

# A computational model of bilingual inhibitory control in a lexical decision task

**Andrew P. Valenti (andrew.valenti@tufts.edu)**

Department of Computer Science, 161 College Avenue  
Medford, MA 02155 USA

**Matthias J. Scheutz (mscheutz@cs.tufts.edu)**

Department of Computer Science, 107B Halligan Hall, 161 College Avenue  
Medford, MA 02155 USA

## Abstract

Results from empirical studies in bilingual research demonstrate that “code-switching” in bilinguals, i.e., the change from one language to another, comes at a cost. What is still unclear is the exact nature of the cost. In this paper, we develop a connectionist model of Green’s Inhibitory Control (IC) hypothesis for a bilingual lexical decision task (Green, 1998, 74), which hypothesizes that the cost of switching between two languages in this task is due to inhibitory forces among different task schema and word forms. We were able to fit the model to the empirical data from Von Studnitz and Green (1997, Experiment 1), thus providing a proof-of-concept model that the IC hypothesis can account for the observed language switching costs and for the L1 language advantage effect in a lexical decision task. More extensive, systematic explorations of the model and its parameters in related but different tasks are, however, required to determine the extent of the IC applicability. Although the IC hypothesis is simple, it advanced earlier hypothesis by suggesting that the major source of language switching costs lies not in the lexico-semantic system but in a task/decision system; this notion is incorporated by Dijkstra and Van Heuven (2002) in their BIA+ hypothesis. By modelling the IC hypothesis, we suggest that not only is it a plausible explanation for the locus of control but also the reasonableness of the BIA+ hypothesis as the basis for further research in code switching and language learning.

**Keywords:** Bilingual; Second language; Inhibitory control; Computational model

## Introduction

Over the years, a general consensus has emerged in bilingual research that both of a bilingual speaker’s languages are active simultaneously and that there is a lexicon unified across both languages. Experimental results have repeatedly shown that the bilingual speaker’s two languages compete in terms of phonological representations (i.e., accents) and word meaning, but proficient bilingual speakers rarely confuse words from the competing language and, despite the intrusion of an accent, can usually be readily understood. Thus, there must be a cognitive process that allows bilinguals to control what they are saying and understand what they are hearing, when they are speaking, reading, or listening in the target language.

One theory proposes the existence of a “language switch” which, when set in the correct position, effectively blocks the other language (Macnamara & Kushnir, 1971). An implication of the language switch theory is that there would be a cost of switching between languages (i.e., the cost of “throwing the switch”) and, moreover, that the switching costs between L1 and L2 would be symmetrical (due to the very nature of a “switch”). However, asymmetrical costs were found in an experiment by Kroll and Stewart (1994), thus prompting

the question of what cognitive process might account for this asymmetry in switching costs? The Inhibitory Control (IC) hypothesis of Green (1998) attempts to explain this asymmetry by proposing that there is an increase in time needed to resolve competition among activated word forms (i.e., lemmas) in L2.

In this paper, we set out to provide evidence for the IC hypothesis by constructing a computational “proof-of-concept” model that implements the hypothesized inhibitory mechanism in the context of a lexical decision task. We start by explaining the IC hypothesis and discussing some of its predictions as they relate to this paper. Next, we review the Von Studnitz and Green (1997) study, the empirical data used, and the experiment’s procedure. We then introduce the model framework used to construct a computational simulation of language switching predicted by the IC hypothesis, and report the model’s results, comparing them to the empirical data and discussing the model’s advantages and disadvantages. Finally, we point to areas in which the model might be extended in order to account for more of the effects reported by Von Studnitz and Green and how the model might be generalized to language switching in speech production.

## Background

In the IC hypothesis, a set of language-specific processes and language task schemas, operating under the control of a general cognitive supervisory process reactively inhibit competitors at the lemma level of the lexico-semantic system using its language tags. A lemma is a representation in the lexico-semantic system that contains syntactic information which Green (1998) identifies as the locus of language membership. IC extends the Kroll and Stewart (1994) revised hierarchical model (RHM) which proposes that a bilingual’s first and second languages (L1 and L2) are connected bi-directionally through links whose strengths vary as a function of the language. However, this model has some limitations. For example, RHM does not specify how a bilingual engaged in a language translation task avoids naming the word to be translated and the IC model suggests a plausible mechanism. Kroll and Stewart (1994) found that when asking individuals to translate words that were blocked by category, for forward translations (i.e., L1 to L2) participants took longer to translate those words than when they were randomly presented. No such effect was observed for backward translations (i.e., L2 to L1). This suggests that in forward translations, according to Kroll and Stewart (1994, 168), blocking words by category

activates the conceptual element, creating difficulty in selecting a single lexical entry for production. Green (1998, 73) hypothesizes that there is an increase in time needed to overcome competition between L2 lemmas which have become activated and suggests the presence of a control mechanism to account for the observed effects.

Green (1998) develops the idea for a control mechanism by building upon the observations of Grosjean (1997) that bilinguals operate in different language modes. They may be speaking in L1, L2 or, in the appropriate context, mixing both their languages. Green hypothesizes that there must be a regulatory mechanism that is both sensitive to external input and has the capacity for internal control. Building upon Green’s prior research derived from the “contention scheduling model” proposed by Norman and Shallice (1986), Green developed the IC hypothesis. Norman and Shallice argue that most attentional conflicts occur in the initiation of an action rather than its execution and propose a two-level control mechanism. The first level is a contention scheduling process that selects from competing schemas; the second is a supervisory attentional component that oversees and biases the selection process. Incorporating this theory, IC hypothesizes that the intention to perform a specific language task is executed by a supervisory attentional system (SAS) which affects the activation of language task schemas that are themselves in competition to control the output. Thus, a set of language-specific processes and general cognitive skills determines how the bilingual responds to language tasks.

Green’s IC hypothesis predicts that language switching may take time because it involves a change in language schema for a given task and because of the time it takes to overcome the inhibition of the previously activated language. IC predicts that there will be such costs when switching among language tasks (i.e., translation and naming) as well as within specific tasks (i.e., language reception and production). The specific task investigated in this paper is regulatory processing in a lexical decision task for which there are empirical results from a study conducted by Von Studnitz and Green (1997, Experiment 1). In this study, German-English bilinguals are asked to decide whether or not a presented letter string (may be a word or non-word) was a word in L1 or in L2 using an alternating runs paradigm (i.e., there is predictable switching between languages). In the study, the color of the background on which the word was presented served as an external cue informing participants of the required language for decision. Figure 1 illustrates the relationship between this cue and two lexical decision schemas inhibiting one another, and the lexico-semantic system.

The SAS establishes the schemas which map an output of the lexico-semantic system (e.g., presence of an L2 tag) to a response (e.g., press left key if L2 word). The control mechanism is driven bottom-up and once established, the SAS monitors it to ensure desired performance. In the case of a new switch trial, a new schema is triggered by the external cue and suppresses the previously active schema. Moreover, a

new word in a different language has to overcome the inhibition on its language tags from the previous trial. Thus, IC proposes two areas of inhibition: (1) schema-level inhibition and (2) tag inhibition in the lexico-semantic system. IC predicts that inhibiting a previously active schema and overcoming the inhibition of a previously active language will take time, manifesting as a switch cost.

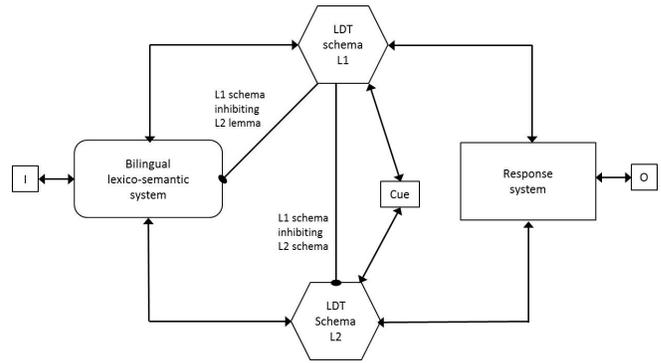


Figure 1: Regulatory processing in an LD task with language switching (Green, 1998). The L1 task schema is suppressing the L2 task schema and inhibiting the L2 lemmas in the lexico-semantic system.

### Empirical Results

In the experiment (Von Studnitz & Green, 1997, Experiment 1), language switches occurred on alternating trials (EEGGEEGGEE, etc.), indicated by a change in the color background which was counterbalanced across participants. Two types of stimuli were used: word and non-word, although both the IC and the computational model used only word stimuli. Each experimental block was preceded by a single filler trial which served to provide a clear designation for the experimental trial and in the case of the computational model, prime the lexico-semantic system and the task schemas to a “resting” state (i.e., activate the control mechanisms associated with either L1 or L2). Two sets of words were constructed with a total of 160 words in each: 80 words were English and 80 words were German in each set. In each case, half the words were high-frequency and half were of low-frequency. The words were matched for syllable length and letter length across the two languages. Words were orthographically possible in either language. Neither cognates (i.e., words that look the same and have the same meaning) nor interlingual homographs, “false friends” (i.e., words that look the same but have different meanings) were included. The experimental procedure is shown in Figure 2.

The experiment found an average switch cost of 118 ms for high-frequency and low-frequency English and German words and that participants were also 63 ms faster responding to German words compared to English words. The results from the experiment are summarized in Table 1.

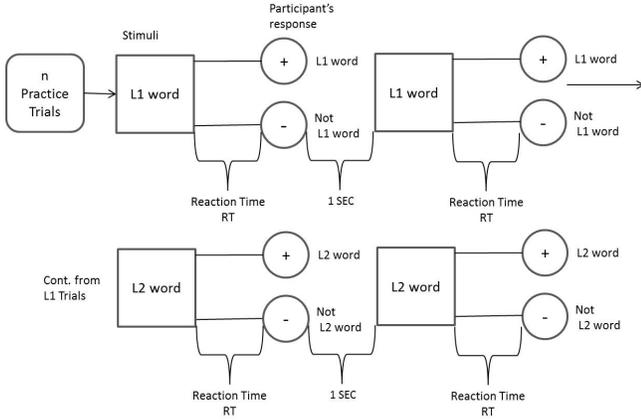


Figure 2: Experiment 1 procedure (Von Studnitz & Green, 1997). After  $n$  practice trials, participants are presented with a letter string and asked to decide whether it is a string in either L1 or L2. Language switches occurred on alternating trials indicated by a change in color background.

Table 1: Experiment 1 results (Von Studnitz & Green, 1997)

Word Type	Switch	Non-switch	Cost
	Mean RT	Mean RT	
L1 German	805 ms	705 ms	100 ms
L2 English	887 ms	752 ms	135 ms
Mean	846 ms	728 ms	118 ms
Avg. cost <i>switch</i> – <i>nonswitch</i> = 118 ms			
L1 RT advantage: 63 ms			

## Model Development

The purpose of the computational model is to verify the inhibitory mechanism proposed by Green (1998) and illustrated in Figure 1. Specifically, the goal is to verify the model’s prediction of a cost when a bilingual switches from a required response in German to an English response for a lexical decision task, as well as from English to German.

The design of the computational model is based on Green’s IC hypothesis for regulatory processing in a lexical decision task as shown in Figure 1. An interactive activation and competition (IAC) connectionist model was built using a neural network simulation tool, NNSIM. One unit per processing pool was allocated since the likely distributed representations in the brain of the regulatory processes were not specified by the IC model and were not the focus of this study. The architecture of the model is shown in Figure 3. The model has three layers of units: an input layer of word representations that are connected to a hidden layer of lemma representations which are in turn connected to an output layer representing the lexical decision task schema. In addition, there are two input units representing the target language response cues used in the experiment (i.e., the background color on which the word is presented) each connected to its respec-

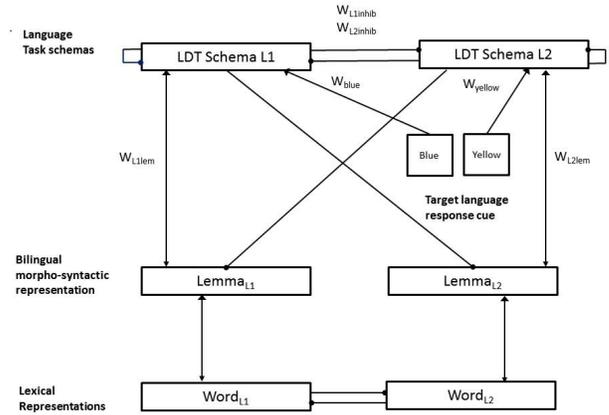


Figure 3: Model design. Stimuli are presented at the word level and activation in bottom up with language “tag” associated with words at the lemma level. The cue activates a task schema indicating whether an LD response should be in L1 or L2. The key IC features are the inhibitory links between the L1/L2 LD schemas and between the L1(or L2) schemas and L2(or L1) lemmas.

tive LDT schema. One instance of this three-tiered structure is provided for L1 and another for L2; the two are connected by inhibitory connections between the L1/L2 schemas, L1/L2 words, and L1 lemmas/L2 schemas, as supposed by the IC hypothesis. In Figure 1, a single cue is seen as connecting to both the LDT L1 Schema and LDT L2 schema, but it has been implemented as two separate color units to more accurately reflect the experimental procedure.

The entirety of the L1/L2 in the bilingual lexico-semantic system is not modelled here. Since the experimental stimuli presented across the trials were an average of high and low frequency English and German words, each L1 or L2 word represents a sample word of average frequency from the trials. The lemma units are the morpho-syntactic representation of the word where Green (1998) posits the language tag is located. As with the word units, only the L1 and L2 lemmas connected to their corresponding L1 and L2 words are modelled. The L1 LDT schema and L2 LDT schema exist outside the bilingual lexico-semantic system and are the units that are monitored for their activation level.

The experiment measured a participant’s reaction time, i.e., from when a word was presented on the computer screen to when the participants press the + or - key. It is apparent that the reaction time (RT) consists of two components: the time it takes to activate the schema plus the time it takes for the participant to move his or her arm and press the key. For the purpose of the output data mapping, only the schema activation time is of interest and the remainder of the reaction time is treated as a constant. Thus in the model only the number of cycles from the resting level of the schema until it settles

at its activation level is measured.

Only the L1 structure was modelled to start. All the weights of the top-down connections were set identically. Three weight groups for the bottom-up excitatory connections were identified: (1) cue unit to LDT Schema, (2) word to lemma, and (3) lemma to LDT Schema. An input was applied to the L1 word unit and to the L1 cue unit and the bottom-up connection weights were adjusted until the L1 LDT schema was strongly activated, i.e., 0.800. This took place at 30 cycles. An identical L2 model was then constructed and the two structures were connected using the inhibitory connections hypothesized by Green. All inhibitory connection weights were set to -0.1. Even without further adjustment of the weights, we noticed a switching cost during a switch trial, but the cost was somewhat greater than what the empirical data suggested. However, by adjusting only the top-down connection weights uniformly, we were able to get a good fit with the empirical results for the language switch cost. In this iteration of the model, the network settles with the L1 schema activation at 0.793.

The symmetrical L1/L2 model however does not account for the L1 advantage observed in the experiment: participants in the study were 63 ms faster in responding to German (L1) words than to English (L2) words. Reasoning that the language effect was located in the lexico-semantic system rather than in the task schema, the weight of the connection between the L2 word and the L2 lemma was adjusted, producing the desired language effect and a good fit to the experimental data. With this change, the network settles with the L2 schema activation at 0.779.

We iterated through process of adjusting the top-down connection weights uniformly as a group and the connection from the L2 word to the L2 lemma until the summed square errors for the switching cost and L1 advantage of the model and experiment were minimized. This resulted in weights of 0.02 and 0.08 respectively; all the connection weights are given in Figure 4.

## Model Results

Table 2: Model Results

Word Type	Switch	Non-switch	Cost
	Mean RT	Mean RT	
L1 German	830 ms	710 ms	120 ms
L2 English	910 ms	770 ms	140 ms
Mean	870 ms	740 ms	130 ms

Avg. cost  $switch - nonswitch = 130$   
L1 RT advantage: 70 ms

The experiment’s reaction times (RT) needed to be mapped onto network update cycles in order to be able to simulate the temporal sequence of reading words and the sequence of internal cognitive processes during the activation of the task schema. This mapping was achieved by dividing the response

		TO							
		L1 wd	L2 wd	L1 lem	L2 lem	L1 sch	L2 sch	L1 cue (Blue)	L2 cue (Yellow)
F R O M	L1 wd		-0.1	0.1					
	L2 wd	-0.1			0.08				
	L1 lem	0.02				0.8			
	L2 lem		0.02				0.8		
	L1 sch			0.02	-0.1	-0.1	-0.1		
	L2 sch			-0.1	0.02	-0.1	-0.1		
	L1 cue (Blue)					0.1			
	L2 cue (Yellow)							0.1	

Figure 4: The connection weights depict the values found during model fitting.

times by 10 and rounding to the nearest whole number, thus one cycle = 10 ms. Using the mappings from the experiment’s reaction times onto update cycles, a sequence of 10 trials was run as shown in Table 3 corresponding to the experimental procedure as shown in Figure 2. There are four types of trials: L1:initialize, L1:non-switch, L1/L2:switch, L2:non-switch, L2/L1:switch. L1:initialize represents a practice trial and it allows the network to settle at its L1 Task schema activation level. Although it is numbered as a trial in Table 3, “blank” is the inter-trial pause. We alternately apply input to the L1 word and Blue cue or to the L2 word and Yellow cue according to whether we want a switch or non-switch trial and then cycle through the network until the corresponding L1 or L2 schema is activated, recording the results.

In Table 3, the number of cycles (i.e., no. cycles) given for the non-switch and switch trials for both L1 and L2 corresponds approximately to the time it takes for the participant to ready a schema for making a lexical decision in the target language indicated by the cue, i.e.,  $LDTsch_t = RT_{avg} - k$ , where  $k$  (i.e., Physical RT) is the time it takes for the response system to initiate the action to press the “+” key, and  $RT_{avg}$  is the average response time as measured in the experiment. Removing the inputs to both the lexical node and the cue for a period of 100 network cycles is the functional equivalent of the experiment’s 1 second pause between trials. During this pause, we want the activated schema, lemma, and word to decay to represent the lower activation of the mental lexical, syntactic, and executive task control processes likely once the stimulus is removed. The resting activation levels of the word, lemma, and schema units represent either the base level from which we wish to return to activation if the next stimulus presented is from the same language as the previous, or the level which will be inhibitory to the rising activation of word, lemma, and schema from the new target language for a language switch trial.

Table 3: Computational Model Trial Runs. Ten trials were conducted, the first being practice. The number of cycles is cumulative and the RT is computed from:  $RT = (EventDuration(Trial_{N,a}) + EventDuration(Trial_{N,b}))$

Trial	Type	No. Cycles	Event Duration	RT (ms)
1	L1 initialize	32	32	
2	blank	132	100	
3a	L1 non switch	160	28	
3b	Physical RT	203	43	710
4	blank	303	100	
5a	L1/L2 switch	351	48	
5b	Physical RT	394	43	910
6	blank	494	100	
7a	L2 non-switch	528	34	
7b	Physical RT	571	43	770
8	blank	671	100	
9a	L2/L1 switch	711	40	
9b	Physical RT	754	43	
10	blank	854	100	830

Trial 1 is a “practice trial”  
One cycle = 10 ms

Comparing the results of the Von Studnitz and Green (1997, Experiment 1) study, Table 1, with the results of the computational model, Table 2, shows a good fit with the significant effect of cost switching as predicted by the IC hypothesis and suggested by the results from the empirical study (i.e., 130 ms for the model, 118 ms for the experiment). The model also exhibits the experiment’s asymmetrical language RTs where participants responded faster to German words compared to English words (i.e., 70 ms for the model, 63 ms for the experiment).

This suggests that we were able to fit the model to the empirical data the best we could to achieve a “proof-of-concept” validating the IC hypothesis. Our results suggest that the inhibitory control mechanism is a possible explanation for the switch costs and asymmetric language RTs seen in the experiment, but the model does not yet provide strong evidence that it is likely the case.

## Discussion

The inhibitory control model is important because it provides a theory for how bilinguals can perform different tasks given different language inputs. In addition, IC explains various effects observed in empirical studies such as, switch costs and unwanted language interference. IC has been an influence on other theories of bilingual word recognition and Dijkstra and Van Heuven (2002) incorporate aspects of Green’s theory in their BIA+ model. However, IC has only been specified descriptively at a functional level, thus lacking the advantages of a computational model. Their disadvantage is that, unlike computational models, functional models rely purely on

behavioral experiments which can only superficially explore cognitive processes, and are not easily generalized, thus limiting their predictive ability.

Our computational model provides a framework for validating the inhibitory control model, and at least captures an important aspect of bilingual word recognition: the regulatory processing mechanism. After additional exploration of the model parameters it can be further developed and generalized to test what IC predicts in other tasks such as language switching in production. One weakness in many computer assisted language learning tools is their ability to train language learners to actually speak a new language. A computational model of regulatory processing in language production would add to our understanding of what inhibits a language learner from producing utterances in the new language.

Although the model provides a proof-of-concept that the IC regulatory mechanism described by Green (1998) is a possible explanation for the switching cost seen empirically, more evidence could be provided if the model incorporated recognition of non-words which were utilized in the empirical study (Von Studnitz & Green, 1997, Experiment 1). The inclusion of non-words makes a lexical decision task more meaningful and also provides a means for testing whether or not the nature of the non-word affected reaction time (e.g., English non-words possible in English only or in both German and English). Showing that participants are affected by the status of a non-word would provide further evidence against the input-switch hypothesis as non-words provide no route to the lexicon and therefore cannot use it to decode a response. Further, Von Studnitz and Green (1997, Experiment 2) conducted an additional experiment using a generalized lexical decision task. In this experiment, participants needed to decide whether or not a letter string was a word in either language; the empirical study found a small but significant switching cost. Green underspecifies the inhibitory mechanism for a generalized LD task (Green, 1998, 74) and this would be an extension to the IC and the computational model. Neither the IC nor the computational model account for frequency effects or cross-language effects demonstrated in other studies of bilingualism and cognition, e.g., (Van Heuven, Dijkstra, & Grainger, 1998). Subsequent models of bilingual word recognition such as BIA+ (Dijkstra & Van Heuven, 2002) incorporate features of Green’s IC hypothesis and include a model of the lexico-semantic system, which is not specified by Green. BIA+ also accounts for more of the empirical results observed in studies of bilingualism (e.g. orthographic neighborhood effects, cross-linguistic effects, non-linguistic context effects, stimulus-response binding). Further research in computation modelling of bilingual cognitive processes may well be better served investigating a more general architecture such as BIA+.

The IC hypothesis also predicts a cost in switching between languages in certain word production tasks such as numeral naming (Green, 1998). Such tasks involve different language schemas and in order to produce speech, the activation of a

new language schema would need to exceed the activation of the current language schema. However, this mechanism for doing so is not fully specified. Models accounting for speech production have been based on models by Levelt and Meyer (1999) and Dijkstra and Van Heuven (2002) discuss generalizing BIA+ to bilingual word production.

Beyond BIA+ there are novel approaches that provide a more dynamic view of the lexicon than the traditional connectionist network and combine localist and distributed properties of processing. One such model is the self-organizing model of bilingual processing, SOMBIP, (Li & Farkas, 2002). It consists of two interconnected self-organizing neural networks, along with a recurrent neural network that computes lexical co-occurrence constraints. SOMBIP captures both bilingual production and comprehension and can account for patterns in the bilingual lexicon without the use of language nodes or language tags. It attempts to answer the question of where the information comes from that allows the bilingual to separate their two languages. A potentially interesting research direction is to examine Grosjean (1997) account of code switching, Levelt's speech production architecture, Dijkstra and van Heuven's BIA+, and Li and Farkas' SOMBIP create a computational model that can account the inhibitory control mechanism involved in code switching.

### Conclusion

Green's IC hypothesis Green (1998), predicts that language switching takes time and as a result, a cost will be incurred. This prediction was supported by evidence from the empirical study (Von Studnitz & Green, 1997, Experiment 1) and the data from the computational model giving a language switching cost of 130 ms for the model vs. 118 ms for the experiment. The model also predicted an advantage of 70 ms when participants responded to a German word compared to an English word compared with the experiment which showed a 63 ms advantage. We built a model that was symmetrical in structure for the L1 and L2 units, connecting the two according to the inhibitory connections postulated by Green. Similar connections were assigned to weight groups and weights were assigned consistently within the groups, providing a minimal fit; dissimilar weights were assigned to the L2 structure's word-to-lemma connection to account for the L1 advantage effect. We adjusted only the L2 word-to-lemma and top-down connection weights until a set of minimum values were found to fit the empirical data. This suggests that it is possible that the IC regulatory mechanism is responsible for the language switch costs seen in Experiment 1, but more extensive systematic exploration of the model and its parameters to account for additional language effects will be required to draw a stronger conclusion.

Although the IC hypothesis is simple as it does not account for many well-known cross-language effects, it advanced earlier hypothesis by suggesting that the major source of language switching costs lies not in the lexico-semantic system but in a task/decision system. The predictive strength of the

model could be improved by extending it to include non-word stimulus and to predict the switching cost observed in a general lexical decision task (Von Studnitz & Green, 1997, Experiment 2). Future research includes developing a computational model of code switching among bilinguals. One path toward this goal is to note that the IC model's task schema and task control concepts have been included in the BIA+ model (Dijkstra & Van Heuven, 2002) and the authors of that model and Green (1998) believe that it can be generalized to bilingual auditory recognition and bilingual word production. Contrasting and comparing this localist connectionist approach with more novel approaches using distributed connectionist architectures such as SOMBIP could lead to a computational model that can account for inhibitory control in bilingual code switching without resorting to the use of language nodes or tags.

### References

- Dijkstra, T., & Van Heuven, W. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, 5(3), 175–179.
- Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, 1, 67–81.
- Grosjean, F. (1997). Processing mixed languages: Issues, findings, and models. In L. Wei (Ed.), *The bilingualism reader*. New York, NY: Routledge.
- Kroll, J., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, 33, 149–174.
- Levelt, W., & Meyer, A. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1–75.
- Li, P., & Farkas, I. (2002). A self-organizing connectionist model of bilingual processing. In R. Heredia & J. Altarriba (Eds.), *Bilingual sentence processing* (Vol. 134). North Holland.
- Macnamara, J., & Kushnir, S. (1971). Linguistic independence of bilinguals: The input switch. *Journal of Verbal Learning and Verbal Behaviour*, 10, 480–487.
- Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation* (Vol. 4). New York, NY: Plenum Press.
- Van Heuven, W. J. B., Dijkstra, A. T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language*, 39, 458–483.
- Von Studnitz, R. E., & Green, D. W. (1997). Lexical decision and language switching. *International Journal of Bilingualism: Cross-Linguistic Studies of Language Behaviour*, 1, 3–24.