Bilingualism: Language and Cognition 5 (3), 2002, 241–251 © 2002 Cambridge University Press DOI: 10.1017/S1366728902003036

The cost of switching language in a semantic categorization task*

ROSWITHA E. VON STUDNITZ
DAVID W. GREEN
University College London

German-English bilinguals decided whether a visually presented word, either German or English, referred to an animate or to an inanimate entity. Bilinguals were slower to respond on a language switch trial than on language non-switch trials but only if they had to make the same response as on the prior trial (e.g., two successive "animate" responses). If they had to change their response from the prior trial, a switch of language speeded responses. This interaction between language switching and response type was abolished for the translations of words previously presented. We consider the implications for theories of language control.

Introduction

Bilinguals name a stimulus such as a numeral (Meuter and Allport, 1999; Jackson, Swainson, Cunnington and Jackson, 2001) or a picture (e.g., Costa, Miozzo and Caramazza, 1999; Hernandez, Dapretto, Mazziotta and Bookheimer, 2001; Kroll, Dietz and Green, in preparation) more slowly when they must switch to a different language. Switching costs may arise as a result of processes within the bilingual lexicon (e.g., suppression of the words in the non-target language on the prior trial) and/or because of processes outside the lexicon associated with task performance (Green, 1998a, b).

There are also switching costs in comprehension. Bilinguals are slower to make a lexical decision about a word on a trial where there is a switch of language compared to a trial where there is not (Grainger and Beauvillain, 1987; von Studnitz and Green, 1997; Thomas and Allport, 2000 – see also Dalrymple-Alford, 1985). The problem is to explain why a switching cost exists when there is strong evidence that lexical access is not restricted to just one language even when only that language is required for the task (Dijkstra, van Jaarsveld and ten Brinke, 1998; Dijkstra, Grainger and van Heuven, 1999; Dijkstra, de Bruijn, Schriefers and ten Brinke, 2000; de Groot, Delmaar and Lupker, 2000; Thomas and Allport, 2000; von Studnitz and Green, 2002). What might give rise to a switch cost? We distinguish sources associated with word recognition from those associated with using accessed information to make a response. We consider word recognition first.

One possibility is that word recognition in one language partially deactivates the representations of words in the other language (e.g., Green, 1998a). A switch cost arises because of the need to overcome this decrement in activation (see also Dijkstra, van Jaarsfeld and ten Brinke, 1998, pp. 63–64; Thomas and Allport, 2000, p. 52).

A second possibility is that switching affects obligatory processes of phonological recoding. Work by Brysbaert, van Dyck and van de Poel (1999) suggests that spelling-sound correspondences in both languages are involved in recognizing words (see Jared and Kroll, 2001 for research showing such effects in word naming).[1] However, languages differ in the extent to which their orthographies show a consistent correlation with sound. For instance, German readers may typically rely on grapheme-phoneme decoding strategies whereas readers of English may supplement such a strategy with the recognition of letter patterns (e.g., Goswami, Ziegler, Dalton and Schneider, 2001). In switching between English and German in word recognition, bilinguals might incur a cost as different types of correspondences must be

* We thank Judy Kroll, Ton Dijkstra and three other anonymous reviewers for many constructive comments on a previous version of this paper. The work reported is part of the research work carried out by the first author for her Ph.D. under supervision of the second author.

[1] Rodriguez-Fornells, Rotte, Heinze, Nösselt and Münte (2002) suggest that bilinguals (in their case, Spanish/Catalan speakers) can strategically control the process of mapping spelling to sound when they are required to ignore words in the non-target language. However in a situation where words from different languages must be responded to there is no obvious benefit in such strategic control.

Address for correspondence
Dr. David W. Green, Department of Psychology, University College London, Gower Street, London WC1E 6BT, UK.
E-mail: *d.w.green@ucl.ac.uk*

https://doi.org/10.1017/S1366728902003036 Published online by Cambridge University Press

emphasised. Inappropriate mappings might delay but not preclude access to the non-target language.[2]

The above possibilities concern what is required in order to achieve word recognition. A number of lines of evidence argue against word recognition as the source of switch costs. First, words with an orthographic pattern specific to one language speed word recognition (Grainger and Beauvillain, 1987) but do not eliminate switch costs when suitable controls are in place (Thomas and Allport, 2000).

Second, if switch costs arise during word recognition then such costs should not vary with the nature of the lexical decision task individuals perform nor with the nature of the response they must make. However, switch costs in a lexical decision task do depend on the precise nature of the task (von Studnitz and Green, 1997; Thomas and Allport, 2000) and do depend on the precise nature of the response individuals must make (Thomas and Allport, 2000). The switch cost is greater in a language-specific lexical decision task that requires individuals to decide if a string of letters is a real word in one language rather than in another, than in a language-general lexical decision task, that requires individuals to decide if a letter string is a word in either language. But this increased switch cost in the language-specific task is not accompanied by any change in the extent to which lexical access is restricted to one language rather than to another. In rejecting nonwords, individuals remain sensitive to orthographic legality of the letter string in the non-target language (cf. Alternberg and Cairns, 1983).

Next, the cost of switching in a lexical decision task interacts with the nature of the response to be made (Thomas and Allport, 2000). We can distinguish two kinds of response. A response might be the same as on the previous trial (a repeated response) or it might be different from the previous trial (a nonrepeated response). Typically, individuals are faster to make repeated compared to nonrepeated responses (Rabbit, 1968; Pashler and Baylis, 1991) but this pattern reverses when individuals are required to switch between two different tasks (Rogers and Monsell, 1995). Thomas and Allport (2000) showed that reaction times (RTs) were considerably slower on switch trials than on non-switch trials for a repeated response but somewhat *faster* than non-switch trials for a nonrepeated response. This interaction arose both for the language-specific, and for the language-general lexical decision tasks.

The effects of task on switch costs and the interaction of language switching and type of response support the notion that switch costs arise outside the bilingual lexicon (von Studnitz and Green, 1997; Green, 1998a, b). Individuals must map outputs from the lexicon onto suitable responses. Green (1998a, b) termed the structures involved "language task schemas". According to this view, language schemas compete to map signals onto responses. In a language-specific lexical decision task two schemas (e.g., one for German and one for English) are in competition. For instance, in the case of a German-specific lexical decision task, real German words must be mapped onto a "yes" response. Any nonwords or real English words must be mapped onto a "no" response. A switch in language requires a remapping. But why should there be a cost in switching between languages in a language-general lexical decision task where only one schema is required? Conceivably, individuals continue to respond by assessing whether or not the stimulus letter string is a word in a given language perhaps because identification of the language of input is normally critical to activating relevant syntactic procedures (Grosjean, 1997). The cost is less than in the language-specific lexical decision task because the competition between the schemas is less.

Alternatively, two processes might contribute to these effects: a remapping process – critical in a language-specific lexical decision task – and a cueing process. A language switch may naturally cue a change in response and lead to faster responses for nonrepeated responses and slower responses for repeated responses (cf. the notion of paralexical decision processes – Monsell, Doyle and Haggard, 1989, and the concept of multiple read-out – Grainger and Jacobs, 1996). In consequence, language switching will interact with response type even if a single schema is involved. An overall switch cost will also be observed if the RT cost of a language switch with no change in response is greater than the RT benefit associated with remaining in the same language and repeating a response.

Finally, research on semantic categorization in two languages seems to rule decisively against word recognition as the source of switch costs. Semantic categorization arguably offers a better test of the source of switching costs than a lexical decision task because it is closer to the normal use of language in that meaning rather than lexical form is central. Any delay in word recognition consequent on a switch of

[2] The use of both languages may also increase competition between the lexical representation of words in the two languages. Dijkstra et al., 1999 showed that word recognition in English is slower for interlingual homographs (words that have the same visual form but a different meaning in two languages) that are phonologically similar in English and Dutch. Such similarity, they suggested, increased competition at the lexical level and so delayed word recognition. It is possible that language switching increases such an effect.

https://doi.org/10.1017/S1366728902003036 Published online by Cambridge University Press

language should delay recognition of word meaning. In fact, existing research suggests that there is no RT cost in switching between languages (Caramazza and Brones, 1980; Shanon, 1982; Potter, So, von Eckhardt and Feldman, 1984) at least in relatively fluent bilinguals (Dufour and Kroll, 1995). However, this null result potentially undermines the task schema account too. As noted above, a change in language may cue a change in response. In line with our account of the lexical decision data, we should find an overall switching cost and an interaction of language switching and response type in a semantic categorization task even though only a single schema maps semantic decisions onto responses.

Prior studies do not refute these predictions because they were not specifically designed to examine the costs of language switching. They used language switching as a means to address the question of whether or not semantics is shared between languages. The aim of the present study is to provide a suitable test and so test the generality of the task schema account.

We asked participants to consider the same categorical question on each trial and so avoided the need to shift the designated category from trial to trial and to present a category name along with an instance as in previously published studies. Participants judged whether a word referred to a living thing (an ''animate'' entity) or to non-living thing (an ''inanimate'' entity). The animacy of an entity appears to be a basic property (e.g., Spelke, Breinlinger, Macomber and Jacobson, 1992) and applies to a wide range of entities. Second, language switch and language non-switch trials were interleaved rather than presented in separate blocks of trials so as to avoid any motivational differences between these two types of trial. We adopted the alternating runs paradigm of Rogers and Monsell (1995), such that the language of input changed regularly on every other trial (see also von Studnitz and Green, 1997; Thomas and Allport, 2000). Third, in order to test whether or not switch costs arise outside the bilingual lexicon, we assessed the effect of language switching on repeated response trials (i.e., where the required response on the current trial, e.g., ''animate'', is the same as that required on the prior trial, i.e., ''animate'') and on nonrepeated response trials (i.e., where the response required on the current trial, e.g., ''animate'' is different from the response required on the prior trial, i.e., ''inanimate'').[3]

[3] Mason (1994) used the alternating runs paradigm to assess the costs of language switching in an animacy categorization task, and reported no cost of language switching, but did not examine the effect of language switching when it coincided with repeated, or with nonrepeated, responses.

A signal about a change in language might arise automatically (cf. Kroll and Dijkstra, 2002). At an early stage of processing, orthography could signal a change in language given that some spellings are unusual or illegal in the other language (e.g., SHIRT and LÄUSE are illegal for German and English, respectively). At a later stage of processing, a signal could arise after contact with the word's lexical representation. For instance, information on language membership may be linked to the word's lemma (see Poulisse, 1997; Green, 1998a; Hermans, 2000 for pertinent discussions). Clearly, a signal of language change can only affect response times if it is available before a decision has been reached. A decision about the animacy of a word's referent requires the retrieval of semantic information and a comparison against some internal criterion for animacy. This process is likely to be lengthier than one associated with determining a language change. However, as individuals come to respond more quickly, signals of language change may cease to affect response times. We explored this idea by contrasting performance in the first half and in the second half of the experiment. The second half comprised translation equivalents of items presented in the first half. We expected that RTs would be faster in the second half than in the first half on two grounds. First, individuals may become more practised in the task. Second, although cross-language repetition priming effects are transient (de Groot and Nas, 1991; Forster and Davis, 1994), participants may speed their responses in the second half of the experiment by retrieving their prior decision about the item.

Method

Participants

Twenty-four right-handed German/English bilingual students from University College London (12 males, 12 females), with an average age of 26.2 (SD = 4.7) years and with normal or corrected vision, participated voluntarily in the study. Participants were remunerated at the standard rate.

Language background and language proficiency

All participants acquired German as their first language (L1) and learned English as a second language (L2) at school, starting at a mean age of 10.1 (2.3) years. They achieved a sense of fluency in English at a mean age of 17.8 (6.9) years but all considered themselves as more proficient in German (scores of 58.3 vs. 41.7 for German and English, respectively, on a comparative rating scale totalling 100). Partici-

https://doi.org/10.1017/S1366728902003036 Published online by Cambridge University Press

Table 1. *Mean values (SD in brackets) of semantic properties (category typicality, Battig and Montague, 1969 (B&M); subjective animacy ratings – see text), word frequencies (Kučera and Francis, 1967 (K&F), for English (E); Ruoff (1990) for German (G); Baayen et al., 1993 (CELEX) for English and German) and word length (letters and syllables), as a function of language, animacy and typicality.*

Language	Animacy	Typicality	Semantic properties		Word frequency		Word length	
			B&M typicality rating	Animacy rating	K&F	CELEX (E)	Letters	Syllables
English	Animate	High	201 (94)	7.2 (0.7)	5 (5)	6 (5)	5 (1)	2 (1)
		Low	25 (17)	6.9 (0.8)	6 (13)	4 (4)	5 (2)	2 (1)
	Inanimate	High	238 (105)	1.7 (0.4)	27 (33)	30 (47)	5 (1)	1 (1)
		Low	18 (15)	1.9 (1.0)	26 (33)	19 (21)	6 (1)	2 (1)
					Ruoff	CELEX (G)		
German	Animate	High	–	6.9 (1.0)	12 (27)	2 (2)	6 (2)	2 (1)
		Low	–	6.7 (1.0)	12 (22)	2 (3)	6 (2)	2 (1)
	Inanimate	High	–	1.9 (0.6)	15 (26)	12 (19)	6 (1)	2 (1)
		Low	–	2.0 (0.9)	15 (19)	15 (41)	6 (1)	2 (1)

pants had used English for daily communication on average for the past 5.8 (5.1) years. On the LLEX tests (Meara, 1994), for which maximum scores are 100 per test, the means were 90.8 (5.8) for English, and 91.0 (7.8) for German. Reported current use of English (for the past month) averaged 76.5% (13.6).

Design

The repeated measures design comprised five factors each with two levels, viz.: session halves (first half, second half), language (English/German), language switching (language switch/language non-switch), animacy (animate/inanimate item) and response type (repeated response – where there was no change in the required response from the prior trial – and nonrepeated response – where there was a change in the required response from the prior trial). The animate and inanimate categories each comprised eight subcategories (with four typical and four less typical exemplars in each), yielding a set of 128 items.

Language switches occurred on alternate trials as per the alternating runs paradigm (Rogers and Monsell, 1995). The first half of the session comprised the full item list of 128 items subdivided into two blocks of 64 trials. A filler trial at the start of each block designated the first measured item as a switch trial or a non-switch trial. The second half of the session consisted of translations of the words seen in the first.

Over the entire experiment each participant was shown the full set of materials in pseudo-random order in both English and German, and, over all participants, each item, by its English name and by its German name, was presented on a language switch and on a language non-switch trial in the context of a repeated, and in the context of a non-repeated, response. Items that appeared in one language in the first half of the session appeared in the other language in the second half of the session and vice versa. Adjacent items were from different subcategories and were not obvious associates.

Materials

Words referring to 64 living things (animate entities) and 64 non-living things (inanimate entities) were selected using the Battig and Montague (1969) norms. The four typical instances in each subcategory had a frequency of mention on average of 100 or more, whereas the less typical instances had a frequency of mention of less than 50. For English, the exemplars of each type for each subcategory were matched in frequency (Kučera and Francis, 1967) as well as in letter length and in syllable length – see Table 1. Ruoff (1990) and CELEX (Baayen, Piepenbrock and van Rijn, 1993) values are given for comparative purposes. CELEX data were available only after the experiment was completed, thus frequency matching was based on Kučera and Francis (1967) norms only. On average, German words were longer and of lower CELEX frequency, though they did not differ in terms of the rated animacy of their referents (see below).

https://doi.org/10.1017/S1366728902003036 Published online by Cambridge University Press

Table 2. *Examples of stimuli in two categories.*

English name	B&M	Mean animacy rating (E)	K&F	CELEX (E)	Letters/ syllables	German name	Mean animacy rating (G)	Ruoff	CELEX (G)	Letters/ syllables
Animate: fruits										
Apple	429	6.3	9	18	5 / 2	Apfel	5.6	27	6	5 / 2
Pear	326	6.3	6	2	4 / 1	Birne	5.8	5	0	5 / 2
Grape	247	6.4	3	2	5 / 1	Traube	5.7	14	1	6 / 2
Plum	167	6.5	1	3	4 / 1	Pflaume	5.9	2	0	7 / 2
Berry	11	6.2	9	2	5 / 2	Beere	5.5	19	0	5 / 2
Fig	16	6.0	12	4	3 / 1	Feige	5.8	0	2	5 / 2
Blueberry	21	6.2	1	2	9 / 2	Blaubeere	5.6	0	1	9 / 3
Raisin	16	6.1	1	1	6 / 2	Rosine	5.5	0	1	6 / 3
Inanimate: toys										
Doll	285	2.5	10	17	4 / 1	Puppe	2.2	4	5	5 / 2
Truck	125	2.0	57	25	5 / 1	Laster	2.0	13	2	6 / 2
Train	113	1.9	82	68	5 / 1	Zug	2.1	66	12	3 / 1
Blocks	98	1.1	37	24	6 / 1	Klötze	1.0	1	0	6 / 2
Football	13	1.9	36	32	7 / 2	Fußball	1.7	24	37	7 / 2
Swing	10	2.0	24	14	5 / 1	Schaukel	2.3	3	0	8 / 2
Cards	12	1.5	36	27	5 / 1	Karten	2.1	35	28	6 / 2
Ship	8	1.6	83	44	4 / 1	Schiff	1.9	27	52	6 / 1

B&M = Battig and Montague (1969) word typicality norms; K&F = Kučera and Francis (1967) word frequency norms for English; CELEX (Baayen et al., 1993) = word frequency norms for English (E) and German (G); Ruoff (1990) = word frequency norms for German; animacy ratings given by English raters (E) and German raters (G), respectively.

Animacy ratings

To ensure that all items fell clearly into either the animate or inanimate category (and that this was comparable for German and English word meanings), an animacy rating study was undertaken. Twenty-two English monolinguals (undergraduates in the Department of Psychology, University College London) received randomised lists of the word stimuli and rated the referents of each word on an eight-point animacy scale, with the end points, suitably rotated over participants, marked as animate (i.e., living) and as inanimate (i.e., non-living). A further twenty-two German/English bilinguals living in London, and from a variety of academic backgrounds, performed the same task in German – none participated in the experiment. For English raters, the average rating for animate items (with higher values indicating higher-rated animacy) was 7.1 (0.8). For inanimate items, the average rating was 1.8 (0.8). German/English bilingual raters showed the same disjoint pattern: M (animate) = 6.9 (1.0), M (inanimate) = 1.9 (0.8). Ratings also showed no effect of typicality for either group (one-way ANOVAs yielded $Fs < 1$), indicating that typicality did not interfere with the crucial animacy categories. We therefore pooled over this factor in subsequent analyses. Table 2 presents sample materials from one animate and one inanimate category. A full list of materials may be obtained from the authors.

Apparatus and procedure

Participants, tested individually in a small cubicle equipped with an IBM compatible PC, were given verbal instructions in German followed by written instructions in English. This was done to ensure that the lexical representations of both languages were active before starting the experiment (e.g., Grosjean, 2001). Participants were informed that they would be shown words, some English and some German, all of which were the names of living or non-living entities. Their task was to decide as quickly and as accurately as possible whether the entity named was animate or inanimate. Participants were told that language was irrelevant to their task. They were also told that they might notice translations (i.e., that a word would be shown in one language at one point and in the other

https://doi.org/10.1017/S1366728902003036 Published online by Cambridge University Press

language at some other point), but that this should not concern them. This was done to stop participants from being surprised at seeing translations, and possibly thinking about this, as this could delay their RTs. Words appeared singly in white uppercase letters centred on a (black) computer screen. The uppercase format allowed the same visual presentation for English and German words, given the need to capitalise the initial letter of German nouns.

A two-button response box registered responses. Participants pressed the '+' button ('animate') with the middle finger of their right hand, and the '−' button ('inanimate') with the index finger of the same hand. Prior literature indicates that individuals reach animate decisions more quickly than inanimate decisions (e.g., de Groot, 1990). We were not concerned with the overall size of this effect and sought to render the decisions more equivalent by assigning the slower decision (inanimate) to the index finger. Response times were measured from the onset of the word's presentation to a button press. Errors elicited a short beep. One second after the button-press, the next word appeared on the screen and remained on the screen until terminated by a response.

After a practice block of 16 trials, none of which appeared in the experimental trials, individuals paced themselves through the experimental trials, counting slowly to twenty between blocks, before initiating the next block of trials. The experiment took on average 20 minutes to complete.

Following the experiment, participants carried out the two LLEX tests (Meara, 1994) and filled in the language background questionnaire. The LLEX tests required an additional 30 minutes. Language background questionnaires were completed at home and returned to the experimenter a week later.

Results

We analysed response times (RTs) for correct responses and error rates. The RT data excluded latencies +/−2 standard deviations (SD) from each participant's cell mean (a total of 4.0% trials of the whole data set).[4] Errors totalled 7.7%, and error analyses were based on the arcsine transformed error proportions. Inspection of the RT and error data revealed no speed/accuracy trade-off. We subjected the RT and the error data to an analysis of variance comprising five factors: session halves, the items' language, language switching, animacy of the item and response type (i.e., whether or not the required response was the same as on the prior trial, a repeated response, or differed from the prior trial, a nonrepeated response). The mean values presented in the text and tables are the means of means (from the correct subject data). Standard deviations are presented in brackets. We describe results critical to the main thrust of the paper first.

The main effect of language switching was significant [$F_1(1,23) = 6.01$, $MSE = 7441.73$, $p < 0.025$; $F_2(1,126) = 14.34$, $MSE = 15448.93$, $p = 0.001$]. Overall, RTs were slower on language switch trials [M = 836ms (165)] than on language non-switch trials [M = 820ms (142)]. In terms of errors, there was no effect of language switching [$p > 0.3$ and $F < 1$]. However, the main effect of language switching was qualified by two interactions. First, there was a marginal two-way interaction with session halves [$F_1(1,23) = 3.44$, $MSE = 6465.98$, $p = 0.08$, $F_2 < 1$] such that there was significant switch cost of 26 ms [M = 865 (173) for switch trials and M = 839 (132) for non-switch trials] in the first half of the session [$F_1(1,23) = 7.49$, $MSE = 8673.87$, $p < 0.025$; $F_2(1,126) = 9.76$, $MSE = 13465.21$, $p < 0.01$] but, contrary to the word recognition account, no switch cost in the second half of the session, Fs < 1, [M = 806 (152) and M = 801 (150) for switch and non-switch trials, respectively]. Second, as predicted, there was a significant three-way interaction of language switching with response type and session halves: $F_1(1,23) = 14.30$, $MSE = 6809.67$, $p = 0.001$; $F_2(1,126) = 7.25$, $MSE = 11541.66$, $p < 0.01$. The relevant data are presented in Table 3.

Critically, and consistently with prediction, in the first half of the session there was a significant interaction between language switching and response type [$F_1(1,23) = 56.05$, $MSE = 3819.64$, $p < 0.001$; $F_2(1,126) = 20.31$, $MSE = 13080.09$, $p < 0.001$]. There was no language switching by response type interaction in the second half of the session [$F_1 < 1$; $F_2(1,126) = 1.05$, $p > 0.3$].

With respect to the predicted contrasts in the first half of the session, for repeated responses, language switch trials were slower than language non-switch trials: $F_1(1,23) = 37.13$, $MSE = 15833.98$, $p < 0.001$;

[4] For trials in English, a coding error, discovered retrospectively, decreased the size of two of the cells (response change, for both animate and inanimate items) and enlarged two others (response repetition, for both animate and inanimate items). Each participant should have received each trial type eight times per half-session. In the reduced cases, they received it six times and in the increased cases ten times per half-session. An analysis using the same factors as the reported analysis, but in which data points were (randomly) eliminated from all complete cells, rendering all cells the same size, showed the same pattern. A further analysis based on the median correct RT, including outliers, also yielded the same pattern. We also analysed mean and median correct RTs excluding trials following a wrong response. These followed the same pattern as the analyses excluding only errors themselves. We are confident therefore that the results reported are robust.

https://doi.org/10.1017/S1366728902003036 Published online by Cambridge University Press

Table 3. *Mean correct RT (ms; SD in brackets) and error percentage per cell, as a function of language switching, response type and session halves.*

Response type	Language switching			
	Language non-switch		Language switch	
	Mean RT (*SD*)	Error %	Mean RT (*SD*)	Error %
Session: first half				
Repeated response	837 (134)	10.4	910 (196)	6.8
Nonrepeated response	842 (131)	8.2	821 (133)	6.8
Session: second half				
Repeated response	810 (164)	7.3	817 (152)	8.1
Nonrepeated response	793 (136)	6.1	795 (153)	8.1

$F_2(1,126) = 23.33$, MSE = 16514.37, $p < 0.001$). A switch in language increased mean reaction time by 73ms [M = 910ms (196) and M = 837ms (134) for language switch and language non-switch trials, respectively]. However, for nonrepeated response, language switch trials were *faster* than language non-switch trials by 21ms [M = 821ms (133) and M = 842ms (131) for language switch trials and language non-switch trials, respectively]. This effect was significant by subjects but not by items: $F_1(1,23) = 18.32$, MSE = 4685.81, $p < 0.001$; $F_2(1,126) = 1.17$, $p > 0.20$. In the case of errors, there was no interaction between language switching and response type [ps > 0.1 or Fs < 1]. For the second half of the session neither predicted contrast was significant on either RT or on errors (all Fs< 1 or ps > 0.05).

Turning now to the other results. There were significant effects of all the other main factors on the RT analysis. Predictably, RTs were slower in the first half of the session than in the second half of the session [$F_1(1,23) = 25.44$, MSE = 17767.00, $p < 0.001$; $F_2(1,126) = 117.91$, MSE = 12203.85, $p < 0.001$]. The mean difference was 48ms [M = 852ms (154) vs. M = 804ms (151)]. There was no difference in error rate [$F_1(1,23) = 1.96$, $p > 0.1$; $F_2 < 1$]. The RT difference is attributable to at least two factors: practice in the task and/or a benefit attributable to the fact that items presented in the second half of the session were translation equivalents of items presented in the first half.

As expected, given their self-ratings, and prior research on reading words aloud (Kroll and Stewart, 1994; La Heij, Hooglander, Kerling and van der Velden, 1996), participants responded more quickly to words in their native language, German, than to words in English [800ms (142) vs. 856ms (162)], though this difference was significant only by subjects [$F_1(1,23) = 35.70$, MSE = 17071.94, $p < 0.001$; $F_2(1,126) = 1.29$, $p > 0.02$]. Participants also made fewer errors on German words [4.7% vs. 10.8%; $F_1(1,23) = 30.79$, MSE = 0.11, $p < 0.001$; $F_2(1,126) = 26.20$, MSE = 0.02, $p < 0.001$].

In line with the literature (e.g., de Groot, 1990), animate decisions were faster than inanimate decisions [M = 798ms (150) vs. M = 858ms (153)], but this effect was significant only by subjects [$F_1(1,23) = 35.64$, MSE = 19166.20, $p < 0.001$; $F_2 < 1$]. Animacy exerted no effect on accuracy [$p > 0.1$ by subjects and by items]. The RT main effect of animacy entered three two-way interactions: one with session halves, one with language and one with response type that we consider after reporting the main effect of response type. The interaction of animacy and session halves was significant only in the subject analysis [$F_1(1,23) = 6.96$, MSE = 7827.27, $p < 0.025$; $F_2 < 1$]; for errors there was no such interaction [Fs < 1]. RTs to animate words were faster by 76ms in the first half of the session [M = 814ms (154) vs. M = 890ms (144)] but by only 46ms in the second half of the session [M = 782ms (144) vs. M = 825ms (155)]. Presumably this reduction reflects the opportunity for individuals to retrieve their previous decisions on translation equivalents to the benefit of the more difficult decisions. The interaction of animacy and language in the RT analysis was again significant only by subjects [$F_1(1,23) = 33.31$, MSE = 6314.97, $p < 0.001$; $F_2(1,126) = 1.08$, $p > 0.3$]. Individuals were much slower to respond to animate items in English and the difference in RT between animate and inanimate words was smaller for English words, 23ms on average [M = 843ms (163) and 869ms (159), respectively] than for German words, 93ms on average [M = 753ms (119) and M = 846ms (147), respectively]. Language and animacy also interacted in terms of errors [$F_1(1,23) = 17.36$, MSE = 0.07, $p < 0.001$; $F_2(1,126) = 9.56$, MSE = 0.02, $p < 0.01$].

https://doi.org/10.1017/S1366728902003036 Published online by Cambridge University Press

Individuals were not only slower responding to animate items in English, they were also more error prone. On English trials, there were more errors for animate (14.2%) than for inanimate items (7.9%) whereas on German trials there was little difference: 3.8% for animate items compared to 5.5% for inanimate items. This pattern of performance suggests that individuals may commonly reach an inanimate decision having failed to find sufficient animacy features by some internal deadline. Individuals may be more likely to reach an internal deadline for response in the case of English words.

Overall, RTs were 31ms *slower* on repeated response trials compared to nonrepeated response trials [M = 843ms (167) vs. M = 812ms (139); $F_1(1,23) = 19.13$, MSE = 9632.32, $p < 0.001$; $F_2(1,126) = 8.94$, MSE = 11021.67, $p < 0.01$], but this effect was qualified by a two-way interaction with animacy, a two-way interaction with session halves and, critically, the three-way interaction with language switching and session halves described above. There were no other significant interactions (Fs < 1, *p*s > 0.10).

The two-way interaction of response type and animacy on RT was significant only in the analysis by subjects [$F_1(1,23) = 16.54$, MSE = 5770.00, $p < 0.001$; $F_2(1,126) = 1.48$, $p > 0.2$]. For animate items, there was little effect of response change – the means for nonrepeated responses and repeated responses were, respectively, M = 794ms (134) and M = 802ms (164). In contrast, for inanimate items, the respective means were: 831ms (142) and 884ms (160). There was no interaction between animacy and response type in the error analysis [F < 1]. This RT pattern may reflect a general tendency to assume that responses alternate. It seems that overcoming such a tendency takes longer when information is less readily available, as in the case of inanimate responses.

The two-way interaction between response type and session halves was significant both by subjects and by item [$F_1(1,23) = 5.25$, MSE = 4495.61, $p < 0.05$; $F_2(1,126) = 6.02$, MSE = 11667.70, $p < 0.025$], The difference between repeated and nonrepeated responses averaged 42ms in the first half of the session [M = 873ms (171) and M = 831ms (132), respectively], but only 20ms in the second half of the session [M = 814ms (157) and M = 794ms (144), respectively]. This interaction, like the overall effect of response type, is most plausibly attributable to the changes in the way individuals respond to language switching.

For errors, there was no effect of response type [$F_1(1,23) = 3.16$, MSE = 0.06, $p > 0.05$; $F_2(1,126) = 1.20$.], and no interaction with session [$p > 0.1$ and F < 1].

Discussion

In contrast with earlier studies (e.g., Caramazza and Brones, 1980; Dufour and Kroll, 1995), which were not specifically designed to assess switching costs, we do find an effect of language switching in a semantic categorization task. But, contrary to the hypothesis that word recognition is the source of such costs, switch costs were restricted to novel items (i.e., items in the first half of the experimental session) and, for these items, the effects of language switching interacted with response type. If individuals needed to reactivate temporarily inhibited lexical representations on a switch trial or to increase the activation of relevant spelling to sound correspondences, then switching language should yield a cost for all items independent of response type. Instead, the results support the notion that costs arise in the course of mapping decisions onto responses.

Language change is the key factor, not simply congruence between a language change and a response change. Inspection of the RT data in Table 3 shows that there is substantial slowing in RT when there is a change in language but no change in the response required compared to when language changes and there is a change in the response required (M = 910 ms vs. M = 821ms, respectively). In contrast, when there is no change in language there is little difference in RT between repeated and nonrepeated responses (M = 837ms vs. M = 842ms, respectively).

The lack of interaction between language switching and response type in the second half of the experimental session (when individuals were responding to translations of items presented in the first half of the session) is open to a number of interpretations. Faster RTs limit the opportunity for a change cue to exert its effect on response times. However, individuals may also come to ignore, or to place less weight on, language change given that it is not predictive of a response change. This possibility could be tested by examining changes in the interaction over successive blocks of novel items.[5]

The interaction we observed may be a task-specific adaptation. At a minimum, this possibility acknowledges that language change can function as a cue to change response and so must be detected early enough to affect responses. However, contrary to our position, in the normal course of events, language change carries no functional significance. It follows that there would be no interaction of language

[5] Conceivably, translation priming effects – typically transient when tapped using a lexical decision task – in fact, endure over longer time intervals when tapped using a semantic decision task (cf. Grainger and Frenck-Mestre, 1998).

https://doi.org/10.1017/S1366728902003036 Published online by Cambridge University Press

switching and response type when language switching is a relatively rare, but not totally unexpected, occurrence in a sequence of trials.

We contend that language switching is of functional significance because of procedures (phonological/syntactic) normally evoked by a word. We cannot know, and do not claim, that language switching is unique in its interaction with the type of response. Other stimulus changes (e.g., font size, colour, location) may also exert at least short-lived effects. Certainly, switching script induces a response cost (Shafiullah and Monsell, 1999) but whether or not script changes interact with the type of response is unknown. What we do contend is that our results are consistent with the *a priori* functional significance of language changes.

If language can function as a cue we should expect it to do so in other tasks. Miller and Kroll (in press) provide supportive data. In a single word translation task, individuals are delayed by the presence of a word semantically related to the target but only when the interfering word is in the language of output. When the target and the interfering word are in the same language, there is no interference. Miller and Kroll propose that in the context of a translation task the language cue blocks same language competitors from interfering with word production.

In sum, we have shown that with novel materials (i.e., prior to item repetition) the effects of a switch in language depend upon the nature of the response. Our results therefore extend the findings of Thomas and Allport (2000) from the context of a lexical decision task to the context of a semantic categorization task, where the language of input is manifestly irrelevant to the task, and where the critical information is of a semantic rather than of a lexical nature. We found, in addition, that these response cueing effects of a language switch were eliminated when individuals were responding to items previously presented in another language. In such circumstances, there was also no overall switch cost.

Conclusion

Is there a cost in switching language? Our data are consistent with the notion that access to the lexicon is non-selective and that costs arise because of the way bilinguals are asked to use their languages. We have established this point now in a lexical decision task and in a semantic task. Conceivably, there are no costs in ordinary circumstances that eschew "arbitrary, task-specific components" (Thomas and Allport, 2000, p. 62). However, if language is a potent cue to language-specific recognition procedures (see Grosjean, 1997) then costs may exist when such expectations are not met. We suggest, then, that the supposition that such costs are irrelevant during natural speech processing or in normal reading is just that – mere supposition.

References

Altenberg, E. & Cairns, H. (1983). The effects of phonotactic constraints in lexical processing in bilingual and monolingual subjects. *Journal of Verbal Learning and Verbal Behavior*, 22, 174–188.

Baayen, H., Piepenbrock, R. & van Rijn, H. (1993). *The CELEX lexical database* (CD-ROM). Philadelphia, PA: University of Pennsylvania, Linguistic Data Consortium.

Battig, W. F. & Montague, W. E. (1969). Category norms of verbal items in 56 categories A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology*, 80(3, Pt. 2), 1–46.

Brysbaert, M., van Dyck, G. & van de Poel, M. (1999). Visual word recognition in bilinguals: Evidence from masked phonological priming. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 137–148.

Caramazza, A. & Brones, I. (1980). Semantic classification by bilinguals. *Canadian Journal of Psychology*, 34, 77–81.

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

Dalrymple-Alford, E. C. (1985). Language switching during bilingual reading. *British Journal of Psychology*, 76, 111–122.

Dijkstra, A. (Ton), de Bruijn, E., Schriefers, H. & ten Brinke, S. (2000). More on interlingual homograph recognition: Language intermixing versus explicitness of instructions. *Bilingualism: Language and Cognition*, 3, 69–78.

Dijkstra, A. (Ton), Grainger, J. & van Heuven, W. J. B. (1999). Recognizing cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language*, 41, 496–518.

Dijkstra, A. (Ton), van Jaarsfeld, H. & ten Brinke, S. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language and Cognition*, 1, 51–66.

Dufour R. & Kroll, J. F. (1995). Matching words to concepts in two languages: A test of the concept mediation model of bilingual representation. *Memory & Cognition*, 23, 166–180.

Forster, K. I. & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 10, 680–698.

Goswami, U., Ziegler, J. C., Dalton, L. & Schneider, W. (2001). Pseudohomophone effects and phonological recoding procedures in reading development in

https://doi.org/10.1017/S1366728902003036 Published online by Cambridge University Press

English and German. *Journal of Memory and Language*, 45, 648–664.

Grainger, J. & Beauvillain, C. (1987). Language blocking and lexical access in bilinguals. *Quarterly Journal of Experimental Psychology*, 39A, 295–319.

Grainger, J. & Frenck-Mestre, C. (1998). Masked priming by translation equivalents in proficient bilinguals. *Language and Cognitive Processes*, 13, 601–623.

Grainger, J. & Jacobs, A. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review*, 103, 518–565.

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

Green, D. W. (1998b). Schemas, tags and inhibition. Reply to commentators. *Bilingualism*: *Language & Cognition*, 1(2), 100–104.

de Groot, A. M. B. (1990). The locus of the associative-priming effect in the mental lexicon. In D. A. Balota, G. B. Flores d'Arcais & K. Rayner (eds.), *Comprehension processes in reading*, pp. 101–122. Hillsdale, NJ: Lawrence Erlbaum Associates.

de Groot, A. M. B., Delmaar, P. & Lupker, S. J. (2000). The processing of interlexical homographs in translation recognition and lexical decision: Support for non-selective access to bilingual memory. *Quarterly Journal of Experimental Psychology*, 53A, 397–428.

de Groot, A. M. B. & Nas, G. J. L. (1991). Lexical representation of cognates and non-cognates in compound bilinguals. *Journal of Memory and Language*, 30, 90–123.

Grosjean, F. (1997). Processing mixed languages: Issues, findings and models. In A. M. B. de Groot & J. F. Kroll (eds.), *Tutorials in bilingualism: Psycholinguistic perspectives*, pp. 225–254. Mahwah, NJ: Lawrence Erlbaum Associates.

Grosjean, F. (2001). The bilingual's language modes. In J. Nicol (ed.), *One mind, two languages: Bilingual language processing*, pp. 1–22. Oxford: Blackwell.

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

Hernandez, A. E., Dapretto, M., Mazziotta, J. & Bookheimer, S. (2001). Language switching and language representations in Spanish-English bilinguals: An fMRI study. *NeuroImage*, 14, 510–520.

Jackson, G. M., Swainson, R., Cunnington, R. & Jackson, S. R. (2001). ERP correlates of executive control during repeated language switching. *Bilingualism: Language and Cognition*, 4, 169–178.

Jared, D. & Kroll, J. F. (2001). Do bilinguals activate phonological representations in one or both of their languages when naming words? *Journal of Memory and Language*, 44, 2–31.

Kroll, J. F., Dietz, F. & Green, D. W. (in preparation). The effects of language switching on picture naming and translation.

Kroll, J. F. & Dijkstra, A. (Ton). (2002). The bilingual lexicon. In R. Kaplan (ed.), *Handbook of applied linguistics*, pp. 301–321. Oxford: Oxford University Press.

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

Kučera, H. & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

La Heij, W., Hooglander, A., Kerling, R. & van der Velden, E. (1996). Nonverbal context effects in forward and backward word translation: Evidence for concept mediation. *Journal of Memory and Language*, 35, 648–665.

Mason, W. T. M. (1994). Language switching in bilinguals: Evidence from the semantic categorization task. Unpublished undergraduate project, University of Cambridge.

Meara, P. (1994). *LLEX: Lingua vocabulary tests v1.4.* Swansea: University of Wales, Centre for Applied Language Studies.

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

Miller, N. A. & Kroll, J. F. (in press). Stroop effects in bilingual translation. *Memory and Cognition*.

Monsell, S., Doyle, M. C. & Haggard, P. N. (1989). Effects of frequency effects in visual word recognition tasks: Where are they? *Journal of Experimental Psychology: General*, 118, 43–71.

Pashler, H. & Baylis, G. (1991). Procedural learning: II., Intertrial repetition effects in speeded-choice tasks. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 17, 33–48.

Potter, M. C., So, K.-F., von Eckhardt, B. & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and fluent bilinguals. *Journal of Verbal Learning and Verbal Behavior*, 23, 23–38.

Poulisse, N. (1997). Language production in bilinguals. In A. M. B. de Groot & J. F. Kroll (eds.), *Tutorials in bilingualism: Psycholinguistic perspectives*, pp. 201–224. Mahwah, NJ: Lawrence Erlbaum Associates.

Rabbit, P. M. A. (1968). Repetition effects and signal classification strategies in serial choice response tasks. *Quarterly Journal of Experimental Psychology*, 20, 232–240.

Rodriguez-Fornells, A., Rotte, M., Heinze, H.-J., Nösselt, T. & Münte, T. F. (2002). Brain potential and functional MRI evidence for how to handle two languages with one brain. *Nature*, 415, 1026–1029.

Rogers, R. D. & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, 124, 207–231.

Ruoff, A. (1990). *Häufigkeitswörterbuch gesprochener Sprache – 2., unveränderte, Auflage*. Tübingen: Max Niemeyer Verlag.

Shafiullah, M. & Monsell, S. (1999). The cost of switching between Kanji and Kana while reading Japanese. *Language and Cognitive Processes*, 14, 567–607.

https://doi.org/10.1017/S1366728902003036 Published online by Cambridge University Press

Shanon, B. (1982). Identification and classification of words and drawings in two languages. *Quarterly Journal of Experimental Psychology*, 34A, 135–152.

Spelke, E. S., Breinlinger, K., Macomber, J. & Jacobson, K. (1992). Origins of knowledge. *Psychological Review*, 99, 605–632.

von Studnitz, R. & Green, D. W. (1997). Lexical decision and language switching. *International Journal of Bilingualism*, 1, 3–24.

von Studnitz, R. & Green, D. W. (2002). Interlingual homograph interference in German-English bilinguals: Its modulation and locus of control. *Bilingualism: Language and Cognition*, 5, 1–23.

Thomas, M. S. C. & Allport, A. (2000). Language switching costs in bilingual visual word recognition. *Journal of Memory and Language*, 43, 44–66.

Received July 19, 2001 Revision accepted May 24, 2002

https://doi.org/10.1017/S1366728902003036 Published online by Cambridge University Press