Brain and Language
doi:10.1006/brln.2001.2544, available online at http://www.idealibrary.com on IDEAL®

Conceptual Representation of Verbs in Bilinguals: Semantic Field Effects and a Second-Language Performance Paradox

Norman Segalowitz and Roberto G. de Almeida

Department of Psychology and Centre for the Study of Learning and Performance, Concordia University, Montreal, Quebec, Canada

It is well known that bilinguals perform better in their first language (L1) than in their second lanaguage (L2) in a wide range of linguistic tasks. In recent studies, however, the authors have found that bilingual participants can demonstrate faster response times to L1 stimuli than to L2 stimuli in one classification task and the reverse in a different classification task. In the current study, they investigated the reasons for this ''L2-better-than-L1'' effect. English–French bilinguals performed one word relatedness and two categorization tasks with verbs of motion (e.g., run) and psychological verbs (e.g., admire) in both languages. In the word relatedness task, participants judged how closely related pairs of verbs from both categories were. In a speeded semantic categorization task, participants classified the verbs according to their semantic category (psychological or motion). In an arbitrary classification task, participants had to learn how verbs had been assigned to two arbitrary categories. Participants performed better in L1 in the semantic classification task but paradoxically better in L2 in the arbitrary classification task. To account for these effects, the authors used the ratings from the word relatedness task to plot three-dimensional ''semantic fields'' for the verbs. Cross-language field differences were found to be significantly related to the paradoxical performance and to fluency levels. The results have implications for understanding of how bilinguals represent verbs in the mental lexicon.

Key Words: bilingualism; semantic fields; verbs; fluency; second-language acquisition.

This article addresses a paradoxical effect sometimes found in word processing by nonfluent bilinguals: Under certain circumstances, performance in the nondominant second language (L2) is actually superior to performance in the dominant first language (L1), an ''L2-better-than-L1'' effect.

Bilinguals who are not fluent in their L2 typically perform *less* well on psychological tasks in L2 compared to similar tasks in L1. This is hardly surprising. Less than native-like fluency in a language usually results from having had less experience and practice with the language, and this obviously has consequences for language

The authors thank Melanie Gampel and Claire Carrière for their special assistance on various parts of this project. We also thank an anonymous reviewer for useful suggestions. This research was funded in part by grants to the first author from the Natural Sciences and Engineering Research Council of Canada and the Quebec Ministry of Education and by a grant to the second author from the Social Sciences and Humanities Research Council of Canada. An earlier version of this article was presented at the Second International Conference on the Mental Lexicon, Montreal, October 2000.

Address correspondence and reprint request to Norman Segalowitz or Roberto G. de Almeida, Department of Psychology, Concordia University, 7141 Sherbrooke Street West, Montreal, Quebec H4B 1R6, Canada. E-mail: segalow@vax2.concordia.ca or almeida@alcor.concordia.ca.

0093-934X/02 $35.00

performance from an information processing point of view (Segalowitz, 2000). For example, Favreau, Komoda, and Segalowitz (1980) found that more processing time in L2 than in L1 was required to obtain a word superiority effect, where identification of a briefly presented letter is superior when it is embedded in a real word than when it is presented alone. This result suggests that there is weaker top-down processing in L2 reading by weak bilinguals. Favreau and Segalowitz (1982) found that bilinguals who read more slowly in L2 than in L1 also required L2 spoken language to be slower in order to comprehend it as well as in L1, which again is evidence of reduced information processing skill in the weaker language. Favreau and Segalowitz (1983), using an adaptation of Neely's (1977) primed lexical decision task, found that readers who were generally slower when reading in L2 than in L1 did not possess automatic single word recognition in L2 but did in L1. Phonological recoding during the reading of single words and sentences was also found to be weaker in L2 than in L1 (Segalowitz & Hébert, 1990). Other authors also have documented important information processing differences between dominant L1 and nondominant L2 (McLaughlin, 1998), differences that may shed light on how bilinguals' lexicons may be organized (e.g., Kroll, 1993; Kroll & de Groot, 1997) and on how individual differences in L2 ability may arise (Miyake & Friedman, 1998).

Recently, we came across an example in our own research in which performance by nonfluent bilinguals in their nondominant L2 was actually *superior* to performance in their L1 (Segalowitz & Frenkiel, in preparation). In this study bilinguals had to make judgments about single words presented in a task modeled after Rogers and Monsell's (1995) alternating runs paradigm for studying task switching. In this paradigm, subjects have to change the focus of their attention by switching from one task instruction set to another during the course of the experiment. In our study, participants were preselected on the basis of performance in a screening task; their reaction times (RTs) to classify 100 different nouns as referring to living or nonliving objects had to be faster in L1 than in L2 to remain in the study. Then, in the main experiment, they had to classify 16 words, presented one at a time, across more than 500 trials in separate L1 and L2 blocks. The classification task here involved deciding whether the word referred to an earlier or later moment in time (time adverbials such as *now, soon, later,* and *never*) or was a word that is often used to refer to the presence or absence of a causal link between two propositions (conjunctions such as *because, consequently, despite,* and *although*). While the main result of that study demonstrated L2 task-switching ability to be related to fluency, it is a secondary result of that study that is of special interest here. All of the participants in that study were slower to classify L2 words than L1 words as referring to living or nonliving things, yet they were *faster,* overall, in the task-switching example in L2 than in L1.

The paradox posed by this finding is the following. If cognitive processing underlying word reading is generally superior in L1 compared to L2 (Favreau & Segalowitz, 1983; Segalowitz & Hébert, 1990; Segalowitz, Poulsen, & Komoda, 1991), then surely performance should be faster in L1 than in L2 in the task-switching example described above. In reviewing all of the cases in which we have obtained this paradoxical effect, the following pattern seems to emerge: L1 performance was faster than L2 performance when the task did not involve repetition of stimulus words and when the stimulus word set was large, and the reverse was true when the stimulus word set was small and the same words were used over and over again. These observations led to the following idea. Because these less-than-fluent bilinguals were, by definition, stronger performers in their L1, perceptual fluency for L1 words was likely to be better than L2 words since they would have encountered L1 words far more frequently in life, both visually and in speech. When, however, the experimental task involved repeated presentations of stimulus words, within the context of the ongoing experiment, word recognition fluency presumably improved markedly, as has been observed in studies of repetition priming (e.g., Franks, Bilbrey, Lien, & McNamara, 2000; Scarborough, Cortese, & Scarborough, 1977). For example, Segalowitz and Segalowitz (1993) found that by the time an L2 stimulus word had been repeated for the sixth time in the context of a lexical decision task, there was evidence of automatic processing that had been absent on performance with the first presentation. So perhaps in the task-switching example described earlier, the repetition of the same small set of words across trials resulted in such improved perceptual fluency that speed of word recognition differences between the languages vanished.

This may explain why word recognition ceased to be faster in L1 than in L2. But why did word recognition in L1 become *slower* than in L2? A possible explanation comes from work reported in an unpublished thesis by Vasos (1983) and described in detail in Segalowitz (1986). In that study, a category name (e.g., CLOTHING) was used to prime a judgment about a pair of line drawings (about whether the two objects pictured belong to the same semantic category). Prime words were either in L1 (English), in L2 (French), or neutral nonwords. The objects pictured were either high or low prototypical members of their respective categories. The most important result came from the condition in which two objects from the same category were depicted. Highly fluent bilinguals showed equally high facilitation effects for both levels of prototypicality and in both L1 and L2 when the appropriate category name prime word was shown. Less fluent bilinguals showed facilitation effects across both levels of prototypicality only in L1 and only with high prototypical items in L2. In a second study, Vasos had subjects judge whether a category name (e.g., CLOTHING) correctly named a pictured item that followed (e.g., a coat). A category repetition trial (e.g., CLOTHING followed by a picture of a hat) could occur after 0, 1, or 2 intervening trials (lags of 0, 1, or 2) involving other category names. She found facilitation effects for lags 0, 1, and 2 in L1 but only for lag 0 in L2. Together, these results suggest that for some bilinguals, a target word in L2 might not influence the processing of concepts that are relatively remote (less prototypical or weakly associated) from the core meaning of the target word and that the influence of a category name is not as enduring in L2 as in L1. In these senses, the representations of L1 words may be thought of as being less restricted in scope than the representations of L2 words. A possible consequence of this difference is that word representations may overlap or interact with each other more in L1 than in L2. In some psychological tasks, such interactions may actually create interference and thereby slow down performance more in L1 than in L2. According to this account, once L1 and L2 words have reached some kind of parity in terms of perceptual fluency, such word interactions in L1 could result in performance actually being slower in L1 than in L2 due to overlap among L1 items, thereby producing the L2-better-than-L1 effect. The goal of the current research was to investigate this possibility by first deliberately producing an L2-better-than-L1 effect and then measuring overlap in word meaning representations to see whether these are associated with the L2-better-than-L1 effect.

The research strategy adopted was the following. To highlight the L2-better-than-L1 effect, we selected two tasks for English-dominant English–French bilinguals to perform, one in which we expected L1 performance to be better and the other in which we expected L2 to be better. The first task involved a simple speeded classification in which words appearing one at a time had to be classified as belonging to one of two semantic categories. The stimulus words were verbs of motion (*jump, run,* etc.) and verbs involving psychological states (*wonder, hate,* etc.). There was a long list of such stimuli—48 in L1 and their translation equivalents in L2—and each word was seen only once in this classification task in a given language.

The second task involved a small subset of these words used over and over again.

This time, the classification was arbitrary; that is, the words were randomly assigned to two categories without regard to meaning. It was expected that in this task the repetition of stimuli would remove any L1 perceptual fluency advantage. To the extent that the semantic representations of L1 words interact more with each other than do the corresponding representations of L2 words, learning the arbitrary category assignments in L1 should be harder, compared to what happens in L2, because of confusions created by the interacting meanings.

Finally, we included a third task to assess separately in each language the degree to which word representations overlap. For this, we used a similarity rating task that permitted us to generate, via multidimensional scaling, separate L1 and L2 ‘‘semantic spaces’’ for each subject in which we could locate the motion and psychological verbs. With verbs located in a semantic space, we could measure the degree to which motion and psychological verbs were, as a group, represented distinctly from each other in each language. These distinctiveness measures were then examined to see whether they could explain the pattern of results in the simple and arbitrary classification tasks and thereby help to account for the L2-better-than-L1 effect.

We chose to use verbs rather than, say, nouns for the following reason. Verbs are combinatorial elements—more so than other syntactic categories (see Grimshaw, 1990). Verbs select the number and nature of their arguments, and so mastering the content of a verb goes beyond determining its referring event; it implies being able to determine the participants and their respective roles in the event labeled by the verb. Fluent L1 and L2 speakers who know the meaning of the verb *kick,* for instance, ought to know that it takes two noun arguments: one bearing the *agent* role and the other bearing the *patient* or *theme* role. In this sense, knowing what *kick* means implies knowing more than the syntactic and semantic properties of the verb in itself; it means knowing something about the very conceptual nature of the grammatical constituents the verb requires. These dimensions of a verb’s meaning provide additional ways in which the mental representations of related verbs can differ in L1 versus L2, such as in terms of how they resemble one another beyond similarities in the basic events to which they refer. By selecting verbs as the stimulus category of focus, we hoped to increase opportunities for observing differences in the ways in which words in L1 and L2 are represented.

It is important to consider how verbs may differ from and resemble each other. Most current linguistic theories about the nature of verbs rely on the assumption that verbs sharing semantic content share conceptual components in addition to similar argument structures. In decompositional theories of verb representation (e.g., Jackendoff, 1990; Rappaport Hovav & Levin, 1997), for instance, researchers make quite explicit the idea that verb classes are determined in large part by the similarity in the semantic templates of the verbs belonging to the class. For the verb *kill,* a conceptual template might take the form [x CAUSE (y BECOME <dead>)], for example. In this template, CAUSE and BECOME are primitive concepts, while x and y stand for the main arguments of the verb *kill* (i.e., the arguments that take the form of the grammatical subject and object of the verb) and *dead* stands for the ‘‘idiosyncratic’’ information pertaining to the resultative event (y’s changing of state) (see, e.g., Rappaport Hovav & Levin, 1997). Verbs that share components such as CAUSE and BECOME would then share semantic content with *kill.* A detailed discussion of this proposal is beyond the purposes of this article. Suffice it to say that linguists who are at the more decompositional end of the lexical–semantic spectrum (e.g., Jackendoff, 1990) have devised ways of accounting for the shared meanings between members of verb classes by proposing series of features that affect the very nature of those conceptual primitives (e.g., CAUSE), thus distinguishing among subtypes of events (e.g., subtypes of CAUSE).

According to this view of lexical–semantic decomposition, the verbs used in the current study are said to share at least some conceptual components within their respective classes. For the category verbs of motion, the shared component would be the primitive concept GO, which denotes an event in which the agent or theme travels through space and the ''idiosyncratic'' information would determine how exactly the traveling was achieved. Although alternative views on lexical–conceptual relations exist (e.g., via meaning postulates [see de Almeida, 1999a]), it has to be acknowledged that the idea that shared meaning implies shared semantic or conceptual constituents represents the dominant view within the cognitive science community. In the case of the psychological verbs, however, it is not clear whether there would be a single conceptual component that could be shared by all of its members; verbs in this class can be intransitive or transitive with different types of roles assigned to their arguments. The verbs we used from this class can be found under different categories in Levin's (1994) classification (''Psychological,'' ''Desire,'' ''Social Interaction,'' etc.). It is clear, however, that the verbs we used all share the property of denoting psychological states of the subject or *experiencer* (they all share the feature +PSYCH, if you will), and thus we can assume that there is conceptual information shared by members of the class that is not captured by the linguistic approaches briefly discussed here. An alternative would be to assume that ''shared content'' is an effect of shared inferences unleashed by verb tokens (see de Almeida, 1999a, b) beyond the purely syntactic or lexical–semantic aspects of a verb's linguistic behavior. Again, a detailed discussion of these classes is beyond the scope of the present exploratory study (for a discussion on different classes of psychological verbs, see Grimshaw, 1990).

We assume that a speaker's ability to distinguish between verb classes in L2 should be at least in part a function of the speaker's representation of the two classes as having distinct semantic components (or templates or sets of inferences). Moreover, it should be expected that the more conceptual components are shared by members of a class, the more ''coherent'' the class is and, therefore, the ''tighter'' its semantic space (i.e., the closer the distribution of the members of the class). We take this view of shared conceptual constitutency as our working definition of ''semantic field'' even if it turns out that verbs are *not* represented by conceptual templates such as the ones proposed in the linguistic literature discussed here. We assumed that multidimensional scaling based on people's judgments of verb relatedness would provide us with a spatial view related to their representation of the verb classes. Assuming our tentative notion of semantic field to be correct, multidimensional scaling should provide a way to measure bilinguals' knowledge of verb classes in terms of overlapping conceptual constituents among members of a given verb class (e.g., among verbs of motion, among psychological verbs) and the degree of separation between classes.

METHOD

Participants

Participants were L1 speakers of English ranging in self-reported knowledge of French 2 to 5 (5 = *fluent*) with a mean of 3.5 collapsed across scales of French reading and speaking ability. No participants rated themselves as a fluent on both scales. All reported learning French as an L2 after first learning English. Only those who performed more slowly on the French than on the English version of the Simple Classification Task described below were retained for the study. The final group consisted of 18 university students (9 females and 9 males), ages 18 to 32 years. Participants were paid or received course credit for taking part in the study.

Stimulus Materials

Stimulus materials consisted of 24 verbs of motion and 24 psychological verbs in English and their translation equivalents in French, distributed into subsets of 6 verbs each for counterbalancing purposes.

In addition, there were 8 filler verbs distributed into two sets of 4 verbs each, and there were 4 warm-up verbs (see Appendix). In English, verbs were always preceded by *to* (*to run, to jump*, etc.) to highlight the verb rather than the noun nature of the words. In selecting these verbs, we ensured through pilot testing that they would be known, in the second language, to typical bilinguals found in Montreal. The verbs chosen tended to be of relatively high frequency within their respective languages. In English, the mean frequency for the verbs of motion was 104 per million ($SD = 129$) and for the psychological verbs was 114 per million ($SD = 110$), according to Francis and Kucera (1982) (difference not significant). In French, the corresponding mean frequencies were 253 ($SD = 269$) and 271 ($SD = 336$), according to the counts provided by Beaudot (1992) (difference not significant). Not too much importance should be attached to the apparent difference between languages in the mean frequencies of the verbs. The databases for the English and French language lists are not highly comparable, and the French data are incomplete insofar as no frequency information was available for the reflexive form of verbs (e.g., *s'approcher*). What these means do show, however, is that on the whole, the verbs are highly frequent in their respective languages.

Tasks

There were three tasks: (1) the Relatedness Task, (2) the Simple Classification Task, and (3) the Arbitrary Classification Task.

In the Relatedness Task, participants saw two verbs at a time in the center of the screen, below which was a Likert-type scale ranging from 1 (*not at all related in meaning*) to 7 (*highly related in meaning*). Participants clicked on the number they felt reflected the appropriate degree of relatedness between the two verbs. For a given participant, the verbs used in the Relatedness Task were 6 motion verbs and 6 psychological verbs taken from the full set of 48 verbs plus 4 filler verbs. These words were combined into pairs to form 120 trials.

In the Simple Classification Task, participants saw one verb at a time in the center of the screen, and they had to press a key to indicate whether the verb referred to motion, defined as action involving displacement of the person from one location to another, or to a psychological or mental state. Participants saw all 48 verbs in this task. The first 4 trials consisted of warm-up verbs. Reaction times and errors were recorded.

In the Arbitrary Classification Task, participants were told that they would see 12 verbs: 6 arbitrarily assigned to Set A and 6 arbitrarily assigned to Set B. Their task was to learn to correctly and quickly classify each verb as belonging to A or B. Stimuli were presented in blocks of 12 trials in which the 12 different verbs appeared one at at a time. Participants pressed keys marked ‘‘A’’ or ‘‘B’’ on each trial according to the category to which they believed the verb belonged. The computer provided auditory feedback each time the participants made an error, enabling gradual learning of correct assignment of verbs to A or B over time. Participants continued receiving such block until they performed 3 blocks in a row with at least 11/12 correct trials on each block or until they completed Block 23. Reaction times and errors were recorded. For a given participant, one of the four subsets of 6 motion verbs and one of the four subsets of 6 psychological verbs were used. Assignment of verbs to Sets A and B was different for each participant and was random with the restriction that each set had exactly 3 motion verbs and 3 psychological verbs in it.

Design

Each participant performed all three tasks, each first in one language and then in the other, with language order counterbalanced.

Participants were assigned to different counterbalancing groups that determined which subsets of stimuli were used in each task. In the Relatedness Task and the Arbitrary Classification Task, these subsets were selected so that (a) the English and French words were not translations of each other and (b) the words used in the Relatedness Task were different from those used in the Arbitrary Classification Task. All of the verbs were seen in the Simple Classification Task.

In the Relatedness Task, the verbs of motion were combined with each other to form a set of 15 unique within-set pairs, the psychological verbs were combined to form another set of 15 within-set pairs, and the motion and psychological verbs were combined to form 36 between-set pairs. Finally, the motion and psychological verbs were combined with filler verbs, and filler verbs were combined with each other to create the remainder of the 120 pairs. Filler verbs were included to encourage participants to think about the verbs more broadly than simply as motion or psychological verbs. Across the 120 pairs, the occurrence of motion and psychological verbs overall was equated, and the occurrence of each particular verb was approximately equal to that of each other verb. The 120 pairs were presented in a different random order for each participant with the following restrictions. First, the initial 12 tirals involved only pairs composed of 1 filler verb and 1 motion or psychological verb, to serve as warm-up trials and to make participants aware of the full range of words involved. Second, the last 4 trials involved only pairs composed of filler verbs, Third, the left–right location on the screen of motion and psychological verbs within a pair was counterbalanced, with each verb occurring approxiamately half of the time on each side.

In the Simple Classification Task, the first 4 trials consisted of the 4 warm-up verbs, and the remaining trials consisted of the 48 motion and psychological verbs presented in a different random order for each participant.

In the Arbitrary Classification Task, the order of presentation of the 12 verbs within each block was randomized from block to block. By way of illustration, 1 participant had to learn to classify as Group A the verbs *hurry, exit, rise, blame, value,* and *recall* and as Group B, the verbs *crawl, travel, step, understand, desire,* and *conclude.*

Procedure

All participants performed the Relateness Task, the Simple Classification Task, and the Arbitrary Classification Task in that order, half performing each task in English first and half performing it in French first. When performing the Relateness Task, no mention was made of the fact that most of the verbs could be classified as either motion or psychological verbs. Each participant was tested individually in a dimly lit room. All tasks were prepared in Apple Hypercard and presented on a Macintosh computer.

RESULTS

Simple Classification Tasks

The 18 participants' reaction times, when correctly classifying words as motion or psychological verbs, were faster in L1 (English) than in L2 (French) (1039 vs 1397 ms), $t(17) = 6.710$, $p < .001$. Their RTs in L2 were, on average, 36.19% ($SD = 22.78$) slower than in L1. Error rates in L1 were 3.2% and in L2 were 11.6%. This RT advantage for L1 was not surprising given that participants were retained in the study only if their L1 RTs were faster than their L2 RTs. In the rest of the analyses, we wished to use performance in the Simple Classification Task as an index of L2 ability on the assumption that the more fluent a person was in the language, the faster he or she would be able to recognize and understand words in that language. However, L1 RTs correlated very strongly with L2 RTs ($r = .907$, $n = 18$, $p < .001$), indicating that about 82% of the variance in L2 performance could be explained by the variance in L1 performance. This presumably reflects the fact the two tasks overlapped considerably in demands, each requiring similar reading skills, decision making, motor responses, and so on. Because of this overlap, raw L2 RTs seemed inappropriate as an index of L2 ability. Therefore, to obtain a measure not reflecting factors related to L1 performance, we partialled out L1 RTs from L2 RTs and used the residuals as the index of a person's L2 ability. Positive residuals indicated that the participant performed more slowly and hence at a lower level than expected relative to others in the group, after taking into account his or her L1 performance. Negative residuals indicated faster and hence better performance in L2 relative to others in the group, after taking into account the participant's L1 performance. It must be remembered, of course, that all of the bilinguals performed less well (i.e., more slowly) in L2 than in L1. These residuals served subsequent analyses as an L2 Ability Index.

Finally, an item analysis was carried out in which we compared the mean RTs for each verb averaged across the 18 participants. The factors in this analysis were Language (English or French), Verb Type (Motion or Psychological), and Frequency (High or Low on a median split within each subcategory). The analysis yielded a significant Language effect, with mean RT for English (L1) being faster (1004 ms) than in French (L2) (1237 ms), $F(1, 87) = 9.211$, $p = .0032$. There was also a marginally significant Language by Verb interaction, $F(1, 87) = 3.703$, $p = .0576$, in which L1 psychological verbs were responded to significantly faster (910 ms) than were motion verbs (1099 ms), whereas in L2 there was no significant difference between the two verb groups (1294 vs 1183 ms). All effects involving frequency were not statistically significant, all Fs < 1, except for the significant three-way interaction, $F(1, 87) = 5.030$, $p = .0275$, that could not be readily interpreted. (The pattern obtained indicated that for L1 psychological verbs, RTs were faster for low-frequency than for high-frequency verbs, and the opposite was true for motion verbs; for L2 psychological verbs, the exact opposite pattern occurred, with low-frequency psychological verbs, and high-frequency motion verbs being responded to faster than the corresponding high-frequency psychological and low-frequency motion verbs.)

Arbitrary Classification Task

This task proved to be very difficult in both languages. Only 1 participant reached the final criterion in L1 before Block 23, and none did so in L2. Reaction times were highly variable throughout, so it was difficult to identify in a consistent fashion when during learning a participant had reached optimal performance. To obtain a useful learning measure on this task that could be compared across L1 and L2, we decided to focus on performance in Blocks 4, 5, and 6. By Block 4, participants had received accuracy feedback three times on their classification of each of the 12 verbs. We reasoned that by Block 4, participants might have begun to classify some of the verbs correctly on a consistent basis. For each participant, we tallied across Blocks 4, 5, and 6, separately in L1 and L2, the number of different verbs that were responded to correctly on all three blocks. On this criterion, 13 of 18 participants performed in the predicted direction ($p < .05$, $n = 18$, sign test), performing correctly on more verbs in L2 (3.2) than in L1 (2.3). It was not possible to compare RTs for L1 and L2 verbs correctly responded to across these three blocks because many participants did not reach criterion in L1 on any verbs. Thus, the main finding in the Arbitrary Classification Task was that, once learning of the classification had begun, performance in L2 was significantly better than in L1 in the sense that more verbs were consistently responded to correctly in L2 than in L1. This provided another example of the paradoxical L2 performance effect described earlier.

We were also interested in whether L2 ability level correlated with performance on the Arbitrary Classification Task. To examine this, we correlated L2 ability—as indexed by the residualized RTs described earlier—with the number of L2 verbs reaching criterion in the Arbitrary Classification Task. This correlation was significant ($r = -.503$, $n = 18$, $p = .033$), indicating that the greater a person's ability in L2 (the smaller residual index), the more L2 verbs that person was able to classify correctly in Blocks 4, 5, and 6. The identical significant correlation ($r = -.503$, $n = 18$, $p = .033$) was obtained when L2 ability was correlated with the difference between L2 and L1 performance (L2-better-than-L1 effect) on the critical blocks in the Arbitrary Classification Task.

Relatedness Task

Each participant's ratings for pairs involving only motion and psychological verbs (no fillers) were converted from similarity ratings (relatedness ratings) to dissimilarity ratings by subtracting each score from 8, and these dissimilarity ratings were submitted to multidimensional scaling analysis in a separate analysis for each participant. The same analysis was performed for each participant, using a three-dimensional solution in every case, to compute the Euclidean distances between the 12 verbs for each participant. The analyses were run in Systat 5.1 using Kruskal scaling. We conducted analyses individual by individual rather than by collapsing the data across all participants because our interest was in how each individual's verb class separation was related to performance on the other tasks. In this study, we did not focus on the specific dimensions underlying verbs' semantic represenations in L1 and L2, although clearly that would be interesting to pursue.

The Euclidean distances obtained from these individual analyses were then used to compute, for each participant and in each language, a mean within-verb group distance, that is, the mean of the distances among the verbs of motion and among the psychological verbs. We also calculated the mean between-verb group distance, that is, the mean of the distances between each motion and each psychological verb. The ratio of the mean between-verb group divided by the mean within-verb group provided a dimensionless Verb Group Separation Index for each subject, indicating how far apart the motion and psychological verbs were from each other relative to that person's mean within-verb group distance.

We had two hypotheses concerning the Verb Group Separation Index. First, we originally predicted that the index would be smaller in L1 than in L2, reflecting the greater semantic richness and hence overlap of L1 verbs. The data did not support this. There was, in fact, a nonsignificant difference in the opposite direction (1.234 vs 1.209 for L1 and L2, respectively, $t(17) = 0.384$, *ns*).

Second, we hypothesized that performance in the Arbitrary Classification Task would reflect semantic field overlap; that is, better classification performance in L2 would reflect greater separation of the verb fields in L2. To test this hypothesis, we correlated performance on the Arbitrary Classification Task with the Verb Group Separation Index. We did this in two ways. We found that the L2 Verb Group Separation Index correlated significantly with L2 performance on the Arbitrary Classification Task after L1 performance on the Arbitrary Classification Task had been partialled out ($r = .656$, $n = 18$, $p = .003$). We also found that the correlation remained significant when we further partialled out the L1 Verb Group Separation index scores from the L2 Verb Group Separation index scores ($r = .669$, $n = 18$, $p = .002$). These results indicate that the better performance seen in L2 compared to L1 in the Arbitrary Classification Task was strongly related to the degree to which the verb categories were separated in L2 and to the degree to which this separation was greater in L2 than in L1.

Finally, we were interested to see whether there was a relationship between L2 ability level, as indicated by the L2 Ability Index, and the degree of semantic separation between verb categories. The correlation between each participant's Verb Group Separation Index difference (L2 minus L1, reflecting the degree to which the semantic fields were more separated in L2 than in L1) and the L2 Ability Index was highly significant ($r = .746$, $n = 18$, $p = .001$). This result indicated that the faster a person was at classifying motion and psychological verbs in L2, after taking his or her ability in L1 into account, the greater was the semantic separation between the two classes of verbs in L2 for that person.

DISCUSSION

The main results of this study are the following. First, we succeeded in producing both the normal L1-better-than-L2 effect and the paradoxical L2-better-than-L1 effect. Bilinguals were significantly faster in L1 than in L2 in the task requiring speeded classification of motion and psychological verbs but were more accurate in L2 than in L1 in the initial phases of learning a random assignment of motion and psychological verbs to different categories in the Arbitrary Classification Task.

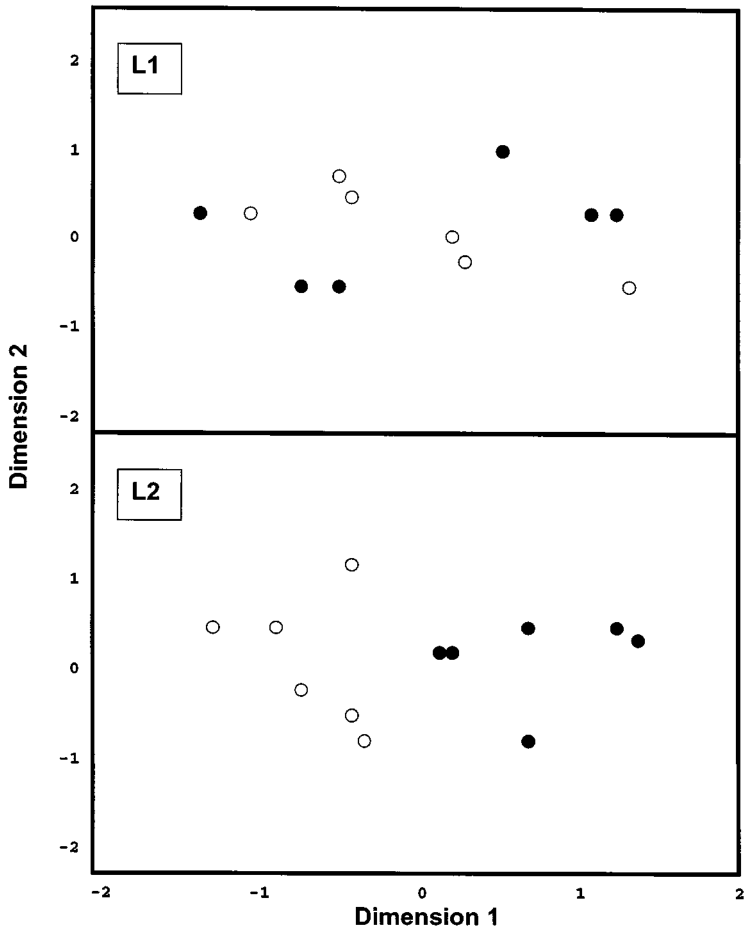

FIG. 1. The semantic space for participant ML in L1 (upper panel) and L2 (lower panel) showing the distribution of verbs of motion (open circles) and psychological verbs (solid circles) in a two-dimensional space.

Second, the Verb Group Separation Index scores derived from the Relatedness Task revealed important and statistically significant relationships to performance on the other two tasks. This finding in itself is interesting because it lends support to the use of multidimensional scaling for obtaining psychologically useful information about how words are semantically represented. Figure 1, for example, depicts two-dimensional solutions in L1 and L2 for 1 participant. In this case, it can be seen that the motion and psychological verbs in L1 have semantic spaces that overlap considerably, at least when viewed in a two-dimensional space. It is likely that with a higher dimensional solution—three or more dimensions—more separation would become evident. In L2, for this individual, the separation between motion and psychological verbs is already evident in the two-dimensional space. Of course, with a sufficient number of dimensions, greater separation can always be achieved. An interesting question, which was not pursued here, is whether one requires solutions of higher dimensionality in L1 than in L2 in general to achieve a given level of separation among different subsets of words. This might be the case because the better known L1 will generally involve richer (more overlapping) semantic representations.

In the current example, the Verb Group Separation Index for this particular was 0.968 in L1 and 1.688 in L2. This greater L2 separation, however, was not reflective of the general pattern obtained across the 18 participants; the Verb Group Separation Index in L2 was not significantly different from that in L1, contrary to what had been hypothesized. Nonetheless, the results did show that the Verb Group Separation Index was significantly correlated with the L2 Ability Index. This indicated that the less overlap there was in a bilingual's representation of motion and psychological verbs in L2, the better was that individual's ability to rapidly classify a stimulus word as a motion or psychological verb in L2. This, of course, was expected because the more distinctive verb groups are, the easier it should be to classify them. Interestingly, the results also revealed that the Verb Group Separation Index was significantly correlated with performance in the Arbitrary Classification Task where the L2-better-than-L1 effect was obtained. Here we found that the better performance was in L2 relative to L1 in the Arbitrary Classification Task (i.e., the greater the L2-better-than-L1 effect was), the greater was the Verb Group Separation Index in L2 than in L1. This pattern of data supports the original hypothesis proposing that the L2-better-than-L1 effect reflects the greater interaction (overlap in semantic representations) between words in L1 than in L2.

This latter point received additional indirect support in the pattern of results obtained in the Simple Classification Task. The item analysis based on verbs instead of subjects revealed that in L1 verbs of motion were responded to significantly more slowly than were psychological verbs, whereas in L2 there was a nonsignificant difference in the opposite direction. This is consistent with the representation overlap hypothesis. According to this hypothesis, in L2 the two verb categories are represented with relatively little overlap between them (either via their conceptual templates or via other mechanisms as discussed before), and so, other things being equal, there is no reason for one group of verbs to be responded to faster than the other. In L1, according to the overlap hypothesis, there should be relatively more overlap between the two categories because word meanings are generally richer in L1. Interestingly, the asymmetry in RTs in L1 may reflect the asymmetrical way in which verbs of motion and psychological verbs may overlap with each other. Verbs of motion tend to have extended meanings that include senses covered by many psychological verbs. That is, while *jump, arrive, come, fall,* and so on, are verbs of motion when understood in their primary senses, these verbs can also have meanings that may overlap with psychological verbs, as in *jump to a conclusion, arrive at a decision, come to believe, fall in love,* and so on. Psychological verbs such as *conclude, decide, believe,* and *love,* however, do not have corresponding extended senses that cover motion.

An important caution must be sounded here, however, regarding the interpretation of the results. Examination of the overall pattern of data points to a potential anomaly. We found in the current study that the greater an individual's L2 ability, with L1 taken as baseline, the greater was the L2 Verb Group Separation Index compared to the L1 Verb Group Separation Index ($r = .748$, $n = 18$, $p < .001$). This correlation could suggest the following odd situation. As a person becomes stronger and stronger in L2 and comes closer to achieving an optimal level of bilingual ability, the L2 Verb Group Separation Index will become even greater relative to the L1 Verb Group Separation Index. However, this clearly cannot be true. At the optimal level of bilingualism, performance in L2 should approximate performance in L1, and hence the Verb Group Separation Index scores in the two languages should become more similar. That is, meanings represented in a fluent bilingual's mental lexicon should be equally rich in the two languages and therefore should overlap with other meanings to a similar extent. In this case, the difference between the two separation index scores should tend toward zero and not become larger as implied by the correlation obtained in this study.

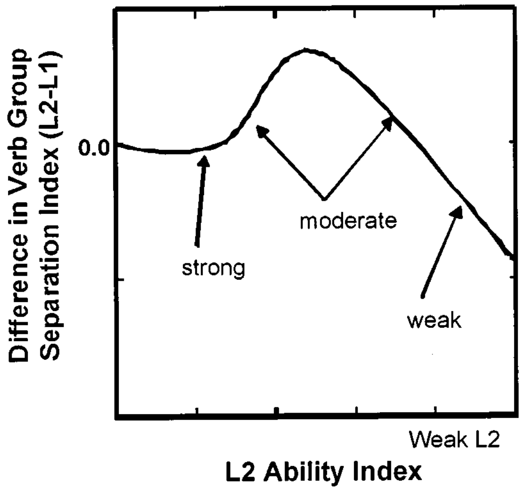

FIG. 2. Hypothesized relationship between L2 minus L1 difference in semantic overlap between categories of words within a given language and L2 performance ability. The curve shows how this relationship might vary for strong, moderate, and weak bilinguals.

This potential anomaly may be more apparent than real. Figure 2 illustrates how L2 ability may really be related to the Verb Group Separation Index differences between L2 and L1. At the high end of the scale of bilingual level, shown in the left-hand part of the graph marked ‘‘strong,’’ the difference between L2 and L1 Verb Group Separation Index scores is shown to be about zero; that is, representations in the two languages should behave in a similar manner. In the middle range of bilingual level, the part marked ‘‘moderate,’’ the relationship is shown as an inverted U. On the weak side of the middle range, the curve is shown as rising (as one moves from right to left), indicating that the Verb Group Separation Index in L2 becomes greater than that in L1, possibly because of the richer semantic connections that exist in L1 that obscure clear category boundaries. As L2 fluency improves, however, this cross-language difference must eventually decrease, with the Verb Group Separation Index scores becoming more similar and their difference approaching zero. It is possible, therefore, that the bilinguals in our study came primarily from the weak–moderate range of L2 fluency, and that is why we obtained a strong positive correlation. If we had included a wider range of bilinguals, then the pattern might have been quite different. A question for further research, then, is whether the following stages of development in bilinguals’ mental representation of verbs may be identified. At first, bilinguals have difficulty in making fine distinctions between verb categories due to their impoverished representations of the meanings of those verbs. Then, as language skills develop, they are better able to differentiate verbs on the basis of their primary meanings, and this provides a good basis for making between-group distinctions. With even greater language skills, however, secondary meanings enter into the picture and previously clearly defined categories of verbs now begin to overlap more, but for reasons different from those for the overlap expected with weaker bilinguals.

Why should the L2 mental representation of basic words such as *run, jump,* and *believe* be relatively impoverished and involve relatively less overlap with other word representations in comparison to L1? Perhaps, more than is the case in L1, L2 lexical representations directly reflect specific communicative experiences with words. That is, as Jiang and Forster (2001) recently argued, perhaps in late L2 learners the mental representation of words is episodically based rather than semantically based. If this is true, then it would not be surprising that even basic words would be less richly represented in L2 given that the learners would have had significantly fewer and less varied encounters with these words. While L2 expressions such as *come to believe* and *jump to conclusions* may be comprehensible to late learners based on semantic considerations, the normal mental representations of *come* and *jump* for these bilinguals might not include these meanings extensions when these expressions have not been directly encountered before in the L2. Thus, to the extent that late learners' L2 communicative experience is impoverished relative to L1, semantic representations—even for basic vocabulary—will be impoverished as well. Note that we are drawing attention here to the low frequency with which the learners encounter the L2 verbs, not to the frequency of the verbs themselves within the language. In L1, for example, some verbs naturally occur less frequently than others, but there is no reason to believe that performance with these verbs should therefore resemble performance with L2 verbs, even though the latter also occur with low frequency in L2 users' experience. This is because L1 users will still have had a great deal of exposure to L1 verbs, so all or most of these verbs' extended senses will be known and readily available, even if some of them happen to occur with low frequency in the language. By contrast, in L2 the overall low rate of exposure will skew the users' experience toward core meanings, because for any given verb the core meaning will tend to be the one most frequently encountered.

Jiang and Forster (2001) recently reported interesting data supporting the idea that late learners represent L2 word meanings episodically. They investigated masked-translation cross-language priming asymmetries with Chinese speakers of English as a second language. In one condition, participants performed a lexical decision task on L1 (Chinese) target words preceded by masked L2 (English) words that were either translation equivalents of the target or control words. Participants had to decide whether a target stimulus was a real word in the language, a task that required them to consult their mental lexicon. Jiang and Forster found no significant facilitation for lexical decision to L1 words by the L2 translation primes compared to control primes. They also conducted an episodic recognition task where instead of judging whether a target stimulus was a real word, participants had to judge whether the target stimulus had been seen in a studied list shown just prior to the main task. In the main task, participants had to consult episodic memory to determine whether a stimulus word had been experienced during the earlier study event. In contrast to the lexical decision results, the authors found that L2 translations did prime L1 target words. Finally, they examined the ability of L1 words to prime L2 words; they found L1 priming effects in lexical decision but not in the episodic memory task. The study manipulated stimulus onset asynchrony and introduced measures to rule out possible confounds due to differences in the tasks and speed with which L1 and L2 words may be processed. Thus, Jiang and Forster reported another interesting example of an L2-better-than-L1 effect (here, in episodic memory priming). Moreover, the explanation they drew from their study—that L2 words are mentally represented in episodic memory rather than in a lexical store—may provide the basis for the effects we observed in our study. Specifically, late learners, such as those who participated in our study, may represent their L2 vocabulary in memory in ways that are strongly related to the experience in which they encounter the words. To the extent that this experience is impoverished in L2 relative to L1, these representations will be impoverished, resulting in an L2-better-than-L1 effect under circumstances such as those created in our study.

The current study supported previous findings from our laboratory and from others of an L2-better-than-L1 effect, even though L2 was the *nondominat* language. The results also suggested that this effect may reflect, in part, differences between the way in which semantic fields in L1 overlap with each other compared to field overlap within L2. It was clear from the results that the relationship between semantic field organization and L2 fluency was complex, and future studies will need to include a far wider range of individuals with respect to L2 ability and other classes of lexical concepts. Future research using the techniques developed here should also explore in greater detail the specific nature of semantic field interactions in L1 and L2 and should further explore how concepts labeled by verbs are represented in the minds of bilinguals.

APPENDIX

Following are verbs of motion, psychological verbs, and filler and warm-up verbs, in English and French (italics), as distributed into subsets for counterbalancing purposes.

Verbs of motion: (1) to walk, to run, to jump, to climb, to leave, to rush; *se promener, courir, sauter, grimper, partir, se presser;* (2) to arrive, to enter, to approach, to return, to come, to bring; *arriver, entrer, s'approcher, retourner, venir, apporter;* (3) to cross, to descend, to follow, to escape, to fall, to swim; *traverser, descendre, surivre, s'évader, tomber, nager;* (4) to crawl, to hurry, to travel, to rise, to exit, to step; *ramper, se dépêcher, voyager, se lever, sortir, marcher*.

Psychological verbs: (1) to doubt, to think, to wonder, to like, to hate, to wish; *douter, penser, se demander; aimer, détester, souhaiter;* (2) to want, to remember, to forget, to decide, to hope, to agree; *vouloir, se souvenir, oublier, décider, espérer, s'accorder;* (3) to believe, to imagine, to forgive, to oppose, to realize, to admire; *croire, imaginer, pardonner, contrarier, réaliser, admirer;* (4) to value, to blame, to understand, to recall, to conclude, to desire; *valoriser, reprocher, comprendre, se rappeler, conclure, désirer.*

Filler verbs (Relatedness Task): (1) to answer, to give, to play, to push; *répondre, donner, jouer, pousser;* (2) to ask, to take, to work, to pull; *demander, prendre, travailler, tirer.*

Warm-up verbs (Simple Classification Task): to go, to drive, to feel, to judge; *aller, conduire, se sentir, juger.*

REFERENCES

Beaudot, Jean (1992). *Fréquences d'utilisation des mots en français écrit contemporain*. Montréal: Presses de l'Université de Montréal.

de Almeida R. G. (1999a). *The representation of lexical concepts: A psycholinguistic inquiry*. Unpublished Ph.D. dissertation, Rutgers University.

de Almeida, R. G. (1999b). What do category-specific semantic deficits tell us about the representation of lexical concepts? *Brain and Language,* **68,** 241–248.

Favreau, M., Komoda, M. K., & Segalowitz, N. (1980). Second language reading: Implications of the word superiority effect in skilled bilinguals. *Canadian Journal of Psychology,* **34,** 370–380.

Favreau, M., & Segalowitz, N. S. (1982). Second language reading in fluent bilinguals. *Applied Psycholinguistics,* **3,** 329–341.

Favreau, M., & Segalowitz, N. S. (1983). Automatic and controlled processes in the first- and second-language reading of fluent bilinguals. *Memory and Cognition,* **11,** 565–574.

Francis, W. N., & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar.* Boston: Houghton Mifflin.

Franks, J., Bilbrey, C. W., Lien, K. G., & McNamara, T. P. (2000). Transfer-appropriate processing (TAP) and repetition priming. *Memory and Cognition,* **28,** 1140–1151.

Grimshaw, J. (1990). *Argument structure.* Cambridge, MA: MIT Press.

Jackendoff, R. (1990). *Semantic structures.* Cambridge, MA: MIT Press.

Jiang, N., & Forster, K. (2001). Cross-language priming asymmetries in lexical decision and episodic recognition. *Journal of Memory and Language,* **44,** 32–51.

Kroll, J. F. (1993). Accessing conceptual representations for words in a second language. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 53–81). Amsterdam: John Benjamins.

Kroll, J. F., & de Groot, A. M. B. (1997). Lexical and conceptual memory in the bilingual: Mapping form to meaning in two languages. In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 169–199). Mahwah, NJ: Erlbaum.

Levin, B. (1994). *English verb classes and alternations.* Chicago: Univ. of Chicago Press.

McLaughlin, B. (1998). Second language learning revisited: The psycholinguistic perspective. In A. Healy & L. Bourne, Jr. (Eds.), *Foreign language learning* (pp. 399–411). Mahwah, NH: Erlbaum.

Miyake, A., & Friedman, N. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A. Healy & L. Bourne, Jr. (Eds.), *Foreign language learning* (pp. 339–364). Mahway, NJ: Erlbaum.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General,* **106,** 226–254.

Rappaport Hovav, M., & Levin, B. (1997). *Building verb meaning.* Unpublished manuscript, Bar-Ilan University and Northwestern University.

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

Scarborough, D., Cortese, C., & Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance,* **3,** 1–17.

Segalowitz, N. (1986). Skilled reading in the second language. In J. Vaid (Ed.), *Language processing in bilinguals: Psycholinguistic and neuropsychological perspectives* (pp. 3–19). Hillsdale, NJ: Erlbaum.

Segalowitz, N. (2000). Automaticity and attentional skill in fluent performance. In H. Riggenbach (Ed.), *Perspectives on fluency,* (pp. 200–219) Ann Arbor: Univ. of Michigan Press.

Segalowitz, N., & Frenkiel, S. (in preparation). Cognitive control and skilled performance: Does language-specific attention control play a role in second language fluency?

Segalowitz, N., & Hébert, M. (1990). Phonological recoding in the first and second language reading of skilled bilinguals. *Language Learning: A Journal of Applied Linguistics,* **40,** 503–538.

Segalowitz, N., Poulsen, C., & Komoda, M. (1991). Lower level components of reading skill in higher level bilinguals: Implications for reading instruction. *AILA Review,* **8,** 15–30.

Segalowitz, N. S., & Segalowitz, S. J. (1993). Skilled performance, practice, and the differentiation of speed-up from automatization effects: Evidence from second language word recognition. *Applied Psycholinguistics,* **14,** 369–385.

Vasos, H. (1983). *Semantic processing in bilinguals.* Unpublished Ph.D. dissertation, Concordia University, Montreal.