Fall 10-25-2007

Phonological facilitation through translation in a bilingual picture-naming task

Paul Amrhein
Montclair State University, amrheinp@montclair.edu

Aimee Knupsky
University of New Mexico, aknupsky@allegheny.edu

Published Citation

Knupsky, Aimee C., and Paul C. Amrhein. "Phonological facilitation through translation in a bilingual picture-naming task." Bilingualism: Language and Cognition 10, no. 3 (2007): 211-223.

Phonological facilitation through translation in a bilingual picture-naming task

Article *in* Bilingualism · November 2007

DOI: 10.1017/S1366728907003033

2 authors:

Aimee Knupsky
Allegheny College

Paul Amrhein
Montclair State University

Bilingualism: Language and Cognition 10 (3), 2007, 211–223 © 2007 Cambridge University Press doi:10.1017/S1366728907003033

Phonological facilitation through translation in a bilingual picture-naming task*

AIMEE C. KNUPSKY
Allegheny College
PAUL C. AMRHEIN
Montclair State University

*We present a critical examination of phonological effects in a picture–word interference task. Using a methodology minimizing stimulus repetition, English/Spanish and Spanish/English bilinguals named pictures in either L1 or L2 (blocked contexts) or in both (mixed contexts) while ignoring word distractors in L1 or L2. Distractors were either phonologically related to the picture name (direct; FISH–*fist*), or related through translation to the picture name (TT; LEG–*milk–leche*), or they were unrelated (*bear–peach*). Results demonstrate robust activation of phonological representations by translation equivalents of word distractors. Although both direct and TT distractors facilitated naming, TT facilitation was more consistent in L2 naming and under mixed contexts.*

A primary question in bilingual language processing concerns the extent to which information from both languages is activated and whether or not this activation spreads across as well as within language systems. This complex question must be examined for both comprehension and production tasks, for each level of the language hierarchy (conceptual, lexical, and sublexical), and with the understanding that a variety of internal and external factors not present in monolingual language processing may inherently influence the answers we derive from our investigations. In comprehension research, a majority of experimental results suggests that languages of a bilingual are activated simultaneously (Beauvillain and Grainger, 1987; Grainger and Dijkstra, 1992; Dijkstra, Grainger and Van Heuven, 1999). Most of this evidence is based on the observation of cross-language influences on performance. For example, the orthographic and phonological information of words in one language can facilitate or inhibit performance with words in another language (Bijeljac-Babic, Biardeau and Grainger, 1997; Brysbaert, Van Dyck and Van de Poel, 1999). In production research, the conditions under which nonresponse languages may become activated are being examined with a focus on determining the extent to which such activation influences production of the response language.

Specifically, the question of whether bilingual production is selective at the conceptual, lexical, or sublexical levels remains unresolved, although Kroll, Bobb and Wodniecka (2006) argue that the most parsimonious account may consist of a generally nonselective system with a selective default under certain processing conditions. While a number of studies demonstrating cross-language semantic interference seem to support this account at a lexical level (Ehri and Ryan, 1980; Chen and Ho, 1986; Miller and Kroll, 2002), less is known about the activation and selection of response and nonresponse representations at the phonological level (Hermans, Bongaerts, De Bot and Schreuder, 1998; Costa, Miozzo and Caramazza, 1999; Costa, Colomé, Gómez and Sebastián-Gallés, 2003). Furthermore, convergence across these studies is often limited by the wide range in proficiency and acquisition histories of participants, differences in experimental contexts, and the fact that within these studies, L1 (first language) and L2 (second language) processing are often compared BETWEEN participants and ACROSS experiments. Conclusions drawn about bilingual production are particularly vulnerable to these factors and would be strengthened by a broader exploration of methodological techniques.

The goal of the present study was to use a within-participants design to manipulate factors that may affect bilingual production in a picture–word interference task. In particular, we were interested in examining the extent to which word distractors are capable of activating the representations of a bilingual's languages at the phonological level. The importance of this question derived from assumptions made in support of a language-specific selection mechanism at the lexical level in models of bilingual production (Costa et al., 1999; Costa et al., 2003; Costa, La Heij and Navarrete, 2006). The study was also designed to avoid the potentially problematic use of repeated stimulus materials in previous bilingual research. Specifically, this repetition method was avoided by having pictures named only once across the experiment.

The language specific selection hypothesis proposes that lexical activation takes place in parallel for both

* This research was conducted while the authors were at the University of New Mexico and was supported in part by a Research, Project and Travel Grant awarded by the Office of Graduate Studies, University of New Mexico.

Address for correspondence:
Aimee Knupsky, Department of Psychology, Allegheny College, 520 North Main Street, Meadville, PA 16335, USA
E-mail: *aknupsky@allegheny.edu*

languages of a bilingual, but that words from only ONE of the bilingual's lexicons compete for selection (Costa et al., 1999), in contrast to a nonspecific selection hypothesis in which words from BOTH of the bilingual's lexicons compete for selection (Green, 1998; Hermans et al., 1998). According to the language specific selection hypothesis, then, selection is specific at the lexical level in that competition among lemmas is limited to the response language. So, when speaking in her L1, words in a bilingual's L2 may receive activation from stimuli in the environment, but this activation would not be considered in response selection because it occurs in the nonselected lexicon. To test their theory, Costa et al. used a picture-naming task in which words of the bilingual's L1 (Catalan) or L2 (Spanish) were presented. The authors examined same (e.g., L1/L1) and cross (e.g., L2/L1) language identity, semantic, and phonological effects. Critically, significant cross-language identity facilitation was found which, the authors contend, would not occur if lemmas from the nonresponse language compete for selection. However, the results also showed that same-language identity effects were almost always LARGER than cross-language identity effects. If lemmas for both languages are initially activated to the same extent, and cross-language competition is excluded via a selection mechanism, then the source of the asymmetry remains to be identified.

Costa et al. suggested that the asymmetry reflects an underlying PHONOLOGICAL component in same-language identity effects not present in cross-language identity effects. As a result, a critical follow-up involved an examination of phonological effects in the task. Distractors were phonologically related to the picture name (BALDUFA-*baralla*) or phonologically related through translation to the picture name (BALDUFA-*pelea*-baralla). For through translation distractors to influence picture-naming, the TRANSLATION of the distractor would have to activate its phonological representations which could then feedback activation to the response lemma. While direct facilitation was found for both same and cross-language pairs, no through translation facilitation was identified. Consequently, the authors concluded that the asymmetrical identity effects were due to a prelexical phonological effect that occurs only in same-language pairs.

Recently, however, Hermans (2004) has suggested that cross-language identity effects may entail a phonological component given certain processing conditions. For example, in the Costa et al. (1999) study, although through translation word distractors served as directly related phonological word distractors in other trials, they were not the names of other pictures in the experiment. Hermans argues that a through translation word distractor that is also the name of a picture in the study would be more likely to activate the phonological representation of its translation equivalent. In such cases, significant through translation facilitation could potentially occur. The results of his study support this argument. In a picture-naming task with Dutch/English bilinguals, phonologically related through translation word distractors (MONKEY-*geld*-money) significantly decreased naming times in comparison to unrelated word distractors when they served as the names of other pictures in the study. Although Hermans points out that this result does not eliminate the possibility of language specific lexical selection, it does question the explanation provided to account for the asymmetry between same and cross-language identity effects. Furthermore, the identification of through translation phonological facilitation provides a new picture–word relationship researchers can use to examine the activation of response and nonresponse languages at the lexeme or sublexical levels. Such facilitation is consistent with, but distinct from, the phono-translation effect, which refers to an *increase* in response latency when a distractor is phonologically related to the translation of the picture name (Costa et al., 2003; Hermans et al., 1998). Most importantly, the result demonstrates the activation of phonological representations by translation equivalents – a finding that complements additional demonstrations of phonological processing in other bilingual production paradigms (Jared and Kroll, 2001; Roelofs, 2003; Sumiya and Healy, 2004).

The present study establishes the reliability of through translation phonological facilitation in a picture–word interference task and extends the generalizability of this effect across a broader range of production environments. These aims were motivated by a consideration of the variety of internal and external factors unique to bilingual production and also by a consideration of previously unexamined methodological issues that may have contributed to the null results reported by Costa et al. (1999). Because mediated effects (e.g., those derived through translation) are almost always smaller than direct effects, and more vulnerable to task constraints (Dell and O'Seaghdah, 1991), differences in methodological paradigms are not unimportant. Therefore, an examination of factors influencing phonological effects in bilingual production is theoretically important as well. Consequently, the word relation between picture names and word distractors was varied so that the distractors were phonologically related directly to the picture name (direct, e.g., FISH-*fist*), through translation to the picture name (TT; e.g., LEG-*milk*-leche) or unrelated (e.g., BEAR-*peach*). Although through translation effects were of most significance, direct effects served as important comparisons. Specifically, we hypothesized that direct effects would be more robust, and that through translation facilitation would be more sensitive to the manipulations of other variables in the study designed to influence the production process.

The first of these manipulations involved response language. Participants were asked to name pictures in both L1 and L2. Given the immediate nature of the direct effect, we hypothesized that it would not vary across response language. In contrast, we hypothesized that the generation of through translation effects would be more likely when participants named pictures in their L2. Essentially, the idea was that the extended time-course often seen in L2 versus L1 production provides greater opportunity for facilitation to develop through a number of mechanisms, both from within the system, but also, potentially, through strategic adjustments on the part of the bilingual. For example, Hermans (2000) proposed that L2 picture-naming may allow for greater activation of lexical candidates via feedback from the phonological level and, in fact, that unbalanced bilinguals may rely on this feedback of activation more during L2 than L1 production. Similarly, Kroll et al. (2006) have suggested that production in L2 may benefit more from priming due to the fact that the L2 is often the less active (and less frequently used) language in comparison to the L1. In translation tasks, performance in this less frequently used language would be more vulnerable to competition during lexical access (p. 128). This difference may be further modulated by degree of proficiency (Kroll, Michael, Tokowicz and Dufour, 2002). Also, Costa et al. (2003) proposed that language specific selection mechanisms may not be fully functional in less proficient bilinguals, and may be the underlying cause for discrepancies seen in past research. Having participants name pictures in both languages in the present study was also important, then, to allow differences in priming across L1 and L2 naming to be compared WITHIN participants. In past studies, differences in naming across L1 and L2 have almost always been compared ACROSS experiments and, consequently, BETWEEN participants. For example, while Costa et al. (1999) report no evidence for through translation phonological facilitation in L1 picture-naming, Hermans (2004) identified through translation facilitation in L2 picture-naming. Although Hermans suggested that the difference may be due to issues of stimulus composition, response language may also be an important contributing factor. The present study examined this possibility.

In addition to varying response language, we also manipulated the production context under which bilinguals named pictures. Specifically, participants named pictures in just one language (blocked context) or in both languages (mixed context). We hypothesized that, in comparison to the blocked context, the mixed context creates a processing environment in which the degree of activation of the nonresponse language is increased, consequently increasing the likelihood of finding through translation effects. However, differences across blocked and mixed conditions could also reflect a STRATEGIC modification of production influenced by the processing constraints under which the bilingual is operating. Amrhein & Sanchez (1997) provide some consideration of performance changes across blocked versus mixed contexts. In their study, bilinguals either drew pictures or wrote words from either Spanish or English word or picture stimuli. The authors found that performance for the bilinguals improved when the production task was known in advance and improved further if the language to be used was known in advance. Very few studies have directly compared bilingual picture-naming under different production contexts. However, Costa et al. (1999) reported no change in performance across blocked versus mixed naming conditions with identically or semantically related distractors (Experiment 2). These results suggest that although through translation facilitation may be generated in mixed contexts only, the direct phonological effect may occur consistently across both blocked and mixed conditions.

Finally, the present study also differed from past research by our decision to control the repetition of concepts used for picture names and word distractors. Specifically, no picture or word distractor was used more than once; no word distractor (or picture name) was later used as the basis for a picture (or word distractor), and the through translation word distractors were not developed by utilizing translation equivalents of distractors from the direct condition. Two important consequences derived from these constraints. First, the breadth of stimuli used for picture-naming was much larger than in past studies, extending the ecological validity of our results and adding depth to our examination of the production process. Second, practice effects on picture-naming were avoided. The potential for practice effects derives from the common methodological practice of having participants name a picture four to six times across the course of a single study (not including practice during training phases). Minimizing these practice effects is important, however, given that the repetition of stimuli may artificially decrease picture-naming times overall (Kroll & De Groot, 1997). This artificial reduction would make it especially difficult to identify any effects through translation. Furthermore, as Kroll et al. (2006) suggest, using results derived from small naming sets may have implications for models of production. Specifically, lexical selection may be constrained by the use of these repeated sets, giving us only a limited understanding of bilingual production, especially in terms of language specific versus nonspecific processes.

In summary, the present study focused on factors that may influence the generation and magnitude of phonological effects (both direct and through translation) in bilingual picture-naming. First, response language varied so that measures of both L1 and L2 naming were collected. Second, production context varied so that naming under conditions of both processing certainty (blocked context) and uncertainty (mixed context) were compared. Third, stimulus language also varied so that both L1 and L2 word distractors were processed. Because a purely within-participants design was utilized, differences among the bilinguals tested did not contribute to differences across conditions. Furthermore, because

Table 1. *Means and standard deviations for L2 measures of self-rated speaking, spoken comprehension, reading, and writing skills (seven-point Likert scale), of a category-naming task (exemplars per 30 sec), and of a reading task (words aloud per min) for the L1English and L1Spanish groups.*

	Speaking	Spoken Comp	Reading	Writing	Cat Naming	Reading Aloud
L1English	5.1 (1.2)	5.8 (0.9)	5.4 (1.0)	4.8 (1.2)	4.0 (2.1)	102.8 (21.9)
L1Spanish	5.6 (0.8)	6.1 (0.6)	5.7 (1.1)	5.3 (1.0)	5.5 (3.0)	141.7 (21.8)

the present study tested both English/Spanish and Spanish/English bilinguals, potential differences due to dominant language could be identified as well. We hypothesized that direct phonological effects would consistently occur across conditions of the study while through translation effects would be more sensitive to experimental manipulations. Specifically, we expected to see through translation more consistently in L2 picture-naming and under mixed production contexts.

Method

Participants

Thirty-five bilinguals from the student community at the University of New Mexico received either course credit or monetary compensation for their participation. Participants were grouped by L1 on the basis of responses to a language history questionnaire. Twenty-two participants were identified as L1English bilinguals while 13 were identified as L1Spanish bilinguals. Both groups had an average age of 22 years (ranging from 18 to 36). Table 1 presents measures of L2 fluency derived from the two groups. Participants rated themselves on measures of L2 speaking, spoken comprehension, reading, and writing (using a seven point Likert scale). Participants also completed two objective language-use tasks including a category-naming task and a reading task (see Soares & Grosjean, 1984 for other examples of these measures). For the category-naming task, participants were given 30 seconds to produce as many exemplars as possible for English and Spanish categories. For the reading task, participants were asked to read aloud passages from English and Spanish editions of *National Geographic*. The number of words read per minute was recorded – incorrectly pronounced words were not counted. Planned mean comparisons of the self-rated fluency measurements and the category naming task showed no significant differences between groups (ps > 0.05). However, a planned mean comparison of the reading task did show a significant difference with the L1Spanish group reading more L2 words aloud per minute than the L1English group, $F(1,33) = 25.77$, $MSE = 479.9$, $p < 0.001$. Furthermore, on the language history questionnaire, the L1English group reported an average 9 years of L2 use at home, work or both while the L1Spanish group reported an average 14 years of L2 use at home, work, school, or all three. Experimental analyses are reported separately for each group.

Materials

Target stimuli consisted of 192 pictures with noncognate names. All pictures were black and white line drawings with a horizontal visual angle between 2° and 8° and were selected either from Snodgrass & Vanderwart (1980) or from a clipart website (ArtToday.com) of which the first author was a member. Each target picture was paired with one visually presented word distractor that was either phonologically related (direct condition; FISH-*fist*), phonologically related through translation (through translation condition; LEG-*milk*-leche), or unrelated (BEAR-*peach*). Phonological relatedness was defined as the sharing of at least the first two phonemes between picture name and distractor. In all cases, picture–word pairs were semantically unrelated. Distractors (capitalized in bold, 27 point Helvetica) were superimposed with an opaque background onto the center of the pictures so that the text was clearly visible, with minimum disruption to picture details. The majority (91%) of the distractors were nouns (80% concrete nouns and 11% abstract nouns). The remaining distractors (9%) were adjectives. Finally, less than 4% of the distractors utilized special orthographic markings (e.g., cinturón).

Figure 1 provides examples of the picture–word pairs and displays the visual relationship between pictures and distractors.[1] The figure also demonstrates that for each of the three word relations, four response language and stimulus language relationships were possible – English picture name with English distractor (EE); English picture name with Spanish distractor (ES); Spanish picture name with English distractor (SE); and Spanish picture name with Spanish distractor (SS). Each of the resulting 12 conditions had 16 unique picture–word pairs which were then RANDOMLY sampled for each participant into two production contexts. Consequently, for any given participant, eight of the picture–word pairs from a condition would be presented in a blocked context and the

[1] A complete list of picture–word pairs is available upon request from the first author.

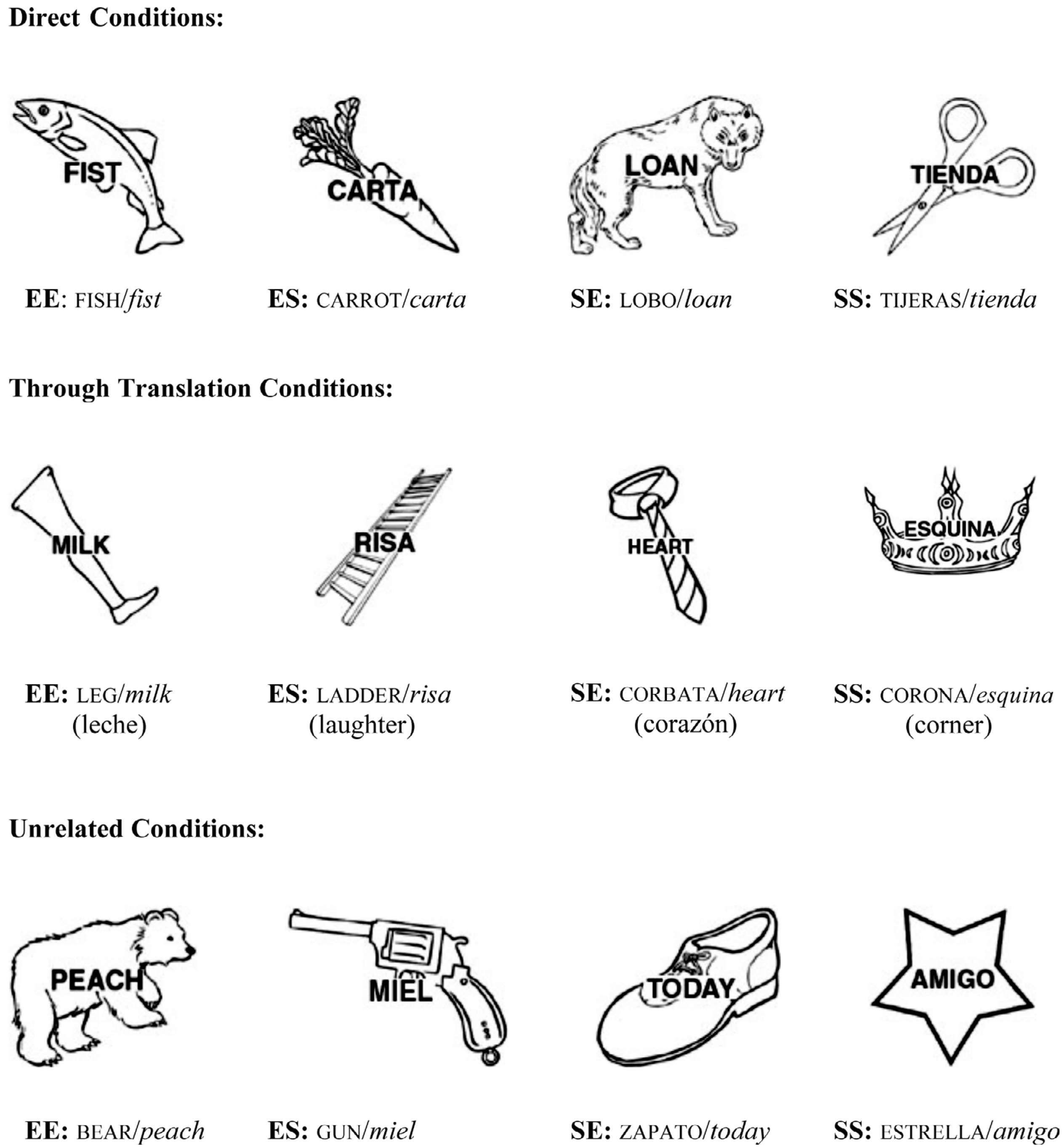

Figure 1. Example stimuli from each of the 12 unique picture–word pair conditions.

remaining eight picture–word pairs would be presented in a mixed context. Therefore, while all participants saw the same set of picture–word pairs, the production context for the picture–word pairs varied across participants.

Because each of the picture–word conditions consisted of unique picture–word pairs, it was necessary to guard against any unanticipated influence of stimulus selection on the results. For picture targets, we examined simple picture-naming reaction times (obtained at the end of the second session), age of acquisition (AoA), printed frequency rate, picture name length, and number of syllables. For distractors, we examined printed frequency rate, length, and number of syllables. AoA scores for the English-named pictures were derived from Snodgrass & Yuditsky (1996) and Carroll & White (1973). The AoA scores for the Spanish-named pictures were derived from Cuetos, Ellis & Alvarez (1999). Using these sources, AoA scores for 61% of the picture names in the present study were identified. Frequencies for the English stimuli were taken from Kučera & Francis (1967). Frequencies for the Spanish stimuli were taken from Sebastián-Gallés (2000). In order to make meaningful comparisons across sources of different corpus size, frequency rates were generated by dividing frequency counts by their respective corpus sizes and then presenting them as counts per 100,000. Each stimulus variable was analyzed using a $3 \times 2 \times 2$ ANOVA with the between-items factors of word relation, response language, and stimulus language. A significant effect of word relation for distractor frequency rate was found, $F(2,180) = 4.67$, $MSE = 301.16$, $p < 0.05$, indicating higher frequency rates in the through translation condition (13.76) than in the direct (5.43) or unrelated (5.88). However, this difference was driven mainly by the inadvertent inclusion of two items (1% of the 192 distractor word ensemble) with higher relative frequency (*back* and *hombre*) in two of the four through translation conditions. When these two items were removed (and replaced by new values that were maximum-likelihood estimated from the remaining items of their respective conditions), the effect for word relation was no longer present (11.48 for through translation conditions), $F(2,180) = 2.99$, $MSE = 246.9$, $p > 0.05$. Furthermore, with all items retained, experimental analyses conducted with distractor frequency rate as a covariate generated the same pattern of findings as those reported in the results section. Consequently, all items were retained in the analyses presented in this report. For all other stimulus variables examined, no significant difference was identified across the 12 unique picture–word conditions that could serve as an alternate explanation for the experimental results. Finally, for picture names, no significant differences were found for initial phoneme manner or voicing.[2]

Procedure

Two separate sessions were conducted during which participants were tested individually. In the first session, participants were asked to complete the L2 fluency measures and one portion of the picture-naming experiment (blocked or mixed). The order of blocked versus mixed production context was counterbalanced across participants. In the second session, participants completed the remaining portion of the picture-naming experiment and then completed a series of post-tests. Overall, the first session took an hour to complete and the second session an hour and a half. All testing took place in a soundproof, dimly lit room with participants seated 100 cm from a computer screen, with a microphone directly in front of them for timing of picture-name responses. Responses were recorded and assessed for accuracy. For the picture-naming task, participants were told that a picture would appear on the computer screen with a word written in front of it. They were instructed to ignore the word and to name the picture as quickly and accurately as possible.

In the blocked condition, participants completed two sets of trials (one for responses in English and one for responses in Spanish). The order of these sets was counterbalanced across participants. For each participant, picture–word pairs from the appropriate conditions were randomly selected by the computer for each blocked set (i.e., from the EE and ES conditions for the English-blocked set; from the SE and SS conditions for the Spanish-blocked set). Before beginning the experimental trials for a given set, participants were familiarized with the selected picture–word pairs. The pictures were presented on the computer screen with appropriate response language names (e.g., a picture of a BEAR would be presented with the word *bear* for the English-blocked set, or with the word *oso* for the Spanish-blocked set). Participants were instructed NOT to name the pictures, but to simply notice the names of the pictures to be used in the experiment. After the familiarization period, participants were reminded of the response language for the current block of experimental trials. A complete trial began with a language prompt presented for 1500 ms ("Speak English" or "Habla Español"), followed by a fixation stimulus for 1000 ms, followed by a picture–word pair that remained on the screen until response. The inter-trial interval was 3000 ms. Naming latencies were measured from the onset of the picture–word pair and were recorded using a voice-activated timing mechanism. Each blocked set consisted of 48 experimental trials preceded by 14 practice trials.

In the mixed condition, the basic procedure was the same except participants were told they would be using both languages to name pictures. Specifically, they were informed to pay close attention to the language prompt which would indicate the language they should use to name each picture. In accordance with the mixed nature of the production context, the language prompt could appear in either language (e.g., "Speak English" or "Habla Ingles"; "Speak Spanish" or "Habla Español"). During familiarization, picture names in both languages were presented so the language of response for experimental trials would not be anticipated. The remaining eight picture–word pairs from each condition (EE, ES, SE, and SS) were used for a total of 96 experimental trials, preceded by 28 practice trials. At the end of the second session, participants were asked to complete a translation test of the distractors. Responses were used to ensure participants were able to correctly translate the through translation distractors.

Results

True production errors comprised 16.1% of the data and included trials on which participants gave an incorrect response (e.g., *don't know*, 12.1%), used correct responses

[2] Full picture–word pair analyses are available upon request from the first author.

that differed from those designated for the study [e.g., *cabello* "hair" for *pelo* "hair", 1.4%], used the wrong language (e.g., *casa* for *house*, 2.4%), read the distractor (e.g., *tienda* for *tijeras*, <1%), or gave the translation of the distractor (e.g., *triste*, the translation of *sad*, for *tree*, <1%). In addition, trials on which equipment failures occurred (e.g., a premature or delayed activation of the voice key) comprised an additional 1.8% of the remaining trials. Finally, trials for which participants did not know the correct translation of the distractor were also removed (4.9%). Using a criterion of two and a half standard deviations, no outliers were removed. Error analyses presented below were conducted for true production errors only.

Naming latencies and error data from each group (L1English and L1Spanish) were subjected to a series of $2 \times 2 \times 2 \times 3$ ANOVAs with production context (blocked, mixed) as a within-participants, within-items variable and response language (L1, L2), stimulus language (L1, L2) and word relation (direct, TT, unrelated) as within-participants, between-items variables. Results for the L1English group are presented first, followed by the results for the L1Spanish group. Only those results critical to the focus of the paper are presented.

Inspection of the RT and error data revealed no speed/accuracy trade-off. In most cases, the results across the two groups were quite similar. Specifically, across both L1English and L1Spanish participants, ERROR ANALYSES demonstrate significant main effects of production context and response language with more errors in mixed than in blocked conditions and more errors when responding in L2 than in L1. No significant differences in error patterns were seen across stimulus language. Finally, error patterns across word relation conditions were not consistent (these patterns are presented in the analyses). Analyses of the NAMING LATENCIES also show similar patterns across L1English and L1Spanish participants. Both groups showed main effects of production context (faster naming in blocked than mixed conditions), of response language (faster naming with L1 than L2), and of word relation (demonstrating direct and through translation facilitation). Most importantly, however, an examination of the WORD RELATION BY RESPONSE LANGUAGE interaction shows that through translation effects were more consistent when responding in L2 than in L1. Furthermore, an examination of the WORD RELATION BY PRODUCTION CONTEXT interaction shows that through translation effects occurred in mixed, but not in blocked contexts.

L1English group

The error analysis revealed a significant main effect of production context with more average errors per condition in the mixed (1.56) than in the blocked (1.11) context, $F_1(1,21) = 13.76$, $MSE = 2.01$, $p < 0.01$; $F_2(1,180) = 17.18$, $MSE = 2.29$, $p < 0.001$. The main effect of response language was also significant with more errors when naming in L2 (2.14) than in L1 (0.53), $F_1(1,21) = 99.26$, $MSE = 3.479$, $p < 0.001$; $F_2(1,180) = 53.72$, $MSE = 8.18$, $p < 0.001$. The main effect of stimulus language was NOT significant, with similar errors in L1 distractor (1.33) and L2 distractor (1.35) conditions, $Fs < 1$. The main effect of word relation was significant in the participant analysis only, $F_1(2,42) = 11.52$, $MSE = 1.152$, $p < 0.001$; $F_2(2,180) = 1.77$, $MSE = 8.18$, $p > 0.1$. Overall, fewer errors tended to occur in direct (1.02), than in through translation (1.52), and unrelated (1.48) conditions. However, the effect of word relation was qualified by an interaction with response language, $F_1(1,21) = 40.41$, $MSE = 0.89$, $p < 0.001$; $F_2(2,180) = 4.99$, $MSE = 8.18$, $p < 0.01$. Planned mean comparisons indicated that when responding in L1, fewer errors occurred in the unrelated (0.25) than in the direct (0.69), $F_1(1,21) = 18.18$, $MSE = 0.12$, $p < 0.001$; $F_2(1, 62) = 5.22$, $MSE = 0.77$, $p < 0.05$ or through translation (0.65) conditions, $F_1(1,21) = 18.47$, $MSE = 0.09$, $p < 0.001$; but $F_2 < 2$, $p > 0.1$. The direct and through translation conditions were not significantly different, Fs < 1, ps > 0.1. In contrast, when responding in L2, fewer errors occurred in the direct (1.35) than in the through translation (2.39), $F_1(1,21) = 30.81$, $MSE = 0.38$, $p < 0.001$; $F_2(1, 62) = 3.83$, $MSE = 7.02$, $p = 0.055$ or unrelated (2.71) conditions, $F_1(1,21) = 57.44$, $MSE = 0.35$, $p < 0.001$; $F_2(1,62) = 6.41$, $MSE = 7.37$, $p < 0.05$. The through translation and unrelated conditions were not significantly different, $Fs < 3$, ps > 0.1. However, an overall absence of a speed-accuracy trade-off suggests that this interaction does not compromise the examination of naming latencies.

The analyses of naming latencies revealed a significant main effect of production context, $F_1(1,21) = 20.47$, $MSE = 534571.06$, $p < 0.001$; $F_2(1,180) = 41.05$, $MSE = 187770.45$, $p < 0.001$. Overall, picture-naming was faster in the blocked (1389 ms) than in the mixed context (1677 ms). The main effect of response language was also significant, $F_1(1,21) = 30.60$, $MSE = 597407.16$, $p < 0.001$; $F_2(1,180) = 35.30$, $MSE = 417977.04$, $p < 0.001$. Overall, picture-naming was faster in L1 (1346 ms) than in L2 (1719 ms). The main effect of stimulus language was NOT significant, $F_1(1,21) = 1.16$, $MSE = 181511.05$, $p > 0.2$; $F_2 < 1$. Picture-naming was similar with L1 (1513 ms) and L2 (1553 ms) distractors. The main effect of word relation was significant, $F_1(2,42) = 22.34$, $MSE = 479689.37$, $p < 0.001$; $F_2(2,180) = 12.08$, $MSE = 417977.04$, $p < 0.001$. Planned mean comparisons revealed that picture-naming in the direct condition (1280 ms) was significantly faster than in the unrelated condition (1773 ms), $F_1(1,21) = 39.76$, $MSE = 538086.23$, $p < 0.001$; $F_2(1,126) = 18.66$, $MSE = 270083.89$, $p < 0.001$. Picture-naming in the direct condition was also faster than in the TT condition (1546 ms), $F_1(1,21) = 24.63$, $MSE = 31600.82$, $p < 0.001$; $F_2(1,126) = 8.80$, $MSE = 165176.29$, $p < 0.01$. Picture-naming in the TT condition tended to be significantly

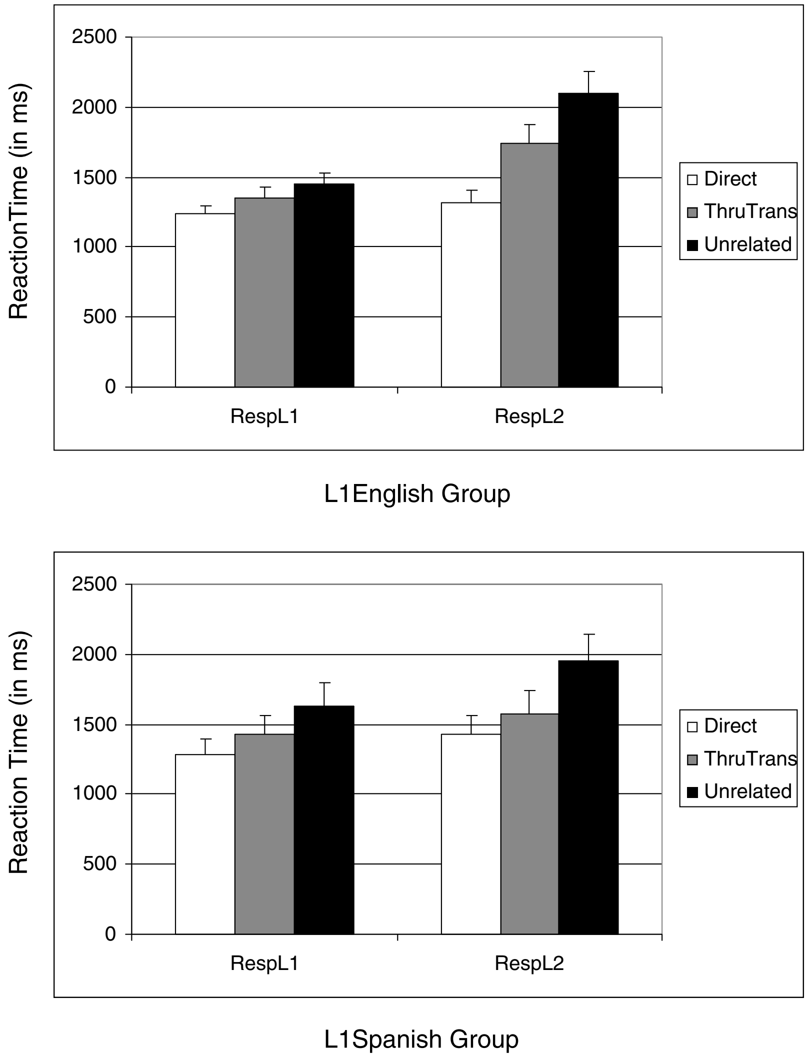

Figure 2. Mean naming onset latencies (in ms) as a function of response language and word relation for the L1English group (top panel) and the L1Spanish group (bottom panel).

faster than in the unrelated condition, $F_1(1,21) = 6.99$, $MSE = 648175.30$, $p < 0.02$; but $F_2(1,126) = 3.19$, $MSE = 338009.37$, $p < 0.08$.

The main effect of word relation was qualified by an interaction with response language, $F_1(2,42) = 8.74$, $MSE = 400088.92$, $p < 0.01$; $F_2(2,180) = 5.97$, $MSE = 417977.04$, $p < 0.01$. The top panel of Figure 2 presents mean naming latencies (with standard error bars) for word relation across response language. Planned mean comparisons revealed that when responding in L1, the difference between the direct and unrelated conditions (211 ms) tended to be significant, $F_1(1,21) = 17.91$, $MSE = 27552.70$, $p < 0.001$; but $F_2(1,62) = 3.58$, $MSE = 61740.52$, $p = 0.06$. However, when responding in L2, this difference (774 ms) was significant in both participant and items analyses, $F_1(1,21) = 37.05$, $MSE = 177982.50$, $p < 0.001$; $F_2(1,62) = 20.02$, $MSE = 365399.39$, $p < 0.001$. Similarly, when responding in L1, the difference between the direct and TT conditions (108 ms) was significant in the participant analysis only, $F_1(1,21) = 6.37$, $MSE = 20120.60$, $p = 0.02$; $F_2(1,62) = 1.29$, $MSE = 82165.04$, $p = 0.26$. In contrast, when responding in L2, this difference (424 ms) was significant in both participant and items analyses, $F_1(1,21) = 17.57$, $MSE = 112512.35$, $p < 0.001$; $F_2(1,62) = 8.97$, $MSE = 212340.76$, $p < 0.01$. Also, when responding in L1, the difference between the TT and unrelated conditions (104 ms) was significant in the participant analysis only, $F_1(1,21) = 4.43$, $MSE = 26728.57$, $p < 0.05$; $F_2 < 1$. However, when responding in L2, this difference (350 ms), was significant in both participant and items analyses, $F_1(1,21) = 4.57$, $MSE = 294936.98$, $p < 0.05$; $F_2(1,62) = 3.79$, $MSE = 462604.97$, $p = 0.05$.

Finally, the interaction between word relation and context was NOT significant, $F_1 < 3$, $p = 0.11$; $F_2 < 2$, $p = 0.29$. The top panel of Figure 3 presents

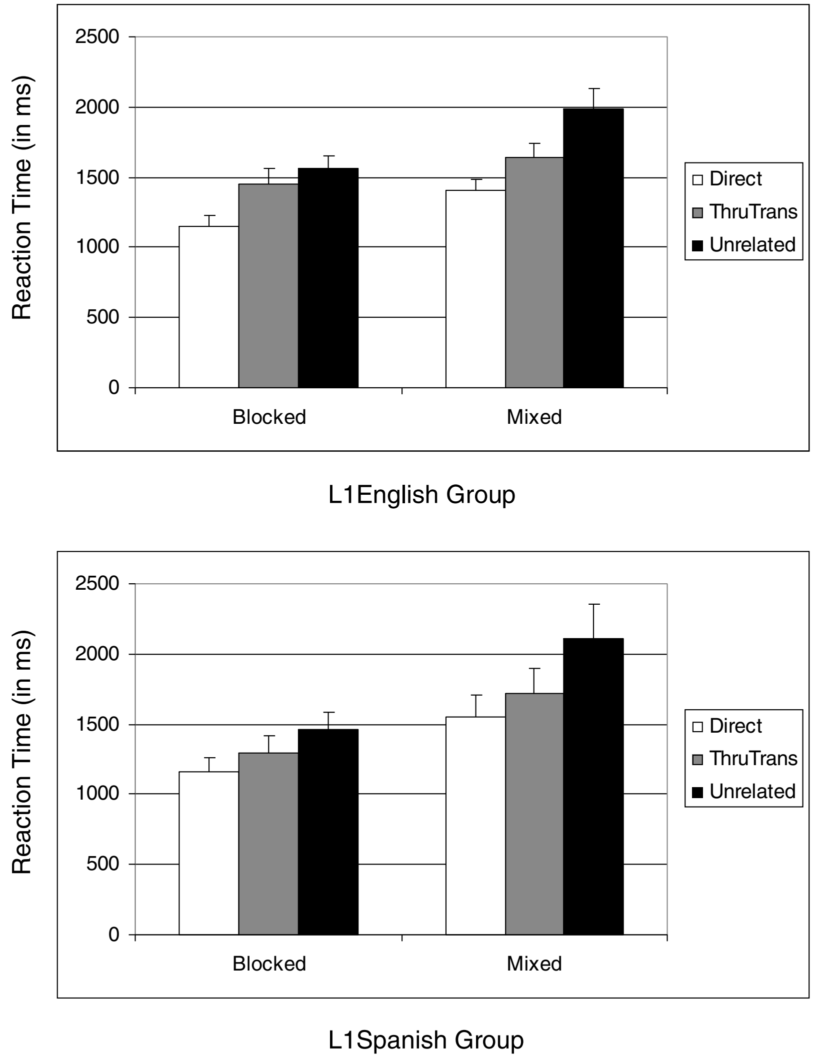

Figure 3. Mean naming onset latencies (in ms) as a function of production context and word relation for the L1English group (top panel) and the L1Spanish group (bottom panel).

mean naming latencies (with standard error bars) for word relation across production context. However, planned mean comparisons revealed that while the difference in naming times was significant between the direct and unrelated condition in both the blocked (406 ms), $F_1(1,21) = 30.90$, $MSE = 58622.02$, $p < 0.001$; $F_2(1,126) = 15.66$, $MSE = 251078.83$, $p < 0.001$, and mixed (580 ms) contexts, $F_1(1,21) = 24.13$, $MSE = 153478.59$, $p < 0.001$; $F_2(1,126) = 13.10$, $MSE = 480133.04$, $p < 0.001$, the difference between the TT and unrelated conditions was significant in the mixed (347 ms), $F_1(1,21) = 7.91$, $MSE = 167232.17$, $p = 0.01$; $F_2(1,126) = 4.21$, $MSE = 547451.47$, $p < 0.05$, but not the blocked (107 ms) context, Fs < 2, ps > 0.2. In addition, picture-naming was significantly faster in the direct than in the TT conditions in both the blocked (298 ms difference), $F_1(1,21) = 14.72$, $MSE = 66492.05$, $p < 0.01$; $F_2(1,126) = 7.84$, $MSE = 257960.15$, $p < 0.01$, and in the mixed context (233 ms difference) $F_1(1,21) = 14.24$, $MSE = 42158.89$, $p < 0.01$; $F_2(1, 126) = 4.54$, $MSE = 215892.30$, $p < 0.05$.

L1Spanish group

The error analysis revealed a significant main effect of production context with more errors in the mixed (1.48) than in the blocked (0.92) condition, $F_1(1,12) = 13.51$, $MSE = 1.75$, $p < 0.01$; $F_2(1,180) = 20.44$, $MSE = 0.92$, $p < 0.001$. The main effect of response language was significant in the participant analysis, but only a trend in the items analysis with a tendency for more errors in the L2 (1.40) than in the L1 (1.00) naming conditions, $F_1(1,12) = 7.60$, $MSE = 1.61$, $p < 0.05$; $F_2(1,180) = 2.89$, $MSE = 2.92$, $p = 0.09$. The main effect of stimulus language was not significant, with errors similar in L1 distractor (1.24) and L2 distractor (1.17) conditions,

Fs < 1. The main effect of word relation was significant in the participant analysis only, $F_1(2,24) = 5.33$, $MSE = 0.43$, $p < 0.05$; $F_2 < 1$, $p > 0.5$. Overall, fewer errors tended to occur in unrelated (1.06), than in direct (1.35), and through translation (1.20) conditions. Given that this main effect was not consistent across the two participant groups (despite being comprised of the same stimulus materials), we think this reflects an unsystematic pattern that is not problematic to an interpretation of the naming latencies.

The analyses of naming latencies revealed a significant main effect of context, $F_1(1,12) = 15.57$, $MSE = 1187880.32$, $p < 0.01$; $F_2(1,180) = 74.22$, $MSE = 322311.73$, $p < 0.001$. Overall, picture-naming was faster in the blocked (1306 ms) than in the mixed (1793 ms) context. The main effect of response language was also significant with picture-naming faster in L1 (1447 ms) than in L2 (1652 ms), $F_1(1,12) = 12.45$, $MSE = 262860.62$, $p < 0.01$; $F_2(1,180) = 7.93$, $MSE = 379797.75$, $p < 0.01$. The main effect of stimulus language was *not* significant, Fs < 1. Picture-naming was similar with L1 (1564 ms) and L2 (1537 ms) distractors. The main effect of word relation was significant, $F_1(2,24) = 20.74$, $MSE = 244262.66$, $p < 0.001$; $F_2(2,180) = 13.43$, $MSE = 379797.75$, $p < 0.01$. Planned mean comparisons revealed that picture-naming in the direct condition (1356 ms) was significantly faster than in the unrelated condition (1790 ms), $F_1(1,12) = 29.59$, $MSE = 41387.22$, $p < 0.001$; $F_2(1,126) = 19.35$, $MSE = 241588.50$, $p < 0.001$. Picture-naming in the direct condition was faster than in the TT condition (1503 ms) in the participant analysis only, $F_1(1,12) = 5.87$, $MSE = 24105.81$, $p < 0.05$; $F_2(1,126) = 2.25$, $MSE = 118163.39$, $p > 0.1$. Picture-naming in the TT condition was also significantly faster than in the unrelated condition, $F_1(1,12) = 20.44$, $MSE = 26105.45$, $p < 0.01$; $F_2(1,126) = 11.53$, $MSE = 235157$, $p < 0.01$.

The interaction between word relation and response language was *not* significant, $F_1 < 2, p > 0.3$; $F_2(2,180) = 2.45$, $p = 0.09$. The bottom panel of Figure 2 above presents mean naming latencies (with standard error bars) for word relation across response language. Planned mean comparisons revealed that the difference between the direct and unrelated conditions was significant both when responding in L1 (352 ms), $F_1(1,12) = 29.14$, $MSE = 27587.21$, $p < 0.001$; $F_2(1,62) = 6.11$, $MSE = 174753.95$, $p < 0.05$ and when responding in L2 (516 ms), $F_1(1,12) = 18.72$, $MSE = 92608.94$, $p < 0.01$; $F_2(1,62) = 14.77$, $MSE = 277302.70$, $p < 0.001$. When responding in L1, the difference between the direct and TT conditions (147 ms) was significant in the participant analysis only, $F_1(1,12) = 6.59$, $MSE = 21407.84$, $p < 0.05$; $F_2(1,126) = 2.37$, $MSE = 115983.93$, $p > 0.1$. When responding in L2, the difference between direct and TT conditions was not significant (148 ms), $F_1 < 2$, $p > 0.12$; $F_2 < 1$, $p > 0.5$. When responding in L1, the difference between the TT and unrelated conditions (204 ms) was significant in the participant analysis only, $F_1(1,12) = 8.02$, $MSE = 33846.13$, $p < 0.05$; $F_2(1,62) = 1.52$, $p > 0.2$. However, when responding in L2, the difference between the TT and unrelated conditions (369 ms) was significant in both the participant and items analyses, $F_1(1,12) = 9.73$, $MSE = 90758.34$, $p < 0.01$; $F_2(1,62) = 12.17$, $MSE = 271994.66$, $p < 0.01$.

Finally, the interaction between word relation and context was *not* significant, $F_1 < 2, p > 0.1$; $F_2 < 1, p > 0.3$. The bottom panel of Figure 3 above presents mean naming latencies RT (with standard error bars) for word relation across production context. However, planned mean comparisons revealed that while the difference in naming times was significant between the direct and unrelated condition in both the blocked (305 ms), $F_1(1,12) = 25.89$, $MSE = 23493.50$, $p < 0.001$; $F_2(1,126) = 12.27$, $MSE = 214977.96$, $p < 0.01$ and mixed (562 ms) contexts, $F_1(1,12) = 16.29$, $MSE = 126060.06$, $p < 0.01$; $F_2(1,126) = 11.31$, $MSE = 644231.17$, $p < 0.01$, the difference between the TT and unrelated conditions was significant in the mixed (398 ms), $F_1(1,12) = 16.12$, $MSE = 63938.98$, $p < 0.01$; $F_2(1,126) = 6.87$, $MSE = 599326.82$, $p = 0.01$, but was only marginal in the blocked (174 ms) context, $F_1(1,12) = 4.2$, $MSE = 47287.84$, $p = 0.063$; $F_2(1,126) = 6.57$, $MSE = 242819.62$, $p < 0.05$. Picture-naming was not significantly faster in the direct than in the TT condition in the blocked (131 ms), Fs < 2, ps > 0.3, nor in the mixed context (164 ms) $F_1(1,12) = 3.48$, $MSE = 50124.79$, $p = 0.086$; $F_2 < 2, p > 0.2$.

Discussion

The results of the present study demonstrate both direct and through translation phonological facilitation in a picture–word interference task. As predicted, the direct effects were almost always greater in magnitude than the through translation effects and were less vulnerable to task constraints. Specifically, direct phonological facilitation was identified in both L1 and L2 naming, under both blocked and mixed production contexts. In contrast, the generation of through translation facilitation was more consistent in L2 naming than in L1 naming, and occurred in mixed, but not blocked production contexts. The results are compelling given that the same general patterns were seen with both L1English and L1Spanish bilinguals. Furthermore, these priming effects were identified using a method that was designed to avoid repetition of picture-naming and to increase the scope of concepts utilized in stimulus generation.

The primary aim of this study was to explore the conditions under which the activation of phonological representations by distractors could influence picture-naming.

In doing so, we have identified two factors that determine when phonological effects through translation occur. First, while Costa et al. (1999) originally failed to find such effects with L1 picture-naming, Hermans (2004) demonstrated through translation facilitation with L2 picture-naming. The outcome of the present study clarifies these contrasting results, showing that through translation facilitation is sensitive to the language of response and is more likely to occur in L2 production. Perhaps this difference is due to the extended time-course of production in L2, especially for less proficient bilinguals. As recently suggested, this extended time-course may provide greater opportunity for feedback from the phonological level (Hermans, 2000). Specifically, a delay of the lexical selection process would provide a window of opportunity for the feedback of activation from translation equivalents. In such cases, the generation of translation-mediated phonological effects would be more likely.

Second, production context appears to play a role in the development of through translation effects. Specifically, uncertainty regarding the response language increased the generation of these effects. There are two potential explanations for this difference. First, in keeping with the discussion of the influence of response language on the generation of through translation effects, mixed contexts may simply reflect a processing condition supportive of the time-course needed to activate translation equivalents. The slower production process under mixed contexts would provide a greater opportunity for priming by delaying the lexical selection process and allowing for the consideration of feedback from translation equivalents. Second, greater through translation facilitation in mixed contexts could reflect a more STRATEGIC modification of the production process. When production environments become more uncertain, bilinguals may begin to process language input (i.e., distractors) more deeply. Deeper processing may support stronger activation of translation equivalents, allowing for the emergence of through translation effects. Future examinations of bilingual processing under varying contexts will help determine the most effective description of their influence on the production process.

Researchers should also consider other methodological manipulations that may impact the generation of through translation effects. First, an important consideration in this research is the degree to which the design of the study is conducive to translation. If there is no potential for translation to occur, then through translation facilitation effects should not be expected. For example, Costa et al. (1999) presented picture–word pairs for brief, 400 ms durations, a constraint that limited encoding time, decreasing the likelihood that participants were able to process the distractors deeply enough for translation to occur. A related concern is whether participants CAN translate the distractors used in these studies. The present study is unique in that participants completed a translation test to ensure they were able to do so. Second, stimulus composition may influence the generation of through translation effects. In fact, this factor helps explain why Hermans (2004) identified through translation facilitation in a blocked production context. He argued that the generation of this facilitation was sensitive to the fact that distractors served as the names of other pictures in the study. Similarly, Costa et al. (2003) argued that the ratio of related pairs influenced the identification of the phono-translation effect. Third, the degree of phonological overlap between picture–word pairs in these studies could also be an important consideration. For example, Costa et al. (2006) suggested that in a phoneme-monitoring task (such as the one conducted by Colomé, 2001), a SINGLE activated target phoneme may be able to feed activation back to the lexical level, having the potential to actually INITIATE activation of a nonresponse language (p. 142). While most studies utilizing phonological similarity describe the degree of overlap in general terms (e.g., the first two or three phonemes), more specific evaluations may be necessary. For example, an examination of the picture–word pairs in the present study (all noncognates) revealed an average overlap of 2.80 phonemes (with a median of 3 phonemes).

Two potential limitations of the present study should be considered. First, overall naming times were slower than those in other bilingual naming studies. However, longer naming times were not surprising given that 1) pictures were only named once, 2) participants did not receive practice NAMING the pictures during familiarization, and 3) the range of stimuli used in the present study was greater than in previous investigations. In addition, it should be noted that the same general pattern of results were identified when analyses were conducted with medians rather than averages. Second, stimulus selection influences on the results were a concern. However, the results of extensive examinations of stimulus attributes argue against any real impact of stimulus selection on the experimental outcomes. We also emphasize that the same general pattern of results was found regardless of dominant language. Recall that the same picture–word pairs were utilized across the two participant populations. So, for example, picture–word pairs with English responses and English distractors (e.g., LEG-*milk*-leche) served as L1/L1 stimuli for L1English participants and as L2/L2 stimuli for the L1Spanish participants. That through translation facilitation was more consistent for both groups when naming in L2 (despite the stimuli used) argues against a stimulus selection interpretation of the results. In addition, production context was a within-items variable meaning that performance across blocked and mixed contexts in these analyses could not be due to stimulus selection. Furthermore, a visual inspection of reaction-time data across different form classes for word distractors (i.e., nouns versus adjectives) showed no obvious disparity between the two classes. Consequently, we are confident that the results of the present study reflect the nature of the picture–word pair relations. Finally, we argue that while longer naming times and stimulus selection concerns are limitations to consider when deciding NOT to reuse pictures in these studies, the alternative also entails limitations. Specifically, if researchers continue to repeatedly present pictures in bilingual naming studies, we may develop accurate models of production, but ones that are limited in how applicable they are to naturally occurring speech (Kroll et al., 2006). Given the number of potentially important variables in bilingual production, issues of design will always be a balancing act. We contend that the solution to this problem is convergence – future replications and extensions of through translation effects should verify these results.

In conclusion, the present study provides an important demonstration of factors that influence the generation and magnitude of phonological effects in bilingual picture-naming. The results also establish the reliability of through translation facilitation. Specifically, visually presented distractor words are capable of activating the phonological representations of their translation equivalents especially when production occurs in L2 or under conditions of processing uncertainty. These results have important implications to current explanations of the asymmetry seen between same and cross-language identity effects in bilingual picture-naming tasks (Costa et al., 1999). Although this asymmetry has been attributed to an underlying phonological component in the same-language identity effect not present in cross-language identity effects, the results of the present study in conjunction with those presented by Hermans (2003) question whether this explanation is sufficient to capture the complete nature of the asymmetry. Both studies demonstrate that a phonological contribution to the cross-language identity effect may be possible when the activation of translation equivalents is supported (e.g., due to stimulus composition, response language, or production context).

In future research, we hope to utilize the through translation relation in a variety of ways to explore the activation of response and nonresponse languages at the lexeme or sublexical levels. In addition, it will be interesting to examine whether identically-related and semantically-related effects show the same sensitivity to manipulations of response language and production context. Such investigations will provide an important element to the design of bilingual production models.

References

Amrhein, P. C. & Sanchez, R. (1997). The time it takes bilinguals and monolinguals to draw pictures and write words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 1439–1458.

Beauvillain, C. & Grainger, J. (1987). Accessing interlexical homographs: Some limitations of a language-selective access. *Journal of Memory and Language*, 26, 658–672.

Bijeljac-Babic, R., Biardeau, A. & Grainger, J. (1997). Masked orthographic priming in bilingual word recognition. *Memory & Cognition*, 25, 447–457.

Brysbaert, M., Van Dyck, G. & Van de Poel, M. (1999). Visual word recognition in bilinguals: Evidence from masked phonological priming. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 137–148.

Carroll, J. B. & White, M. N. (1973). Age-of-acquisition norms for 220 picturable nouns. *Journal of Verbal Learning and Verbal Behavior*, 12, 563–576.

Chen, H. C. & Ho, C. (1986). Development of Stroop interference in Chinese–English bilinguals. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 12, 397–401.

Colomé, A. (2001). Lexical activation in bilinguals' speech production: Language-specific or language-independent? *Journal of Memory and Language*, 45, 721–736.

Costa, A., Colomé, A., Gómez, O. & Sebastián-Gallés, N. (2003). Another look at cross- language competition in bilingual speech production: Lexical and phonological factors. *Bilingualism: Language and Cognition*, 6, 167–179.

Costa, A., La Heij, W. & Navarrete, E. (2006). The dynamics of bilingual lexical access. *Bilingualism: Language and Cognition*, 9, 137–151.

Costa, A., Miozzo, M. & Caramazza, A. (1999). Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection? *Journal of Memory and Language*, 41, 365–397.

Cuetos, F., Ellis, A. W. & Alvarez, B. (1999). Naming times for the Snodgrass & Vanderwart pictures in Spanish. *Behavior Research Methods, Instruments & Computers*, 31, 650–658.

Dell, G. S. & O'Seaghdha, P. G. (1991). Mediated and convergent lexical priming in language production: A comment on Levelt et al. (1991). *Psychological Review*, 98, 604–614.

Dijkstra, T., Grainger, J. & van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language*, 41, 496–518.

Ehri, L. C. & Ryan, E. B. (1980). Performance of bilinguals in a picture–word interference task. *Journal of Psycholinguistic Research*, 9, 285–302.

Grainger, J. & Dijkstra, T. (1992). On the representation and use of language information in bilinguals. In R. J. Harris (ed.), *Cognitive processing in bilinguals*, pp. 207–220. New York: Elsevier.

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

Hermans, D. (2000). Word production in a foreign language. Ph.D. dissertation, University of Nijmegen.

Hermans, D. (2004). Between-language identity effects in picture–word interference tasks: A challenge for language-nonspecific or language-specific models for lexical access? *International Journal of Bilingualism*, 8, 115–125.

Hermans, D., Bongaerts, T., De Bot, K. & Schreuder, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language? *Bilingualism: Language and Cognition*, 1, 213–229.

Jared, D. & Kroll, J. F. (2001). Do bilinguals activate phonological representations in one or both of their languages when naming words? *Journal of Memory and Language*, 44, 2–31.

Kroll, J. F., Bobb, S. C. & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and Cognition*, 9, 119–135.

Kroll, J. F. & De Groot, A. M. B. (1997). Lexical and conceptual memory in the bilingual: Mapping from form to meaning in two languages. In A. M. B. De Groot & J. F. Kroll (eds.), *Tutorials in bilingualism: Psycholinguistic perspectives*, pp. 169–200. Mahwah, NJ: Lawrence Erlbaum.

Kroll, J. F., Michael, E., Tokowicz, N. & Dufour, R. (2002). The development of lexical fluency in a second language. *Second Language Research*, 18, 137–171.

Kučera, H. & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown Univeristy Press.

Miller, N. A. & Kroll, J. F. (2002). Stroop effects in bilingual translation. *Memory & Cognition*, 30, 614–628.

Roelofs, A. (2003). Shared phonological encoding processes and representations of languages in bilingual speakers. *Language and Cognitive Processes*, 18, 175–204.

Sebastián-Gallés, N. (2000). LEXESP: *Léxico informatizado del español*. Edicions de la Universitat de Barcelona.

Snodgrass, J. G. & Vanderwart, 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.

Snodgrass, J. G. & Yuditsky, T. (1996). Naming times for the Snodgrass and Vanderwart pictures. *Behavior Research Methods, Instruments & Computers*, 28, 516–536.

Soares, C. & Grosjean, F. (1984). Bilinguals in a monolingual and a bilingual speech mode: The effect on lexical access. *Memory & Cognition*, 12, 380–386.

Sumiya, H. & Healy, A. F. (2004). Phonology in the bilingual Stroop effect. *Memory & Cognition*, 32, 752–758.

Received November 26, 2002
Revision received August 7, 2006
Accepted October 2, 2006