Retrieval of Bilingual Autobiographical Memories:

Effects of Cue Language and Cue Imageability

Linda Mortensen, Dorte Berntsen, Ocke Schwen-Bohn*

Center on Autobiographical Memory Research,

Department of Psychology and Behavioral Sciences, Aarhus University, Denmark

* Corresponding author (engosb@hum.au.dk)

NOTICE: this is the author's version of a work that was accepted for publication in Memory. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Memory, 23(2):138-56, DOI: 10.1080/09658211.2013.873809

Abstract

An important issue in theories of bilingual autobiographical memory is whether linguistically encoded memories are represented in language-specific stores or in a common language-independent store. Previous research has found that autobiographical memory retrieval is facilitated when the language of the cue is the same as the language of encoding, consistent with language-specific memory stores. The present study examined whether this language congruency effect is influenced by cue imageability. Danish-English bilinguals retrieved autobiographical memories in response to Danish and English high- or low-imageability cues. Retrieval latencies were shorter to Danish than English cues and shorter to high- than low-imageability cues. Importantly, the cue language effect was stronger for low- than high-imageability cues. To examine the relationship between cue language and the language of internal retrieval, participants identified the language in which the memories were internally retrieved. More memories were retrieved when the cue language was the same as the internal language than when the cue was in the other language, and more memories were identified as being internally retrieved in Danish than English, regardless of the cue language. These results provide further evidence for language congruency effects in bilingual memory and suggest that this effect is influenced by cue imageability.

The present study was motivated by the finding that autobiographical memories that are encoded in a particular language are retrieved more easily when the language of the retrieval cue is the same as the language used during encoding than when the languages of encoding and retrieval are different (e.g., Marian & Neisser, 2000). The present study examined whether the strength of this language congruency effect depends on the type of word cue used, in particular whether it depends on the imageability of the word cue. Before turning to the present experiment, we briefly review first the autobiographical memory studies that motivated the present research and then previous lexical access and autobiographical memory studies that have reported cue imageability effects.

Language-Congruency Effects

The language congruency effect has been observed in the retrieval of episodic memory in general, such as in the learning of word lists (e.g., Sahlin, Harding, & Seamon, 2005) and in the retrieval of semantic knowledge (e.g., Marian & Fausey, 2006; Marian & Kaushanskaya, 2007). However, it has been found most consistently in the retrieval of autobiographical memories. A central issue in current theories of bilingual autobiographical memory is whether autobiographical memories are represented in language-specific stores depending on the language in which they were encoded, or whether they are represented in a common language-independent store. Most of the studies investigating this issue have all used the cue word paradigm (Crowitz & Schiffman, 1974), in which spoken or written words are presented to cue autobiographical memories, and each of the cued memories is then subsequently described orally or in writing. The participants are typically consecutive bilinguals who are individuals who were monolingual for many years, but then became bilingual, mostly as a result of immigrating to another country. Given the participants' bilingual background, the language used at recall can be varied by varying the language in which the cue words are presented and the language that is spoken in the recall session, and the hypothesis that

memory recall is language-dependent can be tested. If the language used during memory encoding is accessed during later recall of the memory, this should be reflected in an effect of the language (first or second) used during recall on which memories are remembered, with cue words presented in the first language (L1) biasing bilinguals towards recalling memories originally encoded in that language, and cue words presented in the second language (L2) biasing them towards recalling memories originally encoded in that language.

Whereas the language used during memory recall can be varied experimentally, it is more difficult to accurately measure previous language use, and hence to determine the language used when a memory was encoded. Various methods have been used: Bilinguals have been asked about their age at the time of the remembered event, about the language in which the memory came to them at the time of recall, and about the language spoken at the time of the remembered event. Below, we review four studies that used one or more of these methods.

Schrauf and Rubin (1998) tested Spanish-English bilinguals who had immigrated to the US at a mean age of 28 years and who had a mean age of 65 years at the time of testing. The cue words used were English nouns and their Spanish translation equivalents. The participants saw the English words on one day and the Spanish words on another day and were asked for each word to write down a description of the first autobiographical memory that came to mind using the language used on that day. After having completed the recall task, they were asked to report which of the memories just described were retrieved internally (“came to them”) in the language opposite to the language used to describe the memory, so-called crossover memories, and were asked to estimate their age at the time when each of the remembered events occurred. Contrary to the prediction, L1 Spanish words did not cue earlier memories than L2 English words: For events recalled in response to Spanish and English words, the mean reported age of the participant at the time of the event was 39.8 and 40.6 years, respectively. This suggests that the language of the cue words did not provide

privileged access to memories originally encoded in that language. However, participants identified 20% of their memories as being retrieved internally in the language not being used on that day, and the distribution of these crossover memories across the participants' life span was affected by the language of internal retrieval: Memories retrieved internally in L1 Spanish were earlier than memories retrieved internally in L2 English (mean reported age of participant at time of event: 29.7 and 46.5 years, respectively). Thus, the language of inner (but not overt) speech appears to provide privileged access to memories originally encoded in that language.

Schrauf and Rubin (1998) reasoned that memories can be retrieved internally in one language, and that this language of inner speech is indicative of the language of encoding. The memories can then subsequently be recoded and described in another language. Thus, they concluded that the identification of crossover memories and their temporal distribution were consistent with linguistic encoding of autobiographical memories. They further proposed that the failure to find an effect of the language of the cue words on the age of the recalled memories was related to language proficiency: The bilinguals in their study were highly proficient speakers of both English and Spanish who could access their personal memories equally well in either language and, if needed, translate the language in which the memory was retrieved into the language used in the session. This can explain the high number of early memories retrieved in response to English cue words.

The results obtained by Schrauf and Rubin (1998) are consistent with linguistic encoding of autobiographical memories, but they do not show that memory recall is language-dependent, i.e., that a congruency between the languages used during encoding and during recall facilitates recall. This is because only crossover, and not congruent, memories, were reported. Congruent memories are memories cued in one language and retrieved internally in the same language. If the language of encoding is retrieved internally at recall and if language congruency between encoding and recall facilitates recall, bilinguals should recall more congruent than crossover memories.

This prediction was confirmed in a follow-up study by Schrauf and Rubin (2000). They tested a subgroup of the bilinguals from the earlier study and used the same procedure, except that they asked the participants to report not only which of the memories described were retrieved internally in the language opposite to the language used to describe the memory (crossover memories), but also which of the memories were retrieved internally in the same language as was used to describe the memory (congruent memories). Averaged across the two cue languages, participants retrieved more congruent than crossover memories (47 vs. 25%), but this difference was larger in response to L1 Spanish cues (59 vs. 21%) than L2 English cues (36 vs. 30%). This shows that memories were easier to recall when the language of the cue word was the same as the language of internal retrieval, but that this language-congruency effect was boosted for the native language. Replicating the results of the earlier study, the temporal distribution of the recalled memories was affected by the internal language: Memories retrieved internally in L1 Spanish were earlier (mean reported age of participant at time of event: 28.6 and 26.9 years for congruent and crossover memories, respectively) than memories retrieved internally in L2 English (48.2 and 52.9 years) (see also Larsen, Schrauf, Fromholt, & Rubin, 2002).

A follow-up study by Schrauf (2009) examined whether autobiographical recall in Spanish monolinguals and Spanish-English bilinguals was different for nonlinguistic picture cues (from Snodgrass & Vanderwaart 1980) as opposed to (linguistic) word cues, and whether the language of the cue word had an effect on autobiographical recall. Schrauf (2009) reported no significant differences between memories cued by pictures vs. memories cued by words for any of the variables examined (RT, visual and auditory detail, emotional intensity, significance of the event, frequency of rehearsal) for both the monolingual and the bilingual participants. This suggests that autobiographical retrieval is independent of cue type (non-linguistic visual vs. orthographic linguistic). The Spanish-English bilinguals in the Schrauf (2009) study were further examined for

language congruency effects between the language of the cue word and the “language of the memory” according to the participants’ self report. Schrauf (2009, 40) summarizes his findings by stating that “words in the L1 are more likely to trigger memories in L1. Words in L2 are more likely to trigger memories in L2.”

Schrauf and Rubin (1998, 2000) used the participant’s mean reported age at the time of the remembered event to determine the language of encoding. The validity of using this measure was questioned by Marian and Neisser (2000) who argued that bilinguals’ age at the time of the remembered event does not necessarily indicate the language that was used when the event was encoded, especially not for fluent bilinguals who have used both languages concurrently for most of their lives. Schrauf and Rubin’s (1998) Spanish-English participants had spent at least 30 years in the US at the time of testing and reported speaking English and Spanish about equally frequently. According to Marian and Neisser (2000), this can explain the failure to find significant differences in the ages of the memories retrieved by participants when cued with Spanish and when cued with English words (see also Matsumoto & Stanny, 2006, for evidence supporting this proposal).

Marian and Neisser (2000) tested Russian-English bilinguals who immigrated to the US at a mean age of 14 years and who had a mean age of 22 years at the time of testing. Instead of using age at the time of the remembered event, they explicitly asked their L1 Russian L2 English participants about what language was spoken at the time of the remembered event and used this as a measure of the language of encoding. The participants were interviewed and cued in both languages on the same day, but in separate sessions. They were asked for each spoken word to describe orally the first autobiographical memory that came to mind using the language of that session. After having completed the recall task, they were asked to indicate the language spoken during each remembered event and to estimate their age at the time. The participants retrieved more memories that were encoded in the same language as that used during the interview than memories that were

encoded in the other language. This was true both for memories encoded in Russian (means of 5.2 vs. 2.9 out of possible 8 memories) and for memories encoded in English (means of 3.4 vs. 1.3 memories). Thus, the language of the interview and of the cue words biased participants towards retrieving memories that had occurred in that language. The memories retrieved were earlier when interviewed in L1 Russian than when interviewed in L2 English (mean reported age of participant at time of event: 13.1 vs. 16.1 years, respectively). These findings are consistent with language-dependent recall of autobiographical memory. In a second experiment, Marian and Neisser showed that the interview language and the cue language contributed independently to this effect. This finding of an independent effect of interview language led them to conclude that language functions similarly to other forms of context, so that reinstating the language of encoding at recall facilitates recall (see also Matsumoto & Stanny, 2006; for reviews, see Schrauf, 2000; Schrauf, Pavlenko, & Dewaele, 2003; Schrauf & Rubin, 2003).

In sum, studies of autobiographical memory have demonstrated that bilinguals' access to their memories is influenced by the language spoken at the time of recall: When their L1 is spoken, they tend to access earlier memories, relative to when their L2 is spoken (Marian & Neisser, 2000; Matsumoto & Stanny, 2006; but see Schrauf & Rubin, 1998), to retrieve the memories internally in their L1 (Larsen et al., 2002; Matsumoto & Stanny, 2006; Schrauf & Rubin, 1998, 2000; Schrauf 2009), and to access memories of events that occurred when their first language was spoken (Marian & Neisser, 2000; Matsumoto & Stanny, 2006). The preferred interpretation of these findings is that linguistically encoded memories of bilinguals are stored in a language-specific way depending on the language in which they were encoded.

Concreteness Effects¹

A central issue in theories of bilingual lexical access is whether each of the words in a translation pair has its own semantic representation or whether the two words share a common semantic

representation. Available evidence suggests that this depends on particular word properties. One word property that has been studied extensively is word concreteness. Studies of lexical access within a single language have demonstrated that concrete words are processed faster and more accurately than abstract words – an advantage called the concreteness effect. An early demonstration of this effect was provided by James (1975) who used a lexical decision task and showed that, at least when word frequency is low, concrete words are identified as words faster than are abstract words. Using a correlational approach, Whaley (1978) and Rubin (1980) showed that there is a significant negative correlation between concreteness ratings and lexical decision latencies for the same words. This concreteness advantage has since been replicated in both lexical decision (de Groot, 1989; Kroll & Merves, 1986), word naming (Schwanenflugel & Stowe, 1989), word association (de Groot, 1989), and word recall (e.g., Allen & Hulme, 2006; Romani, McAlpine, & Martin, 2007; Walker & Hulme, 1999).

The two most influential accounts of the concreteness effect are the dual coding theory (Paivio, 1971, 1986, 2007) and the context availability model (Schwanenflugel, 1991; Schwanenflugel, Harnishfeger, & Stowe, 1988; Schwanenflugel & Shoben, 1983). In both of these accounts, words referring to concrete concepts are assumed to have richer semantic representations than words referring to abstract concepts. According to the dual coding theory, both concrete and abstract words are represented in a verbal linguistic system, but concrete words alone are also represented in a nonverbal, image-based system. The cognitive advantage for concrete words is attributed to the availability of both a verbal and an imagistic representation for this type of words. According to the context availability model, the types of representation that are available are less important than the strength of the connections between them. Accessing the meaning of a word involves accessing a network of associated semantic information, and the advantage for concrete words arises because they have stronger and denser associations to contextual knowledge than do abstract words. In sum,

these two accounts disagree on exactly how concrete and abstract words differ (in the type of representations that are available for them or in the density of their associative networks), but they both imply that concrete words are associated with a richer set of semantic features than are abstract words.

This assumption, that concrete and abstract words differ in semantic richness, has been extended to theories of cross-language processing. There are two theoretical views on how this difference is reflected in the memory representations of bilinguals. One is an extension of the dual coding theory to bilingual memory (Paivio, Clark, & Lambert, 1988; Paivio & Desrochers, 1980). This theory assumes that word meanings are represented in two verbal systems, one for each of a bilingual's two languages, and in one image system. Concrete words are represented in both verbal systems and in the image system, whereas abstract words are only represented in the verbal systems. Translations of words that have a representation in the image system are assumed to share this representation. The other view is formulated in the distributed feature model (de Groot, 1992a) according to which the difference between concrete and abstract words is not in the type of representations available, but in the semantic density of their representations. In this model, word meanings are represented as sets of features and the semantic overlap of two translations is determined by the number of features that are shared. The meanings of concrete words have more semantic features than the meanings of abstract words, and therefore have more semantic features to share with their translations across languages. This proposal is based on the assumption that the functions of the objects to which concrete words refer are quite similar across languages. By contrast, abstract words tend to be used in different contexts across languages and hence will be less similar in meaning.

Studies investigating this issue of bilingual lexical access have used a variety of experimental tasks. One task that is often used is word association. In the bilingual version of this task, stimulus

words are typically presented in one or both of the bilingual's two languages, and associative responses are produced in either the language of the stimulus words (within-language response) or in the other language (between-language response). The important issue is whether the within- and between-language responses are the same or different. Responses that are translations of each other are classified as 'same' responses. A high proportion of same responses is taken as evidence that conceptual representations are shared between languages, whereas a high proportion of different responses is taken to indicate that conceptual representations are language-specific. A first study of this issue was conducted by Kolers (1963). Bilinguals who had English as their L2 and German, Spanish, or Thai as their L1 wrote down within- and between-language associative responses to each of a series of stimulus words. A comparison of the within- and between-language responses to each stimulus word showed that there were more different than same responses. Kolers concluded that word meanings are represented in language-specific stores. However, a considerable proportion of the within- and between-language responses was translations of each other (e.g., *king-queen*, *king-reina*), which suggests that at least some word-translation pairs share a conceptual representation in bilingual memory. Importantly, concrete words more often elicited same responses within and across languages than did abstract words (see also Kolers & Gonzalez, 1980; Taylor, 1976). In Paivio's dual coding theory (e.g., Paivio & Desrochers, 1980), this has been interpreted to indicate that the associative responses to concrete words are mediated partly by representations in the image system that is shared between the translation equivalents in each of the bilingual's two verbal systems, whereas associative responses to abstract words are determined primarily by the associative network within each of the two verbal systems. The shared image representation of concrete words can explain the higher proportion of same responses to concrete words than to abstract words.

This difference in the response pattern for concrete and abstract words was replicated in a more recent study by van Hell and de Groot (1998a). Dutch-English bilinguals produced spoken word associates to stimulus words that were either concrete or abstract. A comparison of the within- and between-language responses to each stimulus word revealed that concrete stimulus words were more likely to elicit equivalent responses across languages (e.g., “skirt” eliciting “dress” and “jurk”) than were abstract stimulus words (e.g., “revenge” eliciting “anger” and “boosheid”). This finding suggests that the equivalence in meaning across languages depends on concreteness, with translations of concrete words being more equivalent than translations of abstracts words (for a review see de Groot, 1992b). In line with the distributed feature model (de Groot, 1992a), van Hell and de Groot (1998a) proposed that the meanings of concrete words have more semantic features than the meanings of abstract words, and therefore have more semantic features to share with their translations. This interpretation is consistent with the results of studies using other tasks, including lexical decision and word translation (Jin, 1990; de Groot, Dannenburg, & van Hell, 1994; de Groot & Poot, 1997; van Hell & de Groot, 1998b). It is also consistent with the finding that in bilinguals’ autobiographical narratives, cross-linguistic transfers (i.e., the use of semantic or syntactic structures from the other language without overtly switching to that language) occur more often for concrete than abstract words (Marian & Kaushanskaya, 2007).

However, other studies have shown equivalent cross-linguistic priming effects for concrete and abstract words. For instance, Francis and Goldmann (2011) found that semantic classifications of words (into concrete or abstract) in one language at encoding led to shorter semantic classification latencies at test for items that were repeated in the other language than for new items. Importantly, this cross-linguistic priming effect was similar in size for concrete and abstract nouns. This finding can be interpreted to suggest that both concrete and abstract words have shared conceptual representations across languages and that abstract words do not share fewer conceptual features

than concrete words. This interpretation is supported by the results of word translation studies showing that concreteness effects across languages are eliminated when the concrete and abstract words are matched on context availability (van Hell & de Groot, 1998b) and when limited to words with only a single translation equivalent (Tokowicz & Kroll, 2007). These results suggest that concreteness effects might reflect differences in the context availability of the two types of words, with context being more easily available for concrete than abstract words, or differences in the number of translations across languages, with concrete words having fewer translations across languages than abstract words (e.g., “fødselsdag” has only a single translation equivalent in English, namely “birthday”, whereas “valg” has at least two translation equivalents, namely “choice” and “election”).

The lexical access studies reviewed above invite an alternative view of the relationship between memory and language in bilinguals than the one offered by the autobiographical memory studies reviewed earlier. In this view, the semantic representations of L1 and L2 translation equivalents are not exclusively language-specific or exclusively shared. Instead, different types of word overlap in meaning to different degrees.

Autobiographical memory studies have examined effects of imageability, but only within a single language. Rubin and Schulkind (1997) showed that imageability correlated with the time taken to retrieve autobiographical memories (with concrete words having shorter retrieval times) and with the age of the memories (with concrete words cuing older memories). An effect of cue-word imageability has since been demonstrated in several experimental studies. Williams, Healy, and Ellis (1999) used the cue-word technique to examine the effect of varying the imageability of English cue words on the speed of retrieving autobiographical memories and on the specificity, vividness, and pleasantness of the memories retrieved. They found that compared to low-imageability cue words (e.g., *interest*), high-imageability cues (*lake*) elicited memories that were

retrieved faster, were more specific, more vivid, and more pleasant. These results suggest that the imageability of the retrieval cue determines the degree of specificity of the retrieved memory. Subsequent studies have shown that this cue imageability effect is enhanced when memories are retrieved under dual-task conditions (Anderson, Dewhurst, & Nash, 2012; Williams, Chan, Crane, Barnhofer, Eade, & Healy, 2006). Other studies have compared high-imageability words and emotion words on their efficiency in cuing memories and have consistently found faster retrieval in response to the former than the latter cues (Conway & Bekerian, 1987; Larsen & Plunkett, 1987; Robinson, 1976; Uzer, Lee, & Brown, 2012).

To conclude, the mental representations of high- and low-imageability (or concrete and abstract) words differ in a number of ways (see also Altarriba & Bauer, 2004; Altarriba, 2006), and identifying which of these differences underlies the observed effects in bilingual language processing is difficult. For the purposes of the present study, knowing which difference is the critical one is not important. What is important is to construct two sets of words that differ in imageability or in some highly correlated variable (e.g., concreteness, context availability, or number of word meanings).

The Present Study

The lexical access studies reviewed above investigated effects of word imageability to examine how individual words are represented in the memory of bilinguals. A few studies of autobiographical memory have investigated the same effect, but within a single language. Several autobiographical memory studies have concerned the representations of memories across languages and have observed language congruency effects, but these studies did not vary cue-word imageability. To the best of our knowledge, no studies have examined whether word imageability modulates the effect of language congruency on bilinguals' retrieval of autobiographical memories.

The goal of the present study was to shed further light on the influence of language on the representations of autobiographical memories by examining effects of both word imageability and language congruency on bilinguals' retrieval of autobiographical memories.

The independent variables were cue-word imageability (high or low) and cue-word language (Danish or English). We predicted an advantage when the language of the cue was Danish compared with when it was English, as reflected in shorter recall latencies in the former than the latter condition. This prediction is based on the assumption that the participants, whose dominant language was Danish, were likely to have encoded their memories in Danish, yielding language-congruent (and thus fast) recall when the cue word was in Danish and language-incongruent (and thus slower) recall when it was in English. Based on the existing evidence (see the studies reviewed in the introduction), we also predicted an advantage for the high-imageability words, with shorter recall latencies when the cue word was high than when it was low in imageability.

The question of interest was whether these two effects would interact, with the effect of cue language being stronger when low-imageability words were used as cues than when high-imageability words were used. Recalling a memory in response to a high-imageability cue should be only mildly affected by the language of the cue. This is because high-imageability words tend to have very similar meanings across languages, and hence the language of the cue word should be less salient as a cue to guide memory retrieval. In other words, the word "birthday" should be almost as efficient a cue as the Danish translation "fødselsdag" at retrieving memories encoded in Danish. By contrast, recalling a memory in response to a low-imageability cue should show a stronger effect of cue language, with much slower recall when the cue was in English than in Danish.

Method

Participants. The experiment was conducted with 24 students of Aarhus University (mean age: 22.5 years, $SD = 4.7$) who were all native speakers of Danish. Their language proficiency (understanding, speaking, and reading) as well as age of acquisition and current use were assessed using the Language Experience and Proficiency Questionnaire (LEAP-Q, Marian, Blumenfeld, & Kaushanskaya, 2007). Self-reported ratings of proficiency were obtained on a scale from 0 (none) to 10 (high) for both Danish and English. The participants' responses showed that they began learning English at a mean age of 8.5 years ($SD = 2.0$). Their mean current exposure was 59.5% ($SD = 12.8$) and 39.4% ($SD = 13.7$) of their time for Danish and English, respectively. Their mean proficiency speaking English was 7.9 ($SD = 1.0$). Speaking proficiency was used because self-reported proficiency in speaking a second language is a reliable predictor of behavioral performance on standardized tests of second-language ability. Therefore, if questionnaire data are used to make general inferences about language function, speaking proficiency should be used to index second-language ability (Marian et al., 2007).

Materials. Sixty-four words were used on experimental trials. Half of these were English words and the other half were the Danish translations. The English words were selected from various published corpora of imageability ratings (see footnote to Table 1). Half of the English items were high and the other half were low in imageability (mean of 597.8, $SD = 82.0$, and 285.4, $SD = 47.6$, respectively, $F(1, 60) = 347.41$, $p < .001$). We did not collect imageability ratings for the Danish words, because it has been shown that imageability ratings correlate strongly across languages (see de Groot et al., 1994, for correlations between English and Dutch imageability ratings).³ It is therefore to be expected that the imageability of the Danish words correlates strongly with that of their English translations.

The high- and low-imageability items were matched on word length and on log word frequency in Danish and English. The length of the words was measured in number of phonemes and in

spoken durations. For the English words, the phoneme counts were derived using the N-Watch program (Davis, 2005), and for the Danish words, the corresponding counts were derived using a Danish pronunciation dictionary (Brink, Lund, Heger, & Jørgensen, 1991). The log frequencies of the Danish words were derived from the frequency counts of Bergenholz (1992) and the corresponding frequencies of the English words were derived from the N-Watch program (Davis, 2005), which uses the counts of the CELEX database. Similarly, the items in the Danish and English set were matched on word length and log word frequency. Mean values and standard deviations of the properties of the experimental stimuli are shown in Table 1. A list of the experimental stimuli is shown in Appendix A. An additional eight words were selected to be used on warm-up trials, with equal numbers of Danish and English and high- and low-imageability items. The stimuli were digitally recorded by a balanced early Danish-English bilingual.

Table 1

A pilot study, in which three participants heard each word that had been selected to be used as cue words and were asked to describe the first memory that came to mind, showed that all words were successful in eliciting autobiographical memories. However, two Danish words were semantically ambiguous (“vælge”, which activated both the “elect” and the “choose” meaning, and “himmel” which activated both the “sky” and the “heaven” meaning), and these words were replaced.

Design. The design included two crossed independent variables each with two levels: cue-word language (Danish or English) and cue-word imageability (high or low). The presentation of the cue words was blocked by language, such that the participants heard cue words in only one language during a given block of trials. The high- and low-imageability words were intermixed within each

language block. The order of presenting the test blocks was counterbalanced across participants.

The order of presenting the words within a block was random and different for each participant.

Two sets of words were constructed, with an equal number of Danish and English words in each set and with translation-equivalent words appearing in different sets. Each participant was assigned to one of the sets. The two sets were matched on word length, log word frequency, and word imageability and had an equal number of high- and low-imageability words (see Table 1). Thus, in each item set, there were two test blocks, each consisting of two warm-up trials and 16 experimental trials.

Apparatus. The experiment was programmed using Experiment Builder software provided by SR Research. The cue words were played through Sennheiser headphones. The participants' button-press latencies were recorded with a Cedrus seven-button response box.

Procedure. The participants were tested individually and were seated in front of a monitor screen. At the beginning of the session, they completed the LEAP language proficiency questionnaire. They were then given written instructions about the memory recall experiment. They were told that they would hear words, some of which would be in Danish and some in English. They were instructed for each word to think of a memory of a specific event in their personal past that was related to the word, that did not extend over more than one day, and that was at least a week old. These instructions were given to ensure that specific events were retrieved and to avoid the retrieval of very recent events. The participants were carefully instructed that they had to retrieve a memory from their personal past, and they were provided with relevant examples to illustrate this request as well as to illustrate the distinction between specific and non-specific autobiographical events. They were instructed to press a button as soon as they had retrieved a suitable memory. After the button press, they were asked to make a brief description of the memory retrieved, which would allow them to later identify the retrieved memory and answer a more

extensive questionnaire. The instructions included examples of a correct memory (i.e., the word “birthday” makes you think of the day when you turned 18 and of the birthday party your parents gave you) and an incorrect memory (i.e., the word “birthday” makes you think of all the many birthday parties that have been held in your summer cottage at the sea).

Before the beginning of the test blocks, the participants put on the headphones. At the beginning of each experimental trial, a sound beep was played for 1000 ms after which a word was played. As soon as a suitable memory came to mind the participants were told to press a prespecified button on a push-button panel placed in front of them. This was immediately followed by two questions concerning the presence of language elements in the memory (both instructions and questions were taken from Schrauf & Rubin, 2000, and translated into Danish). The first question concerned the linguistic nature of the retrieval process, and the second question concerned the intensity of the language experience. In response to the first question, the participants pressed one of four keys to indicate whether the memory had come to them in Danish, English, in both languages, or in no language. In response to the second question, they pressed one of seven buttons to indicate the intensity of the language experience. The latency to retrieve each memory was recorded from the offset of the spoken cue word to the onset of the participants’ button-press response.² If no button press was registered within 60 seconds after the cue word had been played, the computer terminated the trial automatically, and the participants were asked to begin a new trial when they were ready.

In addition to rating each memory on the two language dimensions, after the retrieval of all the memories, the participants completed a rating task in which they received a booklet that listed additional memory dimensions to be rated. These were vividness (i.e., the visual clarity and visual intensity of the retrieved memory), sense of re-living, imageability, sensory detail, emotional intensity, accessibility, significance in life, and centrality in life. Each memory was rated on these dimensions on a scale ranging from 1 to 7, with a higher value indicating a higher level of the rated

dimension. The participants were also asked to indicate with a “yes” or “no” response whether or not they had retrieved a specific memory (i.e., a memory of a specific event that occurred on a particular day in the past) and to report their age at the time when each of the remembered events occurred. Knowing the age of the memories was important to ensure that any difference between the high- and low-imageability cue words was not due to a difference in the recency of the memories activated by the two types of cues, with high-imageability cues activating less recent memories than low-imageability cues (Rubin, 1980). An English translation of the questionnaire is presented in Appendix B. An experimental session lasted approximately 2 hours.

Results

Data analyses. Trials on which the participants failed to press the response button (12 trials, 1.6% of all experimental trials) or failed to provide one or more of the memory ratings (6 trials, 0.8% of the remaining trials) were excluded from the analyses of their button-press latencies, their responses to the language questions, and their memory ratings. Trials on which the button-press latencies were too long or too short (i.e., above or below 3 *SDs*, respectively, from that participant’s mean latency across all conditions) were also excluded (17 trials, 2.3% of the remaining trials). We trimmed out the extreme latencies because these latencies are unlikely to reflect the cognitive processes under consideration, and hence are unlikely to be critical in testing our hypotheses. Trimming latencies relative to the mean latency for each subject across experimental conditions is appropriate for a dataset like this one where the range in button-press latencies across participants was large (mean latencies for fastest and slowest participant prior to trimming were 7031 and 24096 ms, respectively) compared with the average differences between conditions for individual participants, see Ratcliff (1993).

The data were entered into analyses of variance with cue-word language (Danish vs. English) and cue-word imageability (high vs. low) as fixed within-participants and between-items variables and either participants (F_1) or items (F_2) as random variables. In the analysis by participants, we used the mean latencies of each participant averaged across the items within each condition. In the analysis by items, we used the mean latencies to each cue word across the participants who responded to that word. Both participants and items were entered as random variables because this allowed us to generalize our results beyond both the group of participants and the word materials selected for the present experiment (Clark, 1973). Only the participant means are reported. All results are shown in Table 2.

Table 2

Missing responses. As Table 2 shows, the mean rate of missing responses was lower when the language of the cue word was Danish than when it was English, and this difference approached significance, $F_1(1, 23) = 3.43, p < .08$; $F_2(1, 60) = 3.29, p < .08$. The mean rate was significantly lower when the cue word was high than when it was low in imageability, $F_1(1, 23) = 10.04, p < .01$; $F_2(1, 60) = 12.65, p = .001$. This suggests that high-imageability words were more efficient than low-imageability words in eliciting memories. The effect of cue language was confined to the low-imageability cue words, but the interaction of the two effects only approached significance, $F_1(1, 23) = 3.43, p = .08$; $F_2(1, 60) = 3.29, p = .08$.

Button-press latencies. The mean button-press latencies for the two item sets did not differ significantly from each other (13558 vs. 12035 ms), $F_1(1, 44) = 1.09$; $F_2(1, 56) = 2.73$. The interaction of item set with cue language was only significant in the analysis by items, $F_1(1, 44) = 2.45, p > .05$; $F_2(1, 56) = 7.96, p < .01$, and the interaction of item set with cue imageability was not

significant in any of the analyses, $F_1(1, 44) = 1.27$; $F_2(1, 56) < 1$. Importantly, there was no three-way interaction of item set with cue language and cue imageability either, both $F < 1$. Therefore, the latencies were collapsed across the two item sets.

As expected, the mean button-press latency was significantly shorter, by 2962 ms, when the language of the cue word was Danish than when it was English, $F_1(1, 23) = 24.63$, $p < .001$; $F_2(1, 60) = 10.23$, $p < .01$. The button-press latency was significantly shorter, by 3945 ms, when the cue word was high than when it was low in imageability, $F_1(1, 23) = 29.81$, $p < .001$; $F_2(1, 60) = 18.20$, $p < .01$.

The most important question was whether the effects of cue language and cue imageability interacted. As expected, the effect of cue language was numerically stronger for the low- than high-imageability cue words (4252 vs. 1672 ms). The interaction of the two effects was only significant in the analysis by participants, $F_1(1, 23) = 6.16$, $p < .05$. It was not significant in the analysis by items, $F_2(1, 60) = 1.77$, $p = .19$, which may reflect the relatively low number of items in each condition.

Responses to the language questions. Immediately following the recall of an event from memory, the participants first identified the language in which the event seemed to come to them, whether it came to them in Danish, English, in both languages, or in no language, and then rated each of the recalled memories on the intensity of the language experience. The identification responses were each assigned to one of the four categories (Danish, English, mixed, or non-linguistic).

A comparison of the memories identified as linguistic (Danish, English, or mixed) and the memories identified as non-linguistic showed that the two types of memories were equally frequent both when cued in Danish (25.3 vs. 23.9%) and when cued in English (25.4 vs. 23.8%), $F_1(1, 23) < 1$; $F_2(1, 120) < 1$. For both the Danish and English memories, there was an effect of cue language

on the participants' identification of their memories, with more Danish memories being retrieved when the cue language was Danish than when it was English (56.8 vs. 36.3%), $F_1(1, 23) = 11.03$, $p < .01$; $F_2(1, 60) = 26.02$, $p < .001$, and similarly, more English memories being retrieved when the cue language was English than when it was Danish (29.1 vs. 10.4%), $F_1(1, 23) = 10.65$, $p < .01$; $F_2(1, 60) = 39.03$, $p < .001$. Thus, for both types of memories, there was a congruency effect: memories were easier to retrieve when the language of the cue was the same as the language in which the memory was internally retrieved (congruent memories) than when the cue was in the participants' other language (incongruent memories). For the mixed and non-linguistic memories, there was no effect of cue language. There was no effect of cue imageability and no interaction with cue language. This was true for all four memory categories.

To investigate further the pattern observed for the Danish and English memories, analyses were carried out that only included these two memory categories and that included language congruency as an additional independent variable. As expected, there were more Danish than English memories (46.5 vs. 19.7%), $F_1(1, 23) = 19.63$, $p < .001$; $F_2(1, 120) = 117.68$, $p < .001$, and more congruent than incongruent memories (42.9 vs. 23.3%), $F_1(1, 23) = 11.81$, $p < .01$; $F_2(1, 120) = 61.12$, $p < .001$. However, importantly, there was no interaction between memory category and language congruency, $F_1(1, 23) < 1$; $F_2(1, 120) < 1$. These results suggest that the retrieval advantage for memories that were cued in the same language as the internal language of the memory over memories that were cued in the other language was unaffected by the language of internal retrieval, i.e., whether this was Danish or English.

In addition to investigating the distribution of congruent and incongruent memories across Danish and English memories, we analysed the distribution across the Danish and English cue-word conditions. There was a large difference in the proportion of congruent and incongruent memories retrieved in response to Danish cue words (56.8 vs. 10.4%), but this difference was much reduced,

and went in the opposite direction, in response to English cue words (29.1 vs. 36.3%). This was reflected in a significant interaction between language congruency and cue language, $F_1(1, 23) = 19.63, p < .001$; $F_2(1, 120) = 117.68, p < .001$. Separate analyses of the two cue languages showed that the retrieval advantage for congruent memories was confined to the Danish cue words, $F_1(1, 23) = 37.57, p < .001$; $F_2(1, 60) = 213.99, p < .001$. With the English cue words, the difference between the proportions of congruent and incongruent memories was not significant in the analysis by participants and only approached significance in the analysis by items, $F_1(1, 23) < 1$; $F_2(1, 60) = 3.87, p < .06$. These results show that Danish memories, whether cued in Danish or English, were easier to elicit than English memories. Thus, these results show that in addition to the retrieval advantage for memories internally retrieved in the same language as the cue, there was a retrieval advantage for memories internally retrieved in the native language.

Memories cued in Danish and English were rated as inducing language experiences of equal intensity as were memories retrieved in response to high- and low-imageability cue words (mean rating of 3.2 in all experimental conditions). This contrasts with the advantage of the native language and the advantage of the high-imageability cues found for the retrieval latencies.

Memory ratings. After the memory retrieval task had been completed, the participants were asked to indicate whether or not they had retrieved a specific memory and to report their age at the time when each of the remembered events occurred. They were also asked to rate each memory for vividness, sense of re-living, imageability, sensory detail, emotional intensity, accessibility, significance in life, and centrality in life.

The age of each retrieved memory was computed by subtracting the participant's age at the time of the remembered event from the age at the time of testing. There was an effect of cue language on the age of the retrieved memories, with Danish cues eliciting memories from an earlier time in life than English cues (mean reported age of 4.0 and 3.0 years, respectively), $F_1(1, 23) = 1.82, p < .05$;

$F_2(1, 60) = 11.10, p = .001$. This was expected given that the participants acquired Danish earlier in life than English, and given that such an effect has been observed in previous studies (e.g., Marian & Neisser, 2000). By contrast, high-imageability cues did not elicit earlier memories than low-imageability cues (mean reported age of 3.6 and 3.4 years), $F_1(1, 23) < 1$; $F_2(1, 60) < 1$.

Memories cued in Danish and English were almost equally specific (87.2 and 89.1% of memories scored as specific, respectively), $F_1(1, 23) = 1.14$; $F_2(1, 60) < 1$. Memories retrieved in response to high-imageability cue words were slightly more specific than memories retrieved to low-imageability cue words (90.3 and 86.1%), but this difference failed to reach significance in both analyses, $F_1(1, 23) = 3.68, p < .07$; $F_2(1, 60) = 3.67, p < .07$. However, an examination of the specificity scores for each language separately showed that the difference in the scores obtained for memories cued with high- and low-imageability words was modulated by cue language: Memories cued in Danish were more specific when the cue word was high in imageability than when it was low in imageability (91.6 and 82.8%). By contrast, the specificity of the memories cued in English was independent of cue imageability, with the memories elicited by high-imageability cue words being no more specific than those elicited by low-imageability cue words (88.9 and 89.4%). This was reflected in a significant interaction between cue language and cue imageability, $F_1(1, 23) = 6.95, p < .05$; $F_2(1, 60) = 4.74, p < .05$. Separate analyses of the two cue languages showed that the cue imageability effect was confined to the Danish cue words, $F_1(1, 23) = 6.74, p < .05$; $F_2(1, 30) = 10.98, p < .01$. With the English cue words, the effect was absent, $F_1(1, 23) < 1$; $F_2(1, 30) < 1$.

We predicted that the advantage of the native language and of the high imageability cues found for the button-press latencies would also appear on the participants' ratings of their memories for phenomenological qualities. An effect of cue language appeared on the accessibility ratings, but the direction of the effect was unexpected: Memories cued in Danish were rated as being *less* accessible than memories cued in English (4.2 vs. 4.9), $F_1(1, 23) = 8.15, p < .01$; $F_2(1, 60) = 37.17$,

$p < .001$. This effect could be due to a difference in the recency of the events activated by the Danish and English cue words: As mentioned earlier, the events activated by the Danish cue words were less recent than the events activated by the English cue words (mean age of 4.0 and 3.0 years, respectively). Alternatively, it may reflect a greater distinctiveness of the memories cued by English words, in part due to the fact that a greater proportion of them had English as the internal language. Cue language and cue imageability had no effects on any of the other memory ratings.

Discussion

The main goal of the present study was to investigate how bilingual language use influences the ease of retrieving autobiographical memories. The role of language has been investigated in numerous previous studies of autobiographical memory, with some studies varying the types of words used to cue memories and others varying the language of the word cues. Both variables have been found to affect access to autobiographical memories when tested separately. In the present study, we combined these two variables to examine whether cue type modulates the effect of cue language. We predicted an effect of cue language when words that are highly similar in meaning across languages (high-imageability words) were used as cues, but predicted an enhancement of the effect when words that are less similar in meaning across languages (low-imageability words) were used. We found support for this assumption by showing that the effect of cue language was stronger when low-imageability words were used as cues as compared to when high-imageability words were used.

Using a cued-recall task, we found that autobiographical memories were easier to retrieve when the cue words were presented in the native (L1) than in the second language (L2) and when they were high than low in imageability. Autobiographical memories were particularly difficult to

retrieve in response to English low-imageability cue words. However, despite a large numerical increase in the size of the cue language effect for low-imageability compared to high-imageability cue words, the interaction of the two effects was only significant in the participant analysis.

The participants' responses to the questions concerning the linguistic characteristics of the retrieved memories yielded two main findings. First, memories were easier to retrieve when the language of the cue was the same as the language in which the memory was internally retrieved than when the cue was in the participants' other language. Second, memories that were internally retrieved in L1 Danish were easier to elicit than memories internally retrieved in L2 English, regardless of the language of the cue words. The participants' ratings of the specificity of their memories showed that more specific memories were retrieved in response to high- than low-imageability cues, but only when the cue language was Danish. Their ratings of the accessibility of their memories showed that memories cued in English were rated as being more accessible than memories cued in Danish. In the remainder of the discussion, we will first discuss the results obtained for the retrieval latencies and the internal language questions, and then turn to the results of the memory specificity and accessibility measures.

Cue language and cue imageability affected the ease of autobiographical memory retrieval. These results replicate findings by Marian and Neisser (2000) and by Schrauf (2009), who also presented retrieval cues in first and second languages, and by Williams et al. (1999), who also used cues that varied in imageability. Cue language and cue imageability have been shown to affect both processes involved in lexical access and processes involved in memory retrieval. We argue that the effects observed in the present experiment arose during memory retrieval. Specifically, the delay in responding to English cues and to low-imageability cues (relative to Danish and high-imageability cues, respectively) reflects the additional time needed to retrieve from memory an event encoded in a language different from the language of the cue, and the additional time needed to retrieve a

memory representation of a specific event in response to a retrieval cue that is associated with few semantic (or visual) features. However, it could be argued that, at least part of, the observed cue language effect arose during word translation and that it reflects the time needed to translate an L2 English cue word into L1 Danish. Bilinguals may or may not engage in a translation process depending, for instance, on their level of L2 proficiency (e.g., Kroll & Stewart, 1994). However, the reported backward-translation (from L2 to L1) latencies are fairly short (around 1 second, e.g., Kroll & Stewart, 1994; Tokowicz & Kroll, 2007; van Hell & de Groot, 1998b) compared with the delay in memory retrieval in response to English relative to Danish cues (almost 3 seconds) observed in our experiment. Thus, participants may have engaged in a translation process when presented with the English cues and this may have contributed to the cue language effect, but it is unlikely to be solely responsible for the effect. Similarly, imageability effects on word translation latencies have been observed in several studies (e.g., van Hell & de Groot, 1998b, Tokowicz & Kroll, 2007), but the reported effects are small (less than 200 ms) compared with the effect on memory retrieval latencies (almost 4 seconds) observed here.

If the cue language effect arises during memory retrieval, as argued above, the question is whether this effect is consistent with an effect of language congruency. Relevant here is the assumption that our participants, whose dominant language was Danish, were likely to have encoded most of the retrieved memories in Danish. On this assumption, the shorter retrieval latencies in response to Danish than English cues suggest that recall was facilitated when the language of the cue was the same as the language used during encoding (both Danish) compared with when the two languages were different (one Danish and the other English). However, this effect of cue language on the retrieval latencies cannot be taken as evidence for an effect of language congruency per se. This is because the language of encoding was not experimentally manipulated, but instead was assumed to be the participants' native language. Therefore, we were

only able to test for an effect of language congruency when the native language was used and were not able to determine whether an effect of language congruency on memory retrieval latencies extends beyond the native language. Thus, the latency results alone do not provide evidence for a language congruency effect.

However, strong evidence for a language congruency effect is provided by the participants' responses to the questions concerning the linguistic characteristics of the retrieved memories. The language of the cue words affected the participants' identification of their memories, with more memories being identified as (i.e., internally retrieved in) Danish when the cue language was Danish than when it was English, and similarly more memories being identified as English when the cue language was English than when it was Danish. This result suggests that a congruency between the language of internal retrieval and the language of the recall cue facilitated recall. On the assumption that the language of internal retrieval is the language in which the memory was originally encoded (cf. Schrauf & Rubin, 1998), this results is consistent with the view that congruency between the language of encoding and the language of recall facilitates recall of autobiographical memories. The additional finding that memories that were internally retrieved in L1 Danish were easier to elicit than memories internally retrieved in L2 English, regardless of the cue language, shows that this effect of language congruency was moderated by the internal language of retrieval, with a retrieval advantage for memories that were internally retrieved in the native language.

These results suggest that the strength of language congruency effects in autobiographical memory is likely to vary depending on the language used during encoding and during recall, in particular on whether the bilingual's native or non-native language is used. These results replicate those of Schrauf and Rubin (2000) who found a larger congruency effect for the native than non-native language. We found a reversal of the congruency effect for the non-native language (29 and

36% of congruent-English and incongruent-Danish memories, respectively, in response to English cues), whereas they found a congruency advantage (36 and 30% of congruent-English and incongruent-Spanish memories in response to English cues). This difference probably reflects that the bilinguals tested in the two studies were quite different. The Danish-English bilinguals in the present study lived in Denmark and mostly used Danish in everyday life. By contrast, the Spanish-English bilinguals in the Schrauf and Rubin study had lived in the US for at least 30 years at the time of testing and reported speaking English as frequently as Spanish. This can explain the greater retrieval advantage for the native language in the present than in the earlier study. Similar influences of the native language were found by Marian and Neisser (2000) and by Matsumoto and Stanny (2006).

The cue imageability effect found on the retrieval latencies was also observed on the memory specificity measure, but only for the native language: For the Danish cues, more specific memories were retrieved in response to high- than low-imageability cues. This result replicates earlier findings (e.g., Williams et al., 1999, 2006) in showing that high-imageability cues facilitate access to memories of specific events. However, the present results also extend the earlier findings by suggesting that the cue imageability effect does not generalize to bilingual situations in which cues are presented in the non-native language: For the English cues, the low-imageability cues elicited memories that were as specific as those elicited by the high-imageability cues. Given previous findings (Anderson et al., 2012; Williams et al., 2006) of reductions in the specificity of memories that are retrieved under conditions in which low-imageability cues are used and a cognitive load is imposed, and given that cues presented in L2 English should be more difficult to process for L1 speakers of Danish, and thus impose a higher cognitive load, than cues presented in Danish, one might have expected memories to be less specific in response to low-imageability cues when these were presented in English. However, this was not what we found. One explanation of this lack of a

reduction in the specificity of the memories cued in English is that these memories, even those cued with low-imageability words, stood out in some way that made them more distinct than the memories cued in Danish, and that this masked the cue-imageability effect. This distinctness account can also explain the results obtained for the memory accessibility measure which showed that the memories cued in English were rated as being more accessible than the memories cued in Danish.

In the introduction, we outlined two competing views on the organization of bilingual memory. The view favored in theories of autobiographical memory is that linguistically encoded memories of bilinguals are stored exclusively in a language-specific way depending on the language in which they were originally encoded. By contrast, the prevalent view in theories of lexical access is that the semantic representations of first and second language translation equivalents are not exclusively language-specific or exclusively shared. Instead, different types of words overlap in meaning to different degrees, such that, for instance, high imageability words have a larger semantic overlap with their translation equivalents in other languages than do low imageability words that have more language-specific meanings. In this view, semantic representations are thus more shared across languages for high-imageability than for low-imageability translation equivalents. The present finding of a larger cue language effect in response to low- than high-imageability cue words supports this latter view. However, it is important to keep in mind that this interaction between the language and imageability of the retrieval cues was significant only in the by-participant analysis, and not in the item analysis, probably because of the relatively low number of items in each condition. For this reason, and given that such an interaction has not been investigated previously, a replication is needed before any firm conclusions can be drawn concerning the representations of autobiographical memories in bilinguals.

Endnotes

1. Some of the studies reviewed in this section varied word imageability, whereas others varied word concreteness. These two measures are closely related, and for most words they are quite similar in that words that refer to easily imageable concepts typically are concrete, whereas words referring to concepts that are hard to image tend to be abstract. In this section, we will use the same label to refer to this measure (concreteness or imageability) as was used in the study that is reviewed. Because the experimental stimuli used in the present study were selected from corpora of word imageability ratings, the present measure will be referred to as the imageability variable.

2. We measured button-press latencies, and not speech onset latencies, because people vary in their speech fluency with some initiating speech in a very disfluent and hesitant way and others speaking more fluently. Such variation is expected particularly when speakers use a non-native language. As a consequence, the onset of the memory description is likely to vary depending on whether speakers are fluent or disfluent in their speech. By measuring button-press latencies, we avoided any influence of speech fluency on the response latencies (for discussion see Marian & Neisser, 2000).

3. We are not aware of any published study reporting imageability ratings for Danish words.

Acknowledgments

This research was supported by the Danish National Research Foundation (DNRF93). We thank Marie Laursen Canter for help in recording the stimuli, and Niels Peter Nielsen and Stine Breum Ramsgaard for help in conducting the experiment.

References

- Allen, R., & Hulme, C. (2006). Speech and language processing mechanisms in verbal serial recall. *Journal of Memory and Language, 55*, 64–88.
- Anderson, R. J., Dewhurst, S. A., & Nash, R. A. (2012). Shared cognitive processes underlying past and future thinking: The impact of imagery and congruent task demands on event specificity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*, 356–365.
- Altarriba, J. (2006). Cognitive approaches to the study of emotion-laden and emotion words in monolingual and bilingual memory. In A. Pavlenko (ed.) *Bilingual minds: Emotional experience, expression, and representation* (pp. 232-256). Clevedon, UK: Multilingual Matters.
- Altarriba, J., & Bauer, L. (2004). The distinctiveness of emotion concepts: A comparison between emotion, abstract and concrete words. *American Journal of Psychology, 117*, 389-410.
- Bergenholtz, H. (1992). *Dansk frekvensordbog*. Copenhagen, Denmark: Gads Forlag.
- Brink, L., Lund, J., Heger, S. & Jørgensen, J. N. (1991). *Den Store Danske Udtaleordbog*. Copenhagen, Denmark: Munksgaard.
- Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior, 12*, 335–359.
- Conway, M. A., & Bekerian, D. B. (1987). Organization in autobiographical memory. *Memory & Cognition, 15*, 119–132.
- Cortese, M. J., & Fugett, A. (2004). Imageability ratings for 3,000 monosyllabic words. *Behavior Research Methods, Instruments, & Computers, 36*, 384–387.
- Francis, W. S., & Goldmann, L. L. (2011). Repetition priming within and between languages in semantic classification of concrete and abstract words. *Memory, 19*, 653–663.
- Davis, C. J. (2005). N-Watch: A program for deriving neighborhood size and other psycholinguistic

- statistics. *Behavior Research Methods*, 37, 65–70.
- de Groot, A. M. B. (1989). Representational aspects of word imageability and word frequency as assessed through word association. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 824–845.
- de Groot, A. M. B. (1992a). Determinants of word translation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 1001–1018.
- de Groot, A. M. B. (1992b). Bilingual lexical representation: A closer look at the conceptual representations. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 389–412). Amsterdam: Elsevier.
- de Groot, A. M. B., Dannenburg, L., & van Hell, J. G. (1994). Forward and backward word translation by bilinguals. *Journal of Memory and Language*, 33, 600–629.
- de Groot, A. M. B., & Poot, R. (1997). Word translation at three levels of proficiency in a second language: The ubiquitous involvement of conceptual memory. *Language Learning*, 47, 215–264.
- Dewhurst, A., & Conway, M. (1994). Pictures, images, and recollective experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 1088–1098.
- Eade, J., Healy, H., Williams, J. M. G., Chan, S., Crane, C., & Barnhofer, T. (2006). Retrieval of autobiographical memories: The mechanisms and consequences of truncated search. *Cognition & Emotion*, 20, 351–382.
- Francis, W. S., & Goldmann, L. L. (2011). Repetition priming within and between languages in semantic classification of concrete and abstract words. *Memory*, 19, 653–663.
- James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance*, 1, 130–136.
- Jin, Y.-S. (1990). Effects of concreteness on cross-language priming in lexical decision. *Perceptual*

and Motor Skills, 70, 1139–1154.

Kolers, P. A. (1963). Interlingual word associations. *Journal of Verbal Learning and Verbal Behavior*, 2, 291–300.

Kolers, P. A., & Gonzalez, E. (1980). Memory for words, synonyms, and translations. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 53–65.

Kroll, J., & Merves, J. S. (1986). Lexical access for concrete and abstract words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 92–107.

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

Larsen, S. F., Schrauf, R. W., Fromholt, P., & Rubin, D. C. (2002). Inner speech and bilingual autobiographical memory: A Polish-Danish cross-cultural study. *Memory*, 10, 45–54.

Larsen, S. F., & Plunkett, K. (1987). Remembering experienced and reported events. *Applied Cognitive Psychology*, 1, 15–26.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, & Hearing Research*, 50, 940–967.

Marian, V., & Fausey, C. M. (2006). Language-dependent memory in bilingual learning. *Applied Cognitive Psychology*, 20, 1–23.

Marian, V., & Kaushanskaya, M. (2007). Cross-linguistic transfer and borrowing in bilinguals. *Applied Psycholinguistics*, 28, 369–390.

Marian, V., & Neisser, U. (2000). Language-dependent recall of autobiographical memories. *Journal of Experimental Psychology: General*, 129, 361–368.

O'Neill, W. (2005). Word-imagery effects on recollection and familiarity in recognition memory.

Perceptual and Motor Skills, 100, 716–722.

Paivio, A., & Desrochers, A. (1980). A dual-coding approach to bilingual memory. *Canadian Journal of Psychology*, 34, 388–399.

Paivio, A., Clark, J. M., & Lambert, W. E. (1988). Bilingual dual-coding theory and semantic repetition effects on recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14, 163–172.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, 114, 510–532.

Robinson, J. A. (1976). Sampling autobiographical memory. *Cognitive Psychology*, 8, 578–595.

Romani, C., McAlpine, S., & Martin, R. C. (2007). Concreteness effects in different tasks: Implications for models of short-term memory. *Quarterly Journal of Experimental Psychology*, 61, 292–323.

Rubin, D. C. (1980). 51 properties of 125 words: A unit analysis of verbal behavior. *Journal of Verbal Learning and Verbal Behavior*, 19, 736–755.

Rubin, D. C., & Schulkind, M. D. (1997). Properties of word cues for autobiographical memory. *Psychological Reports*, 81, 47–50.

Sahlin, B. H., Harding, M. G., & Seamon, J. G. (2005). When do false memories cross language boundaries in English–Spanish bilinguals? *Memory & Cognition*, 33, 1414–1421.

Schock, J., Cortese, M. J., & Khanna, M. M. (2012). Imageability estimates for 3,000 disyllabic words. *Behavior Research Methods*, 44, 374–379.

Schrauf, R. W. (2000). Bilingual autobiographical memory: Experimental studies and clinical cases. *Culture & Psychology*, 6, 387–417.

Schrauf, R. W. (2009). The bilingual lexicon and bilingual autobiographical memory: The

neurocognitive basic systems view. In A. Pavlenko (ed.) *The bilingual mental lexicon:*

Interdisciplinary approaches (pp. 26-51). Buffalo, NY: Multilingual Matters.

Schrauf, R., Pavlenko, A., & Dewaele, J.-M. (2003). Bilingual episodic memory: An introduction.

International Journal of Bilingualism, 7, 221–233.

Schrauf, R. W., & Rubin, D. C. (1998). Bilingual autobiographical memory in older adult

immigrants: A test of cognitive explanations of the reminiscence bump and the linguistic

encoding of memories. *Journal of Memory and Language*, 39, 437–457.

Schrauf, R. W., & Rubin, D. C. (2000). Internal languages of retrieval: The bilingual encoding of

memories for the personal past. *Memory & Cognition*, 28, 616–623.

Schrauf, R. W., & Rubin, D. C. (2003). On the bilingual's two sets of memories. In R. Fivush & C.

Haden (eds.), *Autobiographical memory and the construction of a narrative self:*

Developmental and cultural perspectives (pp. 121–145). Mahwah, NJ: Erlbaum.

Schwanenflugel, P. J., Harnishfeger, K. K., & Stowe, R. W. (1988). Context availability and lexical

decisions for abstract and concrete words. *Journal of Memory and Language*, 27, 499–520.

Schwanenflugel, P. J., & Shoben, E. J. (1983). Differential context effects in the comprehension

of abstract and concrete verbal materials. *Journal of Experimental Psychology: Learning,*

Memory, and Cognition, 9, 82–102.

Schwanenflugel, P. J., & Stowe, R. W. (1989). Context availability and the processing of abstract

and concrete words in sentences. *Reading Research Quarterly*, 24, 114–126.

Snodgrass, J., & Vanderwaart, M. (1980). A standardized set of 260 pictures: Norms for name

agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental*

Psychology: Human Learning and Memory, 6, 174–215.

Stadthagen-Gonzalez, H., & Davis, C. J. (2006). The Bristol norms for age of acquisition,

imageability, and familiarity. *Behavior Research Methods*, 38, 598–605.

- Taylor, I. (1976). Similarity between French and English words – A factor to be considered in bilingual language behavior? *Journal of Psycholinguistic Research*, 5, 85–94.
- Tokowicz, N. & Kroll, J. F. (2007). Number of meanings and concreteness: Consequences of ambiguity within and across languages. *Language and Cognitive Processes*, 22, 727–779.
- van Hell, J. G., & de Groot, A. M. B. (1998a). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition*, 1, 193–211.
- van Hell, J. G., & de Groot, A. M. B. (1998b). Disentangling context availability and concreteness in lexical decision and word translation. *The Quarterly Journal of Experimental Psychology*, 51A, 41–63.
- Walker, I., & Hulme, C. (1999). Concrete words are easier to recall than abstract words: Evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 1256–1271.
- Whaley, C. P. (1978). Word-nonword classification times. *Journal of Verbal Learning and Verbal Behavior*, 17, 143–154.
- Williams, J. M. G., Healy, H. G., & Ellis, N. C. (1999). The effects of imageability and predictability of cues on autobiographical memory. *The Quarterly Journal of Experimental Psychology* 52A, 355–579.
- Uzer, T., Lee, P. J., & Brown, N. R. (2012). On the prevalence of directly retrieved autobiographical memories. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 38, 1296–1308.

Appendix A

Materials used in the experiment

Set 1

High imageability stimuli

English Danish

bike cykel

dinner middag

gift gave

letter brev

ticket billet

forest skov

message besked

bottle flaske

Low imageability stimuli

English Danish

duty pligt

mood humør

noise støj

proud stolt

guilt skyld

honest ærlig

boredom kedsomhed

leisure fritid

Set 2

High imageability stimuli

English Danish

dog hund

wedding bryllup

autumn efterår

cottage hytte

goodbye farvel

colour farve

breakfast morgenmad

angry vred

Low imageability stimuli

English Danish

eager ivrig

visit besøg

loss tab

threat trussel

theft tyveri

luck held

knowledge viden

fatigue træthed

Appendix B

Phenomenological characteristics of the retrieved memories

Instruction: For each of the events that you have just remembered, I will ask you to answer the questions listed below. To remind you of each remembered event, the word you heard and that made you think of the event is printed at the top of each response sheet.

1) How old were you when the event took place? _____ years

1a) If you have stated you current age in response to the above question, how old is the event then when counting the number of days since it took place? Approximately _____ days

2) Is the event about a specific event that occurred on a particular day in the past?

Yes _____ No _____

3) My memory for this event is vivid (that is, visually clear and detailed)

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

4) When I remember the event, I feel that I experience it again

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

5) When I remember the event, I see in my mind what happened

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

6) When I remember the event, I hear in my mind, feel, smell, or taste what happened

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

7) My emotions concerning the event are intense

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

8) When I read the cue word, this event immediately came to mind

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

9) The remembered event is important to my life

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

10) I feel that this event has become a central part of my life

Strongly disagree 1 2 3 4 5 6 7 Strongly agree

Table 1

Mean values (and standard deviations) of lexical properties of the experimental stimuli.

	Set 1		Set 2	
	Imageability		Imageability	
	High	Low	High	Low
Imageability ^a		613.3 (56.2)	283.4 (48.7)	582.3 (103.4) 287.4 (49.7)
Danish log frequency		1.5 (0.4)	1.5 (0.4)	1.5 (0.4) 1.4 (0.4)
English log frequency		1.7 (0.4)	1.6 (0.2)	1.6 (0.2) 1.6 (0.5)
Danish phoneme length		4.6 (0.9)	5.3 (1.7)	4.6 (1.4) 4.6 (1.4)
English phoneme length		4.6 (0.9)	4.3 (0.9)	5.0 (1.4) 4.1 (0.8)
Danish spoken duration (ms)		393.8 (90.9)	436.3 (89.9)	427.5 (132.8) 414.6 (84.7)
English spoken duration (ms)		428.5 (119.3)	496.0 (84.8)	493.9 (63.6) 538.8 (84.6)

^aThe imageability ratings were obtained using published norms (Cortese & Fugett, 2004; Schock, Cortese, & Khanna, 2011; Stadthagen-Gonzalez & Davis, 2006). These norm sets were collected using a seven-point rating scale ranging from 1 (words arousing images with great difficulty or not at all) to 7 (words arousing images most readily). These norms were then scaled to produce a range from 100 to 700.

Table 2

Mean missing button presses (%), button-press latencies (in ms), memory categorization according to internal language (%), and memory qualities. Standard deviations are given in parentheses.

	Cue type			
	Danish		English	
	High imageability	Low imageability	High imageability	Low imageability
<i>Cued recall</i>				
Missing responses	0.0 (0.0)	1.58 (5.65)	0.0 (0.0)	4.83 (7.36)
Button-press latencies	9988 (5132)	12643 (5342)	11660 (6177)	16895 (5616)
<i>Internal language</i>				
Danish memories	57.75 (28.08)	55.75 (30.68)	35.83 (29.27)	36.79 (21.47)
English memories	8.00 (13.39)	12.75 (19.14)	30.92 (30.12)	27.25 (24.04)
Mixed memories	8.50 (13.80)	10.42 (16.01)	10.54 (14.25)	12.17 (13.25)
Non-linguistic memories	26.25 (26.82)	21.63 (24.17)	23.25 (27.44)	24.29 (21.67)
<i>Intensity of language experience</i>	3.05 (0.91)	3.32 (1.08)	3.27 (1.14)	3.16 (1.02)
<i>Memory qualities</i>				
Age of memory ^a	4.35 (2.96)	3.65 (2.74)	2.87 (2.16)	3.11 (1.20)
Specificity ^b	91.58 (16.26)	82.87 (25.24)	88.92 (17.32)	89.38 (18.53)
Vividness	5.10 (0.68)	5.14 (0.65)	5.36 (0.75)	5.16 (0.81)
Re-living	4.64 (0.66)	4.75 (0.87)	4.84 (1.20)	4.57 (1.00)

Imageability	5.23 (0.78)	5.38 (0.70)	5.59 (0.61)	5.43 (0.70)
Sensory detail	4.26 (0.81)	4.36 (0.89)	4.22 (1.66)	4.14 (1.54)
Emotional intensity	4.42 (0.89)	4.66 (1.01)	4.76 (0.78)	4.77 (0.89)
Accessibility	4.25 (0.79)	4.06 (0.72)	4.89 (1.00)	4.90 (1.00)
Significance	3.54 (0.94)	3.47 (1.07)	3.48 (0.89)	3.54 (1.17)
Centrality	2.91 (1.00)	2.96 (1.23)	2.96 (0.94)	3.06 (1.04)

^a The difference between current age and age at time of event.

^b The proportion of “yes” responses out of all responses in that condition.