Second Language Research 18,2 (2002); pp. 137–171

The development of lexical fluency in a second language

Judith F. Kroll *Pennsylvania State University,* **Erica Michael** and **Natasha Tokowicz** *Carnegie Mellon University* and **Robert Dufour** *Dickinson College*

A goal of second language (L2) learning is to enable learners to understand and speak L2 words without mediation through the first language (L1). However, psycholinguistic research suggests that lexical candidates are routinely activated in L1 when words in L2 are processed. In this article we describe two experiments that examined the acquisition of L2 lexical fluency. In Experiment 1, two groups of native English speakers, one more and one less fluent in French as their L2, performed word naming and translation tasks. Learners were slower and more error prone to name and to translate words into L2 than more fluent bilinguals. However, there was also an asymmetry in translation performance such that forward translation was slower than backward translation. Learners were also slower than fluent bilinguals to name words in English, the L1 of both groups. In Experiment 2, we compared the performance of native English speakers at early stages of learning French or Spanish to the performance of fluent bilinguals on the same tasks. The goal was to determine whether the apparent cost to L1 reading was a consequence of L2 learning or a reflection of differences in cognitive abilities between learners and bilinguals. Experiment 2 replicated the main features of Experiment 1 and showed that bilinguals scored higher than learners on a measure of L1 reading span, but that this difference did not account for the apparent cost to L1 naming. We consider the implications of these results for models of the developing lexicon.

I Introduction

How do adult second language (L2) learners acquire lexical representations for L2 words and then connect them to existing representations within the cognitive network for words in the first language (L1) and their meanings? Past research on the development of the bilingual lexicon has considered two alternatives (Potter *et al.*, 1984). According to the word association model, L2 words are mediated via direct connection to their translation equivalents in L1. Alternatively, the concept mediation model proposes that L2 words are connected directly to their

Address for correspondence: Judith F. Kroll, Department of Psychology, 641 Moore Building, Pennsylvania State University, University Park, PA 16802, USA; email: jfk7@psu.edu

 10.1191/0267658302sr201oa

meanings without L1 mediation. Potter *et al.* initially argued for the concept mediation model, even for learners at relatively early stages of acquisition, because a study of translation and picture naming revealed a similar pattern of performance for learners and fluent bilinguals. However, subsequent research suggested a developmental transition from word association to concept mediation. During initial phases of L2 learning, individuals appear to perform in accordance with the predictions of the word association model (Kroll and Curley, 1988; Chen and Leung, 1989). As fluency in L2 increases, there is a corresponding increase in the degree to which meaning can be accessed directly for L2 words, and individuals begin to perform in accordance with the predictions of the concept mediation model (e.g., Talamas *et al.*, 1999). (For reviews of this literature, see Chen, 1992; De Groot, 1993; Kroll, 1993; Kroll and De Groot, 1997; Kroll, 1998; Kroll and Tokowicz, 2001.)[1]

Evidence for transfer from L1 to L2 during second language learning has been documented across many aspects of language processing, including phonology, morphology and syntax (e.g., Kilborn, 1989; Hancin-Bhatt and Nagy, 1994; MacWhinney, 1997). However, the claim that initial reliance on L1 gives way to independent access for L2 at the lexical level has also been challenged by recent evidence for the continuing role of L1 during L2 processing in highly fluent bilinguals. If L1 is simply a temporary crutch to enable L2 to bootstrap its way into the cognitive system, then the same sort of L1 activity that appears to be used by learners might be absent or reduced once an individual becomes a relatively fluent bilingual.

1 The revised hierarchical model

Two lines of research provide evidence which shows that lexical candidates in L1 are active when L2 is processed by highly fluent

[1] We use the term 'bilingual' to refer to anyone who uses a second language at a relatively high level of proficiency. Because most of the research we describe concerns individuals who are late L2 learners, few of them are balanced in their use of the two languages. Furthermore, we do not distinguish between the terms 'semantic' and 'conceptual' for present purposes; however, for recent discussions concerning this distinction, the reader might see Francis, 1999; Pavlenko, 1999. Finally, we use the characterization of individuals as 'fluent' or 'proficient' bilinguals interchangeably. Because the performance measures used in this article focus on production (i.e., word naming or translation), the consequences of developing L2 skill are assessed using the speed and accuracy of speaking. The degree of L2 proficiency is determined by a set of measures, including the number of years of L2 experience and self-assessed ratings for L2. See De Bot and Kroll (in press) and Kroll and Sunderman (in press) for recent tutorial reviews of psycholinguistic approaches that relate the performance of bilinguals and second language learners.

bilinguals. One approach has been to examine the consequences of the developmental sequence described above for the eventual form of the bilingual lexicon. Kroll and Stewart (1994) proposed the revised hierarchical model (Figure 1) to capture the implications of early reliance on L1 for the form of word-to-concept connections. The model merges the word association and concept mediation alternatives into a single model in which the strength of the connections between words in L1 and L2 and concepts is proposed to take on different values. The initial dependence on L1 to mediate access to meaning for L2 words is assumed to create strong lexical-level connections from L2 to L1. However, at a lexical level, the connections from L1 to L2 are not assumed to be particularly strong because there is little need for the learner to use L2 in this way. Likewise, the model assumes that connections between words and concepts are stronger for L1 than for L2.

A number of empirical findings support the predictions of the revised hierarchical model. Kroll and Stewart (1994) demonstrated that the translation performance of highly fluent Dutch–English bilinguals was slower when they translated from L1 to L2 (the direction of translation more likely to rely on conceptual processing) than from L2 to L1 (the direction of translation more likely to take advantage of direct lexical-level connections). More critically, only translation from L1 to L2 was influenced by the presence of semantic information. The absence of semantic effects in the L2 to L1 direction of translation suggests that it was possible for bilinguals to translate directly at a lexical level.

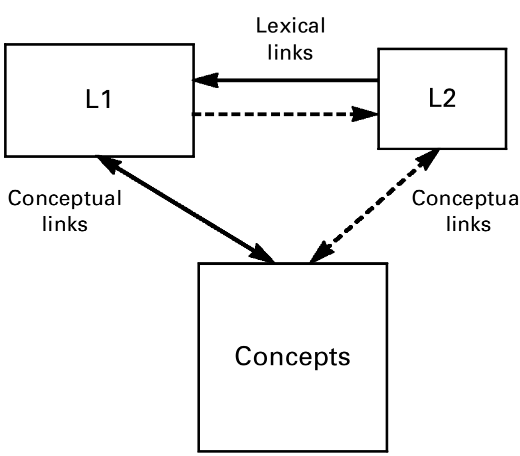

Figure 1 Revised hierarchical model
Source: adapted from Kroll and Stewart, 1994

A subsequent study with relatively fluent English–Spanish bilinguals (Sholl *et al.*, 1995) used a transfer paradigm to test the predictions of the revised hierarchical model. Sholl *et al.* had bilinguals first name pictures in each language, a task hypothesized to require conceptual processing, and then translate words from one language to the other. Some of the words in the translation task had previously been produced in the picture naming tasks and others were new. The question was whether processing the same set of concepts in picture naming would facilitate later translation of those same concepts, relative to the new items. The results showed that initial picture naming transferred only to the L1 to L2 direction of translation, suggesting that only the L1 to L2 direction of translation benefited from prior conceptual processing. Because only L1 to L2 translation was facilitated by prior picture naming, the findings were taken as support for the predictions of the revised hierarchical model.[2]

2 The bilingual interactive activation model

Within the revised hierarchical model, the manifestation of L1 activation during L2 processing consists of direct access to the L1 translation equivalent.[3] However, a second line of research on word recognition suggests that it is not the translation equivalent itself that is active, but rather lexical form relatives. Van Heuven *et al.* (1998) demonstrated that when fluent Dutch–English bilinguals perform lexical decision in L2, words that are orthographic neighbours of the L2 word in both L1 and L2 influence performance. In an extensive series of experiments, Dijkstra and his colleagues have shown that the activation of lexical form information from L1 to L2 extends to the processing of cognates (words that share lexical form and meaning) and interlingual homographs (words that share form but not meaning across languages); e.g., see Dijkstra and Van Heuven, 1998; Dijkstra *et al.*, 1998a; 1998b; Dijkstra *et al.*, 1999; Dijkstra *et al.*, 2000; Van Hell and Dijkstra, in press). Furthermore, these effects appear to be driven in a bottom-up manner so that they are relatively uninfluenced by factors such as instruction about which language is likely to appear (Dijkstra *et al.*, 2000).

To account for the apparently nonselective nature of lexical

[2] For a discussion of evidence that appears contrary to the predictions of the revised hierarchical model, see Kroll and De Groot, 1997; Kroll and Tokowicz, 2001.

[3] Recent articles by Costa *et al.* (1999) and Hermans *et al.* (1998) on language production in bilinguals also show that when fluent bilinguals name pictures in L2, there is activation of L1 candidates – including the translation equivalent – well into the process of lexicalization.

access across languages, a bilingual version of the McClelland and Rumelhart (1981) interactive activation model was proposed (Dijkstra and Van Heuven, 1998; Dijkstra *et al*., 1998a; Van Heuven *et al*., 1998). The central claim of the bilingual interactive activation model (or BIA) is that the bilingual's lexicon is integrated and that lexical access is nonselective such that lexical candidates in both languages are activated whenever the input shares features with alternatives in each of the languages (Figure 2).

Is the activation of lexical form relatives across languages a consequence of fluent bilingualism or – like the claims of the revised hierarchical model concerning the activation of translation equivalents – a residual reflection of the manner in which L2 was acquired? Only a few studies have examined the development of lexical form activation with increasing L2 fluency. Bijeljac-Babic *et al*. (1997) showed that the presence of cross-language inhibitory priming between form-related words was greater the more proficient the bilingual. In a study of translation recognition, in which participants had to decide whether the second of two words was the correct translation of the first, Talamas *et al*. (1999) found that form-similar foils within L2 produced more interference for less fluent than for more fluent bilinguals. Finally, a recent study by Jared and Kroll (2001) showed that whether the naming performance of native English speakers was affected by the presence of neighbours in French that were phonological 'enemies' of the English words depended on their L2 fluency and on the relative activation of L2 within the task. When the participants were relatively proficient in French, the explicit activation of French resulted in a significant naming cost for cross-language enemies. When they were less proficient in French, only the naming of the enemies themselves resulted in an effect on L1. Each of these studies suggests that the nature of lexical form activation across languages changes with increasing proficiency in L2. However, very little information from the existing literature tells us how the influence of L1 on L2 develops over the course of acquisition.

3 The present study

The goal of the present work was to examine the process of lexical access for both L1 and L2 during second language acquisition. In each of the two experiments that are reported, the performance of less and more fluent bilinguals was compared in language production tasks that required simple word naming or single word translation. The revised hierarchical model (Kroll and Stewart, 1994) predicts that tasks that utilize the direct lexical connections that are hypothesized to be in place during early stages of L2 learning should be similar for less and more fluent bilinguals. In contrast, tasks that require conceptual processing and subsequent lexicalization into L2 should be especially sensitive to changes in

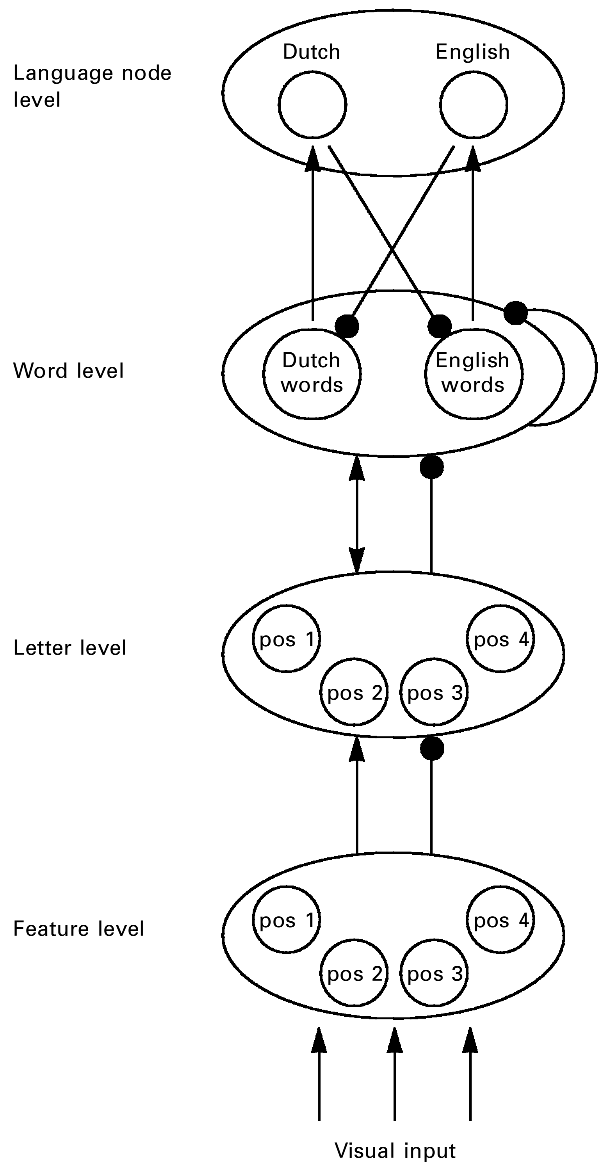

Figure 2 Bilingual interaction activation model
Source: adapted from Dijkstra *et al.*, 1998a

the organization and strength of connections within the lexicon as fluency develops. The experiments reported here were not designed to test the predictions of the BIA model (Dijkstra and Van Heuven, 1998). However, it is also not entirely clear from the existing research on fluent bilinguals how lexical form activation might be expected to change with increasing L2 proficiency. On one hand, it seems certain that early in learning the more dominant language, L1, will influence the less dominant language, L2, more than the reverse. An effect of L2 on L1 would be expected for only the most fluent bilinguals. Exactly how the activation of L1 form relatives influences the processing L2 target words is also likely to depend on assumptions that are made about the nature of the lexicon, and in particular on whether the lexicon is assumed to be integrated from the start, regardless of the number and type of L2 words that are known. In both experiments we examined the issue of lexical form in a preliminary manner by considering how the lexical transparency of cognate translations was processed at different levels of proficiency.[4]

II Experiment 1

In the first experiment we compared the performance of two groups of learners who were native English speakers but differed in their fluency in French. Each participant performed a simple word naming task (i.e., pronounce a visually presented word aloud) and also a single word translation task. In the word naming task, a word was presented on a computer screen and the participant simply had to read the word aloud as quickly as possible in the language in which it was presented. In the translation task a word was presented on a computer screen and the participant was asked to translate the word into the other language. Past research on word naming within a single language suggests that word naming reflects primarily lexical-level processing. With a few exceptions, only when words are long, unfamiliar or ambiguous is word naming affected by semantic variables (e.g., Lupker, 1984; Strain, *et al.*, 1995; La Heij *et al.*, 1996; Pecher, 2001). However, reliable effects of lexical variables that reflect the ease of spelling-to-sound computations (e.g., Jared *et al.*, 1990; Peereman, and Content, 1997) and word frequency (e.g., Forster and Chambers, 1973; Frederiksen and Kroll, 1976) are typically observed in naming. Naming performance is,

[4] Sunderman (in preparation) has directly compared the predictions of the two models for lexical development and specifically addressed the issue of whether it is possible to inhibit L1 activation during L2 learning.

thus, likely to reveal the ease with which learners access lexical information in each of their languages. For L2, the naming task also provides a measure of the difficulty of accessing and producing the phonology.

Unlike word naming, translation requires that words are known and that, at least under some circumstances, their meanings are accessed. Although the processes engaged in translation have been the focus of debate (see the discussion above concerning the claims of the revised hierarchical model that translation from L2 to L1 is lexically mediated), what is clear from past research is that at least translation in the forward direction (from L1 to L2) requires lexicalization in a form that appears similar to other production tasks such as picture naming. That is, when a word in L1 is presented for translation into L2, the meaning of the word is accessed before a lexical candidate is chosen and its phonology specified (e.g., De Groot, 1992; Sánchez-Casas *et al*., 1992; Kroll and Stewart, 1994; Sholl *et al*., 1995; La Heij *et al*., 1996). We return later to the issue of whether the two translation directions have different processing requirements.

1 Method

a Participants: Fifty-nine undergraduate female students at Mount Holyoke College participated in the experiment. They were native speakers of English enrolled in French classes to meet course requirements. Participants received class credit or were paid for their participation. All participants completed a language history questionnaire in which they provided information about their experience in French.

b Materials: A set of 120 French words and their translations was compiled from an elementary French textbook. The words were divided into four lists to be used in each of the four tasks (word naming in L1 and L2 and word translation into L1 and into L2). The lexical properties of the 30 words assigned to each of the four lists are given in Table 1. These values include word frequency in English (Francis and Kučera, 1982) and word length, in number of letters, for both the French words and their English translations.[5] One-way analyses of variance on each of these measures revealed no significant differences across the four lists (all p values $> .10$).

[5] English rather than French word frequency was used because the learners were highly English dominant. Although word frequency is highly correlated across languages, the tabulated values in French are unlikely to provide an accurate estimate of frequency for learners for whom the subjective familiarity of the L2 is low relative to L1.

Table 1 Lexical properties of the 120 words used in Experiment 1

	List 1 (n = 30)	List 2 (n = 30)	List 3 (n = 30)	List 4 (n = 30)	Mean (n = 120)
English word frequency (per million) (Francis and Kučera, 1982)	70.6	69.4	70.6	71.1	70.4
Word length (number of letters) in English	5.7	5.5	5.7	5.5	5.6
Word length (number of letters) in French	6.2	5.9	6.0	6.1	6.1
Context availability (scale of 1 to 7)	5.6	5.7	5.4	5.4	5.5
Familiarity (scale of 1 to 7)	4.9	5.1	4.8	4.9	4.9

The French words were significantly longer than the English words, t (119) = –3.4, $p < .01$. The four lists were equated on two other variables that were assessed in English only. Independent groups of monolingual English speakers rated the English words (on a scale with '1' as low and '7' as high) for context availability (Schwanenflugel *et al.*, 1988) and item familiarity (Gernsbacher, 1984). Context availability is the ease with which a context can be generated for a word and is typically highly correlated with concreteness (e.g., Schwanenflugel *et al.*, 1988). Analyses of variance on these measures also revealed no differences across the four stimulus lists (p values > .10). The order of the four lists was counterbalanced across participants so that an individual viewed a particular word only once in one of the four tasks of the experiment. An additional set of practice items was presented prior to each of the tasks.

c Procedure: Participants performed each of the four tasks, i.e., word naming in English (L1) and French (L2) and word translation from English to French (L1 to L2) and from French to English (L2 to L1). The tasks were blocked and their order was counterbalanced across participants. Words were presented one at a time for 300 ms at the centre of an IBM PC screen and participants were instructed to name them aloud in the language in which they were presented or to speak aloud the translation. The order of presentation of words within a block was randomized individually for each participant. Prior to the presentation of each target word, a fixation cross was presented. Participants were instructed to respond as quickly and accurately as possible, and to indicate when unaware of the correct translation of the presented word by saying 'no'. Reaction time (RT) was recorded by the computer program in milliseconds from the onset of the presentation of the word to the onset of articulation. The onset of spoken responses was recorded by a voice-activated relay connected to a microphone. All responses were tape-recorded for later coding of accuracy. Following the completion of the four experimental blocks, participants were given the language history questionnaire.

2 *Results and discussion*

a Fluency measures: The 59 participants were divided into two fluency groups based on the number of years they had studied French. The less fluent group consisted of 26 learners who had studied French for 5 years or less and the more fluent group consisted of 33 learners who had studied French for more than 5 years. The less fluent group had a mean of 3.42 years of French study whereas the more fluent group had a mean of 8.39 years of French study.

b Performance measures: Analyses of variance were performed separately on mean correct response times (RTs) and percentage accuracy for the performance of the two groups on the word naming and translation tasks. Data were excluded from trials in which errors were made and trials on which RTs exceeded a criterion of 2.5 standard deviations for an individual participant's mean.

We first look at word naming. The results for the word naming task are shown in Table 2, where mean naming latencies (in milliseconds) and percentage accuracy are given for less and more fluent participants naming in English (L1) and French (L2). The analysis on mean RTs revealed significant main effects of fluency group, with the less fluent group slower overall than the more fluent group, F (1, 57) = 7.34, $p < .01$, and language of naming, with slower overall naming latencies for the French than for the English words, F (1, 57) = 92.14, $p < .01$. However, the main effects were also qualified by a significant interaction between fluency group and language of naming, F (1, 57) = 5.50, $p < .05$. Simple effects tests showed that the differences between all conditions were significant (p values < .01) with the exception of the difference in the time to name words in English, which was marginally significant, F (1, 57) = 3.61, $p < .07$. The large cost to L2 naming for the less fluent group is not surprising given their experience in French, but the marginally significant difference in English is surprising, since both groups of participants were native English speakers.

An analysis on the percentage accuracy for naming conditions produced a pattern of results that was similar to that reported for

Table 2 Mean response latencies (in milliseconds) and percentage accuracy (in parentheses) for less and more fluent participants to name words in L1 and L2 and translate words into L1 and L2 in Experiment 1

Task	Language of production		Language difference
	L1	L2	
Word naming			
Less fluent	622 (.96)	773 (.69)	151 (.27)
More fluent	568 (.98)	661 (.83)	93 (.15)
Fluency difference	54 (.02)	112 (.14)	
Translation			
Less fluent	1223 (.65)	1511 (.44)	288 (.21)
More fluent	964 (.75)	1104 (.59)	140 (.16)
Fluency difference	259 (.10)	407 (.15)	

the RTs. Both main effects of fluency group and language were significant for group, $F\ (1, 57) = 21.59$, $p < .01$, and for language, $F\ (1, 57) = 171.82$, $p < .01$, indicating lower accuracy for the less fluent group compared to the more fluent group and lower accuracy for French than for English. Again the main effects were qualified by a significant interaction between fluency group and language, $F\,(1,57) = 13.42$, $p < .01$. A series of simple effects tests showed that all differences between conditions were significant (p values $< .01$) and that the effect of fluency group was even significant for naming in English ($p < .05$). Percentage naming accuracy was lower for conditions in which RTs were long, suggesting that the RT differences reflect difficulty of processing and not a speed–accuracy trade off.

Overall, performance on the naming task shows that individuals become faster to name words in their L2 as proficiency increases. But even for the more fluent group in this experiment, there was still a strong effect of language, with faster latencies to name words in the native language. However, an unexpected result in the present experiment was that the less fluent group was also somewhat slower and less accurate to name words in their L1 than the more fluent group. The result is particularly surprising given that the words in Experiment 1 were presented in separate blocks. There are some claims in the literature that bilinguals are slower to perform speeded tasks than monolinguals (e.g., Ransdell and Fischler, 1987). If the degree of competition across languages increases with increasing expertise in L2, then we might have expected to see a greater cost for the more rather than the less fluent group. The data show the opposite pattern.

There are at least two alternative explanations for why we observe a cost to L1 naming for the less fluent group. One possibility is that there are differences in the cognitive abilities of more and less fluent bilinguals. Not all learners become fluent bilinguals, and there is ample evidence in the psycholinguistic literature to suggest that individual differences in the allocation of cognitive resources affect language processing (e.g., Daneman and Green, 1986; Just and Carpenter, 1992). The difference in L1 naming may thus be a reflection of a self-selection effect; only some individuals – those with high cognitive abilities – go on to become fluent bilinguals. Learners at early stages of L2 acquisition may include a mixture of abilities, whereas more fluent individuals may only include high ability survivors. A second possibility is that initial L2 learning incurs a cost to L1, due to changes in the nature of lexical representation and/or in the control structures that mediate access to the output to the lexical system. We investigate these alternatives in Experiment 2.

Turning now to the translation data, a set of analyses similar to those performed on the naming data was computed for this data. The mean translation latencies (in milliseconds) and percentage accuracy are shown in Table 2. As in the analyses of the naming data, both main effects and the interaction between fluency group and language were significant. The more fluent group was faster to translate than the less fluent group, $F\ (1, 57) = 20.12$, $p < .01$, and performance for both groups was faster in the L2 to L1 (backward) direction than in the L1 to L2 (forward) direction of translation, $F\ (1, 57) = 41.84$, $p < .01$. The interaction between fluency group and direction of translation, $F\ (1, 57) = 5.39$, $p < .025$, reflects the larger translation asymmetry for the less than the more fluent group. Simple effects tests showed that all differences between conditions were significant (all p values $< .01$). The larger translation asymmetry for the less fluent group and the larger fluency difference for the L1 to L2 direction of translation provide support for the predictions of the revised hierarchical model in that proficiency had a larger effect on translation in the direction of translation hypothesized to require conceptual processing.

Like the pattern of naming data, translation accuracy was lower in conditions with longer RTs. Translation performance was more accurate for the more fluent than for the less fluent group, $F\ (1, 57) = 26.89$, $p < .01$, and more accurate in the L2 to L1 than the L1 to L2 direction of translation, $F\ (1, 57) = 168.60$, $p < .01$. The interaction between fluency group and direction of translation was marginally significant, $F\ (1, 57) = 3.52$, $p < .07$.

c Cognate status: A final series of analyses was performed on the data from Experiment 1 to examine the effect of cognate status on naming and translation. Because cognates share lexical form as well as meaning with their translation equivalents, they provide a means to evaluate the role of lexical form in processing. In past research cognate status has been defined in many different ways (e.g., De Groot, 1993; Dijkstra *et al.*, 1999; Friel and Kennison, 2001). Using the procedure described by Kroll and Stewart (1994), cognate status in the present experiment was operationalized on the basis of a subjective measure of lexical transparency. A group of monolingual English speakers guessed the translations of the French words. Any French word whose translation could be guessed by 50% or more of the monolinguals was identified as a cognate. By this criterion, of the original 120 words, 47 were identified as cognates and 73 as noncognates. A subset of 40 cognates and 40 noncognates was then selected to form sets of words that were matched on word frequency in English, word length in English and French, familiarity, and context availability. The lexical properties of the cognate and noncognate subsets are shown in Table 3.

Looking first at the word naming data, these data for the matched items are shown in Figure 3. Mean naming latencies are plotted as a function of language of production (L1 or L2), cognate status of the word (cognate or noncognate) and fluency of the learner (less or more fluent). The graph shows clearly that the data for this subset of items replicate the main features seen earlier for the full set of materials. The less fluent group was slower than the more fluent group for both languages, and naming in L2 was slower than naming in L1. To examine the effects of cognate status, separate analyses of variance were performed for the less and more fluent groups. In each of these analyses there was a significant effect of

Table 3 Lexical properties of the 40 cognate and 40 noncognate translations selected from the 120 words in Experiment 1 to form a matched subset

Measure	Word type		
	Cognate (n = 40)	Noncognate (n = 40)	p values (t-test)
English word frequency (per million)	53.9	55.1	n.s.
Word length (number of letters) in English	6.2	5.7	n.s.
Word length (number of letters) in French	6.6	6.1	n.s.
Context availability (scale of 1 to 7)	5.4	5.5	n.s.
Familiarity (scale of 1 to 7)	4.9	4.9	n.s.
Cognate status (percentage of guesses)	79.0	12.5	.01

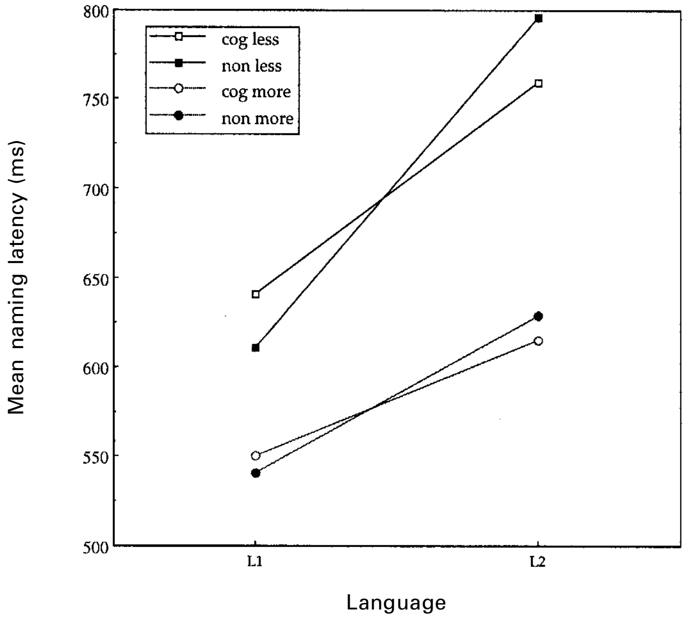

Figure 3 Mean naming latencies (in milliseconds) for less and more fluent learners in Experiment 1 to name words that were cognates (cog) or noncognates (non) as a function of the language of production (L1 or L2)

language of naming (p values $< .01$). For the less fluent participants, there was no main effect of cognate status ($p > .10$) but a significant interaction between language and cognate status, $F\ (1, 78) = 6.17, p < .025$. The nature of the interaction can be seen clearly in the graph. Cognates were named more slowly than noncognates in L1, but more quickly than noncognates in L2. For the more fluent participants, the same interaction was much smaller and only marginally significant, $F\ (1, 78) = 3.62, p < .07$.

The naming data show that less fluent participants were indeed influenced by lexical form relations across their two languages. In L2, naming was facilitated when the L2 word bore a similarity to its form in L1. Although the phonology of cognates is rarely identical, the benefit in the overlap of the orthography and phonology appears to outweigh the costs associated with the computation of an alternative pronunciation. In L1, the presence of the new L2 vocabulary – perhaps together with an experimental context in which participants knew they would have to speak L2 words – may have been sufficient to activate the weaker L2 alternative to a level at which it functioned as a competitor. Jared and Kroll (2001) reported a similar result for English–French bilinguals naming English words with French enemies (i.e., lexical neighbours in French with different pronunciations). When participants named words in English, their L1, in an experiment which appeared to be in English only, there was no effect of the French enemies. When they were subsequently required to also name words in French, a later block of English naming trials revealed a significant inhibitory effect for English words with French enemies. The result is similar to the inhibitory result for cognates in L1 naming in the present experiment.

The reduction of the cognate effects in L2 for the more fluent group might be taken to suggest that with increasing fluency there is decreasing reliance on lexical form. That claim is consistent with observations of lexical transfer during early stages of acquisition and with the predictions of the revised hierarchical model. However, the reduction of cognate effects in naming for the more fluent group is more difficult to understand from the perspective of recent research showing that highly fluent bilinguals, who are presumably well past these early stages of acquisition, are quite sensitive to the orthographic and phonological similarity of lexical forms across languages even when they are processing in one language only (e.g., Dijkstra *et al*., 1998a; 1998b; Van Heuven *et al*., 1998). In a recent word naming study, Schwartz *et al*. (2000) had English–Spanish bilinguals name cognates and noncognates in both languages. Like the results for the more fluent learners in Figure 3, there was little overall effect of cognate status on naming latencies. However, when the data were analysed to take into account the cross-language similarity of lexical features, it turned out that the orthographic-to-phonological correspondences across languages had a clear effect on naming performance, with faster naming latencies when orthography and phonology matched (i.e., were both similar or both different) than when they mismatched. Because the present materials were not designed to examine the fine details of the orthography and phonology, it is not possible to determine whether the more fluent group in Experiment 1 would have been sensitive to these variations.

The suggestion from these analyses is that there are two different ways in which lexical form manifests itself during acquisition. One effect is based on global lexical transparency, whereby similarity to the L1 form on any dimension provides a clue to the identity of the L2 word. The other is more subtle, reflecting the interplay between lexical features during the identification process. Our hypothesis is that the latter may reflect the manner in which the lexical system is retuned as L2 vocabulary is acquired whereas the former is a

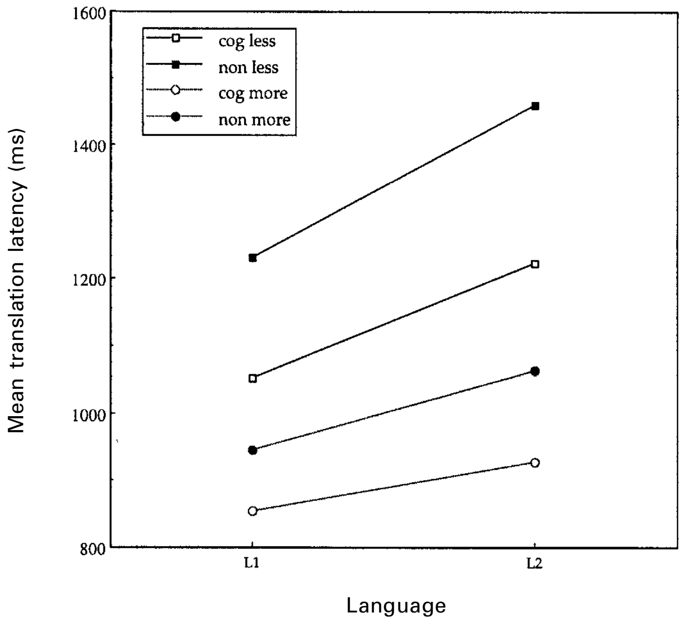

Figure 4 Mean translation latencies (in milliseconds) for less and more fluent learners in Experiment 1 to translate words that were cognates (cog) or noncognates (non) as a function of the language of production (L1 or L2)

means to initiate L2 activity when very little L2 information is known.

Turning now to the translation data, Figure 4 presents the latency data for translation in both directions as a function of fluency group and cognate status. Again, the graph reveals the main features of translation observed for the larger set of materials (Table 2). Furthermore, we see large effects of cognate status, with faster translation latencies for cognate than for noncognate translations, replicating past studies (e.g., Sánchez-Casas *et al.*, 1992; De Groot, 1993; Kroll and Stewart, 1994). The cognate effect was larger for the less fluent (209 ms) than for the more fluent learners (114 ms), but significant in both cases: F (1, 78) = 39.3, $p < .01$ for the less fluent group and F (1, 78) = 14.43, $p < .01$ for the more fluent group. In neither case did cognate status interact with direction of translation ($p > .05$ for both groups). The very large benefit of cognate status for the less fluent learners is consistent with the interpretation we assigned to the effects observed in naming. During early stages of acquisition, lexical transparency provides a means to utilize highly available information in L1 to facilitate the processing of L2.

d Summary: The following aspects of these results are theoretically important for claims about the development of lexical fluency.

1) As predicted by the revised hierarchical model, translation was faster and more accurate from L2 to L1 than from L1 to L2.
2) The difference in simple naming times for L1 and L2 was smaller than the difference between the two translation directions. Although the naming and translation tasks share only some processing components, the smaller language difference in naming suggests that simple aspects of production and articulation cannot account for the translation asymmetry.
3) The translation asymmetry was noticeably larger for the less fluent than for the more fluent group, with increased proficiency having a more dramatic effect on the L1 to L2 direction of translation.
4) There was an unexpected cost to L1 word naming for the less fluent group.
5) Less fluent learners appear to rely more on form relations across languages than more fluent learners.

In Experiment 2 we sought to replicate the main features of these results with a different set of participants and new materials. Furthermore, a measure of reading span was included in Experiment 2 to assess the effects of cognitive capacity on performance in naming and translation. If the unanticipated cost to L1 naming for the less fluent group in Experiment 1 was due to the relatively lower abilities of some learners at early stages of L2 acquisition, then we predicted that a similar cost would be observed in Experiment 2 but only for low span learners.

III Experiment 2

In the second experiment we again compared the performance of two groups of second language speakers on word naming and word translation tasks. In Experiment 1 the two groups were learners at different levels of L2 proficiency. In Experiment 2 the two groups consisted of nonfluent learners at very early stages of L2 learning and relatively fluent bilinguals. The proficiency difference was greater in Experiment 2 than it was in Experiment 1, and we hoped that it would provide a more sensitive context for observing the hypothesized changes in lexical processing and their relation to performance on the span task. If the critical problem for the second language learner is to accomplish sufficient activation of L2 to permit concept mediation, then we expected that tasks that include conceptual processing as a mandatory component would be particularly difficulty for the nonfluent group and most likely to reflect differences in processing capacity measured by the memory span task.

1 Method

a Participants: Thirty-one students enrolled in an intensive summer language institute at Pennsylvania State University served as the nonfluent participants. Eighteen were students in the Spanish program and 13 were students in the French program. Seventeen graduate-level instructors in Spanish and French served as the bilingual participants. Nine of them were fluent in Spanish and eight in French. All participants completed a language history questionnaire in which they provided information about their experience and self-assessed proficiency in either French or Spanish. They also completed a lexical decision task in L2 to provide a converging performance measure of their fluency.

b Materials: A set of 120 words and their translations in English were selected from elementary textbooks in Spanish and French to increase the likelihood that the nonfluent participants would be able to recognize and translate the words presented in each task. The words were divided into separate matched versions so that 60 of the critical words appeared in the word naming task and 60 in the translation task. Within each task, the 60 words were further divided into two versions which were counterbalanced across language of production. A different set of 30 words and 30 nonwords was selected to be used in the lexical decision task. Because the nonfluent participants in Experiment 2 were less skilled than the learners in Experiment 1, the words in Experiment 2 were, on average, more frequent than those in Experiment 1. The mean word frequency in English (Francis and Kučera, 1982) was 89.3 times per million in word naming and 86.1 per million in translation. As in Experiment 1 the words across versions were closely matched in length, with the L2 words slightly longer overall than the L1 words.

A version of the reading span task used by Waters and Caplan (1996) was included to assess individual differences in cognitive capacity. In this task, participants are presented with a series of English sentences. They first have to decide whether a given sentence is plausible by pressing a 'yes' or 'no' key on the computer, and then attempt to remember the final word of the sentence for later recall. After a series of two, four or six sentences they are asked to recall as many of the final words as possible. The span measure is based on the number of final words recalled. The latency data for the plausibility judgements can be used to determine whether participants traded speed for accuracy in processing the sentence. The span task consisted of 80 test sentences, half semantically plausible and half semantically implausible. The two sentence types were matched for number of words, word frequency, concreteness and imageability.

Additional practice items were presented prior to each of the tasks.

c Procedure: The nature of the summer program in which the learners were enrolled and constraints on the number of tasks that could be completed in a given session required that tasks be spread out over two sessions. Tasks at the first time of testing included the language history questionnaire, the lexical decision task to assess fluency, the reading span task, and translation. The word naming task was administered in the second session. Because there was some attrition between the two times of testing, we report data for the full set of participants for translation, and for the remaining subset for word naming.

Looking first at word naming and translation, participants performed each of the critical tasks, i.e., word naming in English (L1) and French or Spanish (L2) and word translation from English to French or Spanish (L1 to L2) and from French or Spanish to English (L2 to L1). The tasks were blocked, with the translation task presented at the first time of testing and the word naming task presented approximately two weeks later at the second time of testing. Words were presented one at a time at the centre of a Macintosh microcomputer screen and participants were instructed to name them aloud in the language in which they were presented or to speak aloud the translation. The order of presentation of words within a block was randomized individually for each participant. Prior to the presentation of each target word, a fixation cross was presented. Participants were instructed to respond as quickly and accurately as possible, and to indicate when unaware of the correct translation of the presented word by saying 'no'. Each word remained on the screen until a response was recorded or for 4000 ms. Reaction time (RT) was recorded by the computer program in milliseconds from the onset of the presentation of the word to the onset of articulation. The onset of spoken responses was recorded by a voice-activated relay connected to a microphone. All responses were tape-recorded for later coding of accuracy.

The procedure for lexical decision was similar to that for the production tasks with the following exceptions. Only words in L2 (French or Spanish) or nonwords derived from L2 words were presented. On each trial a word was presented and the participant had to judge whether or not the letter string was a real word by pressing a 'yes' or 'no' key. RT was measured from the onset of the presentation of the word to the button press.

Turning to reading span, prior to the presentation of each set of sentences a 'ready' prompt was presented and remained on the screen until the subject initiated the trial by pressing a key. For each sentence within the series, a fixation point was presented for 300 ms and followed by the target sentence which remained on the screen until the participant made a response or for 5000 ms. RTs were measured from the onset of the sentence to the onset of the button press. Each series ended when the word 'recall' appeared on the screen, indicating that participants should write down as many of the final words as they could remember. The span task was administered in English, the L1 of all of participants.

2 *Results and discussion*

a Fluency measures: In Experiment 1 the assignment of participants to fluency groups was made after they participated in the study, based on their number of years of L2 experience. In Experiment 2, the two groups were identified in advance, based on their status as learners or second language instructors. Most of the learners were in a first or second semester course in Spanish or French. The instructors had studied Spanish or French for an average of 13 years. A series of *t*-tests was performed to determine whether the two groups differed on any factors other than their L2 fluency. The means for each of these comparisons is shown in Table 4, where the two groups are compared for age, number of years in the USA, number of years in US schools, lexical decision performance, and self-assessed ratings of reading, writing and conversation in English and in either French or Spanish. The results showed that the groups were closely matched on all but their ratings of proficiency in L2 and on their L2 lexical decision performance. In each case, the relatively fluent bilingual participants rated themselves significantly higher on all measures and judged letter strings as real words in L2 faster and more accurately than the nonfluent learners. The difference between the ratings for L1 and L2 suggest that the bilinguals were highly fluent in L2 but still dominant in L1.

Table 4 Characteristics of participants in Experiment 2, including self-assessed proficiency ratings on reading, writing and conversation in L2 (either French or Spanish)

Measure	Proficiency group		
	Nonfluent learners (n = 31)	Fluent bilinguals (n = 17)	p values (t-test)
Age (years)	24.0	25.3	n.s.
Years in the US	21.7	19.8	n.s.
Years in US schools	16.7	15.2	n.s.
L2 lexical decision (Yes trials)			
RTs (accuracy)	1124 ms (79%)	748 (94%)	.01
Self-ratings L1 (English)			
Reading	9.5	9.6	n.s.
Writing	9.1	9.2	n.s.
Conversation	9.4	9.6	n.s.
Self-ratings L2 (French or Spanish)			
Reading	3.5	8.6	.01
Writing	3.0	7.8	.01
Conversation	2.7	8.5	.01

Note: (Each scale was rated from 1 to 10 where 1 was not very fluent and 10 was highly fluent.)

b Performance measures: As in Experiment 1, analyses of variance were performed separately on mean correct response times (RTs) and percentage accuracy for the performance of the two groups on the word naming and translation tasks. Data were excluded from trials in which errors were made and trials on which RTs exceeded a criterion of 2.5 standard deviations for an individual participant's mean.

The results for the word naming task are shown in Table 5, where mean naming latencies (in milliseconds) and percentage accuracy are given for the learners and bilinguals for naming in English (L1) and French or Spanish (L2). Like the results of Experiment 1, the analysis revealed main effects of fluency group, $F(1, 36) = 22.13$, $p < .01$, language, $F(1, 57) = 46.05$, $p < .01$, and an interaction between fluency group and language, $F(1, 57) = 24.75$, $p < .01$. The bilinguals were faster overall than the learners and the time to name words in L2 was generally slower than in L1. However, the interaction shows that there was a larger difference between the groups for L2 than for L1, and that like the results of Experiment 1, the learners were also slower to name words in L1 than the bilinguals. Simple effects tests on the interaction showed that the bilinguals were indeed highly fluent; the 19 ms difference in the time they took to name words in L1 and L2 was not significant ($p > .10$). The same comparison for the 168 ms difference for the learners was highly significant ($p < .01$). These results demonstrate that it is possible for individuals who are highly fluent in an L2 to speak words as quickly in L2 as in L1, a result that is particularly impressive given that the L2 words were slightly longer than the L1 words.

An analysis on the percentage accuracy for the naming conditions produced a pattern of results that was generally similar to that reported for the RTs. The main effects of fluency group was significant, $F(1, 36) = 7.28, p < .05$, indicating lower accuracy for the learners than for the bilinguals. The main effect of language was not reliable, $F(1, 36) = 2.47, p > .10$, but the significant interaction between fluency group and language was significant, $F(1, 36) = 14.30, p < .01$. A series of simple effects tests showed that the bilinguals were highly accurate in naming words in both languages but that the learners were more accurate in L1 than L2 ($p < .01$).

We turn now to translation. The fluent bilinguals in this experiment failed to produce evidence for a language difference in simple naming. If the asymmetry in translation direction observed in Experiment 1 and in past studies is a reflection of the ease with which individuals can assemble L2 phonology, then the translation asymmetry might also be expected to be absent for this group. The mean translation latencies (in milliseconds) and percentage accuracy are shown in Table 5. These data show that both groups produced a translation asymmetry, with slower RTs in the forward

Table 5 Mean response latencies (in milliseconds) and percentage accuracy (in parentheses) for less and more fluent participants to name words in L1 and L2 and translate words into L1 and L2 in Experiment 2

	Language of production		
Task	L1	L2	Language difference
Word naming			
Learners (n = 21)[a]	589 (.95)	757 (.89)	168 (.06)
Bilinguals (n = 17)	537 (.95)	556 (.97)	19 (.02)
Fluency difference	52 (.00)	201 (.08)	
Translation			
Learners (n = 31)	1232 (.43)	1343 (.44)	111 (.01)
Bilinguals (n = 17)	902 (.82)	950 (.83)	48 (.01)
Fluency difference	330 (.39)	393 (.39)	

Note: [a]The lower n for the learners in the word naming task is a reflection of attrition in participation from the first to the second time of testing than in the backward direction of translation. Although the magnitude of the asymmetry was again larger for the learners than for the fluent bilinguals, an analysis of variance revealed only main effects of fluency group and translation direction; the interaction between them was not significant ($p > .10$). The bilinguals were faster to translate than the learners, F (1, 46) = 7.28, $p < .01$, and translation was faster for both groups in the L2 to L1 direction than in the L1 to L2 direction, F (1, 46) = 10.18, $p < .01$.

An analysis on the translation accuracy data revealed only a main effect of fluency group, F (1, 46) = 85.27, $p < .01$. Unlike the results of Experiment 1, there was little difference in accuracy across the two translation directions ($p > .05$), and the interaction between fluency group and direction was not significant ($p > .05$).

The pattern of results in Experiment 2 for both tasks generally replicated the findings of Experiment 1. Despite differences in the nature of the participant groups in the two experiments and in the materials, in both experiments there was a larger naming difference and translation asymmetry for the less fluent participants and an indication that naming performance in L1, the native language, was slower for learners than for more fluent L2 speakers. A further question we wished to investigate in this study was whether the cost to L1 naming observed in both experiments is due to individual differences in cognitive abilities among learners and fluent bilinguals.

c Individual difference measures: reading span: We first report the results for performance on the reading span task and then examine how differences in span relate to performance on the naming and translation tasks described above. The mean recall score on the reading span task was 44.83 words out of a possible 80 words. The recall data and corresponding RTs and accuracy for the plausibility judgements are shown in Table 6 for the learners and fluent bilinguals. One participant in the learner group did not complete the span task, thus reducing that group to 30.

The mean reading span score was approximately 10 points higher for the bilinguals than for the learners, and that difference was statistically significant, t (45) = 2.80, $p < .01$. Given the similarity of the two groups on all measures other than L2 performance (see Table 4), it is surprising to observe such a large difference when all individuals performed the task in their L1.

To determine whether the span advantage for the bilinguals reflected a difference in the strategy they used to perform the task, speed and accuracy was compared for the plausibility judgements. An analysis of variance was computed for mean RTs with fluency

Table 6 Mean number of words recalled in the reading span task and mean RTs and percentage accuracy to make plausibility judgements for learners and bilinguals in Experiment 2

Measure	Fluency group: Learners (n = 30)		Fluency group: Bilinguals (n = 17)	
Mean number of words recalled	41.0		51.2	
Plausibility	yes	no	yes	no
RT (ms)	3215	3319	2818	2992
Accuracy (%)	86.8	75.7	91.8	81.9

group and plausibility (yes vs. no) as the factors. Both main effects were significant. The bilinguals were faster to make plausibility judgements than the learners, F (1, 45) = 7.24, $p < .01$, and plausible sentences were judged more quickly than implausible sentences, F (1, 45) = 4.88, $p < .05$. The interaction between fluency group and plausibility was not significant ($p > .10$). The analysis on the accuracy of the plausibility judgements corresponded closely to the RTs. The bilinguals were more accurate than the learners, F (1, 45) = 4.41, $p < .01$, and decisions overall were more accurate for plausible than for implausible sentences, F (1, 45) = 38.66, $p < .01$. The interaction was not significant ($p > .10$). The plausibility results allow us to rule out the alternative that the higher span scores for the bilinguals were due to a strategy in which sentences were processed more slowly to achieve greater accuracy. To the contrary, the bilinguals processed the sentences more quickly and accurately than the learners and also produced higher span scores.

Why do the bilinguals have higher reading span than the learners? The span advantage may reflect a self-selection factor based on verbal and/or cognitive abilities. That is, only individuals with relatively high cognitive abilities may succeed in second language learning and go on to become fluent bilinguals. Alternatively, the bilingual span advantage may be a positive cognitive consequence of bilingualism (for related arguments concerning childhood bilingualism, see Bialystok, 1997). On the basis of the data we have reported, it is impossible to determine which of these alternatives is correct or even whether they are mutually exclusive. However, it is possible to examine the consequences of span for bilingual performance in the lexical tasks used in the present experiment.

Table 7 Mean number of words recalled in the reading span task, age of participants and self-reported rating of proficiency in L2 (on a scale from 1 to 10 with 10 as the most fluent) as a function of fluency level and span group

Fluency group	Mean words recalled (span)	Age (years)	Mean L2 rating
Bilinguals			
Low span (n = 5)	35.2	22.4	8.2
High span (n = 12)	58.4	26.5	8.4
Learners			
Low span (n = 17)	33.1	25.1	3.2
High span (n = 13)	51.1	23.3	3.1

d Effects of span on performance: We divided the learners and bilinguals into lower and higher memory span groups based on a criterion of 40 words recalled (50% of the maximum 80 words). The mean number of words recalled for the four groups that resulted is shown in Table 7 along with their mean age, and rating of self-assessed proficiency in L2. What is clear is that the learners included a more even distribution of high and low span individuals whereas the bilinguals were predominantly high span. There were no significant differences among the high and low span groups in age or within fluency groups on ratings of proficiency (p values > .05).

Looking at effects of span on word naming, the immediate question of interest was whether the cost to L1 naming observed for the learners was due entirely to the performance of the low span participants.[6] The time to name words in L1 for learners and bilinguals was compared for the high and low span groups. For the low span groups, the nonfluent participants were slower to name words in L1 than the fluent participants (588 vs. 562 ms, respectively), but the difference was not significant ($p > .05$). Contrary to the self-selection hypothesis, the same difference for the high span groups was significant, t (21) = –2.05, $p < .05$. High span learners were reliably slower to name words in L1 than high span bilinguals (589 vs. 527 ms respectively). It is important to note that the results of these analyses must be regarded as tentative because of the small sample size in the groups when they are broken down in this way. However, they do provide a preliminary answer to the question of why learners exhibit a cost to L1 naming.

[6] Because the word naming data were based on a smaller sample of learners due to attrition at the second time of testing, it was necessary to reanalyse the breakdown of the span data by group to be sure that the high and low span groups were equivalent for the learners and bilinguals. This was, indeed, the case. For the bilinguals, the high vs. low span scores were 35.2 and 58.4 respectively, while the same values for the learners were 33.7 and 54.0.

These data show that span does not account for the observed difference. Instead they suggest that the process of L2 learning affects the processing of L1. Like recent connectionist approaches (e.g., Thomas, 1997), they can be understood as a reflection of changes to the lexicon as new information is acquired.

Looking at effects of span on translation, a final set of analyses was conducted to examine the consequence of span differences for translation. These analyses focus on the learners because only in that group was there a sufficient number of participants at each level of span to provide a legitimate estimate of the effect of span on translation performance. The mean translation RTs for the learners is shown in Table 8 broken down by span group (low or high), direction of translation (L1 to L2 or L2 to L1), and by cognate status (cognate or noncognate translation). As discussed in Experiment 1, translation is reliably faster for cognates than noncognates (e.g., Sánchez-Casas *et al.*, 1992; De Groot, 1993; Kroll and Stewart, 1994), presumably because of their lexical transparency. However, the presence of interlingual homographs, and the predominance of noncognates, makes a lexical-level strategy potentially risky. The goal of the present analysis was to see whether span had any consequences for the speed of translation and for sensitivity to lexical-level factors.

An analysis of variance was performed on learners' translation data. There was a significant main effect of cognate status, $F\ (1, 28) = 18.41$, $p < .01$, replicating past studies reporting faster translation performance for cognates than for noncognates. There was also a significant main effect of translation direction, $F\ (1, 28) = 6.48$, $p < .05$, with faster RTs in the backward than forward direction, reflecting the overall asymmetries reported in the earlier analysis (see Table 5). In addition, the interaction between cognate status

Table 8 Mean translation RTs (in milliseconds) for low and high span learner groups as a function of direction of translation and cognate status of the target word

	L1 to L2 translation (forward)			L2 to L1 translation (backward)		
Group	Cognate	Noncognate	Cognate effect	Cognate	Noncognate	Cognate effect
Low span (n = 17)	1179	1457	278	1087	1377	290
High span (n = 13)	1323	1409	86	1194	1267	73
Span effect	–144	48		–107	110	

and span group was significant, F (1, 28) = 5.77, $p < .05$. The interaction can be understood by examining the results in Table 7. In both directions of translation, the magnitude of the cognate advantage was dramatically smaller for high span than for low span learners. Indeed, in both cases the high span learners were slower than the low span learners to translate cognates. The reverse was true for the noncognates, in which case the high span learners were faster to translate than the low span learners. These results suggest that high span learners are less likely to rely on word form cues than their low span counterparts. They provide support for the hypothesis that second language learners with high span may allocate their mental resources to generate strategies that increase conceptual processing, even when those strategies produce a processing cost.

IV General discussion

1 Summary of main results

The two experiments reported here compared the performance of individuals at different stages of L2 acquisition on lexical-level naming and translation tasks. In each case our interest was to identify the lexical-level changes that occur with increasing proficiency in L2. In Experiment 1, the performance of two learner groups was compared. In Experiment 2, the performance of a group of learners was compared to the performance of fluent bilinguals. The main results of these experiments show that learners become faster to name words in L2 and to translate words from one language to the other with increasing L2 proficiency. In both experiments, and for all participant groups, there was a translation asymmetry supporting the predictions of the revised hierarchical model. Translation was faster from L2 to L1 than from L1 to L2. Furthermore, the magnitude of the asymmetry was larger for the less proficient learners than for more proficient learners or bilinguals. According to the model (Kroll and Stewart, 1994), the translation asymmetry reflects two distinct routes to translation. Whereas backward translation can be achieved, at least in principle, on the basis of direct lexical links to the translation equivalent, forward translation requires activation of meaning and subsequent lexicalization and selection of an L2 word. The latter process has been hypothesized to be vulnerable to lexical competition of the sort that is typically observed in production tasks such as picture naming (e.g., Hermans *et al.*, 1998; Costa *et al.*, 1999). Although both directions of translation became faster and more accurate with increasing fluency, the difference was always greater for forward than for backward translation.

In Experiment 1 we obtained the unexpected result that less fluent learners were slower to name words in English than more fluent learners. As noted earlier, the result is surprising because English was the native language of all participants and the word naming task has been thought to be a task that reflects automatic aspects of lexical processing (e.g., Lupker, 1984). In Experiment 2 we replicated the L1 naming cost during early stages of L2 acquisition and tested the hypothesis that it occurs because learners represent a mix of cognitive abilities, with only a subset of them likely to eventually succeed in becoming reasonably fluent in L2. The learners and fluent bilinguals in Experiment 2 were given a reading span task to assess their ability to allocate cognitive resources to processing and storage. The task was administered in English and, as such, should be independent of their L2 proficiency. The bilinguals scored reliably higher on the span task than the learners, therefore providing support for the hypothesis that the cognitive abilities of learners vary over a larger range than those of fluent bilinguals. However, when the L1 naming data were reanalysed to determine whether lower span on the part of learners accounted for longer naming latencies, the results showed clearly that span differences did not provide an explanation for the observed effect. If anything, the L1 cost in word naming was larger for the higher span learners than for the lower span learners, suggesting that it occurs as a consequence of L2 learning.

In future research it will be of interest to compare the performance of learners on L1 in the bilingual context such as the one used in these experiments, with a more monolingual task environment. Although we blocked the language of naming, it is possible that the expectation that L2 will have to be used is sufficient to produce the observed cost (for an account of ‘language mode’ that might accommodate these results, see Grosjean, 2001). That is, there may be a processing cost for a less fluent individual to assume ‘bilingual’ mode in which both languages must be reasonably active, and that cost may be most apparent for L1. It will also be of interest to determine whether the result is truly a cost to L1 naming for learners, or a benefit for the more fluent bilinguals. If bilingualism confers cognitive benefits to language processing, then even performance on highly skilled L1 tasks might be expected to improve with increasing fluency. Comparisons with individuals who have similarly high cognitive abilities but who are genuinely monolingual could provide important data to evaluate this alternative hypothesis. Unlike fluent bilinguals, who may have internalized the control mechanisms required to monitor the output of lexical processing, second language learners at low levels of skill in L2 may be more dependent on external cues to language.

2 Developing lexical fluency or control?

There is no doubt that with increasing expertise in a second language, learners acquire a richer lexical network for words in L2 that is at least partly responsible for the increasing speed and accuracy we observed in the experiments reported here for the naming and translation tasks. However, as Green (1998) notes, the lexical system itself may not be capable of sorting out the resulting activation of information in both languages; a control process needs to be implemented to allow regulation of the selected language output.

One way in which the operation of such a control process has been studied is to observe the consequences of language switching. For example, Meuter and Allport (1999) found that when bilinguals had to switch between languages in naming numbers, there were larger switch costs into L1 than into L2. The observation of an asymmetric switch cost across languages has been interpreted as support for Green's (1998) inhibitory control model. The model assumes that the more active language, L1, will have to be inhibited in order for L2 to be produced. When, on a switch trial, bilinguals must produce in L1 after inhibiting L1 to produce L2, there will be a cost in processing speed relative to a nonswitch trial. Although to our knowledge there have been no developmental studies of language switching, it provides a useful context for thinking about the development of L2 fluency. In some respects, when a nonfluent learner is placed in an experiment in which there is an expectation that both languages will have to be used, they are in a functionally mixed or switch-type environment. The L1 cost to word naming in both experiments in the present article may be understood in the same way. It may be a reflection of the inability to attend selectively to L1 when there is a possibility that L2 will also have to be processed. The observation that the cost to L1 naming occurred even for high span learners in Experiment 2 further suggests that the ability to effect controlled processing alone is not sufficient at very early stages of learning.

3 Towards a model of the developing lexicon

In this article we consider two models of the lexicon, the revised hierarchical model (Figure 1) and the BIA model (Figure 2).

Whereas the revised hierarchical model provides an account of the development of interlanguage connections with increasing L2 expertise, the BIA model provides an account of the bilingual lexicon based on how lexical forms are activated within and across languages during visual word recognition. As noted in the introduction, very little has been said about how the architecture of the BIA model – proposed to capture lexical processing in the fluent bilingual – changes and acquires its ultimate state during acquisition. Likewise, within the revised hierarchical model, very little has been said about how lexical and semantic forms are activated and interact with one another. Neither model provides details concerning the manner in which the output of the lexical system is controlled to achieve skilled performance or how that control might develop. A goal of future work will be to provide such an integrated account.

The empirical results we have reported, together with other recent studies, provide some structure in constraining the model that we eventually want to propose. These findings show that although a general pattern of improvement with increasing skill is observed, there are a number of aspects of the results that require additional assumptions to be made about lexical development. First, and perhaps most important, is that the representations and processes associated with L1 are not constant across L2 development. Connectionist accounts of L2 syntactic development have made a similar point (e.g., Kilborn, 1989; MacWhinney, 1997). L1 not only affects L2, but is also affected by the new L2 representations and by the increasing need to negotiate potential competition across representations, both within and across languages.

A second point is that lexical development is affected by cognitive abilities, but in rather specific ways. Our results show that learners with relatively high span perform differently from learners with low span. However, these are not across-the-board differences; they depend on the nature of the task and on the nature of the lexical input. Cognitive abilities assessed in this way appear to be revealed in tasks that involve greater computation and processing (e.g., span differences are important for translation but not for simple word naming) and for words that differ in their lexical transparency (e.g., high span learners appear to actively avoid using lexical information that is potentially an unreliable cue to meaning).

In future research it will be important to address these issues and also to understand how lexical development operates in context and out of context. Almost all of the psycholinguistic work on this topic has focused on the perception and production of isolated words. We believe that the theoretically motivated approach taken to the study of how words are understood and spoken will provide a framework for approaching this important problem.

Acknowledgements

This research was supported in part by NSF Grants BCS-9905850 and BCS-0111734, a sabbatical grant from the Nijmegen Institute for Cognition and Information, The Netherlands, and an RGSO grant from The Pennsylvania State University to Judith F. Kroll. We thank Ton Dijkstra, Tamar Gollan, David Green, Christina Schelletter, and two anonymous reviewers for helpful comments on an earlier version of the manuscript, and Amber Leslie, Bill Lucas, Yolanda Rivera, and Patricia Roufca for research assistance. Portions of this research were presented at the first International Symposium on Bilingualism, Newcastle upon Tyne, England and at a workshop on L2 processing at the University of Paderborn, Germany. Correspondence concerning this article should be addressed to Judith F. Kroll, Department of Psychology, 641 Moore Building, The Pennsylvania State University, University Park, PA 16802, USA. Electronic mail may be sent to jfk7@psu.edu.

V References

Bialystok, E. 1997: Effects of bilingualism and biliteracy on children's emerging concepts of print. *Developmental Psychology* 33, 429–40.

Bijeljac-Babic, R., Biardeau, A. and **Grainger, J.** 1997: Masked orthographic priming in bilingual word recognition. *Memory and Cognition* 25, 447–57.

Chen, H.-C. 1992: Lexical processing in bilingual or multilingual speakers. In Harris R.J., editor, *Cognitive processing in bilinguals*. Amsterdam: Elsevier, 253–64.

Chen, H.-C. and **Leung, Y.-S.** 1989: Patterns of lexical processing in a nonnative language. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15, 316–25.

Costa, A., Miozzo, M. and **Caramazza, A.** 1999: Lexical selection in bilinguals: do words in the bilingual's two lexicons compete for selection? *Journal of Memory and Language* 41, 365–97.

Daneman, M. and **Green, I.** 1986: Individual differences in comprehending and producing word in context. *Journal of Memory and Language* 25, 1–18.

De Bot, K. and **Kroll, J.F.** in press: Psycholinguistics. In Schmitt, N., editor, *Introduction to applied linguistics*. Arnold Publishers.

De Groot, A.M.B. 1992: Determinants of word translation. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18, 1001–18.

—— 1993: Word-type effects in bilingual processing tasks: support for a mixed representational system. In Schreuder, R. and Weltens, B., editors, *The bilingual lexicon*. Amsterdam: John Benjamins, 27–51.

—— 1995: Determinants of bilingual lexicosemantic organization. *Computer Assisted Language Learning* 8, 151–80.

Dijkstra, A., De Bruijn, E., Schriefers, H. and **Ten Brinke, S.** 2000: More on interlingual homograph recognition: language intermixing versus explicitness of instruction. *Bilingualism: Language and Cognition* 3, 69–78.

Dijkstra, A., Grainger, J. and **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., Timmermans, M. and **Schriefers, H.** 2000: On being blinded by your other language: effects of task demands on interlingual homograph recognition. *Journal of Memory and Language* 42, 445–64.

Dijkstra, A. and **Van Heuven, W.J.B.** 1998: The BIA model and bilingual word recognition. In Grainger, J. and Jacobs, A., editors, *Localist connectionist approaches to human cognition*. Hillsdale, NJ: Lawrence Erlbaum, 189–225.

Dijkstra, A., Van Heuven, W.J.B. and **Grainger, J.** 1998a: Simulating cross-language competition with the Bilingual Interactive Activation model. *Psychologica Belgica* 38, 177–97.

Dijkstra, A., Van Jaarsveld, H. and **Ten Brinke, S.** 1998b: Interlingual homograph recognition: effects of task demands and language intermixing. *Bilingualism: Language and Cognition* 1, 51–66.

Forster, K.I. and **Chambers, S.** 1973: Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior* 12, 627–35.

Francis, W.N. 1999: Cognitive integration of language and memory in bilinguals: semantic representation. *Psychological Bulletin* 125, 193–222.

Francis, W.N. and **Kučera, H.** 1982: *Frequency analysis of English usage*. Providence, RI: Brown University Press.

Frederiksen, J.R. and **Kroll, J.F.** 1976: Spelling and sound: approaches to the internal lexicon. *Journal of Experimental Psychology: Human Perception and Performance* 2, 361–79.

Friel, B.M. and **Kennison, S.M.** 2001: Identifying German–English cognates, false cognates, and non-cognates: methodological issues and descriptive norms. *Bilingualism: Language and Cognition* 4, 249–74.

Gernsbacher, M.A. 1984: Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General* 113, 256–81.

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

Grosjean, F. 2001: The bilingual's language modes. In Nicol, J.L., editor, *One mind, two languages: bilingual language processing*. Cambridge, MA: Blackwell, 1–22.

Hancin-Bhatt, B. and **Nagy, W.** 1994: Lexical transfer and second language morphological development. *Applied Psycholinguistics* 15, 289–310.

Hermans, D., Bongaerts, T., De Bot, K. and **Schreuder, R.** 1998: Producing words in a foreign language: can speakers prevent interference from their first language? *Bilingualism: Language and Cognition* 1, 213–29.

Jared, D. and **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.

Jared, D., McRae, K. and **Seidenberg, M.S.** 1990: The basis of consistency effects in word naming. *Journal of Memory and Language* 29, 687–715.

Just, M.A. and **Carpenter, P.A.** 1992: A capacity theory of comprehension: individual differences in working memory. *Psychological Review* 99, 122–49.

Kilborn, K. 1989: Sentence processing in a second language: the timing of transfer. *Language and Speech* 32, 1–23.

Kroll, J.F. 1993: Accessing conceptual representations for words in a second language. In Schreuder, R. and Weltens, B., editors, *The bilingual lexicon*. Amsterdam: John Benjamins, 58–81.

Kroll, J.F. and **Curley, J.** 1988: Lexical memory in novice bilinguals: the role of concepts in retrieving second language words. In Gruneberg, M., Morris, P. and Sykes, R., editors, *Practical aspects of memory*, Volume 2. London: John Wiley, 389–95.

Kroll, J.F. and **De Groot, A.M.B.** 1997: Lexical and conceptual memory in the bilingual: mapping form to meaning in two languages. In De Groot, A.M.B. and Kroll, J.F., editors, *Tutorials in bilingualism: psycholinguistic perspectives*. Mahwah, NJ: Lawrence Erlbaum, 169–99.

Kroll, J.F., Michael, E. and **Sankaranarayanan, A.** 1998: A model of bilingual representation and its implications for second language acquisition. In Healy, A.F. and Bourne, L.E., editors, *Foreign language learning: psycholinguistic experiments on training and retention*. Mahwah, NJ: Lawrence Erlbaum, 365–95.

Kroll, J.F. and **Stewart, E.** 1994: Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language* 33, 149–74.

Kroll, J.F. and **Sunderman, G.** in press: Cognitive processes in second language acquisition. In Doughty, C. and Long, M., editors, *Handbook of second language acquisition*. Cambridge, MA: Blackwell.

Kroll, J.F. and **Tokowicz, N.** 2001: The development of conceptual representation for words in a second language. In Nicol, J.L., editor, *One mind, two languages: bilingual language processing*. Cambridge, MA: Blackwell, 49–71.

La Heij, W., Kerling, R. and **Van der Velden, E.** 1996: Nonverbal context effects in forward and backward translation: evidence for concept mediation. *Journal of Memory and Language* 35, 648–65.

Lupker, S.J. 1984: Semantic priming without association: a second look. *Journal of Verbal Learning and Verbal Behavior* 23, 709–33.

MacWhinney, B. 1997: Second language acquisition and the competition model. In de Groot, A.M.B. and Kroll, J.F., editors, *Tutorials in* *bilingualism: psycholinguistic perspectives*. Mahwah, NJ: Lawrence Erlbaum, 11342.

McClelland, J.L. and **Rumelhart, D.E.** 1981: An interactive activation model of context effects in letter perception, Part 1: an account of basic findings. *Psychological Review* 88, 375–405.

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

Pavlenko, A. 1999: New approaches to concepts in bilingual memory. *Bilingualism: Language and Cognition* 2, 209–30.

Pecher, D. 2001: Perception is a two-way junction: feedback semantics in word recognition. *Psychonomic Bulletin and Review* 8, 545–51.

Peereman, R. and **Content, A.** 1997: Orthographic and phonological neighborhoods in naming: not all neighbors are equally influential in orthographic space. *Journal of Memory and Language* 37, 382–410.

Potter, M.C., So, K.-F., Von Eckardt, B. and **Feldman, L.B.** 1984: Lexical and conceptual representation in beginning and more proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior* 23, 23–38.

Ransdell, S.E. and **Fischler, I.** 1987: Memory in a monolingual mode: when are bilinguals at a disadvantage? *Journal of Memory and Language* 26, 392–405.

Sánchez-Casas, R.M., Davis, C.W. and **García-Albea, J.E.** 1992: Bilingual lexical processing: exploring the cognate-noncognate distinction. *European Journal of Cognitive Psychology* 4, 293–310.

Schwanenflugel, P.J., Harnishfeger, K.K. and **Stowe, R.W.** 1988: Context availability and lexical decisions for abstract and concrete words. *Journal of Memory and Language* 27, 499–520.

Schwartz, A., Kroll, J.F. and **Diaz, M.** 2000: Reading Spanish words with English word bodies: activation of spelling-to-sound correspondences across languages. Paper presented at the Second International Conference on the Mental Lexicon, Montreal, Canada.

Sholl, A., Sankaranarayanan, A. and **Kroll, J.F.** 1995: Transfer between picture naming and translation: a test of asymmetries in bilingual memory. *Psychological Science* 6, 45–49.

Strain, E., Patterson, K. and **Seidenberg, M.S.** 1995: Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21, 1140–54.

Sunderman, G. in preparation: Lexical development in a second language: can the first language be suppressed? Doctoral dissertation, Pennsylvania State University, University Park, PA.

Talamas, A., Kroll, J.F. and **Dufour, R.** 1999: From form to meaning: stages in the acquisition of second language vocabulary. *Bilingualism: Language and Cognition* 2, 45–58.

Thomas, M. 1997: Connectionist networks and knowledge representation: the case of bilingual lexical processing. DPhil thesis, University of Oxford, UK.

Van Hell, J. and **Dijkstra, T.** in press: Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin and Review*.

Van Heuven, W.J.B., Dijkstra, A. and **Grainger, J.** 1998: Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language* 39, 458–83.

Waters, G.S. and **Caplan, D.** 1996: Processing resource capacity and the comprehension of garden path sentences. *Memory and Cognition* 24, 342–55.