Semantic access in second-language visual word processing: Evidence from the semantic Simon paradigm

WOUTER DUYCK AND JAN DE HOUWER
Ghent University, Ghent, Belgium

The present study investigated whether visually presented second-language words activate their meaning during low-level word processing, just as native-language words do. Using the semantic Simon paradigm (De Houwer, 1998) with a letter-case judgment task, Dutch–English bilingual participants were instructed to classify targets' letter case using verbal labels (e.g., by saying "animal" to uppercase targets or "occupation" to lowercase targets). Results showed that both native-language and second-language targets yielded faster responses if the verbal response corresponded to the targets' semantic category (e.g., the response "animal" to the target LEEUW or LION) than when it did not (e.g., LAWYER), even though the meaning of target words was irrelevant for the task. These results show that second-language word forms may quickly and automatically activate their meaning through strong form-to-meaning mappings, which is consistent with theories of bilingual lexicosemantic organization, such as that of Duyck and Brysbaert (2004).

Ever since the growth of interest in bilingual language processing, a great deal of research has focused on the similarities between language processing and language representation in the native language (L1) and in a second language (L2). This research, mostly focused on word recognition, has now yielded a growing consensus that lexical access is similar in L2 and L1, because lexical entries from both languages are represented in a unified lexicon, which is characterized by cross-lingual interactions during early stages of word recognition. Indeed, it has been shown that L1 knowledge influences L2 word recognition and vice versa (see, e.g., Duyck, 2005). There is more debate, however, over how these lexical representations interact with their underlying semantic representations. This debate originates from a central assumption of the most influential model of bilingual lexicosemantic organization—namely, the revised hierarchical model (RHM) of Kroll and Stewart (1994). This model assumes that L2 word forms (lexical representations) do not have direct access to semantics, unlike L1 words, but instead only indirectly activate their meaning through lexical associations with their L1 translation equivalents. The RHM assumes that strong form-to-meaning mappings for L2 words develop only at very high levels of L2 proficiency. Recently, however, other models of bilingual lexicosemantic organization, such as the model proposed by Duyck and Brysbaert (2004), have argued that neither the speed of development nor the strength of L2 form-to-meaning mappings depends primarily on general subject characteristics, such as L2 proficiency. Instead, it is assumed that the mapping of form to meaning in L2 is also dependent on word variables. More specifically, this model predicts that L2 word forms may develop strong connections with semantic representations early if word meaning is well confined and largely overlapping across languages, as is the case for number words.

At present, evidence for either of these theoretical viewpoints comes exclusively from translation tasks. Using these tasks, researchers have investigated whether translation from L2 to L1 (backward translation) and vice versa (forward translation) is sensitive to semantic manipulations. Typically, if backward translation does not yield semantic effects but forward translation does, this is attributed to the fact that L2 words do not have strong form-to-meaning mappings and activate their L1 translation equivalents through lexical associations, providing evidence for the asymmetric lexicosemantic organization of the RHM. This line of reasoning was adopted, for example, in the original formulation of the RHM by Kroll and Stewart (1994), who observed that blocking the translation targets by semantic category inhibited performance in forward translation, but not in backward translation.

Recently, however, there have been a number of demonstrations of semantic effects in backward translation, suggesting that strong lexicosemantic links may exist for L2 words as well, as was assumed by Duyck and Brysbaert (2004). First, using a bilingual Stroop task, La Heij, Hooglander, Kerling, and van der Velden (1996) found that congruent color words (for which the ink color corresponded to the word) were translated faster than incongruent color words. This was true for both directions of translation, suggesting that L2 color words quickly activated

W. Duyck, wouter.duyck@ugent.be

their underlying meaning during translation. Second, this bilingual Stroop effect was replicated by Altarriba and Mathis (1997) in a group of monolinguals who were trained on a new set of English–Spanish translation equivalents, showing that these form-to-meaning mappings may develop early during L2 word learning. Third, other evidence for semantic mediation in backward translation comes from number-word translation studies. Duyck and Brysbaert, for example, applied the monolingual number magnitude effect, commonly investigated in the numerical cognition domain, as a semantic marker to a word-translation task. They found that it takes longer to translate number words (in both directions) representing larger quantities (e.g., *eight*) relative to number words representing smaller quantities (e.g., *four*), independently of word length and frequency effects. Because such a number magnitude effect reveals semantic access, Duyck and Brysbaert argued that L2 number words are mapped strongly onto their underlying semantic representation. This effect was also found with a set of artificial number words acquired just before the translation task, again providing evidence for the fast development of lexicosemantic connections.

Although the question of whether word translation as a process yields semantic activation is certainly interesting on its own, it should not be forgotten that the translation paradigm was initially just a tool to investigate the architecture of the bilingual lexicosemantic system. However, currently all of the evidence against (e.g., Kroll & Stewart, 1994) or supporting (e.g., Duyck & Brysbaert, 2004) the existence of L2 form-to-meaning mappings relies on the translation paradigm, even though these theories also yield different predictions for semantic access in other L2 word-processing tasks. Therefore, it is the aim of the present study to investigate whether L2 form-to-meaning mappings exist and whether these connections may also be quickly activated during low-level L2 word processing, not just during backward translation (which requires deeper lexical processing). Finding such evidence would constitute stronger support for the existence of strong L2 form-to-meaning mappings.

To investigate whether processing L2 visual word forms may quickly yield semantic access, the present study made use of the semantic Simon paradigm, proposed by De Houwer (1998). This paradigm originated from the classical Simon paradigm, in which participants are required to make a spatial response (e.g., to press the right key) to a nonspatial stimulus characteristic (e.g., when the stimulus is green). The crucial point is that an irrelevant spatial stimulus characteristic (e.g., its position to the left or right of the fixation location) has to be ignored. Typically, this results in faster responses to congruent targets, in which the response corresponds to the spatial feature of the target (e.g., faster responses to green stimuli presented on the right), even though this is irrelevant for the task at hand. De Houwer proposed a new variant of this classical paradigm, in which the irrelevant stimulus property concerned the meaning of word targets. In one of the experiments, he presented words from two semantic categories (names of animals and occupations) in either uppercase or lowercase letters (the relevant stimulus feature). Par-

ticipants had to make letter-case judgments by responding with one of the semantic category's names (e.g., by saying "occupation" for uppercase targets). He observed that responses were faster if the response meaning was congruent with the task-irrelevant target meaning (e.g., saying "occupation" for PILOT) than when it was not (e.g., saying "occupation" for LION). From this, he concluded that targets automatically activated their underlying meaning and the verbal response associated with the respective semantic category. Subsequently, this facilitated (congruent) or competed with (incongruent) the verbal response to the relevant target feature, causing the Simon effect. Because this semantic Simon paradigm allows integration of both the relevant (case) and irrelevant (meaning) stimulus features in the same stimulus, and because these features are completely unrelated, this paradigm offers an interesting way to investigate automatic semantic processing. This is less the case in the semantic-priming paradigm and in Stroop studies, two probable alternatives to investigate such issues, in which semantically related targets and distractors are often also related on nonsemantic dimensions (e.g., associatively or perceptually; De Houwer, 1998). A few earlier studies have, therefore, already successfully adopted the semantic Simon paradigm to investigate semantic access in monolingual word processing (e.g., Ghyselinck, Custers, & Brysbaert, 2004).

To date, no study has ever used this paradigm to investigate bilingual word processing. To investigate the existence and the time course of activation of L2 form-to-meaning mappings, the present study therefore adopted a bilingual version of De Houwer's (1998) semantic Simon paradigm. More specifically, we investigated whether a low-level processing task, performed by bilinguals on visually presented L2 words, yields a semantic congruency effect similar to the monolingual effects described above. This would offer strong evidence for the assumption that L2 form-to-meaning mappings exist and may indeed be quickly activated during word processing, just as in L1.[1] The case-judgment task (see also Ghyselinck et al., 2004) provides a particularly strong test of this assumption, because it is a task that may be performed solely on the basis of visual stimulus features, not requiring any lexical processing (unlike the translation task or grammatical word category judgment and language classification; De Houwer, 1998).

METHOD

Participants

Sixteen university students participated for course requirements. Mean age was 20.2 years (SD = 2.7). They were all late Dutch–English bilinguals and were living in an L1-dominant environment, speaking Dutch at home, at school, with parents, with friends, and so on. All participants had learned English and French in a scholastic setting (formal English and French courses are mandatory in Belgian high schools). In everyday life, Belgian teenagers are much more exposed to English than to French through Belgian popular media and entertainment, such as music, the Internet, films, television, and so on. This high exposure, and the larger similarity between Dutch and English, relative to French, explains why participants reported English as their L2. Participants were asked to rate their L1 and L2 proficiency with respect to several skills (reading, writing, speaking, general proficiency) on a 7-point Likert scale ranging from *very bad*

Table 1
Self-Assessed Ratings of L1 and L2 Proficiency on a 7-Point Likert Scale Ranging From 0 (*Very Bad*) to 7 (*Very Good*)

Skill	L1 (Dutch)		L2 (English)	
	M	*SD*	*M*	*SD*
Writing	6.5	0.6	4.5	1.0
Speaking	6.6	0.5	4.9	1.0
Reading	6.6	0.5	5.2	0.9
General proficiency	6.6	0.5	4.9	0.9

to *very good*. Means are reported in Table 1. Self-reported, general L1 ($M = 6.5$) and L2 ($M = 4.9$) proficiency differed significantly [$F(1,15) = 50.86, p < .001$]. Ratings for specific language skills also differed significantly between L1 and L2 (all $ps < .001$).[2]

Stimuli

The materials and procedure for the semantic Simon paradigm were as similar as possible to those of De Houwer (1998, Experiment 2).

Stimuli consisted of 40 L1 target words and their 40 L2 translation equivalents, half of which were names of animals and half of which were occupations (see the Appendix). We used translation equivalents to ensure that any difference between L1 and L2 processing could not be confounded with the specific concepts tested. Cognates (i.e., words that are identical across languages with respect to meaning and orthography; e.g., Dutch–English: *film*) and interlingual homographs (i.e., words that share orthography but not meaning; e.g., *room*, which means "cream" in Dutch) were excluded. As can be seen in Table 2, for each semantic category, L1 targets were matched with their L2 translation equivalents with respect to number of letters, number of syllables, neighborhood size, and log frequency per million words (all $ps > .25$), calculated using the WordGen stimulus generation program (Duyck, Desmet, Verbeke, & Brysbaert, 2004). This ensures that any difference in the semantic Simon effect between L1 and L2 may not be attributed to any inherent lexical word characteristic confounded with language. Participants saw each target word only once, either in uppercase or in lowercase letters. Therefore, two stimulus lists were used, counterbalanced over participants. Hence, across participants, for each category and language, each target appeared an equal number of times in uppercase and lowercase. Again, for each semantic category and for each language, these lists were also matched on number of letters, number of syllables, neighborhood size, and log frequency per million words (all $ps > .43$).

Table 2
Mean Stimulus Characteristics

	Language		
	L1	L2	*p*
Animals			
Number of letters	4.7	4.8	.70
Number of syllables	1.3	1.5	.27
Word frequency[a]	1.32	1.32	.98
N[b]	7.6	6.0	.37
Occupations			
Number of letters	6.4	6.0	.26
Number of syllables	2.1	1.9	.33
Word frequency[a]	1.32	1.35	.49
N[b]	2.8	2.5	.55

Note—Reported *p* values indicate significance levels of dependent samples comparisons between L1 and L2 translation equivalents. [a]Logarithm of word frequency per million words according to the CELEX Lexical Database (word lemmata) (Baayen, Piepenbrock, & van Rijn, 1993). [b]Neighborhood size (i.e., the number of existing words in a given language differing by only a single letter), calculated using the WordGen program (Duyck, Desmet, Verbeke, & Brysbaert, 2004) on the basis of the CELEX lemma database (Baayen et al., 1993).

Procedure

The experiment consisted of two experimental sessions, with a minimum of 5 days in between. In one session, participants performed the case-judgment task with L1 targets; in the other, with L2 targets. The order of sessions was counterbalanced across participants. Half of the participants were instructed to respond "dier" ("animal") to words written in uppercase letters and "beroep" ("occupation") to words written in lowercase letters; the reverse was true for the other participants. Responses were always in L1, providing the most conservative test of semantic access during L2 word processing. Mapping of responses to case was also counterbalanced across participants, together with stimulus list and order of experimental sessions. It remained constant across sessions within participants, to avoid long-term task interference effects. As in De Houwer (1998), instructions mentioned that stimuli would correspond to names of animals or occupations, but it was emphasized that the meaning of the word should be ignored. Before the actual experiment, participants completed 16 practice trials. As in the actual experiment, half of the practice targets were presented in uppercase letters, the other half in lowercase letters. In these trials, half of the targets were occupations, half were animals. Each trial started with the presentation of a warning tone (1000 Hz) lasting 200 msec, followed by a fixation cross (for 500 msec), a blank screen (for 500 msec), and the target. Targets stayed on the screen until the participant's response triggered a highly accurate voice key (Duyck et al., in press), or until the maximal response time (RT) was reached (3,000 msec). During the intertrial interval (1,500 msec), the experimenter noted whether the voice key had functioned correctly. The order of experimental trials within a session was randomized. After the two experimental sessions, participants completed the proficiency questionnaire and a short translation test of the L2 targets, which revealed that all participants knew the meaning of the L2 targets.

RESULTS

Practice trials, trials with background noise, and trials for which a response was missing were not included in the analyses. The proportion of case-judgment errors on the remaining trials was very low (2.11%). All RTs that deviated more than 2.5 standard deviations from the participant's overall RT were considered outliers and were removed from the analyses (2.54% of the remaining data). Mean RTs on correct trials were analyzed by means of a repeated measures ANOVA with language (L1 vs. L2) and congruency (incongruent vs. congruent) as independent variables.[3] Analyses were run with participants (F_1) and items (F_2) as the random variables. Because accuracy was virtually at a maximum in all conditions, analyses of errors showed no significant effects (all $ps > .21$) and will therefore not be reported below. Mean RTs and accuracy (% errors) across participants as a function of language and congruency are displayed in Table 3.

The main effect of language was not significant (both $Fs < 1$). If anything, participants made somewhat faster case judgments for L2 words ($M = 546$) than for L1 words ($M = 560$). The absence of a language effect supports the assumption that case judgment is a low-level task that is not sensitive to lexical processing. Importantly, the effect of congruency was significant, in the analyses both by participants and by items [$F_1(1,15) = 8.65, p < .01$; $F_2(1,78) = 10.13, p < .01$]. Congruent responses— trials in which the vocal response corresponded to the target's semantic category—were faster ($M = 540$) than incongruent responses ($M = 566$). Planned comparisons

Table 3
Mean Reaction Times (RTs, in Milliseconds) and
Accuracy (% Errors) Across Participants As a
Function of Language and Congruency

	L1				L2			
	RT		Accuracy		RT		Accuracy	
	M	SD	M	SD	M	SD	M	SD
Incongruent	574	32.3	1.25	0.72	557	23.5	2.57	0.94
Congruent	546	24.7	1.67	0.94	534	23.4	2.96	0.96
Effect	28*		−0.42		23*		−0.39	

*$p < .05$ in analyses across participants and across items.

showed that in L1, the difference between incongruent ($M = 574$) and congruent ($M = 546$) trials was significant [$F_1(1,15) = 5.67$, $p < .05$; $F_2(1,39) = 5.17$, $p < .05$]. Most important, in L2, the difference between incongruent ($M = 557$) and congruent ($M = 534$) responses was also significant [$F_1(1,15) = 5.75$, $p < .05$; $F_2(1,39) = 5.10$, $p < .05$]. Interestingly, the effect of congruency did not interact with language (both Fs < 1). The 28-msec semantic congruency effect for L1 was not significantly stronger than the 23-msec effect for L2.[4]

DISCUSSION

Letter-case judgments were faster if the associated verbal response (e.g., saying "animal" to uppercase targets) was congruent with the semantic category that the target belonged to (e.g., target DOG), even though the meaning of targets was irrelevant for the task. Interestingly, this congruency effect was as strong in L1 as it was in L2. This provides strong evidence that L2 word forms may automatically access their underlying semantic representation very quickly, just as L1 words do (De Houwer, 1998). Note that the letter-case judgment task constitutes a very conservative test of this assumption, because it may accurately be performed solely on the basis of visual feature analysis, not requiring any lexical or semantic processing.

These results are consistent with earlier findings of semantic mediation in backward translation of L2 words (e.g., Altarriba & Mathis, 1997; Duyck & Brysbaert, 2004; La Heij et al., 1996). Importantly, the present study offers an important extension of these findings, because at present, the existence of L2 form-to-meaning mappings has exclusively been investigated using the translation paradigm. The semantic Simon paradigm (De Houwer, 1998) adopted here revealed not only that L2 form-to-meaning mappings exist, but also that they may be quickly activated during low-level L2 word processing, not only during backward word translation (which requires deeper lexical processing). Note that the present study was not limited to a very specific class of words that may be represented somewhat differently in the bilingual language system, such as number words (Duyck & Brysbaert, 2004) or color words (La Heij et al., 1996), but instead used common words of two broad semantic categories, just as in the original formulation of the RHM (Kroll & Stewart, 1994). As such, these findings support the generalizability of models of bilingual lexicosemantic organization that do not assume qualitatively different representations for L1 and L2 words,

such as that of Duyck and Brysbaert. In this view, it will be interesting to see whether this L2 semantic Simon effect may be replicated with languages that have different writing systems or with a third language. Also, given that the model of Duyck and Brysbaert allows influences of word variables on lexicosemantic organization, unlike the RHM, future research might explore whether this L2 Simon effect may also be obtained with abstract words, for example.[5]

To our knowledge, there is currently only one other study in the literature that has looked at L2 word comprehension to make inferences about bilingual lexicosemantic organization. La Heij and Akerboom (2007) showed that semantic categorization of L2 words (e.g., *horse*) was not influenced by L1 distractors (e.g., *paal*, "pole") that are phonologically related to the targets' L1 translation equivalents (e.g., *paard*). Because speeding up retrieval of L1 translation equivalents should facilitate L2 word processing if L2 words indeed access semantics through L1, La Heij and Akerboom concluded that L2 words have direct access to meaning without activation of their L1 translation equivalents. These findings are consistent with the present study but offer somewhat less strong evidence. First, these conclusions rely on a null effect of the phonological manipulation in the semantic categorization task. Second, the present semantic congruency effect was obtained with a low-level visual analysis task (case judgment), whereas semantic categorization explicitly requires access to meaning. Third, case judgment yields much faster RTs ($M = 546$ msec) than the semantic categorization task of La Heij and Akerboom ($M = 979$ msec), providing strong evidence for the fast time course by which L2 words may activate meaning.

Interestingly, there are also two other indications of automatic semantic activation in L2 word processing, originating from the literature on the processing of emotional stimuli. Very recently, Eilola, Havelka, and Sharma (2007) observed that ink color judgments by Finnish–English bilinguals were delayed for negative and taboo words, relative to neutral words (the so-called emotional Stroop effect), in both L1 and L2. Note, however, that this paradigm involved repeated presentation of all targets (four times), which may have boosted semantic activation, relative to neutral language contexts. Also, this task (L2, $M = 717$ msec) was again much slower than case judgment (probably because there were four possible manual responses) and may therefore not reveal much about the particularly fast time course by which L2 words may activate meaning. This emotional Stroop effect has been replicated by Sutton, Altarriba, Gianico, and Basnight-Brown (2007) using early Spanish–English bilinguals living in an L2 environment.

Finally, it is also important to consider an alternative account of the L2 Simon effect within the framework of the RHM (Kroll & Stewart, 1994). One could assume that L2 word forms accessed their L2 lexical representations, which then activated the L1 lexical representations of their translation equivalents, which subsequently activated semantics. In this view, the L2 Simon effect would emerge from intermediate, internal backward translation, and would indeed not reflect the existence of strong and direct L2 form-to-meaning mappings. However, because responses took only about 550 msec in both the L2 and

L1 conditions, this account would also need to make the additional assumptions that such an internal backward translation step occurs very rapidly (i.e., within a few hundred milliseconds) and does not interfere with the time and resources needed to perform the primary task (because latencies were, if anything, shorter in the L2 than in the L1 condition). Under these assumptions, it is indeed difficult to exclude with absolute certainty that such a fast and efficient translation process occurred during the 550 msec that participants needed to perform a case judgment on L2 words. However, we believe that this is not a very plausible or theoretically valuable account. First, we know that translation processes, even disregarding their input comprehension and response production components, require resources and cost time (see, e.g., Cheung & Chen, 1998; Duyck & Brysbaert, 2004). Hence, one would at least expect that they should interfere with the primary task. This was not confirmed by the data, since case judgment was equally fast for L2 as for L1. Second, the assumption of a quick and efficient intermediate translation process without time cost for the primary task contradicts the RHM model itself: If it is indeed the case that L2 words may access semantics through intermediate L1 activation, virtually as fast as L1 words do directly, and even simultaneously with other tasks, there would be no point in having an asymmetric lexicosemantic organization with strong form-to-meaning mappings for L1 but not for L2, because such a model would behave in a functionally equivalent manner for L1 and L2. In this view, L2 words should always yield semantic activation as fast as L1 words do. Consequently, one would certainly expect backward translation (which is much slower than case judgment) to be sensitive to semantic manipulations, just as forward translation is. And this is exactly what the RHM does not predict. Finally, note that the low-level case-judgment task that we used yielded particularly fast RTs compared with the tasks of earlier studies that have argued for semantic access in L2 word processing (see above). As such, the present L2 Simon effect is probably least vulnerable to this alternative account with simultaneous intermediate translation and, therefore, offers the strongest evidence available in the literature for the existence of strong L2 form-to-meaning mappings.

To summarize, using a bilingual version of the semantic Simon paradigm, we observed that bilinguals making case judgments to L2 targets (e.g., DOG) were faster if the required verbal response corresponded to the target's semantic category (e.g., "animal" for uppercase targets) than when it did not. These results suggest that L2 word forms may quickly and automatically activate their underlying meaning.

REFERENCES

ALTARRIBA, J., & MATHIS, K. M. (1997). Conceptual and lexical development in second language acquisition. *Journal of Memory & Language*, **36**, 550-568.

BAAYEN, R. H., PIEPENBROCK, R., & VAN RIJN, H. (1993). *The CELEX Lexical Database* (Release 1) [CD-ROM]. Philadelphia: University of Pennsylvania, Linguistic Data Consortium.

CHEUNG, H., & CHEN, H.-C. (1998). Lexical and conceptual processing in Chinese–English bilinguals: Further evidence for asymmetry. *Memory & Cognition*, **26**, 1002-1013.

DE HOUWER, J. (1998). The semantic Simon effect. *Quarterly Journal of Experimental Psychology*, **51A**, 683-688.

DUSCHERER, K., HOLENDER, D., & MOLENAAR, E. (2008). Revisiting the affective Simon effect. *Cognition & Emotion*, **22**, 193-217.

DUYCK, W. (2005). Translation and associative priming with cross-lingual pseudohomophones: Evidence for nonselective phonological activation in bilinguals. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **31**, 1340-1359.

DUYCK, W., ANSEEL, F., SZMALEC, A., MESTDAGH, P., TAVERNIER, A., & HARTSUIKER, R. (in press). Improving accuracy in detecting acoustic onsets. *Journal of Experimental Psychology: Human Perception & Performance*.

DUYCK, W., & BRYSBAERT, M. (2004). Forward and backward number translation requires conceptual mediation in both balanced and unbalanced bilinguals. *Journal of Experimental Psychology: Human Perception & Performance*, **30**, 889-906.

DUYCK, W., DESMET, T., VERBEKE, L. P. C., & BRYSBAERT, M. (2004). WordGen: A tool for word selection and nonword generation in Dutch, English, German, and French. *Behavior Research Methods, Instruments, & Computers*, **36**, 488-499.

EILOLA, T. M., HAVELKA, J., & SHARMA, D. (2007). Emotional activation in the first and second language. *Cognition & Emotion*, **21**, 1064-1076.

GHYSELINCK, M., CUSTERS, R., & BRYSBAERT, M. (2004). The effect of age of acquisition in visual word processing: Further evidence for the semantic hypothesis. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **30**, 550-554.

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

LA HEIJ, W., & AKERBOOM, S. P. (2007). Word comprehension in a second language: A direct or an indirect route to meaning? *Psychological Reports*, **100**, 838-846.

LA HEIJ, W., HOOGLANDER, A., KERLING, R., & VAN DER VELDEN, E. (1996). Nonverbal context effects in forward and backward word translation: Evidence for concept mediation. *Journal of Memory & Language*, **35**, 648-665.

SUTTON, T. M., ALTARRIBA, J., GIANICO, J. L., & BASNIGHT-BROWN, D. M. (2007). The automatic access of emotion: Emotional Stroop effects in Spanish–English bilingual speakers. *Cognition & Emotion*, **21**, 1077-1090.

NOTES

1. De Houwer (1998) also tested for a semantic Simon effect in a language judgment task (Dutch vs. English). This implied that both L1 and L2 targets were presented, and that these existing data may therefore already reveal something about the semantic congruency effect in the present study, even though bilingual language processing was not of interest to De Houwer. A new analysis on these existing data showed that his L2 targets indeed also yielded a congruency effect [$F(1,15) = 7.44, p < .05$].

2. Our bilinguals were very similar to those of Kroll and Stewart (1994), who were also unbalanced Dutch–English bilingual university students, with a similar English schooling background and a very similar L2 proficiency rating (4.9 and 4.8, transformed to a 7-point scale).

3. Additional analyses showed that the counterbalancing variables did not interact with any of the effects reported below.

4. Both these L1 and L2 Simon effects were replicated in an experiment that also included 30 semantically neutral filler stimuli, although both effects were somewhat smaller (similar to earlier monolingual findings; Duscherer, Holender, & Molenaar, 2008).

5. We thank an anonymous reviewer for these suggestions.

APPENDIX
Stimuli As a Function of Language
and Semantic Category

Animals		Occupations	
L1 (Dutch)	L2 (English)	L1 (Dutch)	L2 (English)
konijn	rabbit	slager	butcher
arend	eagle	soldaat	soldier
dolfijn	dolphin	boer	farmer
aap	monkey	koning	king
vos	fox	pastoor	priest
leeuw	lion	tuinman	gardener
paard	horse	duiker	diver
eend	duck	postbode	postman
gans	goose	spion	spy
hond	dog	advocaat	lawyer
koe	cow	komiek	comedian
vogel	bird	matroos	sailor
schaap	sheep	tandarts	dentist
varken	pig	rechter	judge
spin	spider	ober	waiter
schildpad	turtle	leraar	teacher
uil	owl	piloot	pilot
duif	pigeon	schilder	painter
muis	mouse	jager	hunter
kip	chicken	dichter	poet