Second Language Research

http://slr.sagepub.com

# First language polysemy affects second language meaning interpretation: evidence for activation of first language concepts during second language reading

Kerrie E. Elston-Güttler and John N. Williams

*Second Language Research* 2008; 24; 167
DOI: 10.1177/0267658307086300

The online version of this article can be found at:
http://slr.sagepub.com/cgi/content/abstract/24/2/167

Published by:
SAGE Publications
http://www.sagepublications.com

**Additional services and information for *Second Language Research* can be found at:**

**Email Alerts:** http://slr.sagepub.com/cgi/alerts

**Subscriptions:** http://slr.sagepub.com/subscriptions

**Reprints:** http://www.sagepub.com/journalsReprints.nav

**Permissions:** http://www.sagepub.com/journalsPermissions.nav

**Citations** (this article cites 19 articles hosted on the
SAGE Journals Online and HighWire Press platforms):
http://slr.sagepub.com/cgi/content/refs/24/2/167

Downloaded from http://slr.sagepub.com by JOHN WILLIAMS on March 19, 2008
**© 2008 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.**

*Second Language Research* 24,2 (2008); pp. 167–187

# First language polysemy affects second language meaning interpretation: evidence for activation of first language concepts during second language reading

**Kerrie E. Elston-Güttler** *Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany* and **John N. Williams** *University of Cambridge*

Received August 2005; revised March 2007; accepted March 2007

The present study investigates the influence of first language (L1) lexicalization patterns on the processing of second language (L2) words in sentential contexts by advanced German learners of English. The focus was on cases where a polysemous word in the L1 is realized by independent words in the L2, e.g. German *Blase* realized by English *bubble* and *blister*. An anomaly detection task was used in which participants had to indicate whether a target word formed an acceptable completion to a sentence. The critical condition was where the other sense (*blister*) of the translation equivalent *Blase* was appropriate, but the word (*bubble*) did not complete the sentence meaningfully, e.g. 'His shoes were uncomfortable due to a bubble.' This was compared to a control condition in which neither sense of the L1 translation made sense, e.g. 'She was very hungry because of a bubble.' Factors of word type (noun vs. verb) and degree of relatedness of L1 senses (high vs. moderate) were also manipulated. Relative to native speakers of English, advanced German learners made more errors and displayed longer correct response times in the critical condition compared to the control condition. An effect of meaning relatedness was obtained for nouns but not verbs. The results are discussed in terms of the role of lexical-level translation connections in activating L1 concepts from L2 words, even in highly proficient learners and in all-L2 tasks.

Address for correspondence: Kerrie E. Elston-Güttler, Max Planck Institute for Human Cognitive and Brain Sciences, Neurocognition of Language, PO Box 500 355, Stephanstraße 1a, D-04103 Leipzig, Germany; email: guettler@cbs.mpg.de

© 2008 SAGE 10.1177/0267658307086300

Downloaded from http://slr.sagepub.com by JOHN WILLIAMS on March 19, 2008
© 2008 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

**Keywords:** second language processing, polysemy, lexicalization, conceptual processing, semantic processing, on-line reading

## I Introduction

Lexicalization patterns – the mapping between words and concepts – differ across languages. For example, German uses one word *Tasche* to refer to both bags and pockets, whereas English uses different lexical items to distinguish these concepts. Of course, speakers of German appreciate the conceptual distinction between a bag and a pocket; it is just that they can use the same word *Tasche* to refer to both concepts. Words like *Tasche* are referred to as polysemous words because, unlike homonyms, native speakers rate their various senses as highly or moderately related.

Cross-linguistic differences in lexicalization patterns point to language-specificity in the mapping between words and concepts. In a so-called two-level semantics, non-linguistic concepts are distinct from sets of language-specific features, or semantic forms, which control the mapping from words to non-linguistic concepts (Bierwisch and Schreuder, 1992; see also Pavlenko, 1999). Alternatively, in a one-level semantics, words can be conceived as linking directly to atomic lexical concepts (Levelt *et al.*, 1999; Roelofs, 2000). Non-linguistic concepts are mapped onto language-specific lexical concepts at a stage of conceptual preparation in a process that as been called 'thinking for speaking' (Slobin, 1987; 1991; 1996). Because in the present study it is not possible to distinguish the effects of lexical-semantic and non-linguistic conceptual information, we shall simply refer to the 'meaning interpretation' activated by a word.

Given the language-specificity of lexicalization patterns, the question arises whether differing patterns in the bilingual/multilingual's languages are kept distinct or interfere with each other in processing. Arguments can be made for either hypothesis. In favour of interference, one could appeal to a model of bilingual lexical memory such as the Revised Hierarchical Model (RHM) of Kroll and Stewart (1994). According to this model, second language (L2) words are produced and understood predominantly via lexical-level translation connections to first language (L1) words. Direct connections from L2 words to conceptual representations strengthen only

Downloaded from http://slr.sagepub.com by JOHN WILLIAMS on March 19, 2008
© 2008 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

gradually as proficiency increases (Chen and Leung, 1989; Kroll and Tokowicz, 2001; Hernandez *et al.*, 2005). This would mean that when a German learner of English reads the word *bag*, a lexical-level link would activate the German translation equivalent *Tasche*. If *Tasche* in turn activates its L1 meaning, then inappropriate conceptual features associated with the concept 'pocket' would become activated. This would make it hard for the learner to distinguish between the L1 and L2 lexicalization patterns, leading to interference in L2 semantic processing tasks.

But one could equally make arguments in favour of L1 and L2 independence. It has been argued that even at low levels of proficiency, L2 learners have acquired direct connections between L2 forms and concepts. For example, there is considerable evidence for efficient conceptual access from L2 words (Altarriba and Mathis, 1997; Frenck-Mestre and Prince, 1997). Kroll and Tokowicz (2001) concede that this argues against the lexical-mediation hypothesis, at least in comprehension tasks. Besides, as soon as one considers cases of lexicalization differences across languages it is clear that, for example, if the German learner of English is going to comprehend *bag* and *pocket* correctly at all, then they have to develop direct mappings from these forms to the conceptual level and to suppress the tendency to rely on the L1 translation equivalent.

Few studies have examined the effects of differing lexicalization patterns on L2 semantic processing. Jiang (2002) examined L1 polysemy in the context of Chinese learners of English, for whom, for example, the English words *problem* and *question* share a common L1 translation equivalent, *wenti*. He used meaning relatedness tasks in which the participants had to indicate whether they thought pairs of words were related in meaning. Relative to related word pairs with different translations, pairs with common translations such as *problem* and *question* had higher ratings of relatedness and decreased relatedness decision times. Jiang (2000; 2002) argues that after an initial stage of relying on lexical-level translation connections, learners do actually access conceptual information directly from L2 words. However, they do so using a copy of the semantic specifications of the L1 translation equivalent. In fact, he is rather pessimistic about the prospects of ever acquiring native-like form–meaning mappings. In part this is because the semantic differences may be slight or non-existent for the majority of the learner's vocabulary.

Downloaded from http://slr.sagepub.com by JOHN WILLIAMS on March 19, 2008
© 2008 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

A different conclusion is suggested by Elston-Güttler *et al*. (2005), who examined the effects of translations of L1 homonyms in German learners of English (e.g. German *Kiefer* corresponds to English *pine* and *jaw*) using event-related brain potentials (ERPs). They showed that for low proficiency learners the prime word *pine* interfered with lexical decisions on the target word *jaw*, even when the prime occurred in a biasing sentence context. They argued that this must have been due to activation of the L1 translation equivalent via translational form-level links. However, high proficiency learners did not show this effect. In support of the effect of form-level links in the low proficiency group, the ERP recording showed a modulation of the N200 component as the target word was processed. Since the N200 is associated with lexical-level processing, it can be argued that the interference effects in the low proficiency group originated at the lexical level, consistent with the RHM. The lack of such effects in the higher proficiency group suggests the acquisition of independent form–meaning mappings. These results are consistent with the idea that direct connections from L2 words to conceptual representations strengthen only gradually as proficiency increases (Chen and Leung, 1989; Kroll and Tokowicz, 2001; Hernandez *et al*., 2005).

It could be argued that translations of L1 homonyms are a special case because the lack of any relationship between the meanings of the L1 homonym encourages the learner to develop independent form–meaning mappings in the L2. In the present experiments, we examine polysemous words in an attempt to see whether – even at high levels of proficiency – there is interference from L1 translations or whether independent form–meaning mappings have been acquired. We also attempt to evaluate whether interference effects are purely a result of lexical-level translation connections, or whether they reflect the activation of L1 concepts.

We examined processing of translations (*bag*, *pocket*) of L1 polysemous words such as *Tasche* in a semantic anomaly judgement task. (It was not possible to use a priming task – such as that in Elston-Güttler *et al*., 2005 – because of the high meaning-relatedness of the two senses of the words.) Using sentences such as 'On his shoulder, he carried a large leather *pocket*' and 'On his trousers was a large *bag*', both groups of learner and native participants were asked to judge whether the final

Downloaded from http://slr.sagepub.com by JOHN WILLIAMS on March 19, 2008
© 2008 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

word could complete the sentence (YES button) or whether the final word was incongruous (NO button). For the trials of interest, the correct answer was always NO, and the question was whether, compared to natives, the learners would show slower response times (RTs) or increased errors due to the felicity of the L1 translation equivalent word (*Tasche*) in that context.

There are essentially three processing routes that could lead to interference:

- Route 1: when a German learner of English reads the word *bag*, a lexical-level link would activate the German translation equivalent *Tasche*, which would in turn directly activate the word *pocket*. If this word *pocket* then activates the associated concept, a learner would experience interference in recognizing the anomaly in 'On his trousers was a large *bag*'. According to this model, the learner has acquired the correct form–meaning mappings for *bag* and *pocket*. The source of the interference is a translation connection at the lexical level.
- Route 2: *Bag* activates the translation equivalent *Tasche* via a lexical-level translation connection as in Route 1, but now *Tasche* activates the associated L1 'pouch, container' concept from German, which causes interference.
- Route 3: *Bag* activates the German 'pouch, container' concept directly. This would occur if, as argued by Jiang (2000; 2002), L2 words merely inherit the semantic specifications of their L1 translation equivalents.

The essential difference between Route 1 on the one hand and Routes 2 and 3 on the other is that only in Routes 2 and 3 is interference a result of activation of an L1 concept. In the present experiment we attempt to distinguish these two possibilities by manipulating the level of relatedness of the two English polysemous words. The polysemous L1 words were divided into highly related (*bag*, *pocket* for *Tasche*) and moderately related (*snake*, *queue* for *Schlange*). According to Route 1, an alternative L2 concept should become equally active in both cases because that concept becomes active through lexical-level translation connections. Therefore, whilst there should be an overall effect of whether the L1 translation is polysemous, the degree of relatedness of the word senses should not affect performance (or at least not more than

Downloaded from http://slr.sagepub.com by JOHN WILLIAMS on March 19, 2008
© 2008 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

for the native control group). On the other hand, if L1 concepts are involved (Routes 2 and 3) then meaning relatedness should have an effect. For example, the general L1 'Tasche' concept contains a large number of features that are common to 'bag' and 'pocket'. When *bag* occurs in a context that is more relevant to *pocket,* the large number of relevant features activated in the 'Tasche' concept could produce interference. In contrast, a general concept like 'Schlange' contains fewer shared features for 'snake' and 'queue' and so it produces less interference. We will not be able to tell whether the L1 concept becomes active directly from the L2 word (Route 3), or via the L2 translation equivalent word (Route 2); but, by manipulating meaning relatedness, we will at least be able to tell whether L1 conceptual structures become active at all.

We also included both nouns and verbs in the experiment, as verbs have been argued to be more polysemous by nature than nouns and therefore more problematic for learners (see Miller and Fellbaum, 1992; Källkvist, 1997). It has also been suggested that nouns and verbs may differ with respect to their reliance on lexical-level translation connections in processing (Van Hell and De Groot, 1998). By conducting the same norming procedure and obtaining comparable relatedness ratings on polysemous nouns and verbs, we could test whether nouns and verbs were still processed differently by learners.

## II Experiment

The experiment contained the main between-participant factor of Group (learners vs. natives) and the within-participant factors of Relatedness (*test sentences* anomalous to natives but possibly acceptable to learners due to L1 lexicalization vs. *control sentences* anomalous for both natives and learners), High/Moderate (highly related vs. moderately related polysemous meanings) and Noun/Verb (nouns vs. verbs). See Table 1 for examples of stimuli used.

As discussed above, all of the three processing accounts predict L1 interference for the learners, and if Routes 2 or 3 apply, then we predict an effect of degree of semantic relatedness. The English group is likely to show some minimal interference for highly related polysemous words only if more shared features cause stronger interference effects. However, this

Downloaded from http://slr.sagepub.com by JOHN WILLIAMS on March 19, 2008
© 2008 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

**Table 1** Examples of critical items

| German word | Related sentences | Control sentences |
|---|---|---|
| *Highly related*: | | |
| Noun: *Blase* | His shoes were uncomfortable due to a *bubble*. | She was very hungry because of a *bubble*. |
| | With the chewing gum the child made a *blister*. | With a desk chair the secretary made a *blister*. |
| Verb: *tragen* | The heavy suitcase was difficult to *wear*. | The hard work was difficult to *wear*. |
| | In winter a heavy warm coat is practical to *carry*. | The new car was comfortable to *carry*. |
| *Moderately related*: | | |
| Noun: *Schlange* | A frightening thing to see in the forest is a *queue*. | A nice thing to see hanging on the wall is a *queue*. |
| | She had to wait several minutes due to the *snake*. | She worked very hard due to the *snake*. |
| Verb: *reiben* | Mary's back hurt so the massage expert began to *grate*. | Mary's hair was too long so the woman at the salon began to *grate*. |
| | The hard cheese was easy to *rub*. | The new dining room table was easy to *rub*. |

would not change our predictions substantially, as the added influence of the L1 lexicalization pattern on processing should still cause significant differences between learners and natives.

## *1 Methods*

*a Participants*: The two participant groups consisted of 32 German learners (18 female, mean age 24) and 20 English natives (12 female, mean age 23) with minimal or no knowledge of German. Highly and moderately related conditions were presented in two separate experimental blocks, both involving the exact same procedure, but with the moderately related stimuli always presented first. All German participants were highly advanced speakers of English, started English at age 10–11, and were either living in England (12 connected with the University of Cambridge) or were studying English at University (20 from the Heinrich-Heine Universität in Düsseldorf) and had at least two months English language experience abroad. All learner participants characterized their learning situation as classroom/real-world exposure, and the mean percentage of time each day they felt they read or thought in English was 16%. Of the learners, 12.5% considered themselves

Downloaded from http://slr.sagepub.com by JOHN WILLIAMS on March 19, 2008
© 2008 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

advanced in French, while 46.8% considered themselves intermediate in French, Spanish or Italian. Native English participants were either students at the University of Cambridge (10) or were University graduates in the Milwaukee/Chicago area (10), none of whom knew any German. Of the natives, 25% considered themselves fluent in either French or Spanish, while 40% had basic knowledge of either French, Spanish or Italian.

*b Stimuli*: The meanings of polysemous words in German were divided into High/Moderate and Noun/Verb. To control for relatedness over nouns and verbs and to define our High/Moderate factor, we administered a norming questionnaire to 71 native German participants (mean age 29) who were professionals in the Augsburg/Munich area. The questionnaires were entirely in German. For the experiment, a list of 184 German polysemous words (taken partly from Götz *et al.*, 1993) was generated and reviewed for semantically related, multiple meanings, and then two sentences were constructed for each meaning.

The 71 participants rated half of the items (92) in one of two randomized questionnaire versions, yielding 35/36 observations per item. For each item, participants read two German sentences containing the polysemous German word as in 'Nach einem langen Marsch hatte er sich eine *Blase* am Fuß gelaufen' ('After the long march he got a *blister* on his foot') and 'Mit dem Kaugummi hat das Kind eine riesige *Blase* gemacht' ('With the chewing gum, the child blew a huge *bubble*'). After reading the sentences, participants had to judge on a scale of 0 to 4 how related the meanings in the respective sentences were: 0 = the same meaning; 1 = very similar; 2 = fairly similar; 3 = fairly different; 4 = very different. The sentences were constructed using examples based on the learner's dictionary entries (Götz *et al.*, 1993) and with the help of a German native speaker.

The mean relatedness judgement and SD for each polysemous word was calculated, then the 10 best items per condition were chosen with the following criteria:

- simple nouns and verbs;
- meanings clearly within the 1 to 2.5 range for highly related and 2.5 to 4 range for moderately related words;

Downloaded from http://slr.sagepub.com by JOHN WILLIAMS on March 19, 2008
© 2008 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

- SDs (informant variation) as low as possible;
- words items with unambiguous English translations;
- items conducive to naturalistic sentence construction.

This yielded:

- 10 items in the 'highly related nouns' condition with a range of 1.11 to 2.36 and a relatedness rating mean of 1.75;
- 10 'highly related verbs' with a range of 1.17 to 2.47 and a mean of 1.88;
- 10 'moderately related nouns' with a range of 2.69 to 3.64 and a mean of 3.16; and
- 10 'moderately related verbs' with a range of 2.56 to 3.60 and a mean of 3.10.

For each of the 40 items chosen, we constructed sentences for the actual experiment. First, we constructed two anomalous test sentences (80 in total) for each word. Sentences were constructed carefully by using an English analogue version of the sentences used in the questionnaire, but with the two English translations of the polysemous German word as sentence-final words. The sentence-final words were then switched across sentences, yielding the two anomalous sentences – e.g. 'With the chewing gum the child made a *blister*' and 'His shoes were uncomfortable due to his *bubble*' – each used on different presentation lists. The rationale for this word switching was to create a context in which the other translation was acceptable in order to test the ability to distinguish between two words that are non-specific in the L2. In a pen-and-paper pre-test, 10 native English speakers rated the sentences as ending with a semantically incongruent word (see description below). See Table 1 for examples of sentences in each condition.

Next, 80 control sentences were constructed by mirroring the syntax of the test sentences and having the same sentence-final words as the test sentences. Anomalous control contexts were designed to be semantically unacceptable and, crucially, the German translation of the sentence-final word had to be unacceptable in the context. See Table 1 for examples. As above, the sentences were given to 10 native speakers of English in a pen-and-paper task. If any respondent indicated that an anomalous word ending the sentence was acceptable, or if a control word was judged as

Downloaded from http://slr.sagepub.com by JOHN WILLIAMS on March 19, 2008
© 2008 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

unacceptable, the sentence was modified. Last, 60 non-anomalous filler sentences were constructed, e.g. 'The large lorry nearly hit the *cyclist*'.

Highly and moderately related items were presented in separate blocks of trials (see Procedure). Within each block, 20 test and 20 control sentences were balanced by including one test sentence per polysemous meaning and a control sentence for the other meaning, i.e. *bubble* used in the test sentence and *blister* used in a control sentence. Another list was then formed with these assignments reversed. With 30 non-anomalous filler sentences included within each block, there was a semantically intact to anomalous ratio of .75 and a total list length of 100 trials. The length of sentence measured by number of words and the letter length and CELEX frequency (Baayen *et al.*, 1993) of sentence-final words were carefully balanced for highly and moderately related critical items, yielding the means listed in Table 2.

*c Procedure*: The Experiment was administered as part of a battery of six experiments (not all reported here; see Elston-Güttler, 2000). The moderately related block was the third in this battery, and the highly related block was the fifth experiment. The experiment intervening the moderately and highly related blocks concerned homonym processing and lexical decision of target words after reading sentences that ended with a prime. This block design was chosen to prevent participants from deriving the intent of the study, which was more likely to become apparent in the highly related trials.

**Table 2** Characteristics of the critical items

| Presentation list/condition | Test sentence word count | Control sentence word count | Sentence-final word frequency* | Sentence-final word letter length |
|---|---|---|---|---|
| *Highly related:* | | | | |
| Nouns | 9.6 | 9.3 | 28.3 | 6.0 |
| Verbs | 9.8 | 9.9 | 72.1 | 4.8 |
| Mean | 9.7 | 9.6 | 50.2 | 5.4 |
| *Moderately related*: | | | | |
| Nouns | 9.8 | 9.8 | 50.8 | 6.8 |
| Verbs | 10.9 | 10.9 | 55.1 | 5.3 |
| Mean | 10.4 | 10.4 | 53.0 | 6.1 |

*Note*: *Frequency per million of meaning used in the sentence using the English lemma frequency dictionary in the CELEX lexical database (Baayen *et al.,* 1993).

Downloaded from http://slr.sagepub.com by JOHN WILLIAMS on March 19, 2008
© 2008 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

Stimuli were presented under computer control with automatic randomization of trials using Tscop (Norris, 1984). All participants were tested individually seated at a laptop computer placed at a comfortable reading distance, with a two-button panel placed before them. The panel was positioned such that the participant's dominant hand made the YES response, while the other hand made the NO response. Directions, with examples, were presented in written form on the screen in English. For each trial, a sentence (all but the last word) was presented in white courier font on a black background centred on the screen. Participants read the sentence, then pressed YES when ready. At that point, the sentence disappeared from the screen and the sentence-final word appeared centred on the screen after a stimulus-onset-asynchrony (SOA) of 250 ms. Once the sentence-final word appeared, the participant responded YES if the word made sense as a completion of the sentence and NO if it did not. The word disappeared when the participant made a response (or automatically after 4000 ms). The inter-trial-interval was 1000 ms. After the session, each participant completed a checklist of all the words used in the experiment and had to indicate whether the word and its meaning were familiar.

### *2 Results*

Some participants had trouble with the task in the first moderate block, so three learners and two natives had to be discarded and replaced. There were 32 learners and 20 natives in the high and moderate blocks, but with slightly different participants in each block. In the statistical analyses, we therefore defined the High/Moderate factor as a between-participants factor. Note that the participants in the two blocks were completely comparable in terms of proficiency and profile. Equal numbers of participants were assigned to the two presentation lists in each block.

Unsuitable data outside the cut-off of 4000 ms and the range of 2.5 SD from the participant mean were discarded from analysis; over all conditions, this accounted for 3.6% of data for learners and 3.4% of data for natives. Removed items due to low familiarity (as established by the checklist) were *rim* and *soar* in the highly related conditions, and *grate* and *starched* in the moderately related conditions. ANOVAs were

Downloaded from http://slr.sagepub.com by JOHN WILLIAMS on March 19, 2008
© 2008 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

performed on the remaining error data and correct NO RTs. Participants and items were considered random variables, yielding F1 and F2 statistics, respectively. In the analyses, Relatedness and Noun/Verb were treated as within-participants and between-items factors. High/Moderate was treated as a between-participants and between-items factor and Group was a between-participants and within-items factor. For transparency, only theoretically relevant main effects and interactions are reported.

### *3 Analysis of errors*

Table 3 shows the effect of Relatedness in each condition under the heading 'difference' (starred if significant), and Figure 1 provides a graphical comparison of the interference effects. The critical interaction between the factors Group and Relatedness was significant; $F1(1,96) = 15.76$, $p < .0001$, $F2(1,72) = 10.39$, $p < .01$. The 11.9% overall interference effect obtained for the learner group was highly significant while the 3.8% effect obtained for natives was smaller, although still significant. Second, the three-way interaction between Group, High/Moderate and Relatedness was significant ($F1(1,96) = 9.80$, $p < .01$, $F2(1,72) = 5.03$, $p < .001$), meaning that the learners were affected more by degree of L1 meaning relatedness (23.5% vs. 2.9%) than natives were (6.3% vs. 1.2%), as assumed in our predictions. Next, this interaction was qualified by Noun/Verb with a significant four-way interaction between Group, High/Moderate, Noun/Verb and Relatedness significant by participants; $F1(1,96) = 5.75$, $p < .025$, $F2(1,72) = 1.50$. To follow up on the four-way interaction, we performed breakdown analyses of the Noun and Verb conditions to explore the critical two-way and three-way interactions we obtained above.

When Noun error rates were analysed, the interaction between Group and Relatedness was significant; $F1(1,96) = 13.99$, $p < .001$, $F2(1,36) = 11.15$, $p < .01$. The 9.8% interference effect obtained for learners was highly significant while the natives showed a non-significant 0.7% effect. The three-way interaction between Group, High/Moderate and Relatedness was also significant ($F1(1,96) = 16.61$, $p < .0001$, $F2(1,36) = 13.32$, $p < .001$), reflecting the fact that for learners, Relatedness effects were larger for highly related than for moderately related nouns (20% vs. –0.3%), while for the natives the effects for nouns were not significant in either condition (1% and 0.5% respectively).

Downloaded from http://slr.sagepub.com by JOHN WILLIAMS on March 19, 2008
© 2008 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

**Table 3** Mean RT in ms and percent errors by condition and participant group

| | Related | | | Control | | | Difference | |
|---|---|---|---|---|---|---|---|---|
| | RT | SD | Error % | RT | SD | Error % | Error[1] | RT[2] |
| *Learners:* | 1290 | 432 | 15.8 | 1150 | 369 | 3.9 | 11.9** | +140** |
| Nouns | 1275 | 409 | 13.9 | 1154 | 362 | 4.1 | 9.8** | +121** |
| Verbs | 1304 | 497 | 17.7 | 1145 | 397 | 3.8 | 13.9* | +159** |
| Highly related | 1364 | 469 | 24.9 | 1204 | 420 | 4.1 | 23.5** | +160** |
| Nouns (*bag–pocket*) | 1414 | 509 | 25.0 | 1208 | 403 | 5.0 | 20.0** | +206** |
| Verbs (*carry–wear*) | 1314 | 476 | 24.7 | 1199 | 457 | 3.1 | 21.6** | +115** |
| Moderately related | 1215 | 395 | 6.7 | 1095 | 318 | 3.8 | 2.9* | +120* |
| Nouns (*queue–snake*) | 1136 | 308 | 2.8 | 1099 | 321 | 3.1 | –0.3 | +37 |
| Verbs (*rub–grate*) | 1294 | 518 | 10.6 | 1091 | 336 | 4.4 | 6.2* | +203** |
| *Natives:* | 923 | 176 | 7.7 | 871 | 148 | 3.9 | 3.8* | +52 |
| Nouns | 918 | 210 | 4.0 | 877 | 159 | 3.3 | 0.7 | +41 |
| Verbs | 929 | 163 | 11.3 | 865 | 151 | 4.5 | 6.8* | +64 |
| Highly related | 956 | 183 | 9.3 | 861 | 119 | 3.0 | 6.3* | +95* |
| Nouns (*bag–pocket*) | 945 | 215 | 4.5 | 888 | 145 | 3.5 | 1.0 | +57 |
| Verbs (*carry–wear*) | 967 | 177 | 14.0 | 833 | 106 | 2.5 | 11.5** | +134** |
| Moderately related | 890 | 168 | 6.0 | 881 | 176 | 4.8 | 1.2 | +9 |
| Nouns (*queue–snake*) | 890 | 205 | 3.5 | 865 | 172 | 3.0 | 0.5 | +25 |
| Verbs (*rub–grate*) | 891 | 149 | 8.5 | 897 | 196 | 6.5 | 2.0 | –6 |

*Notes*: [1]Starred if the main effect of Relatedness (difference between Related and Control) for error rate is significant in a one-way ANOVA by subjects performed for that condition, *p $<$.05, ** $<$.01. [2] Starred if the main effect of Relatedness for RTs is significant in a one-way ANOVA by subjects performed for that condition, *p $<$.05, **p $<$.01.

When Verb error rates were analysed, the interaction between Group and Relatedness was significant by participants only ($F1(1, 96) = 7.61$, $p < .01$, $F2(1,36) = 2.85$), reflecting a highly significant 13.9% effect for learners whilst the natives showed a less significant 6.8% effect with more item variability. The critical interaction between Group, Relatedness and High/Moderate was not significant; $F1(1,96) = 1.26$, $F2 < 1.0$ (see Figure 1). This reflects the fact that meaning relatedness affected both groups: learners showed a larger

Downloaded from http://slr.sagepub.com by JOHN WILLIAMS on March 19, 2008
© 2008 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

effect for highly related than for moderately related verbs (21.6% vs. 6.2%), and so did natives (11.5% vs. 2.0%).

### *4 Omnibus analyses of correct NO RTs*

Table 3 shows the effect of Relatedness in each condition (starred if significant), and Figure 2 provides a graphical comparison of the interference effects. In the ANOVA of NO RTs over all conditions and both participant groups, the critical two-way interaction between Group and Relatedness was significant; $F1(1,96) = 7.52$, $p < .01$, $F2(1,68) = 7.19$, $p < .01$. The learners showed a highly significant 140 ms interference effect while the natives showed a non-significant 52 ms effect. The three-way interaction between Group, High/Moderate and Noun/Verb was not significant ($F1(1,96) = 2.34$, $F2(1,68) = 1.24$), but there was a four-way interaction between Group, High/Moderate, Noun/Verb and Relatedness; $F1(1,96) = 10.78$, $p < .01$, $F2(1,68) = 3.92$, $p = .051$. The four-way interaction suggests that the groups are differentially affected by meaning relatedness depending on whether words are nouns or verbs.

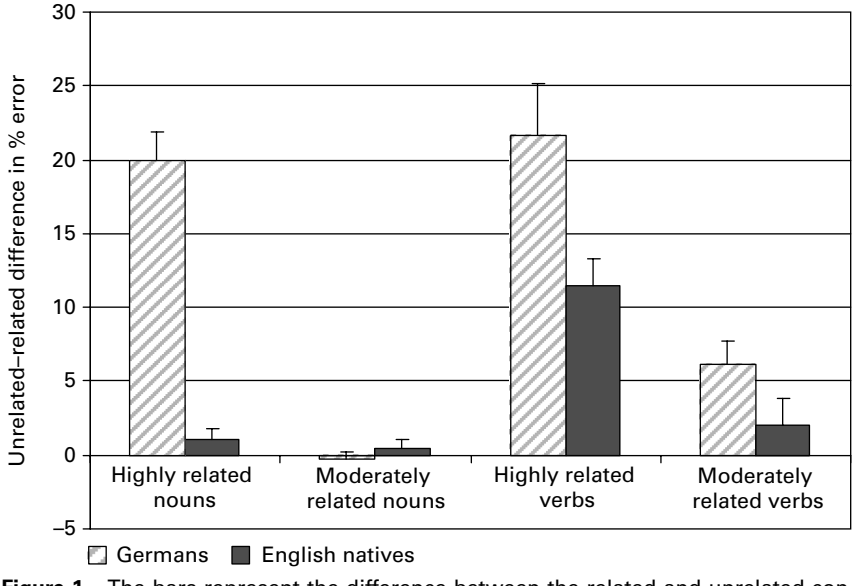


**Figure 1** The bars represent the difference between the related and unrelated conditions in percentage of errors made in the semantic anomaly judgement task, with Standard Error indicated above each bar

Downloaded from http://slr.sagepub.com by JOHN WILLIAMS on March 19, 2008
© 2008 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

Separate analyses of the Noun and Verb conditions were conducted to explore the critical two- and three-way interactions.

When RT Noun data were analysed, the critical two-way interaction between Group and Relatedness was significant; $F1(1,96) = 3.92$, $p = .05$, $F2(1,35) = 4.30$, $p < .05$. The learners showed a significant 121 ms interference effect, while the natives showed a non-significant 41 ms effect. The three-way interaction between Group, High/Moderate and Relatedness was borderline by participants and significant by items ($F1(1,96) = 3.06$, $p = .08$, $F2(1,35) = 4.30$, $p < .05$), reflecting the fact that learners showed larger interference effects for highly related than for moderately related nouns (206 ms vs. 37 ms), whilst natives showed comparable, but non-insignificant effects for highly and moderately related nouns (57 ms and 25 ms respectively).

When RT Verb data were analysed, the critical two-way interaction between Group and Relatedness was significant by participants and borderline by items; $F1(1,96) = 5.23$, $p < .025$, $F2(1,33) = 3.80$, $p = .06$. The three-way interaction between Group, High/ Moderate and Relatedness was significant (but in a form other than predicted) by

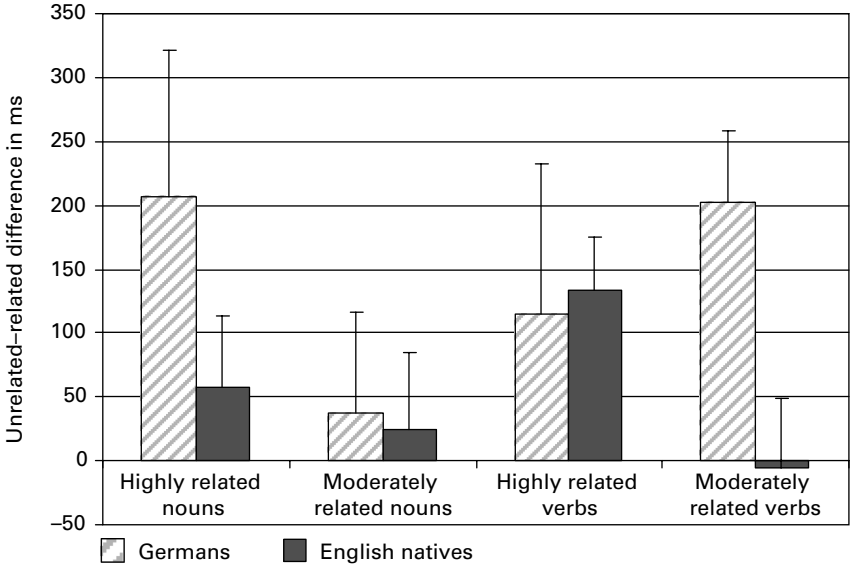


**Figure 2** The bars represent the difference between the related and unrelated conditions in reaction times made in the semantic anomaly judgement task, with Standard Error indicated above each bar

Downloaded from http://slr.sagepub.com by JOHN WILLIAMS on March 19, 2008
© 2008 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

participants only; $F1(1,96) = 6.64$, $p = .025$, $F2(1,33) = 2.09$. This interaction reflects the fact that natives showed a significant interference effect for highly related verbs (134 ms) but not for moderately related verbs (–6 ms), whereas the learners showed significant interference effects in both highly and moderately related verb conditions (115 ms and 203 ms respectively).[1]

## III Discussion

The significant Relatedness differences between the English and German participant groups in both errors and reaction times suggest L1 influence on L2 processing in instances where lexicalization patterns differ between L1 and L2. In addition, the Group by Relatedness interaction was accompanied by a four-way interaction between Group, Relatedness, High/Moderate and Noun/Verb. When we split up the analyses by Noun/Verb, only the nouns showed the three-way interaction that we argued would indicate conceptual influence from the L1. This means that, at least for nouns, access to L1 concepts is likely to be involved in the interference effects, either via a lexical-level translation connection (Route 2) or directly from L2 words (Route 3).

In the case of verbs, both the natives and learners showed highly significant interference in the highly related verbs conditions with pairs such as *scrape–scratch*, but the effects disappeared in the moderately related conditions for the natives and actually increased for the learners. As we argue below, the pattern for the natives is explicable in terms of the relatedness of verb meanings, and the results for the learners are consistent with Route 1, according to which interference is the result of lexical-level translation connections.

[1] Another difference observed between the groups was each groups' response to the block design (moderately related block followed by the highly related block) as observed in the control conditions. As can be seen in Table 3, natives showed a clear practice effect between the first and second blocks as they were faster and more accurate for highly related control trials than for moderately related ones. In contrast, the learners showed a shift between blocks in a possible speed–accuracy trade-off: they were slower, but equally accurate, in the highly related control conditions as compared to the moderately related control conditions. These control condition effects cannot have driven the highly significant Group differences and Relatedness effects, however.

Downloaded from http://slr.sagepub.com by JOHN WILLIAMS on March 19, 2008
© 2008 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

Considering first the natives, their results are explicable in terms of verb processing in general. Källkvist (1997) observed that verbs are generally more polysemous than nouns: according to the *Collins English Dictionary* verbs on the whole have an average of 2.11 senses while nouns have 1.74 senses (Miller and Fellbaum, 1992). Along these lines, Brown (1994: 89) showed with a verbatim recall task that participants are more likely to use the nominal arguments of verbs than the verbs themselves in a retelling. This is not to say that the verbs in the highly related conditions were more highly related than the nouns (the ratings in the two conditions were comparable), but it could mean that – due to the general tendency of verbs to be polysemous and their senses context dependent – the natives had more difficulty trusting their on-line judgements of acceptability for highly related verbs, leading to longer response times.

The learners, in contrast, showed significant inhibition in both highly and moderately related conditions. Verbs appear to be especially problematic for L2 learners. Källkvist (1997: 119, 135) found that verbs are more likely than nouns to be used incorrectly by advanced Swedish learners in written compositions and in retellings of stories. This helps to explain why verbs posed problems for the learners across the board. Another explanation can be found in Van Hell and De Groot (1998), who suggested that verb meanings are less likely than concrete noun meanings to be mediated conceptually because verbs share fewer features than nouns. On this view, words that share a large number of conceptual features are linked at the conceptual level, while words that differ are more likely to be linked via word form. As one might argue that verb meanings have more language-specific features, conceptual overlap is less likely, so the system might rely on L1–L2 lexical links instead. This implies that shared concepts cannot be the primary locus of the interface between the L1 and L2 lexicons, so translational links at the word form level across languages might be particularly strong. For example, *scrape* activates *Kratzen*, which in turn activates *scratch* (Route 1).

Turning to nouns, the results suggest that learners activate L1 conceptual information, perhaps directly from L2 words (Route 3). Jiang (2000; 2002) has argued that even advanced learners might simply utilize copies of L1 lexical-semantic information in order to map L2 words

Downloaded from http://slr.sagepub.com by JOHN WILLIAMS on March 19, 2008
© 2008 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

onto meanings. The immediate problem with this approach, however, is to explain how it is that our participants were able to produce correct negative responses 84.2% of the time, if only more slowly than in the non-polysemous control condition. This shows that most of the time they were able to derive a meaning interpretation of, say *bag*, that was sufficient to reject it as a completion of 'On his trousers was a large...' Jiang (2002) argues that learners can use explicitly available declarative knowledge to distinguish the uses of L2 words that share common L1 translations, whilst at the level of implicitly represented word meanings they could still be utilizing the semantic specifications inherited from L1. On this view, the interference experienced by the learners is a result of explicit knowledge overriding the positive response delivered by implicit knowledge.

Given the nature of the decision task in our experiment we cannot rule out the possibility that performance was contaminated by explicit knowledge. There is no way of knowing exactly how 'on line' a task has to be in order to preclude the intervention of explicit knowledge, and this makes Jiang's proposal difficult to falsify. However, note that the overall interference effect in reaction time was only 88 ms greater for the learners than the natives, and 8.1% greater in errors. One might have expected a larger increase in error rate if the learners were utilizing L1 lexical-semantic information (which would deliver a positive response). Also, a greater reaction time increase might have been expected if learners were using explicitly retrieved declarative knowledge to override implicitly generated positive responses. Also note that it would still be necessary to appeal to lexical-level translation connections (Route 1) to account for the absence of a relatedness effect for verbs.

Therefore, it seems more plausible that the present results for nouns reflect Route 2 rather than Route 3. On this view, the learners utilized correct L2 form–meaning mappings in order to derive a negative response most of the time, but the L1 concept also became somewhat active due to lexical-level translation links, leading to slight increases in reaction times and errors. The co-activation of L1 concepts via lexical-level translation connections could reflect the residual effects of an earlier stage of development at which the learner was more reliant on lexical-level translation. Thus, the effects we observed in this study may reflect an intermediate stage of development towards language-specific lexicalization patterns

Downloaded from http://slr.sagepub.com by JOHN WILLIAMS on March 19, 2008
© 2008 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

(compare MacWhinney, 2005). This result is perhaps rather surprising, given that for the kinds of words investigated here, there is considerable pressure on the learner to develop strong form–meaning connections and not to rely on translation equivalence. From this perspective, the results are consistent with models of the bilingual lexicon in which lexical-level translation connections are particularly strong in the L2-to-L1 direction (Kroll and Stewart, 1994) and continue to affect processing even after direct form–meaning connections have been established.

Taken together the results for both nouns and verbs suggest that L1 lexicalization patterns do indeed influence semantic processing of L2 words. Even in a strong English L2 environment (experimenter, stimuli and task all in English), learners activated feature sets and/or lexical concepts that differed from those activated by natives while reading words in context. However, we have to limit our claim of strong L1 conceptual influence to a certain type of L2 learner. It may be that early acquisition is crucial if cross-language interference in meaning activation is to be avoided. There is ERP evidence to suggest that early age of acquisition is particularly important for native-like word-to-concept links (Kotz and Elston-Güttler, 2004), while word-to-word links can be more easily acquired later in life. Also, parallel bilinguals with an early age of acquisition appear to set the conceptual boundaries for words like *cup* and *bowl,* where lexicalization differs between their two languages, somewhere in between the boundaries exhibited by monolingual speakers of those languages (Ameel *et al.*, 2005). Thus, the kind of reliance on L2 lexicalization patterns observed in the present set of experiments may be limited to late learners of a language. This, of course, remains an issue for future research.

*Acknowledgements*

The authors wish to thank David Singleton and Richard Breheny for helpful comments on an earlier version, and Anette Rosenbach for participant recruitment in Düsseldorf.

## IV References

**Altarriba, J.** and **Mathis, K.M.** 1997: Conceptual and lexical development in second language acquisition. *Journal of Memory and Language* 36, 550–68.

Downloaded from http://slr.sagepub.com by JOHN WILLIAMS on March 19, 2008
© 2008 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

**Ameel, E., Storms, G., Malt, B.C.** and **Sloman, S.A.** 2005: How Bilinguals solve the naming problem. *Journal of Memory and Language* 52, 60–80.

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

**Bierwisch, M.** and **Schreuder, R.** 1992: From concepts to lexical items. *Cognition* 42, 23–60.

**Brown, G.** 1994: The role of nominal and verbal expressions in retelling narratives: implications for discourse representations. *Working Papers in English and Applied Linguistics* 1, 81–90.

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

**Elston-Güttler, K.E.** 2000: An enquiry into cross-language lexical–conceptual relationships and their effect on L2 lexical processing. Unpublished PhD dissertation, University of Cambridge.

**Elston-Güttler, K.E., Paulmann, S.** and **Kotz, S.A.** 2005: Who's in control?: proficiency and L1 influence on L2 processing. *Journal of Cognitive Neuroscience* 17, 1593–1610.

**Frenck-Mestre, C.** and **Prince, P.** 1997: Second language autonomy. *Journal of Memory and Language* 37, 481–501.

**Götz, D., Haensch, G.** and **Wellmann, H.,** 1993: *Langenscheidts Grosswörterbuch: Deutsch als Fremdsprache [Langenscheidts dictionary of German as a foreign language]*. Berlin: Langenscheidt.

**Hernandez, A., Li, P.** and **MacWhinney, B.** 2005: The emergence of competing modules in bilingualism. *Trends in Cognitive Sciences* 9, 220–25.

**Jiang, N.** 2000: Lexical representation and development in a second language. *Applied Linguistics* 21, 47–77.

—— 2002: Form–meaning mapping in vocabulary acquisition in a second language. *Studies in Second Language Acquisition* 24, 617–37.

**Källkvist, M.H.** 1997: Lexical form–class effects in foreign language learning: a study of the English produced by advanced Swedish learners. Unpublished PhD dissertation, University of Cambridge.

**Kotz, S.A.** and **Elston-Güttler, K.E.** 2004: The role of proficiency on processing categorical and associative information in the L2 as revealed by reaction times and event-related brain potentials. *Journal of Neurolinguistics* 17, 215–35.

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

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

Downloaded from http://slr.sagepub.com by JOHN WILLIAMS on March 19, 2008
© 2008 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

**Levelt, W.J.M., Roelofs, A.** and **Meyer, A.S.** 1999: A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22, 1–75.

**MacWhinney,** 2005: Emergent fossilization. In Han, Z. and Odlin, T., editors, *Studies of fossilization in second language acquisition*. Clevedon: Multilingual Matters, 134–56.

**Miller, G.A.** and **Fellbaum, C.** 1992: Semantic networks of English. *Cognition* 41, 197–229.

**Norris, D.** 1984: A computer-based programmable tachtiscope for non-programmers. *Behavior Methods and Instrumentation* 16, 25–27.

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

**Roelofs, A.** 2000: Word meanings and concepts: what do the findings from aphasia and language specificity really say? *Bilingualism: Language and Cognition* 3, 25–27.

**Slobin, D.I.** 1987: Thinking for speaking. In Aske, J., Beery, N., Michaelis J. and Filip, H., editors, *Berkeley Linguistics Society: Proceedings of the thirteenth annual meeting*. Berkeley, CA: Berkeley Linguistics Society.

**Slobin, D.I.** 1991: Learning to think for speaking. *Pragmatics* 1, 7–25.

—— 1996: From 'thought and language' to 'thinking for speaking'. In Gumperz, J.J. and Levinson, S.C., editors, *Rethinking linguistic relativity.* Cambridge: Cambridge University Press, 70–96.

**Van Hell, J.G.** and **De Groot, A.M.B.** 1998: Conceptual representation in bilingual memory: effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition* 1, 193–211.

Downloaded from http://slr.sagepub.com by JOHN WILLIAMS on March 19, 2008
© 2008 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.