

# Language Control and Code-Switching

David W. Green

Division of Psychology and Language Sciences, University College London, Gower Street, London WC1E 6BT, UK; d.w.green@ucl.ac.uk; Tel.: +44-20-7679-7567

Received: 22 November 2017; Accepted: 22 March 2018; Published: 27 March 2018

**Abstract:** Analyses of corpus-based indices of conversational code-switching in bilingual speakers predict the occurrence of intra-sentential code-switches consistent with the joint activation of both languages. Yet most utterances contain no code-switches despite good evidence for the joint activation of both languages even in single language utterances. Varying language activation levels is an insufficient mechanism to explain the variety of language use. We need a model of code-switching, consistent with the joint activation of both languages, which permits the range of language use in bilingual speakers. I treat overt speech as the outcome of a number of competitive processes governed by a set of control processes external to the language networks. In a conversation, the speech of the other person may “trigger” code-switches consistent with bottom-up control. By contrast, the intentions of the speaker may act top-down to set the constraints on language use. Given this dual control perspective, the paper extends the control process model (Green and Wei 2014) to cover a plausible neurocomputational basis for the construction and execution of utterance plans in code-switching. Distinct control states mediate different types of language use with switching frequency as a key parameter in determining the control state for code-switches. The paper considers the nature of these states and their transitions.

**Keywords:** language control processes; control states; code-switching; priming

## 1. Introduction

Experimental research on code-switching in bilingual speakers has established some of the factors that affect the likelihood and type of a code-switch (see Van Hell et al. (2015) for a review). The research indicates that the immediate spoken context affects the likelihood of a code-switch. It has been established that the grammatical structure of a prior utterance can influence use of its counterpart in the other language. Under the conditions examined then, and in line with theoretical proposals (Hartsuiker et al. 2004), there is structural or grammatical priming across languages. A further factor that influences the likelihood of a code-switch is the presence of bilingual homophones. Such an effect is consistent with the trigger hypothesis (Clyne 2003) and its extension (Broersma and De Bot 2006). From the point of view of understanding the cognitive mechanism that mediates code-switching, experimental data indicate that prior utterances can influence the activation of lexico-syntactic representations, making such representations more available for selection. Such representations provide a source of stimulus-driven or bottom-up control. Two matters remain open from experimental data. First, do such results generalise to real world contexts? Corpus data are highly pertinent in establishing the ecological validity of experimental data. Second, code-switching may be sensitive to priming but bottom-up processes of control cannot be sufficient because bilingual speakers may code-switch in order to convey their communicative intentions (e.g., Myers-Scotton and Jake 2017). There is a need for code-switching to reflect the speaker’s communicative intention that, if necessary, overrides bottom-up influences. I refer to intentional control as top-down control. I consider each question.

Fricke and Kootstra's (2016) rich analysis of the Bangor-Miami corpus data (Deuchar et al. 2014) provides good evidence for priming effects. They analysed successive clauses. I note three results from their analysis. Code-switching can be primed by lexical items in a prior utterance that unambiguously belong to the language of the code-switch in a current utterance. Structurally, the presence of a finite verb in one language increased the likelihood that that language would form the matrix language for a code-switch in the current utterance independent of lexical factors. Corpus data then corroborate inferences from experimental research of the importance of bottom-up control (see also Broersma and De Bot 2006), but code-switched utterances accounted for only 5.8% of utterances in the analysis of the Bangor-Miami corpus by Fricke and Kootstra (2016). The bulk of utterances were in a single language. Priming is relevant, but it seems preferable to say that top-down processes of control allow such priming in overt speech in line with the speaker's communicative intention.

The evidence that code-switches can be primed is consistent with the idea that representations in the language network vary in their degree of activation. It is then a simple step to suggest that when a speaker is using just one of their languages that representations in their other language are suppressed (e.g., Muysken 2000). However, are variations in levels of activation a sufficient mechanism to account for the pattern of language use in the corpus data? Experimental evidence indicates that even when only one language is in play, lexical representations and grammatical constructions are active in the other language and lexical representations reach to the level of phonological form (e.g., (Blumenfeld and Marian 2013; Christoffels et al. 2007; Costa et al. 2000; Hoshino and Thierry 2011; Kroll et al. 2015) for a review). Given the corpus data cited above, top-down processes of control must allow speakers to produce words and constructions in just one language in a code-switching context despite parallel activation of words and constructions in the other language (for further discussion of bottom-up and top-down processes of language control, see (Kleinman and Gollan 2016; Morales et al. 2013)). What kind of mechanism might be envisaged? Core to the proposed mechanism is the idea that control processes external to the language network govern entry of items and constructions into a speech plan (Green 1986, 1998). Such an idea is consistent with neuropsychological data where, to take one example, following stroke, a bilingual patient may display intact clausal processing in each of their languages, but switch languages unintentionally when talking to a monolingual speaker (Fabbro et al. 2000). Based on this idea, Green and Wei (2014), proposed a control process model (CPM) of code-switching in which control processes establish the conditions for entry to a speech plan by setting the state of a language "gate". Given that utterances are planned in advance of their production, the CPM also includes a means (a competitive queuing network) to convert the parallel activation of the items in the plan to a serial order required for production. The aim of the present paper is to extend the CPM significantly and yet retain its basic architecture. In the next section, I review the background to the extension, specifically concerning the control processes, before identifying and presenting the extensions in the following section.

## 2. Background to the Proposal

Bilingual speakers use their languages in different ways as a function of the interactional context, and so any language control mechanism must be capable of enabling different patterns of language use (Green and Abutalebi 2013). In a single language context, only one language may be known by the other party in a conversation, and so only that language can be used. In a dual language context, a work environment, for instance, a speaker may converse in one language to one person and in another language to a different interlocutor. Both these contexts require the selection of one language but not the other. Qualitatively, the control regime is competitive in which activated items from the language not in use are blocked from entry into the utterance planning mechanism. The gate for that language is locked. Code-switching contexts can be different and allow both languages to be in play between the same interlocutors. Qualitatively, utterances can reflect a cooperative control process in which the resources of both language networks can be deployed. Different kinds of code-switches, as distinguished by Muysken (2000), may then be accounted for by differences in the nature of a cooperative control process that governs entry into the utterance planning mechanism. The language gates can be in different states, or, equivalently, a single gate can be in different states.

Muysken (2000) distinguished three different kinds of code-switching: alternation, insertion and congruent lexicalisation. Alternation refers to instances of code-switching where stretches of words of one language alternate with those of another within a conversational turn:

1. Moroccan Arabic/Dutch  
 maar 't hoeft niet li-'anna ida Šeft  
 but it need not for when I-see I  
 'But it need not be, for when I see, I...' Nortier 1990, p. 126 (cited in Muysken 2000, p. 5)
2. Spanish/English  
 andale pues and do come again  
 'That's all right then, and do come again.' Gumperz and Hernandez-Chavez 1971, p. 118  
 (cited in Muysken 2000, p. 5)

As the name implies, insertion covers instances where words or constituents from one language are inserted into a syntactic frame or matrix language (Myers-Scotton and Jake 2000) provided by another language.

3. Bolivian Quechua/Spanish  
 chay-ta las dos de la noche-ta chaya-mu-yku  
 that-AC the two of the night-AC arrive-CIS-1PL  
 'There at two in the morning we arrive.' Muysken 2000, p. 63

Congruent lexicalisation refers to code-switching when there is a shared (or at least largely shared) structure between the two languages that can be lexicalised by elements from either language (Muysken 2000, p. 5). In such a situation, words or morphemes from each language can be combined, permitting contributions that, standing alone, would be ungrammatical in each language, as in the following example:

4. Sranan/Dutch  
 wan heri gedeelte de ondro beheer fu gewapende machten  
 one whole part COP under control of armed forces  
 'One whole part is under control of the armed forces.' Bolle 1994, p. 75 (cited in Muysken 2000, p. 139)

Green and Wei (2014) contrasted two forms of cooperative control: coupled control and open control. They proposed that insertion and alternation can be realised by coupled control. Under coupled control, the matrix language temporarily cedes control to the other language to allow the intended insertion or alternation before control is returned back. Coupled control allows the other language gate to be "on the latch" and pushed open by an increase in item appropriateness before closing again after entry into the utterance planning mechanism (see Green and Wei 2014, p. 502, Figure 1, legend). By contrast, code-switches associated with congruent lexicalisation are realised by open control. Under open control, entry into the utterance planning mechanism is determined by whichever items (or constructions) from either language are most active at some moment in time. Both language gates are open. Items are accepted into the speech plan regardless of their membership in one language or the other but only by virtue of their appropriateness at that moment, for example, to meet the conditions for grammaticality. Since shared structure potentially affords copious switching to and from between languages, open control circumvents the awkwardness that would obtain under coupled control where there would be a need for the repeated ceding and taking back of control by one language with respect to another. An everyday physical example of how a control change minimises physical awkwardness is the gait change between walking and running. As stride frequency increases beyond a certain point we run rather than walk and spend more time in "flight" than on the ground.

### 3. The Nature of the Extended Control Process Model

The current paper extends the CPM in Green and Wei (2014) in four ways:

First, I consider the cognitive and neural means by which the utterance plan is actually constructed. In this extension, the gate is an active constructor of the utterance plan, and I propose testable predictions of its operation. Second, Green and Wei (2014) associated differences in control processes with structural (linguistic) distinctions in the type of code-switched utterances. From a control perspective, switching between languages imposes a demand on the mechanisms involved in switching, and so a critical factor is the copiousness of language-switching in a given utterance. Structural descriptions of different types of code-switches do not describe how they are realised in speech—a point recognised by Muysken (2000, p. 8). In Green and Wei (2014), we used the phenomenal/behavioural description “dense code-switching” to refer to utterances containing copious language switches. In that paper, we associated dense code-switching with typologically close language pairs that through congruent-lexicalisation (Muysken 2000, see above) permit particularly copious code-switching. Dense code-switching requires open control that allows items and constructions from either language into the utterance plan on an opportunistic basis (see also Green and Abutalebi 2013). However, typologically rather distinct languages may also elicit dense-code switching.

In Mandarin-English, for instance, items from different lexical categories in the two languages can be threaded together to yield a well-formed sentence. In the following example (Wei and Green 2015), English prepositions, adjectives and nouns are switched (/indicates clause boundaries).

#### 5 Mandarin/English

你	up stairs,	走到底,	向	left,	最	front	那张
ni		zoudaodi	xiang		zui		nazhang
you		go + to + end	towards		most		that + measure word

‘You go upstairs/ walk to the very end/ turn left/ the one at the forefront.’

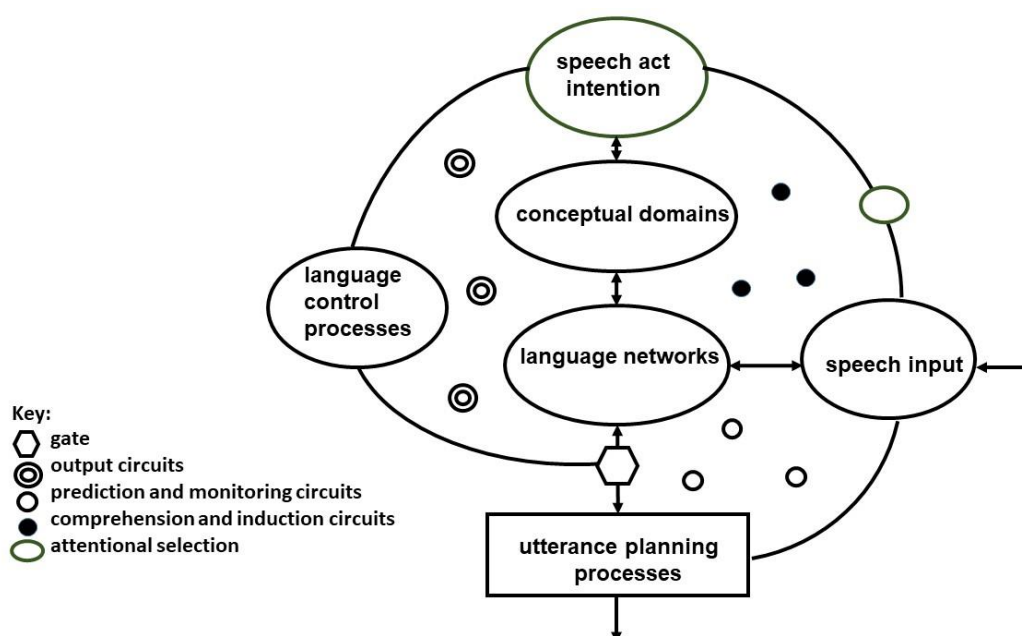
In the extended CPM, I do not limit the association of dense code-switching to congruent-lexicalisation. This means that open control can mediate copious code-switches in typologically distinct languages. In other words, we propose a generalization of the CPM. I therefore and prefer use of the term dense code-switching to refer to the presence of copious intra-clause switches. Third, as noted in the Introduction, the speech of the other conversational partner is a relevant factor in code-switching as it may serve to prime particular words or expressions and this is now taken into account explicitly. Fourth, the CPM distinguished between a competitive control state (in which items and constructions from only one language can be used) and a cooperative control states (in which items and constructions from either language can be used as in open control). The present paper additionally explores in a novel fashion the attentional and neural correlates of these control states and the transitions between them.

### 4. The Extended Control Process Model

The extended CPM retains the architecture of the original model but includes speech input as an explicit component and articulates the construction of a speech plan. Figures 1 and 2 illustrate the extended CPM. Figure 1 provides a schematic of the process of mapping a speech act into the utterance planning process. Figure 2 illustrates the role of the gate in controlling entry of syntactic form and lexical content into the competitive queuing network that yields the serial order required for speech production.

In Figure 1, the lower right quarter of the figure indicates that speech production processes can tacitly support speech perception (Halle and Stevens 1962; Skipper et al. 2017). The upper right quadrant refers to processes that can determine the meaning of what is said and induce significance. A speech act binds thoughts (conceptual content) onto the language networks and leads to the construction of an utterance plan (left half of the figure). In line with conventional proposals, and as described in Green and Wei (2014), nodes in the language network refer to words (lexical concepts), collocations and constructions.

I reprise the description of the network in Green and Wei (2014) so as to flesh out, for the present paper, the nature of the language networks envisaged. Connecting links between the nodes pass activation as a function of link strength. Activation of a given set of items can then reflect not only activation from the unfolding conceptual representation but properties of the network itself and momentary changes arising from the speech of another or oneself. Each lexical concept or construction is identified in terms of the language to which it belongs, that is, it is tagged for language membership (e.g., Albert and Obler 1978; Green 1986; Poullisse and Bongaerts 1994) by a link to a language node (e.g., Dijkstra and Van Heuven 2002). Such linkage allows functionally distinct but interconnected networks (Kroll et al. (2010) for further discussion) and is consistent with evidence of intermingled neuronal populations mediating language representation (Consonni et al. 2013; Green 2003; Paradis 2004) and for wider discussion (Green and Kroll forthcoming). We followed Hartsuiker et al. (2004) in supposing that common syntactic constructions, that underlie congruent lexicalisation, are represented by combinatorial nodes (see also Kootstra et al. 2010). For example, English and Dutch would share common combinatorial nodes for the verb *give* for a prepositional object construction (as in *Jack gave the ball to Jill*) and for the double object construction (as in *Jack gave Jill the ball*). Language pairs from the same language family (e.g., English-Dutch) will share more of such combinatorial nodes.

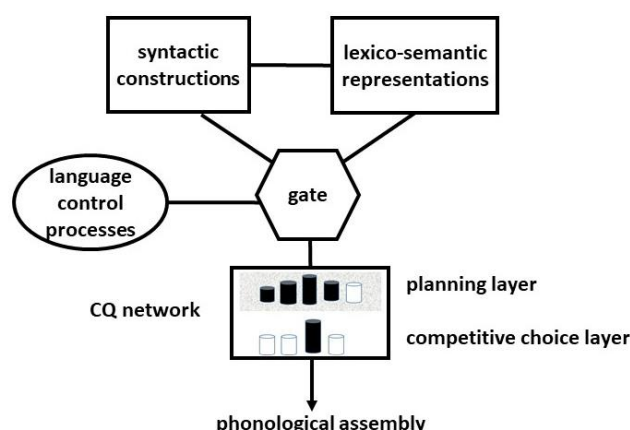


**Figure 1.** Schematic of the process of mapping a speech act into overt speech.

Given experimental evidence that conceptual activation induces activation of the syntactic and lexical inventories of both languages, there is then competition for entry into the utterance plan. The gate selects items and constructions on the basis of language control signals. Such signals reflect the communicative intentions of the speaker and I treat such signals as necessary and sufficient to control the operation of the gate. Here, I assume a single gate that can have different operating properties depending on the language control signals. In line with Green and Wei (2014) and Green and Abutalebi (2013), I suppose that these signals are generated by language task schemas that can be configured in different ways to implement the speaker's intention to use their languages in specific ways (Green 1986, 1998). With repeated use, a given configuration becomes a habit of control that can then be triggered by contextual cues (Green and Abutalebi 2013). Schemas can be configured competitively. In this configuration, activation of one schema suppresses activation of the other and signals the selection of one language but not the other. Alternatively, schemas can be coordinated cooperatively to signal coupled or open control. I describe how the gate operates, given the language control signals, after reviewing the speech planning process. However, the net result is that selection reflects items and constructions that are the most appropriate at that moment in time from the

speaker’s perspective on the basis of pragmatic, semantic, syntactic and collocational considerations and readiness for use (cf. Ward 1992).

Given that we cannot utter all that we have in mind at once, speech demands the imposition of serial order over the set of activated items in the plan (Lashley 1951). Competitive queuing (CQ) networks comprising a planning layer and competitive choice layer provide a neuroanatomically plausible way to achieve this goal (Bohland et al. 2009; Grossberg 1978; Houghton 1990). The gradient of item activations in the planning layer indicates production order (see Figure 2). In the choice layer, the item with the current highest level of activation suppresses all others, allowing that item to be released from the planning layer. Once released, its activation is suppressed in the planning layer, allowing the next most active item to be selected via the choice layer and so the cycle iterates. A plan precedes its execution but this does not mean that it must be fully specified before a person begins to speak. Rather, planning and execution can be interleaved (see MacDonald 2013) enabling the time between conversational turns to be quite short (Levinson 2016).



**Figure 2.** Schematic of utterance planning and execution via a competitive queuing (CQ) network.

Granting this way to control serial order, and that other CQ networks are required in the mapping to overt speech, we need to envisage a neuroanatomically plausible way to construct an utterance plan in the first place and one which enables code-switching. Whatever the mechanism, it must be one that captures a critical aspect of language use: namely the human facility to create diverse utterances using a finite means. A basic requirement is the separation of form (syntactic constructions) from content (lexico-semantic representations).

In the case of a sentence, a constituent can be represented as a variable permitting any word to be assigned or bound to it that meets its specification, as a finite verb, say or as the agent in a sentence. We do not have a deep understanding of how the human brain achieves such variable binding but one way, consistent with known neuroanatomy, has been proposed (Kriete et al. 2013), and I adopt it.

Figure 2 isolates components from Figure 1 and illustrates a partitioning of the language networks. Each syntactic construction is a frame comprising a set of slots or roles (e.g., agent-action-patient). In accordance with Kriete et al. (2013), I suppose that each slot can point to, that is bind, specific lexical content. This means, for example, that a given lexical item can serve as an agent in one sentence but a patient in another. In this way, the same structure can give rise to diverse sentences.

In Figure 2 I treat the gate as the plan constructor. Control signals from the gate select from activated constructions, update their lexico-semantic content, and signal release to the CQ network. Allowing selected structures in turn to affect the control signals from the gate provides a way to build nested structures. Known multiple reciprocal connections between frontal cortex and subcortical structures provide a plausible neural substrate (see also Stocco et al. 2014). Following, Kriete et al. (2013), I identify the gate with the basal ganglia (a subcortical structure) and the syntactic constructions and content with regions (“stripes”) in the prefrontal cortex. We note that neuroimaging data further suggest that one particular region of the basal ganglia, the caudate, is important in selecting a syntactic structure (Argyropoulos et al. 2013). Caudate activation, for

example, increases when experimental participants have to generate a sentence (i.e., produce and execute an utterance plan) rather than merely repeat a presented sentence. Damage to the caudate, following stroke, impairs a person's ability to avoid inappropriate code-switching. However, their code-switched utterances are nonetheless well-formed as in *I cannot communicate con you* (patient, A.H., Abutalebi et al. (2000)).

Kriete et al. (2013) implemented a network model of their proposal and so established proof-in-principle of its viability. The goal here is more circumscribed: it is to establish a conceptual extension of their model for the case of bilingualism and code-switching specifically. In this extension the operation of the gate depends on the nature of the language control signals (Figure 2). In an interactional context where only one language is in use, the gate blocks entry into the speech plan of activated constructions and items from the non-target language. Crucially, on this proposal, activated but non-target, language items are inhibited before they compete for binding.

Code-switching, on the other hand, requires that both languages are in play. Under coupled control, in the case of alternation, the gate opens to select a phrase or clause from the non-matrix language as part of the conversational turn. For insertions, the gate temporarily opens to bind an item from the non-matrix language to a role in the current clause of the matrix language. A further operation is presumably needed if the item has to be adapted to the local context as in the adaptation of the French verb *choisir* ('to choose') to *choisieren* with a German particle *-ieren* (Edwards and Gardner-Chloros 2007, p. 82; see also Gardner-Chloros 2009). We consider insertion as the best example of coupled control because it is intraclausal whereas, alternation, at least interclausal alternation, is consistent with competitive control. For present purposes, the key point is that because speech is planned before it is executed, any processing correlate of an intended insertion should be detectable *before* the actual insertion occurs in the speech stream. In other words, we should look at what is going on, not only at the time of overt switching but at the covert processes before that time.

Could there be a processing cost? One potential cost is associated with switching away from, and then back to, the matrix language. Such switching acts at a global level rather than at the item level (Green 1998) to inhibit any item from the current language. A second potential cost is increased competition in the binding of items from the non-matrix language to unfilled roles in the matrix language frame once the opportunity to do so is briefly available. This possibility arises, in contrast to single language use, because items from both languages are temporarily available for binding. So despite the insertion being maximally appropriate pragmatically (e.g., Myers-Scotton and Jake 2017), there could be a binding cost. Such a cost arises at the item level (e.g., Green 1998). How might these two types of cost be detected?

We know from research on language switching that activation in the caudate increases during an actual language switch (e.g., Abutalebi and Green 2007; Abutalebi et al. 2013; Crinion et al. 2006). Activation reflects control at the level of the language rather than at the level of the item (see DeClerk and Philipp (2015) for discussion of the loci for language switching)<sup>1</sup>. Assume for present purposes that a caudate activation also tracks covert language switching. It follows that any switching cost associated with coupled control *prior* to the actual code-switch should be detectable, all else being equal, as increased caudate activation relative to baseline utterances where there is no code-switching. How might binding competition be detected? Binding competition has yet to be explored but it seems reasonable to suppose that it may increase activation in frontal regions of the brain known to be linked to the suppression of lexical competitors (e.g., de Zubicaray et al. 2006; Shao et al. 2014). If so, any effects of binding competition associated with coupled control *prior* to the actual code-switch should be detectable, all else being equal, as increased frontal activation relative to baseline utterances where there is no code-switching. No current neuroimaging data bear on these predictions. We do know though that speech rate adjusts *prior* to the actual production of an insertional code-switch in a main clause. Miami-Bangor corpus data indicate, that prior to an insertional code-switch, speech rate

<sup>1</sup> Caudate activation also increases when participants, in a colour-word Stroop task, must name the colour in which a colour word is printed and suppress reading the incongruent colour word itself (Ali et al. 2010). Such an increase most likely reflects inhibition of a habitual plan to name the word rather than the colour in which it is printed.

decreases relative to that in matched unilingual control utterances by the same speaker in the same conversation (Fricke et al. 2016). So, for example, in a conversation predominantly in Spanish (Fricke et al. 2016, p. 115) the inserted word involves a switch from Spanish to English as in the example:

6. a. donde siempre tenemos el *delay*  
'Where we always have the delay.'

and the speech rate comparison is with its unilingual control

6. b. yo los vi en casa  
'I saw them at home.'

However, as yet, we cannot be certain whether such slowing really reflects language switching and/or increased binding competition, as predicted, or decreased activation of the matrix language in advance of the insertion, as proposed by Fricke et al. (2016).

Coupled control requires selection of items and constructions by language. Open control, by contrast, does not. The gate selects constructions and items opportunistically from each language network. Under this regime, the local context of the utterance effectively establishes requests for completion (including affixes and suffixes) that can be met by either language network. Under open control, language switching costs should be minimised (Green and Abutalebi 2013; Green and Wei 2014). It is possible though that the process of intertwining the morphosyntax of two languages increases demand on neural regions involved in the temporal control of morphosyntax (Green and Abutalebi 2013). If so, neuropsychological data implicate a circuit connecting the cerebellum and the frontal cortex (Mariën et al. 2001). If this is the case, this circuit will become more active during dense code-switching.

But what about binding costs that were not considered in the previous papers? These are potentially increased. This possibility arises because the gate is free to bind items to all currently active syntactic roles regardless of language membership—an option precluded in selective language use. However, binding costs are reduced as long as the competition for open slots is minimal. This minimal condition arises when there is a marked disparity in the accessibility of competitors and may obtain during dense code-switching though empirical data are lacking. Overall, open control during dense code-switching predicts that caudate activation will be insensitive to the amount of language switching. On the other hand, response in frontal regions will track the degree of binding competition.

From a control perspective, we have emphasised the association of open control with dense code-switching rather than restricting it to code-switches under the structural distinction of congruent lexicalisation. Such an association generalises the account of Green and Wei (2014) but leads one to wonder about the role of other structural designations in determining the control process. Muysken (2000, p. 228, 2nd paragraph), for example, indicates that alternational switches can also be quite copious. Here, the awkwardness argument also surely results (see earlier) because under coupled control there would need to be repeated cycles of ceding and taking back of control by one language with respect to another. The simplest control prediction is that open control will be obtained at some switching frequency (to be determined empirically) beyond a single insertion or alternation.

Of course, the notion of open control may be incorrect. If so, one possibility is that switching between languages will engage the network typically recruited during language switching (e.g., anterior cingulate cortex/pre-supplementary motor area, caudate (Luk et al. 2012)). This alternative possibility makes clear qualitative predictions: relative to those utterances where speakers stick to a single language, speech rate will be greatly slowed and caudate activation, along with other regions in the switching network, dramatically amplified both prior to, and during, dense code-switching.

## 5. Language Control States

Conversations can involve predominantly unilingual turns in one language or another or can involve code-switching. According to the present proposal, code-switching requires a shift from a



competitive to cooperative language control state. Within a code-switching context, code-switched utterance may be quite few in number and so the shift to cooperative control state might be quite transient. I consider two general and interrelated aspects of these control states. The first aspect associates changes in control states with a fundamental trade-off in human foraging and decision-making between exploiting a given resource and exploring an alternative. The second aspect associates changes in control state with changes in attentional breadth.

Viewed as a form of sampling or search, single language production in a bilingual speaker exploits the resources of a single language network and restricts utterances to constructions and items from a single language. The control state involves a narrow state of attention. By contrast, code-switching explores the resources of both language networks. The control state, especially in the case of dense code-switching, involves a broad attentional state.

If dense code-switching is more exploratory, it should increase activation in neural regions such as the frontopolar cortex that mediates exploration over exploitation in decision-making (e.g., Laureiro-Martínez et al. 2013). Remarkably, pupil diameter also provides an index of the trade-off (Aston-Jones and Cohen 2005). Research implicates neuroadrenergic signals from the locus coeruleus (LC) in the exploitation/exploration trade-off. Pupil diameter is sensitive to two classes of these signals, tonic (i.e., baseline) and phasic (i.e., event-related), that are inversely related to one another. Tonic pupil diameter increases with shifts to exploration in decision-making tasks (Jepma and Nieuwenhuis 2011) and with shifts to exploring the inner world as in “mind-wandering” during reading (Franklin et al. 2013). All else being equal, if dense code-switching is more exploratory than single language use then we can predict that tonic pupil diameter should increase during periods of dense code-switching. If, however, contrary to hypothesis, dense code-switching actually involves competitive control then the repeated demands to switch between languages will increase cognitive effort. In this case, the prediction is that phasic pupil diameter will increase because phasic pupil diameter tracks cognitive effort (Kahneman 1973; Kahneman and Beatty 1996).

The second aspect associates differences in language control with differences in breadth of attention. Attention can be focused more or less broadly (e.g., Eriksen and Yeh 1985; Wachtel 1967). I consider possible neural correlates in the following section but note here that language use can affect behavioural responses to interference consistent with the notion that there are attentional correlates of language control states. Competitive control requires a narrowing of attention that is arguably enhanced when participants must use one rather than another language in a dual language context. An ingenious experiment by Wu and Thierry (2013) found that during a dual language context, bilinguals were indeed more effective at resisting non-verbal interference as tested in a non-verbal flanker task (see also (Hommel et al. 2011) for evidence of benefits on a convergent thinking task). Cooperative control, by contrast, is predicted to increase the breadth of attention. It follows that during dense code-switching, in particular, participants should be more susceptible to interference. Such interference might be verbal or non-verbal (e.g., a distracting sound or visual stimulus). With respect to verbal interference, I conjectured above that dense code-switching may increase a specific kind of verbal interference, binding competition. If so, participants who routinely engage in dense code-switching may become adept at resolving such interference. In this case, an increased susceptibility to immediate interference during dense code-switching may be compensated by a greater facility in disengaging from it. Such facility might be general. If so, in a non-verbal, visual flanker task, speakers who routinely engage in dense code-switching, may show reduced effects of the congruency of a prior trial on the performance of a current trial (e.g., see (Grundy et al. 2017) for experiments examining such “sequential congruency effects” in bilinguals and monolinguals).

## 6. Network Measures and Language Control States

I have identified differences in language control states with differences in the breadth of attention. Competitive language control induces a narrow focus of attention. Cooperative language control, especially open control, induces a broader attentional state. Such differences can be captured by two measures of how neural networks work together: network synchrony and network metastability (Shanahan 2010). Network synchrony refers to the coherence of activity in the neural

regions and networks that mediate task performance. Network metastability refers how such coherence changes over time. Leech and Sharp (2014) propose that a narrow focus attentional state requires high synchrony and low metastability so that coactivation amongst the participating networks can be maintained over time to perform the task. By contrast, a broad focus attentional state requires low synchrony and high metastability and does permit capture by other possibilities. My earlier contrast applies: exploitation of a given resource requires high synchrony and so low metastability whereas exploration requires high metastability and so low synchrony.

The shift to a different attentional state during language control would then appear to require a change in the metastability of the participating neural networks. Leech and Sharp (2014) propose that a posterior region of the cingulate cortex may mediate such changes. If so, and identifying different language control states with differences in the breadth of attention, a shift between single language use and dense code-switching will be signalled by increased metastability in participating networks triggered by changes in their pattern of connectivity with the posterior cingulate cortex. Such shifts may precede speech output given that a speech plan precedes execution of it. Further, to the extent that there are spontaneous shifts in metastability, control demands to ensure the use of a single language will fluctuate over time.

## 7. Discussion

Varying the level of activation of different language networks is insufficient to account for the variety of utterance plans in bilingual speakers where speakers can use a single language, switch between two languages within a conversation or code-switch within a clause of the same utterance. Instead, the proposal here is that language control signals external to the language network help construct utterance plans. I briefly review the proposal before offering comment.

In a conceptual extension of an implemented neural network model of utterance production (Kriete et al. 2013), language control signals operate on a subcortical gate that acts as a constructor of utterance plans. The gate interacts with frontal regions to select a syntactic structure and binds roles in that structure to specific lexical content. Plans are constructed in the planning layer of a competitive queuing CQ network. The competitive choice layer of this network allows serial order to emerge from the parallel activation of items in the plan.

A key claim of the proposal is that single language use requires competitive language control whereas the different types of code-switching require cooperative language control. Language control signals in the single language case inhibit the selection of syntactic forms and lexical items from the non-target language. Competitive language control exploits the resources of a single language network and requires a narrow focus of attention. By contrast, cooperative language control explores the resources of both language networks and broadens the focus of attention. Single insertions, and perhaps alternations, are mediated by coupled control, in which control by the matrix language is temporarily ceded to the other language. More copious or dense code-switching, associated with congruent lexicalisation but not restricted to it, is mediated by open control that suspends selection based on language membership. Pertinent neural and behavioural predictions were proposed.

How plausible is it that intra-clausal code-switching, and especially dense code-switching, induces a language control state distinct from that involved in single language use? Parsimony favours a single control state but predicts a computational and energetic cost: dense code-switching will increase demands on regions such as the caudate involved in language switching. However, such a cost does not necessarily weigh the case in favour of open control: the cooperative control state may increase binding competition over and above that generated under a single control state.

A further objection on the grounds of parsimony might be raised. The idea of distinct language control states carries as a corollary that the mind/brain can be in distinct, and co-occurring, attentional states (Green and Wei 2016). For instance, a speaker can be focussed on achieving the goal of a speech act (a narrow attentional state) but use a linguistic means (e.g., dense code-switching) that requires a broad attentional state. There is a hierarchical relationship between these attentional states with the language control state nested under the narrow focussed, sustained attentional state associated with

achieving the communicative goal. Subjective experience does not automatically rule out the co-occurrence of distinct attentional states. We can talk attentively and walk, as it were, on auto-pilot. Typically too, our focus of attention is the goal of what we want to say, rather than on the linguistic means we recruit to do so.

Empirical testing of co-occurring and dissociable attentional states requires distinct neural markers for sustained attention and dense code-switching. There are reasonable grounds for identifying a network of frontal regions of the brain as instrumental in sustaining attention to a goal (e.g., Rosenberg et al. 2016). We lack an agreed, empirically determined marker for dense code-switching but I have proposed two possible indices: changes in pupil diameter and changes in the metastability of participating language networks. The upshot is that we need empirical data to test for differences in language control states. How might such data be collected?

Neuroimaging research indicates that the patterns of neural activation in the listener synchronise with those of the speaker (Stephens et al. 2010). If so, attentional states induced in different forms of language use will entrain the same states in listeners, at least when participants are drawn from the same speech community. Eliciting relevant and extended speech samples seems tractable. Scripted dialogue tasks, such as the map description task, offer a promising approach (Beatty-Martínez and Dussias 2017). Along with corpus data (Fricke and Kootstra 2016), such tasks allow identification of the cues to an upcoming code-switch and the opportunity to examine neural adaptation in the listener. So, for example, in the context of prior stretches of dense code-switching in the discourse, a cue to a code-switch (Fricke et al. 2016; Beatty-Martínez and Dussias 2017) should increase tonic pupil diameter and increase metastability in the networks mediating language control.

On the other hand, scripted dialogue and responses to it, may need inventive designs to capture other aspects of code-switching noted in corpus data. Myslín and Levy (2015) argued that end of clause, single noun, insertional code switches (from Czech to English) in their corpus data met a discourse-functional purpose. Such switches signalled to the listener the need to heed discourse meaning. Similarly, Backus (2001) identified the importance of culturally specific connotations in determining insertional code-switches. In addition to network activity putatively mediating distinct language control states, such results point to the subtlety and complexity of neural response as listeners track and respond to the speech acts of their interlocutor and prepare their own utterances.

In line with the present proposal, future research may identify distinct language control states that affect on-line production and comprehension. If so, different profiles of bilingual language use (e.g., single language use versus extensive code-switching) may, over the longer term, affect resilience to brain atrophy with age (Alladi et al. 2013; Bialystok et al. 2007; Gold 2015). If there are dissociable, and co-occurring, attentional states induced by the type of language use, then, if we are to gain a fuller picture, we should also consider the content of what is being said, as well as how it is said. Conversations are about topics as illustrated by the mapping from conceptual domains to language networks depicted in Figure 1. Different topics, such as describing the cloudscape from a plane window or reminiscing about a shared holiday, pose different attentional demands that vary in their breadth and their reference to external or internal worlds (Green forthcoming). May not the attentional states induced co-occur with those elicited during utterance production in bilingual speakers?

## 8. Summary

Patterns of code-switching within a conversational turn can be characterised structurally as alternations, insertions and congruent lexicalisations (Muysken 2000). Code-switched utterances though occur in the context of conversations that may predominantly involve conversational turns in just one language. Conceivably, representations in the non-selected language are simply suppressed for the duration until they become activated to allow for code-switching. However, experimental work indicates that the non-selected language is active even during single language use and indeed can be primed by the speech of the conversational partner. Such data make it unlikely that the production of different types of code-switches is sufficiently explained by momentary changes in activation within the language network. Instead, and in line with the CPM of code-

switching by Green and Wei (2014), the present paper contends that single language use, as well as code-switching, is attributable to language control processes external to the networks that represent the lexical items and constructions of the two languages. The extended CPM, developed here, explored how these control processes act to construct and execute an utterance plan. I conceptually extended a neurocomputational account of sentence production (Kriete et al. 2013). Their account comprised a subcortical gate, and cortical regions representing syntactic constructions and lexical items. The gate operates to select a syntactic construction and to bind lexical items to open slots in the selected construction. In the bilingual case of the extended CPM, language control signals determine the operation of the gate. Use of a single language involves competitive language control in which the gate allows only constructions and items from a single language to enter the utterance plan. Code-switching, by contrast, involves cooperative language control. Single switches covering intra-clause insertions, for example, can be handled by coupled control in which the control of entry into the utterance plan by one language (the matrix language) is temporarily ceded to another. By contrast, copious switching (associated with congruent lexicalisation) or, dense code-switching more generally, requires open control in which entry into the utterance plan is opportunistic. Open control circumvents the awkwardness of repeatedly ceding and returning control between one language and another and so I suggest that open control will be obtained at some switching frequency (to be determined empirically). As an everyday example of a control change that minimises physical awkwardness, I mentioned how our gait changes from walking to running at some stride frequency. Open control minimises the costs of switching languages but allows that, under certain conditions, items may compete to be bound to open slots. In addition to exploring the control processes involved in utterance production, the paper proposed that cooperative language control, and open control, specifically, induces, at least transiently, a broader attentional state detectable in measures of pupil dilation and the dynamical properties of neural networks. Advances in experimental techniques and analytic methods are needed to test the proposal but are attainable.

**Acknowledgments:** Some of the content in this paper was presented at a meeting entitled *Interdisciplinary perspectives on code-switching*, University of Cambridge, 3–4 October 2016. I thank the organizers and contributors for constructive comments. I am grateful too for the comments and constructive suggestions of the Editors and two anonymous reviewers.

**Conflicts of Interest:** The author declares no conflict of interest.

## References

- (Abutalebi and Green 2007) Abutalebi, Jubin, and David Green. 2007. Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics* 20: 242–75.
- (Abutalebi et al. 2000) Abutalebi, Jubin, Antonio Miozzo, and Stefano F. Cappa. 2000. Do subcortical structures control language selection in bilinguals? Evidence from pathological language mixing. *Neurocase* 6: 101–6.
- (Abutalebi et al. 2013) Abutalebi, Jubin, Pasquale A. Della Rosa, Guosheng Ding, Brendan Weekes, Albert Costa, and David W Green. 2013. Language proficiency modulates the engagement of cognitive control areas in multilinguals. *Cortex* 49: 905–91.
- (Albert and Obler 1978) Albert, Martin L., and Loraine K. Obler. 1978. *The Bilingual Brain: Neuropsychological and Neurolinguistics Aspects of Bilingualism*. New York and London: Academic Press.
- (Ali et al. 2010) Ali, Nilufa, David W. Green, Ferath Kherif, Joseph T. Devlin, and Cathy J. Price. 2010. The role of the left head of caudate in suppressing irrelevant words. *Journal of Cognitive Neuroscience* 22: 2369–86.
- (Alladi et al. 2013) Alladi, Suvarna, Thomas H. Bak, Vasanta Duggirala, Bapiraju Surampudi, Mekala Shailaja, Anuj K. Shukla, Jaydip R. Chaudhuri, and Subhash Kaul. 2013. Bilingualism delays age at onset of dementia, independent of education and immigration status. *Neurology* 81: 1938–44.
- (Argyropoulos et al. 2013) Argyropoulos, Georgios P. D., Pascale Tremblay, and Steven L. Small. 2013. The neostriatum and response selection in overt sentence production: An fMRI study. *NeuroImage* 82: 53–60.
- (Aston-Jones and Cohen 2005) Aston-Jones, Gary, and Jonathan D. Cohen. 2005. An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review Neuroscience* 28: 403–50.

- (Backus 2001) Backus, Ad. 2001. The role of semantic specificity in insertional codeswitching: Evidence from Dutch-Turkish. In *Codeswitching Worldwide II*. Edited by Rodolfo Jacobson. Berlin: Mouton de Gruyter, pp. 125–54.
- (Beatty-Martínez and Dussias 2017) Beatty-Martínez, Anne L., and Paola E. Dussias. 2017. Bilingual experience shapes language processing: Evidence from code-switching. *Journal of Memory and Language* 95: 173–89.
- (Bialystok et al. 2007) Bialystok, Ellen, Fergus I. M. Craik, and Morris Freedman. 2007. Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia* 45: 459–64.
- (Blumenfeld and Marian 2013) Blumenfeld, Henrike K., and Viorica Marian. 2013. Parallel activation and cognitive control during spoken word recognition in bilinguals. *Journal of Cognitive Psychology* 25: 547–67.
- (Bohland et al. 2009) Bohland, Jason W., Daniel Bullock, and Frank H. Guenther. 2009. Neural representations and mechanisms for the performance of simple speech sequences. *Journal of Cognitive Neuroscience* 22: 1504–29.
- (Broersma and de Bot 2006) Broersma, Mirjam, and Kees de Bot. 2006. Triggered code-switching: A corpus based evaluation of the original triggering hypothesis and a new alternative. *Bilingualism: Language and Cognition* 9: 1–13.
- (Christoffels et al. 2007) Christoffels, Ingrid K., Christine Firk, and Niels O. Schiller. 2007. Bilingual language control: An event-related brain potential study. *Brain Research* 1147: 192–208.
- (Clyne 2003) Clyne, Michael G. 2003. *Dynamics of Language Contact: English and Immigrant Languages*. Cambridge: Cambridge University Press.
- (Consonni et al. 2013) Consonni, Monica, Riccardo Cafiero, Dario Marin, Marco Tettamanti, Antonella Iadanza, Franco Fabbro, and Daniela Perani. 2013. Neural convergence for language comprehension and grammatical class production in highly proficient bilinguals is independent of age of acquisition. *Cortex* 49: 1252–58.
- (Costa et al. 2000) Costa, Albert, Alfonso Caramazza, and Nuria Sebastián-Gallés. 2000. The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory and Cognition* 26: 1283–96.
- (Crinion et al. 2006) Crinion, Jenny, Robert Turner, Alice Grogan, Takashi Hanakawa, Uta Noppeney, Joseph T. Devlin, Toshihiko Aso, Shinichi Urayama, Hidenao Fukuyama, Katharine Stockton, and et al. 2006. Language control in the bilingual brain. *Science* 312: 1537–40.
- (de Zubicaray et al. 2006) de Zubicaray, Greig, Katie McMahon, Matthew Eastburn, and Alan Pringle. 2006. Top-down influences on lexical selection during spoken word production: A 4T fMRI investigation of refractory effects in picture naming. *Human Brain Mapping* 11: 864–73.
- (DeClerk and Philipp 2015) DeClerk, Mathieu, and Andrea M. Philipp. 2015. A review of control processes and their locus in language switching. *Psychological Bulletin & Review* 22: 1630–45.
- (Deuchar et al. 2014) Deuchar, Margaret, Peredur Davies, Jon R. Herring, M. Carmen Parafita Couto, and Diana Carter. 2014. Building bilingual corpora. In *Advances in the Study of Bilingualism*. Edited by Enlli M. Thomas and Ineke Mennen. Bristol: Multilingual Matters, pp. 93–110.
- (Dijkstra and Van Heuven 2002) Dijkstra, Ton, and Walter J. B. Van Heuven. 2002. The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition* 5: 175–97.
- (Edwards and Gardner-Chloros 2007) Edwards, Malcolm, and Penelope Gardner-Chloros. 2007. Compound verbs in code switching: Bilinguals making do? *International Journal of Bilingualism* 11: 73–91.
- (Eriksen and Yeh 1985) Eriksen, Charles W., and Yei-Yu Yeh. 1985. Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception & Performance* 11: 583–97.
- (Fabbro et al. 2000) Fabbro, Franco, Miran Skrap, and Salvatore Aglioti. 2000. Pathological switching between languages following frontal lesions in a bilingual patient. *Journal of Neurology, Neurosurgery and Psychiatry* 68: 650–52.
- (Franklin et al. 2013) Franklin, Michael S., James M. Broadway, Michael D. Mrazek, Jonathan Smallwood, and Jonathan W. Schooler. 2013. Window to the wandering mind: Pupillometry of spontaneous thought while reading. *The Quarterly Journal of Experimental Psychology* 66: 2289–94.
- (Fricke and Kootstra 2016) Fricke, Melinda, and Gerrit J. Kootstra. 2016. Primed code-switching in spontaneous bilingual dialogue. *Journal of Memory and Language* 91: 181–201.
- (Fricke et al. 2016) Fricke, Melinda, Judith F. Kroll, and Paola E. Dussias. 2016. Phonetic variation in bilingual speech: A lens for studying the production-comprehension link. *Journal of Memory and Language* 89: 110–37.

- (Gardner-Chloros 2009) Gardner-Chloros, Penelope. 2009. *Code-Switching*. Cambridge: Cambridge University Press.
- (Gold 2015) Gold, Brian T. 2015. Lifelong bilingualism and neural reserve against Alzheimer's disease: A review of findings and potential mechanisms. *Behavioural Brain Research* 28: 9–15.
- (Green 1986) Green, David W. 1986. Control, Activation, and Resource: A framework and a model for the control of speech in bilinguals. *Brain and Language* 27: 210–23.
- (Green 1998) Green, David W. 1998. Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition* 1: 67–81.
- (Green 2003) Green, David W. 2003. Neural basis of lexicon and grammar in L2 acquisition. In *The Lexicon Syntax Interface in Second Language Acquisition*. Edited by Roeland van Hout, Aafke Hulk, Folkert Kuiken and Richard J. Towell. Amsterdam: John Benjamins Publishing, pp. 197–218.
- (Green forthcoming) Green, David W. Forthcoming. Language control and attention: An exploration. In *The Handbook of the Neuroscience of Multilingualism*. Edited by John W. Schwieter. Hoboken: Wiley-Blackwell.
- (Green and Abutalebi 2013) Green, David W., and Jubin Abutalebi. 2013. Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology* 25: 515–30.
- (Green and Kroll forthcoming) Green, David W., and Judith F. Forthcoming. The neurolinguistics of bilingualism: Plasticity and control. In *Oxford Handbook of Neurolinguistics*. Edited by Greig de Zubicaray and Niels O. Schiller. Oxford: Oxford University Press.
- (Green and Wei 2014) Green, David W., and Li Wei. 2014. A control process model of code-switching. *Language, Cognition and Neuroscience* 29: 499–511.
- (Green and Wei 2016) Green, David W., and Li Wei. 2016. Code-switching and language control. *Bilingualism: Language and Cognition* 19: 883–84.
- (Grossberg 1978) Grossberg, Stephen. 1978. A theory of human memory: Self-organization and performance of sensory-motor codes, maps, and plans. In *Progress in Theoretical Biology*. Edited by Robert J. Rosen and Fred M. Snell. New York: Academic Press, pp. 233–74.
- (Grundy et al. 2017) Grundy, John G., Ashley Chung-Fat-Yim, Deanna C. Friesen, Lorinda Mak, and Ellen Bialystok. 2017. Sequential congruency effects reveal differences in disengagement of attention for monolingual and bilingual young adults. *Cognition* 163: 42–55.
- (Halle and Stevens 1962) Halle, Morris, and Kenneth N. Stevens. 1962. Speech recognition: A model and a program for research. *IRE Transactions on Information Theory* 8: 155–59.
- (Hartsuiker et al. 2004) Hartsuiker, Robert J., Martin J. Pickering, and Eline Veltkamp. 2004. Is syntax separate or shared between languages? cross-linguistic syntactic priming in Spanish–English bilinguals. *Psychological Science* 15: 409–14.
- (Hommel et al. 2011) Hommel, Bernhard, Lorenza S. Colzato, Rico Fischer, and Ingrid K. Christoffels. 2011. Bilingualism and creativity: Benefits in convergent thinking come with losses in divergent thinking. *Frontiers in Psychology* 2: 273.
- (Hoshino and Thierry 2011) Hoshino, Noriko, and Guillaume Thierry. 2011. Language selection in bilingual word production: Electrophysiological evidence for cross-language competition. *Brain Research* 1371: 100–9.
- (Houghton 1990) Houghton, George. 1990. The problem of serial order: A neural network model of sequence learning and recall. In *Current Research in Natural Language Generation*. Edited by Robert Dale, Chris S. Mellish and Michael Zock. London: Academic Press, pp. 287–319.
- (Jepma and Nieuwenhuis 2011) Jepma, Marieke, and Sander Nieuwenhuis. 2011. Pupil diameter predicts changes in the exploration–exploitation trade-off: Evidence for the Adaptive Gain Theory. *Journal of Cognitive Neuroscience* 23: 1587–96.
- (Kahneman 1973) Kahneman, Daniel. 1973. *Attention and Effort*. Englewood Cliffs: Prentice-Hall.
- (Kahneman and Beatty 1996) Kahneman, Daniel, and Jackson Beatty. 1996. Pupil diameter and load on memory. *Science* 154: 1583–85.
- (Kleinman and Gollan 2016) Kleinman, Daniel, and Tamar H. Gollan. 2016. Speaking two languages for the price of one: By passing language control mechanisms via accessibility driven switches. *Psychological Science* 27: 700–14.
- (Kootstra et al. 2010) Kootstra, Gerrit J., Janet G. van Hell, and Ton Dijkstrac. 2010. Syntactic alignment and shared word order in code-switched sentence production: Evidence from bilingual monologue and dialogue. *Journal of Memory and Language* 63: 210–31.

- (Kriete et al. 2013) Kriete, Trenton, David C. Noelle, Jonathan D. Cohen, and Randall C. O'Reilly. 2013. Indirection and symbol-like processing in the prefrontal cortex and basal ganglia. *Proceedings of the National Academy of Sciences U.S.A* 110: 16390–95.
- (Kroll et al. 2010) Kroll, Judith F., Janeth G. Van Hell, Natasha Tokowicz, and Green, David W. 2010. The Revised Hierarchical Model: A critical review and assessment. *Bilingualism: Language and Cognition* 13: 373–81.
- (Kroll et al. 2015) Kroll, Judith F., Paola E. Dussias, Kinsey Bice, and Lauren Perrotti. 2015. Bilingualism, mind, and brain. *Annual Review of Linguistics* 1: 377–94.
- (Lashley 1951) Lashley, Karl S. 1951. The problem of serial order in behavior. In *Cerebral Mechanisms in Behavior: The Hixon Symposium*. Edited by Lloyd A. Jeffress. New York: Wiley, pp. 112–36.
- (Laurreiro-Martínez et al. 2013) Laurreiro-Martínez, Daniella, Nicola Canessa, Stefano Brusoni, Maurizio Zollo, Todd Hare, Federica Alemanno, and Stefano F. Cappa. 2013. Frontopolar cortex and decision-making efficiency: Comparing brain activity of experts with different professional background during an exploration-exploitation task. *Frontiers in Human Neuroscience* 7: 927.
- (Leech and Sharp 2014) Leech, Robert, and David J. Sharp. 2014. The role of the posterior cingulate cortex in cognition and disease. *Brain* 137: 12–32.
- (Levinson 2016) Levinson, Stephen C. 2016. Turn-taking in human communication- origins and implications for language processing. *Trends in Cognitive Sciences* 20: 6–14.
- (Luk et al. 2012) Luk, Gigi, David W. Green, Jubin Abutalebi, and Cheryl Grady. 2012. Cognitive control for language switching in bilinguals: A quantitative meta-analysis on functional neuroimaging studies. *Language and Cognitive Processes* 27: 1479–88.
- (MacDonald 2013) MacDonald, Maryellen C. 2013. How language production shapes language form and comprehension. *Frontiers in Psychology* 4: 226.
- (Mariën et al. 2001) Mariën, Peter, Sebastiaan Engelborghs, Franco Fabbro, and Peter P. De Deyn. 2001. The lateralized linguistic cerebellum: A review and a new hypothesis. *Brain and Language* 79: 580–600.
- (Morales et al. 2013) Morales, Julia, Carlos J. Gómez-Ariza, and M. Teresa Bajo. 2013. Dual mechanisms of cognitive control in bilinguals and monolinguals. *Journal of Cognitive Psychology* 25: 531–46.
- (Muysken 2000) Muysken, Petere. 2000. *Bilingual Speech: A Typology of Code-Mixing*. Cambridge: Cambridge University Press.
- (Myers-Scotton and Jake 2000) Myers-Scotton, Carol, and Janice L. Jake. 2000. Testing the 4-M model. *International Journal of Bilingualism* 4: 1–8.
- (Myers-Scotton and Jake 2017) Myers-Scotton, Carol, and Janice L. Jake. 2017. Revisiting the 4-M model: Code-switching and morpheme election at the abstract level. *International Journal of Bilingualism* 21: 340–66.
- (Myslin and Levy 2015) Myslin, Mark, and Roger Levy. 2015. Code-switching and predictability of meaning in discourse. *Language* 91: 871–905.
- (Paradis 2004) Paradis, Michel. 2004. *A Neurolinguistic Theory of Bilingualism*. Amsterdam: John Benjamins Publishing.
- (Poullisse and Bongaerts 1994) Poullisse, Nanda, and Theo Bongaerts. 1994. First language use in second language production. *Applied Linguistics* 15: 36–57.
- (Rosenberg et al. 2016) Rosenberg, Monica D., Emily S. Finn, Dustin Scheinost, Xenophon Papademetris, Xilin Shen, R. Todd Constable, and Marvin M. Chun. 2016. A neuromarker of sustained attention from whole-brain functional connectivity. *Nature Neuroscience* 19: 165–79.
- (Shanahan 2010) Shanahan, Murray. 2010. Metastable chimera states in community-structured oscillator networks. *Chaos* 20: 013108.
- (Shao et al. 2014) Shao, Zeshu, Ardi P. A. Roelofs, Daniel J. Acheson, and Antje S. Meyer. 2014. Electrophysiological evidence that inhibition supports lexical selection in picture naming. *Brain Research* 1586: 130–42.
- (Skipper et al. 2017) Skipper, Jeremy I., Joseph T. Devlin, and Daniel R. Lametti. 2017. The hearing ear is always found close to the speaking tongue. *Brain and Language* 164: 77–105.
- (Stephens et al. 2010) Stephens, Greg J., Lauren J. Silbert, and Uri Hasson. 2010. Speaker-listener neural coupling underlies successful communication. *Proceedings of the National Academy of Sciences USA* 107: 14425–30.
- (Stocco et al. 2014) Stocco, Andrea, Brianna Yamasaki, Rodion Natalenko, and Chantel S. Prat. 2014. Bilingual brain training: A neurobiological framework of how bilingual experience improves executive function. *International Journal of Bilingualism* 18: 67–92.

- (Van Hell et al. 2015) Van Hell, Janet G., Kaitlyn A. Litcofsky, and Caitlin Y. Ting. 2015. Sentential code-switching: Cognitive and neural approaches. In *The Cambridge Handbook of Bilingual Processing*. Edited by John W. Schwieter. Cambridge: Cambridge University Press, pp. 459–82.
- (Wachtel 1967) Wachtel, Paul L. 1967. Conceptions of broad and narrow attention. *Psychological Bulletin* 68: 417–29.
- (Ward 1992) Ward, Nigel 1992. A parallel approach to syntax for generation. *Artificial Intelligence* 57: 183–25.
- (Wei and Green 2015) Wei, Li (Institute of Education London), and David W. Green (University College London). 2015. Personal communication.
- (Wu and Thierry 2013) Wu, Yang Jing, and Guillaume Thierry. 2013. Fast modulation of executive function by language context in bilinguals. *Journal of Neuroscience* 33: 13533–37.



© 2018 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).