogenies. Furthermore, those methods that rely on cognate counts as direct evidence of language relatedness (i.e. more traditional lexicostatistics)—and assume that the presence of more cognates means a more recent common ancestor—will be rendered invalid where there is, or has been, substantial overlap in language communities.

Although the key elements of subgrouping and reconstruction in traditional historical linguistics do not rely on the specific counts of cognates—relying instead on patterns of sound correspondences identified—they are not immune to the problems created by a paucity of cognates. With insufficient cognates, patterns of sound correspondence cannot be established, and consequently, familial relationships cannot be determined.

5.9. THE ANTI-DOPPEL BIAS AND THE INDEPENDENCE OF LEXICAL REPLACEMENT. The anti-doppel bias causes another problem for probabilistic tree-building on the basis of cognates. This problem is that the rates of lexical replacement of doppels in related languages—separate branches in the language tree—are not independent. Consequently, the likelihood of a particular combination of replacement events cannot be evaluated by simply taking the product of the probabilities of those events separately.⁸

This can be illustrated with an abstract example. Suppose two languages A and B have undergone little sound change from their common ancestor. Consequently, what cognates they share are doppels. If language A replaces its lexical form for a doppel with meaning X at time *t*, then the doppel’s realization in language B will be less likely to suffer subsequent lexical replacement than it was before time *t* (see Fig. 10c above). This is because that form is no longer part of a doppel, and so is no longer subject to the anti-doppel bias. From this point on, it is no more likely to suffer replacement through drift than any other nonshared, nondoppel lexical form.

Another interesting consequence is that the new form in language A is also less likely to suffer replacement than the shared form was. So replacement of one doppel form in language A leads to a decrease in the overall rate of lexical replacement in both languages. While this is not necessarily a problem for phylogenetic modeling that allows variable rates of replacement (see e.g. Greenhill et al. 2010), it does mean that a different prior distribution over rates of change is required. As languages in contact lose

⁸ Greenhill and colleagues (2009) have shown success in coping with other forms of interdependence, namely borrowing. While the anti-doppel bias similarly affects the independence of cognate counts across languages, the impact of borrowing versus the anti-doppel bias on individual lexical items is different.

shared vocabulary, the prior distribution over rate of replacement should increasingly favor lower rates, approaching an asymptotic level of replacement through drift.⁹

5.10. DISRUPTING SUBGROUPING. The disruption that the anti-doppel bias can cause to identifying the relationship between languages can spill over into the already contentious task of subgrouping.¹⁰ Subgrouping is the determination of the internal phylogenetic relationships within a language family. In order to be consistent with a strict, deterministic tree model, languages should be assigned to a nested set of subgroups, none of them properly overlapping.

Ellison & Miceli 2015 describes how an anti-doppel bias can lead to difficulties in subgrouping. It uses an iterated Markov model (Keilson 2012) reflecting the action of an anti-doppel bias to show the likely rates of retention of characters in three descendants (A, B, and C) of a common ancestor. If A and B undergo accelerated lexical divergence due to an anti-doppel bias, then the result will be that A and C will share many characters not shared with B, and likewise B and C will have many forms in common not shared with A. Without other evidence showing that these are retentions and not innovations, the evidence would (misleadingly) support an account with two overlapping subgroups: A–C, and B–C. This conclusion conflicts with cladistic models in general, as the subgroups properly overlap. This type of situation is commonly observed, being described as DIALECT/LANGUAGE CONTINUA or LINKAGES (Ross 1988).

Allowing wave models, or re-entrant trees, in the explanation of the language family does model the shared innovations. However, this explanation of the data, while consistent, does not reflect the history that generated the data. The correct explanation is one in which languages undergo accelerated divergence.

6. CONCLUSION. The primary conclusion of this article is that monitoring in bilingual lexical selection results in a cognitive bias against shared word forms, in turn leading—in bilingual populations—to rapid lexical divergence. We argued for a combined model integrating lexical activation with monitoring for language. Experimental evidence supported this combined model, with best-fit parameters showing a high level of language monitoring in bilinguals. This evidence specifically showed the action of an anti-doppel bias, as the bilinguals used forms shared with their other language less frequently than monolingual participants.

The parameters of the model that best fit the experimental data were used as the basis for a number of multi-agent simulations. These showed that repeated bilingual transmission with the anti-doppel bias results in the rapid loss of shared vocabulary—compared to rates of change due to sampling error (drift) alone. The dynamics of this loss follows—in ideal cases—the well-known S-shaped curve for language change. The rates of change and the nature of the outcome depend on the demographic relationship between the two language communities.

We propose monitoring as a mechanism by which social pressures are implemented in individual behavior. The anti-doppel bias arises only when language monitoring is

⁹ We appreciate that variable replacement-rate in phylogenetic models can go some way toward alleviating complications due to an anti-doppel bias.

¹⁰ While computational phylogenetics relies heavily on lexical data, historical linguists traditionally build family trees on the basis of a variety of innovations, especially ordered sound changes and morphological changes. Nevertheless, a sizable set of shared word forms is also necessary, since establishing regular sound correspondences requires recurrence across a number of items. It is on the basis of these correspondences that sound changes are then determined and that the cognacy of phonological material, including forms that fill morphological paradigms, is assumed. See Miceli 2015 for further discussion.

applied, and the motivation for individuals to monitor their speech may be at least partly socially conditioned. We suggest that the anti-doppel bias is a key link in the causal chain from social forces to language change.

REFERENCES

- AIKHENVALD, ALEXANDRA Y. 2002. *Language contact in Amazonia*. Oxford: Oxford University Press.
- ALARIO, F.-XAVIER; ALBERT COSTA; and ALFONSO CARAMAZZA. 2002. Frequency effects in noun phrase production: Implications for models of lexical access. *Language and Cognitive Processes* 17.299–319. DOI: 10.1080/01690960143000236.
- ARNAL, ANTONI. 2011. Linguistic changes in the Catalan spoken in Catalonia under new contact conditions. *Journal of Language Contact* 4.5–25. DOI: 10.1163/187740911X558815.
- ATKINSON, QUENTIN D.; ANDREW MEADE; CHRIS VENDITTI; SIMON J. GREENHILL; and MARK PAGEL. 2008. Languages evolve in punctuational bursts. *Science* 319(5863).588. DOI: 10.1126/science.1149683.
- BADRE, DAVID, and ANTHONY D. WAGNER. 2004. Selection, integration, and conflict monitoring: Assessing the nature and generality of prefrontal cognitive control mechanisms. *Neuron* 41.473–87. DOI: 10.1016/S0896-6273(03)00851-1.
- BARCH, DEANNA M.; TODD S. BRAVER; FRED W. SABB; and DOUGLAS C. NOLL. 2000. Anterior cingulate and the monitoring of response conflict: Evidence from an fMRI study of overt verb generation. *Journal of Cognitive Neuroscience* 12.298–309. DOI: 10.1162/089892900562110.
- BARONI, MARCO, and ALESSANDRO LENCI. 2010. Distributional memory: A general framework for corpus-based semantics. *Computational Linguistics* 36.673–721. DOI: 10.1162/coli_a_00016.
- BERG, THOMAS. 1986. The problems of language control: Editing, monitoring, and feedback. *Psychological Research* 48.133–44. DOI: 10.1007/BF00309161.
- BLYTHE, RICHARD A., and WILLIAM CROFT. 2012. S-curves and the mechanisms of propagation in language change. *Language* 88(2).269–304. DOI: 10.1353/lan.2012.0027.
- BOERO, RICCARDO, and FLAMINIO SQUAZZONI. 2005. Does empirical embeddedness matter? Methodological issues on agent-based models for analytical social science. *Journal of Artificial Societies and Social Simulation* 8(4):6. Online: <http://jasss.soc.surrey.ac.uk/8/4/6.html>.
- BORETZKY, NORBERT. 1984. The Indo-Europeanist model of sound change and genetic affinity, and change in exotic languages. *Diachronica* 1.1–51. DOI: 10.1075/dia.1.1.03bor.
- BOWERN, CLAIRE. 2006. Another look at Australia as a linguistic area. *Linguistic areas: Convergence in historical and typological perspective*, ed. by Yaron Matras, April McMahon, and Nigel Vincent, 244–65. Berlin: Springer. DOI: 10.1057/9780230287617_10.
- BOWERN, CLAIRE. 2013. Relatedness as a factor in language contact. *Journal of Language Contact* 6.411–32. DOI: 10.1163/19552629-00602010.
- BRANIGAN, HOLLY P.; MARTIN J. PICKERING; SIMON P. LIVERSEDGE; ANDREW J. STEWART; and THOMAS P. URBACH. 1995. Syntactic priming: Investigating the mental representation of language. *Journal of Psycholinguistic Research* 24.489–506. DOI: 10.1007/BF02143163.
- BRAUNMÜLLER, KURT; STEFFEN HÖDER; and KAROLINE KÜHL (eds.) 2014. *Stability and divergence in language contact: Factors and mechanisms*. (Studies in language variation 16.) Amsterdam: John Benjamins.
- BROWN, JOSHUA W. 2013. Beyond conflict monitoring: Cognitive control and the neural basis of thinking before you act. *Current Directions in Psychological Science* 22.179–85. DOI: 10.1177/0963721412470685.
- CHRISTOFFELS, INGRID K.; CHRISTINE FIRK; and NIELS O. SCHILLER. 2007. Bilingual language control: An event-related brain potential study. *Brain Research* 1147.192–208. DOI: 10.1016/j.brainres.2007.01.137.
- COLOMÉ, ÀNGELS. 2001. Lexical activation in bilinguals' speech production: Language-specific or language-independent? *Journal of Memory and Language* 45.721–36. DOI: 10.1006/jmla.2001.2793.

- COLOMÉ, ÀNGELS, and MICHELE MIOZZO. 2010. Which words are activated during bilingual word production? *Journal of Experimental Psychology: Learning, Memory, and Cognition* 36.96–109. DOI: 10.1037/a0017677.
- COSTA, ALBERT; ALFONSO CARAMAZZA; and NURIA SEBASTIAN-GALLES. 2000. The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26.1283–96. DOI: 10.1037/0278-7393.26.5.1283.
- CRISTOFFANINI, PAULA; KIM KIRSNER; and DAN MILECH. 1986. Bilingual lexical representation: The status of Spanish-English cognates. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology* 38.367–93. DOI: 10.1080/14640748608401604.
- DE BOT, KEES. 1992. A bilingual production model: Levelt's 'Speaking' model adapted. *Applied Linguistics* 13.1–24. DOI: 10.1093/applin/13.1.1.
- DE BOT, KEES, and ROBERT SCHREUDER. 1993. Word production and the bilingual lexicon. *The bilingual lexicon*, ed. by Robert Schreuder and Bert Weltens, 191–214. Amsterdam: John Benjamins.
- DE GROOT, ANNETTE M. B. 1992. Bilingual lexical representation: A closer look at conceptual representations. *Orthography, phonology, morphology and meaning* (Advances in psychology 94), ed. by Ram Frost and Leonard Katz, 389–412. Amsterdam: North-Holland. DOI: 10.1016/S0166-4115(08)62805-8.
- DE GROOT, ANNETTE M. B. 2002. Lexical representation and lexical processing in the L2 user. *Portraits of the L2 user*, ed. by Vivian Cook, 32–63. Clevedon: Multilingual Matters.
- DE GROOT, ANNETTE M. B. 2011. *Language and cognition in bilinguals and multilinguals: An introduction*. Hove: Psychology Press.
- DECLERCK, MATHIEU, and ANDREA M. PHILIPP. 2015. A review of control processes and their locus in language switching. *Psychonomic Bulletin & Review* 22.1630–45. DOI: 10.3758/s13423-015-0836-1.
- DEDIU, DAN, and STEPHEN C. LEVINSON. 2013. On the antiquity of language: The reinterpretation of Neandertal linguistic capacities and its consequences. *Frontiers in Psychology* 4:397. DOI: 10.3389/fpsyg.2013.00397.
- DELL, GARY S. 1990. Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes* 5.313–49. DOI: 10.1080/01690969008407066.
- DELL, GARY S., and PADRAIG G. O'SEAGHDHA. 1992. Stages of lexical access in language production. *Cognition* 42.287–314. DOI: 10.1016/0010-0277(92)90046-K.
- DEWAELE, JEAN-MARC. 2001. Activation or inhibition? The interaction of L1, L2 and L3 on the language mode continuum. *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives* (Bilingual education and bilingualism 31), ed. by Jasone Cenoz, Britta Hufeisen, and Ulrike Jessner, 69–89. Bristol: Multilingual Matters.
- DIJKSTRA, TON, and JANET G. VAN HELL. 2003. Testing the language mode hypothesis using trilinguals. *International Journal of Bilingual Education and Bilingualism* 6.2–16. DOI: 10.1080/13670050308667769.
- DIXON, ROBERT M. W. 1997. *The rise and fall of languages*. Cambridge: Cambridge University Press.
- DONKERS, FRANC C. L., and GEERT J. M. VAN BOXTEL. 2004. The N2 in go/no-go tasks reflects conflict monitoring not response inhibition. *Brain and Cognition* 56.165–76. DOI: 10.1016/j.bandc.2004.04.005.
- ELLISON, T. MARK, and LUISA MICELI. 2015. Language discontinua: How bilingual-led divergence disrupts cladistics modelling. Paper presented at the 22nd International Conference on Historical Linguistics, Naples.
- ENFIELD, NICHOLAS J. 2001. Remarks on John Haiman, 1999: 'Auxiliation in Khmer: The case of *baan*'. *Studies in Language* 25.115–24. DOI: 10.1075/sl.25.1.05enf.
- ENFIELD, NICHOLAS J. 2003. Demonstratives in space and interaction: Data from Lao speakers and implications for semantic analysis. *Language* 79.82–117. DOI: 10.1353/lan.2003.0075.
- EPPS, PATIENCE. 2007. The Vaupés melting pot: Tukanoan influence on Hup. *Grammars in contact: A cross-linguistic typology*, ed. by Alexandra Y. Aikhenvald and Robert M. W. Dixon, 267–89. Oxford: Oxford University Press.

- EPPS, PATIENCE; JOHN HUEHNERGARD; and NA'AMA PAT-EL. 2013. Introduction: Contact among genetically related languages. *Journal of Language Contact* 6.209–19. DOI: 10.1163/19552629-00602001.
- EVANS, NICHOLAS. 2015. Contact without convergence. Plenary address presented at the 22nd International Conference on Historical Linguistics.
- FAY, NICOLAS, and T. MARK ELLISON. 2013. The cultural evolution of human communication systems in different sized populations: Usability trumps learnability. *PLoS ONE* 8(8):e71781. DOI: 10.1371/journal.pone.0071781.
- FRANÇOIS, ALEXANDRE. 2011. Social ecology and language history in the northern Vanuatu linkage: A tale of divergence and convergence. *Journal of Historical Linguistics* 1.175–246. DOI: 10.1075/jhl.1.2.03fra.
- GARLOCK, VICTORIA M.; AMANDA C. WALLEY; and JAMIE L. METSALA. 2001. Age-of-acquisition, word frequency, and neighborhood density effects on spoken word recognition by children and adults. *Journal of Memory and Language* 45.468–92. DOI: 10.1006/jmla.2000.2784.
- GRAY, RUSSELL D.; QUENTIN D. ATKINSON; and SIMON J. GREENHILL. 2011. Language evolution and human history: What a difference a date makes. *Philosophical Transactions of the Royal Society B: Biological Sciences* 366.1090–1100. DOI: 10.1098/rstb.2010.0378.
- GREEN, DAVID W. 1998. Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition* 1.67–81. DOI: 10.1017/S1366728998000133.
- GREENBERG, JOSEPH H.; CHARLES E. OSGOOD; and SOL SAPORTA. 1954. Language change. *Psycholinguistics: A survey of theory and research problems*, ed. by Charles E. Osgood and Thomas A. Sebeok, 146–63. Baltimore, MD: Waverly.
- GREENHILL, SIMON J.; QUENTIN D. ATKINSON; ANDREW MEADE; and RUSSELL D. GRAY. 2010. The shape and tempo of language evolution. *Proceedings of the Royal Society of London B: Biological Sciences* 277.2443–50. DOI: 10.1098/rspb.2010.0051.
- GREENHILL, SIMON J.; THOMAS E. CURRIE; and RUSSELL D. GRAY. 2009. Does horizontal transmission invalidate cultural phylogenies? *Proceedings of the Royal Society B: Biological Sciences* 276.2299–2306. DOI: 10.1098/rspb.2008.1944.
- GRIFFIN, ZENZI M., and KATHRYN BOCK. 1998. Constraint, word frequency, and the relationship between lexical processing levels in spoken word production. *Journal of Memory and Language* 38.313–38. DOI: 10.1006/jmla.1997.2547.
- GROSJEAN, FRANÇOIS. 1985a. The bilingual as a competent but specific speaker-hearer. *Journal of Multilingual and Multicultural Development* 6.467–77. DOI: 10.1080/01434632.1985.9994221.
- GROSJEAN, FRANÇOIS. 1985b. Polyglot aphasics and language mixing: A comment on Perecman (1984). *Brain and Language* 26.349–55. DOI: 10.1016/0093-934X(85)90048-3.
- GROSJEAN, FRANÇOIS. 1989. Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language* 36.3–15. DOI: 10.1016/0093-934X(89)90048-5.
- GROSJEAN, FRANÇOIS. 1997. Processing mixed language: Issues, findings, and models. *Tutorials in bilingualism: Psycholinguistic perspectives*, ed. by Annette M. B. de Groot and Judith F. Kroll, 225–54. Abingdon-on-Thames: Lawrence Erlbaum.
- GROSJEAN, FRANÇOIS. 2001. The bilingual's language modes. *One mind, two languages: Bilingual language processing*, ed. by Janet Nicol, 1–22. Malden, MA: Blackwell.
- GROSJEAN, FRANÇOIS. 2008. *Studying bilinguals*. Oxford: Oxford University Press.
- GROSJEAN, FRANÇOIS. 2013a. Bilingualism: A short introduction. *The psycholinguistics of bilingualism*, ed. by François Grosjean and Ping Li, 5–21. Chichester: Wiley-Blackwell.
- GROSJEAN, FRANÇOIS. 2013b. Speech production. *The psycholinguistics of bilingualism*, ed. by François Grosjean and Ping Li, 51–59. Chichester: Wiley-Blackwell.
- GULLIFER, JASON W.; JUDITH F. KROLL; and PAOLA E. DUSSIAS. 2013. When language switching has no apparent cost: Lexical access in sentence context. *Frontiers in Psychology* 4:278. DOI: 10.3389/fpsyg.2013.00278.
- GUMPERZ, JOHN, and ROBERT WILSON. 1971. Convergence and creolization. *Pidginization and creolization of languages*, ed. by Dell Hymes, 151–67. Cambridge: Cambridge University Press.
- HARTSUIKER, ROBERT J. 1999. Priming word order in sentence production. *The Quarterly Journal of Experimental Psychology Section A* 52.129–47. DOI: 10.1080/713755798.

- HARVEY, MARK. 2011. Lexical change in pre-colonial Australia. *Diachronica* 28.345–81. DOI: 10.1075/dia.28.3.03har.
- HÖDER, STEFFEN. 2014. Convergence vs. divergence from a diasystematic perspective. *Stability and divergence in language contact*, ed. by Kurt Braunmüller, Steffen Höder, and Karoline Kühn, 39–60. Amsterdam: John Benjamins.
- KASSTAN, JONATHAN R. 2013. Standardisation through the back door? Neo-speakers and ‘distanciation strategies’ in Francoprovençal. Paper presented at the Association for French Language Studies, Université de Perpignan.
- KEILSON, JULIAN (ed.) 2012. *Markov chain models—Rarity and exponentiality*. Berlin: Springer.
- KORMOS, JUDIT. 1999. Monitoring and self-repair in L2. *Language Learning* 49.303–42. DOI: 10.1111/0023-8333.00090.
- KUTEVA, TANIA. 1999. Languages and societies: The ‘punctuated equilibrium’ model of language development. *Language & Communication* 19.213–28. DOI: 10.1016/S0271-5309(98)00024-X.
- LA HEIJ, WIDO. 2005. Selection processes in monolingual and bilingual lexical access. *Handbook of bilingualism: Psycholinguistic approaches*, ed. by Judith F. Kroll and Annette M. B. de Groot, 289–307. New York: Oxford University Press.
- LALOR, ERIN, and KIM KIRSNER. 2000. Cross-lingual transfer effects between English and Italian cognates and noncognates. *International Journal of Bilingualism* 4.385–98. DOI: 10.1177/13670069000040030501.
- LEVELT, WILLEM J. M. 1989. *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- MALKIEL, BRENDA. 2009a. When *idioti* (idiotic) becomes ‘fluffy’: Translation students and the avoidance of target-language cognates. *Meta* 54.309–25. DOI: 10.7202/037683ar.
- MALKIEL, BRENDA. 2009b. Translation as a decision process: Evidence from cognates. *Babel* 55.228–43. DOI: 10.1075/babel.55.3.02mal.
- MATRAS, YARON. 1998. Utterance modifiers and universals of grammatical borrowing. *Linguistics* 36.281–332. DOI: 10.1515/ling.1998.36.2.281.
- MATRAS, YARON. 2007. The borrowability of structural categories. *Grammatical borrowing in cross-linguistic perspective*, ed. by Yaron Matras and Jeanette Sakel, 31–73. Berlin: Walter de Gruyter.
- MEEUWIS, MICHAEL, and JAN BLOMMAERT. 1998. A monolectal view of code-switching: Layered code-switching among Zairians in Belgium. *Code-switching in conversation: Language, interaction and identity*, ed. by Peter Auer, 76–98. London: Routledge.
- MEYER, DAVID E., and ROGER W. SCHVANEVELDT. 1971. Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology* 90.227–34. DOI: 10.1037/h0031564.
- MICELI, LUISA. 2015. Pama-Nyungan. *The Routledge handbook of historical linguistics*, ed. by Claire Bowern and Bethwyn Evans, 713–24. New York: Routledge.
- MUYSKEN, PIETER. 2013. Language contact outcomes as the result of bilingual optimization strategies. *Bilingualism: Language and Cognition* 16.709–30. DOI: 10.1017/S1366728912000727.
- MYERS-SCOTTON, CAROL. 2006. Natural codeswitching knocks on the laboratory door. *Bilingualism: Language and Cognition* 9.203–12. DOI: 10.1017/S1366728906002549.
- NADKARNI, MANGESH V. 1975. Bilingualism and syntactic change in Konkani. *Language* 51.672–83. DOI: 10.2307/412892.
- NOZARI, NAZBANOU; GARY S. DELL; and MYRNA F. SCHWARTZ. 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.
- OLDFIELD, R. C., and A. WINGFIELD. 1965. Response latencies in naming objects. *Quarterly Journal of Experimental Psychology* 17.273–81. DOI: 10.1080/17470216508416445.
- O’SHANNESSY, CARMEL. 2015a. Multilingual children increase language differentiation by indexing communities of practice. *First Language* 35.305–26. DOI: 10.1177/0142723715609227.
- O’SHANNESSY, CARMEL. 2015b. Typological and social factors influencing a new mixed language, Light Warlpiri. *Code-switching between structural and sociolinguistic perspectives*, ed. by Gerald Stell and Kofi Yakpo, 289–304. Berlin: De Gruyter.

- PETRIDES, MICHAEL. 1996. Specialized systems for the processing of mnemonic information within the primate frontal cortex [and discussion]. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 351.1455–62. [With discussion by A. Baddeley.] DOI: 10.1098/rstb.1996.0130.
- PICKERING, MARTIN J., and HOLLY P. BRANIGAN. 1998. The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language* 39.633–51. DOI: 10.1006/jmla.1998.2592.
- POARCH, GREGORY J., and JANET G. VAN HELL. 2012. Cross-language activation in children's speech production: Evidence from second language learners, bilinguals, and trilinguals. *Journal of Experimental Child Psychology* 111.419–38. DOI: 10.1016/j.jecp.2011.09.008.
- RIDDERINKHOF, K. RICHARD; WERY P. M. VAN DEN WILDENBERG; SIDNEY J. SEGALOWITZ; and CAMERON S. CARTER. 2004. Neurocognitive mechanisms of cognitive control: The role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain and Cognition* 56.129–40. DOI: 10.1016/j.bandc.2004.09.016.
- ROBERTS, PATRICIA M., and LOUISE DESLAURIERS. 1999. Picture naming of cognate and non-cognate nouns in bilingual aphasia. *Journal of Communication Disorders* 32.1–23. DOI: 10.1016/S0021-9924(98)00026-4.
- ROSS, MALCOLM. 1988. *Proto Oceanic and the Austronesian languages of Western Melanesia*. (Pacific linguistics C98.) Canberra: Australian National University.
- ROSS, MALCOLM. 2001. Contact-induced change in Oceanic languages in North-West Melanesia. *Areal diffusion and genetic inheritance: Problems in comparative linguistics*, ed. by Alexandra Y. Aikhenvald and Robert M. W. Dixon, 134–66. Oxford: Oxford University Press.
- ROSS, MALCOLM. 2007. Calquing and metatypy. *Journal of Language Contact* 1.116–43. DOI: 10.1163/000000007792548341.
- RUBENSTEIN, HERBERT; LONNIE GARFIELD; and JANE A. MILLIKAN. 1970. Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior* 9.487–94. DOI: 10.1016/S0022-5371(70)80091-3.
- SALAMOURA, ANGELIKI, and JOHN N. WILLIAMS. 2006. Lexical activation of cross-language syntactic priming. *Bilingualism: Language and Cognition* 9.299–307. DOI: 10.1017/S1366728906002641.
- SCHVANEVELDT, ROGER W.; DAVID E. MEYER; and CURTIS A. BECKER. 1976. Lexical ambiguity, semantic context, and visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance* 2.243–56. DOI: 10.1037/0096-1523.2.2.243.
- SILVERSTEIN, MICHAEL. 1981. *The limits of awareness*. Austin, TX: Southwest Educational Development Laboratory.
- SMITH, ELIOT R., and FREDERICA R. CONREY. 2007. Agent-based modeling: A new approach for theory building in social psychology. *Personality and Social Psychology Review* 11.87–104. DOI: 10.1177/1088868306294789.
- SMITH, KENNY; SIMON KIRBY; and HENRY BRIGHTON. 2003. Iterated learning: A framework for the emergence of language. *Artificial Life* 9.371–86. DOI: 10.1162/106454603322694825.
- STANFORD, JAMES N. 2009. 'Eating the food of our place': Sociolinguistic loyalties in multidialectal Sui villages. *Language in Society* 38.287–309. DOI: 10.1017/S0047404509090502.
- SZUBKO-SITAREK, WERONIKA. 2011. Cognate facilitation effects in trilingual word recognition. *Studies in Second Language Learning and Teaching* 1.189–208. DOI: 10.14746/ssllt.2011.1.2.2.
- TAMARIZ, MONICA; T. MARK ELLISON; DALE J. BARR; and NICOLAS FAY. 2014. Cultural selection drives the evolution of human communication systems. *Proceedings of the Royal Society B: Biological Sciences* 281:20140488. DOI: 10.1098/rspb.2014.0488.
- THOMASON, SARAH GREY. 2001. *Language contact*. Edinburgh: Edinburgh University Press.
- THURSTON, WILLIAM R. 1987. *Processes of change in the languages of north-western New Britain*. (Pacific linguistics B99.) Canberra: Australian National University. DOI: 10.15144/PL-B99.
- VAN VEEN, VINCENT, and CAMERON S. CARTER. 2002. The timing of action-monitoring processes in the anterior cingulate cortex. *Journal of Cognitive Neuroscience* 14.593–602. DOI: 10.1162/08989290260045837.

- WÄLCHLI, BERNHARD. 2010. We just lag behind, or phonetics is ahead of semantics, as usual. *Linguistic Discovery* 8.375–76. DOI: 10.1349/PS1.1537-0852.A.382.
- WEINREICH, URIEL; WILLIAM LABOV; and MARVIN HERZOG. 1968. Empirical foundations for theory of linguistic change. *Directions for historical linguistics: A symposium*, ed. by Winfred P. Lehmann and Yakov Malkiel, 95–188. Austin: University of Texas Press.
- ZURAW, KIE. 2003. Probability in language change. *Probabilistic linguistics*, ed. by Rens Bod, Jennifer Hay, and Stefanie Jannedy, 139–76. Cambridge, MA: MIT Press.

Ellison
[m.ellison@anu.edu.au]

[Received 14 December 2015;
revision invited 23 April 2016;
revision received 20 May 2016;
accepted 24 May 2016]

Miceli
Linguistics Discipline Group
School of Social Sciences M257
The University of Western Australia
35 Stirling Hwy
Crawley, WA 6009, Australia
[luisa.miceli@uwa.edu.au]