Integrated Phonological Processing in Bilinguals: 
Evidence from Spoken Word Recognition

Olga V. Boukrina (o-boukrina@northwestern.edu)
Department of Communication Sciences and Disorders, 2240 Campus Drive
Evanston, IL 60208 USA

Viorica Marian (v-marian@northwestern.edu)
Department of Communication Sciences and Disorders, 2240 Campus Drive
Evanston, IL 60208 USA

Abstract
The role of cross-linguistic phonological overlap in native and non-native word recognition was examined using an auditory lexical decision task. The degree of phonological overlap across languages was manipulated. Cross-linguistic overlap facilitated word recognition in the non-native language, but inhibited word recognition in the native language. The observed facilitation and inhibition effects provide evidence for a parallel account of bilingual word recognition and suggest an asymmetry in native and non-native phonological processing.

Introduction
Modern accounts of bilingual lexical representation and processing suggest that the two lexica of a bilingual are intergraded. Parallel activation of target (language of the task) and non-target (irrelevant to the task) languages is supported by empirical data from both the auditory (e.g., Blumenfeld & Marian, 2005; Marian & Spivey, 2003a, b; Spivey & Marian, 1999; Van Wijnendaele & Brysbaert, 2002) and the visual modalities (e.g., Costa, Miozzo, & Caramazza, 1999; Hermans, Bongaerts, De Bot, & Schreuder, 1998; Van Heuven, Dijkstra, & Grainger, 1998), as well as by simulations of cross-language competition (e.g., Dijkstra, van Heuven, & Grainger, 1998). The degree of activation of the non-target lexicon may vary as a function of language dominance and other factors (e.g., modality of presentation, similarity of sensory input in the target language to phonology or orthography of the non-target language, etc).

Parallel activation has been shown to occur more reliably with high-proficiency non-target languages than with low-proficiency non-target languages (Jared & Kroll, 2001; Silverberg & Samuel, 2004; Van Hell & Dijkstra, 2002; Weber & Cutler, 2004). For example, while findings of parallel first-language (L1) activation during second-language (L2) processing have been consistent (Blumenfeld & Marian, 2005, Marian & Spivey, 2003a, b; Weber & Cutler, 2004), findings of parallel L2 activation during L1 processing have been mixed (Ju & Luce, 2004; Marian & Spivey, 2003b; Weber & Cutler, 2004). Using eye-tracking, Marian and Spivey (2003a) tested a group of Russian-English bilinguals when Russian was the non-target language and when English was the non-target language. They found co-activation of the non-target language, both when it was the L1 and when it was the L2. In contrast, Weber and Cutler (2004) tested a group of Dutch-English bilinguals in both Dutch and English, and found co-activation of the non-target language when it was the L1, but not when it was the L2. One explanation for this discrepancy lies in the different levels of second-language proficiency and experience across the two participant groups, with higher L2 proficiency levels in the Russian-English bilinguals (tested in the US) than in the Dutch-English bilinguals (tested in the Netherlands). Another possible explanation for the observed differences in L2 activation across studies lies in the degree to which L2 phonetic characteristics match those of L1. For instance, in another eye-tracking study with Spanish-English bilinguals, Ju and Luce (2004) demonstrated that participants fixated interlingual distractors more frequently than control distractors when voice onset times in Spanish auditory stimuli matched the voice onset times appropriate for English. Consistently, Blumenfeld and Marian (2005) found that when degree of phonetic overlap at word onset is increased, co-activation of L2 increases, but only in cognates. Thus, co-activation of L2 may be amplified by matching L1 and L2 phonetic characteristics. Furthermore, manipulating degree of phonological overlap across languages may produce different levels of non-target language co-activation.

Different patterns of activation observed for L1 and L2 suggest that cross-talk between languages may vary for first and second language processing. This may be the case because proficiency is usually greater in the first language than in the second language, or as a result of differences in age of acquisition and history of language use (e.g., Zevin & Seidenberg, 2002; Jared & Kroll, 2001; Grosjean, 1997). For example, monolingual interlocutors and language settings influence a bilingual’s language choice by increasing the use of one language and decreasing its threshold of activation. As a result, the language used more frequently long-term may become dominant and more readily available for processing, and this variability in individual history of language use may contribute to bilinguals’ asymmetry in word recognition across languages.

The dynamic nature of bilingual lexical representations is captured in Kroll and Stewart’s (1994) Revised Hierarchical Model (RHM). According to the RHM, bilinguals’ proficiency influences first and second language processing.
and underlying representational mechanisms. During initial stages of second language acquisition, L2 words are connected to L1 words via lexical links, and L1 words are in turn connected to semantic information. As bilinguals continue to learn the second language and their proficiency level increases, L2 words begin to form direct links to conceptual representations. At later stages of acquisition, L2 words have established connections with conceptual information, but the links between L2 and L1 at the lexical level are preserved and may be relied upon when processing in a highly-proficient second language. The RHM proposes that the strength of various connections is not the same, with conceptual representations linked more strongly to L1 lexical representations than to L2 lexical representation. At the lexical level, the path from L2 to L1 is stronger than the path from L1 to L2. Connections of different strength suggest an asymmetry in bilingual lexical organization and processing.

While the exact nature of the asymmetry between L1 and L2 phonological processing remains unclear, one possible explanation relies on bilinguals’ lack of fine-grained distinctions in non-native phonological representations. Research with non-native listeners suggests that auditory word recognition is more difficult in the second language than in the first language (e.g., Bradlow & Bent, 2002). The ease of phonological processing may vary with proficiency, similarly to lexical processing (Kroll & Stewart, 1994). Initially, L1 phonological representations may be organized as tightly constrained categories of sounds and include phonological representations for similar L2 categories. For instance, Best and colleagues (2001) suggested that some L2 phonemes can be perceptually assimilated to L1 phonetic categories, based on commonalities in the organs of articulation, and the place and manner of articulation. Empirical evidence supports this account and shows that L2 phonemes similar to a common L1 category are discriminated with more difficulty than L2 phonemes that do not bear resemblance to an L1 category (e.g., Best, McRoberts, & Goodell, 2001). Imai, Walley, and Flege (2005) proposed that with increased L2 word learning and exposure, L2 phonological representations become more fine-grained. In sum, phonological competence in L1 and L2 may influence the extent of parallel language activation. One way to test this hypothesis is to examine how bilinguals process input that varies in degree of cross-linguistic phonological similarity. If lower phonological competence in L2 contributes to asymmetry in L1-L2 processing, then sensitivity to cross-linguistic phonological similarity should differ across the two languages.

The objective of the present study was to examine the role of cross-linguistic phonological overlap during L1 and L2 processing in the auditory domain. The study was modeled after a visual language processing experiment by Jared and Kroll (2001). Jared and Kroll examined the role of proficiency and language context on parallel activation of bilingual lexica. They tested activation of phonological representations in bilinguals’ two languages when reading stimuli with overlapping graphemic form. English-French and French-English bilinguals read aloud words with varying consistency of grapheme-to-phoneme mapping across languages. The stimuli were presented in three phases: an English-words phase, a French-words phase, and another English-words phase. Results varied depending on whether bilinguals were processing words in L1 or L2. French-English bilinguals activated French spelling-to-sound correspondences while reading in English, as indicated by increased error rates and slower naming latencies for words with French competitors (words with different letter-to-sound mappings in French, e.g., lait) than for words with no competitors. During the first English phase, English-French bilinguals did not activate French spelling-to-sound correspondences, even if they were fluent in French. However, in the second English phase after completing the French phase of the experiment, participants experienced interference from French spelling-to-sound correspondences.

Similar to Jared and Kroll (2001), the present study tested the effect of cross-linguistic overlap on first and second language processing and examined the role of language context on parallel activation of bilingual lexica. The design of the study followed that of Jared and Kroll and included three language phases. Alternating between languages across the three phases (second language, followed by first language, followed by second language) made it possible to examine the costs of switching language contexts on parallel language activation. The differences between the two studies were in (1) the modality of processing, and (2) ways in which input was varied. While Jared and Kroll targeted visual word recognition and manipulated spelling-to-sound consistency, the present study targeted auditory word recognition and manipulated phonological overlap. Phonological overlap was defined by the presence of phonemes shared across native and non-native languages. In order to manipulate phonological overlap, phonemes in each language were divided into unique and non-unique (i.e., shared). Uniqueness was established after comparing corresponding phonemes in L1 and L2 on their characteristics. While Jared and Kroll (2001) used words with no competitors, words with competitors in the same language, and words with competitors in the other language, the present study used words that did not overlap phonologically, words that overlapped phonologically for one-third of auditory input, words that overlapped phonologically for two-thirds of auditory input, and words that overlapped phonologically completely. Using four levels of overlap made it possible to manipulate phonological similarity in a gradual manner and perform a more fine-grained analysis of the impact of phonology on bilingual spoken word recognition. The direction and consistency of effects across different degrees of overlap were investigated.

To test the extent to which bilinguals activated phonological representations of both languages simultaneously, Russian-English bilinguals were asked to
perform a lexical decision task and to decide whether auditory input constituted a word or a non-word. It was hypothesized that if phonological representations of two languages were co-activated and accessed simultaneously whenever one of the languages was activated, then response latencies and accuracy rates for auditory stimuli that overlapped phonologically across languages would differ compared to stimuli that did not overlap across languages. Moreover, the role of phonological overlap was predicted to vary across first and second languages. Specifically, overlap with a more proficient language was predicted to affect performance in a less proficient language more than overlap with a less proficient language would affect performance in a more proficient language.

Method

Participants
Twenty-six Russian-English bilinguals (15 females) participated in the study. Their mean age at the time of testing was 22.12 years (SD = 6.26). Participants were students at an American university and had lived in the US for an average of 12.65 years (SD = 9.16). They had known English for an average of 12.75 years (SD = 8.90) and Russian for an average of 20.17 years (SD = 5.18), paired samples t (25) = 4.04, p < .001. Participants reported speaking English on average 7 hours per day (range 0.5-12) and Russian 3.52 hours a day (range 0-3.7), paired samples t (25) = 3.68, p < .001. English was the preferred language for 13 participants; Russian was preferred by 10 participants, while 3 participants reported no language preference. Participants were naïve to the experimental manipulation and were paid for their participation.

Materials
The stimuli were three-phoneme Russian and English words and non-word phoneme-sequences, coded according to the International Phonetic Alphabet (IPA, 1999). All words were unique to Russian and English and no cognates, homophones, or homographs were used. Two-hundred-and-forty stimuli were divided into three sets: Russian set, first English set, and second English set. Each set consisted of 40 words and 40 non-words.

In the Russian set, the words were selected so that 10 were comprised of unique Russian phonemes (0-phoneme overlap), another 10 included two unique and one non-unique Russian phonemes (1-phoneme overlap), a third subset of 10 contained one unique and two non-unique phonemes (2-phoneme overlap), and the last 10 consisted of only non-unique Russian phonemes (3-phoneme overlap). Corresponding sets of non-word stimuli were constructed in the same manner using unique and non-unique Russian phonemes. The words and non-words in the English sets were selected in the same manner using unique and non-unique English phonemes.

Woods were matched for frequency of occurrence within each language, using Sharoff’s (2003) frequency dictionary for Russian and the Kucera and Francis (1967) dictionary for English. A one-way ANOVA on four subsets of Russian words revealed no differences in mean frequencies, F (3, 36) < 1. A 2 x 4 ANOVA (English Set x Phonological Overlap) for English word frequencies showed no main effect of English set [F (1, 68) < 1], no main effect of Phonological Overlap [F (3, 68) = 1.13, p = .34], and no interaction between the two [F (3, 68) < 1]. In addition, words in the Russian phase (M = 50.15, SD = 72.87) did not differ in frequency from words in the first English phase (M = 59.18, SD = 78.80), t (77) = .53, p = .598, or second English phase (M = 62.59.15, SD = 90.05), t (75) = .67, p = .506. English stimuli were recorded by a native speaker of English in a sound-proof booth. Russian stimuli were recorded in a similar manner by a native speaker of Russian.

Design and Procedure
The experiment followed a 3 x 4 x 2 within-subjects design. The first factor, phase, had three levels: first English phase, Russian phase, and second English phase. Phonological overlap included four levels: 0-phoneme overlap, 1-phoneme overlap, 2-phoneme overlap, and 3-phoneme overlap. The third factor, lexical status, had two levels: word and non-word. Latency of response and response accuracy were measured.

At the start of an English phase, instructions were presented in English; at the start of the Russian phase, instructions were presented in Russian. Participants heard the stimuli over standard headphones. The first set of English items was played first; followed by the set of Russian items and the second set of English items. On each trial, participants performed a lexical decision task on a phoneme sequence by pressing either a “word” or “non-word” key on the response box. There was a 1500 ms inter-trial interval, and a self-paced break was offered after every 20 trials. Reaction times were measured from stimulus offset. At the end of the experiment participants completed a questionnaire about their linguistic background.

Coding and Analyses
Items with accuracy rates less than 70% across participants were excluded from analyses, resulting in elimination of 9.2% of word data. In another 0.9% of word data and 3.75% of non-word data, reaction times were greater than 2500 ms and were substituted with 2500 ms, which was equal to about 2.5 SDs above the mean RT across participants.

Results

Reaction Time
A 3-way ANOVA with Phase (first English phase, Russian phase, second English phase), Lexical Status (word, non-word) and Phonological Overlap (0-phoneme overlap, 1-phoneme overlap, 2-phoneme overlap, 3-phoneme overlap) was performed. Results revealed a main effect of Phase [F (2, 50) = 4.41, MSE = 147,497.33, p < .05] and a main effect
of Lexical Status \([F (1, 25) = 56.75, \text{MSE} = 329,919.11, p < .001]\). Participants were faster in the Russian phase \((M = 608.69, SD = 253.57)\) than in the first English phase \((M = 711.25, SD = 246.29), t (25) = 2.95, p < .01, or the second English phase \((M = 698.59, SD = 274.79), t (25) = 2.05, p = .051, and responded faster to words \((M = 499.62, SD = 204.54)\) than to non-words \((M = 846.06, SD = 386.76)\).

Significant interactions were found between Phase and Phonological Overlap \([F (6, 150) = 4.11, \text{MSE} = 12,643.56, p < .01]\), between Lexical Status and Phonological Overlap \([F (3, 75) = 3.42, \text{MSE} = 15,806.49, p < .05]\), and between Phase, Lexical Status and Phonological Overlap \([F (6, 150) = 4.84, \text{MSE} = 14,625.08, p < .01]\).

Follow-up analyses showed a main effect of Phonological Overlap in the first English phase \([F (3, 75) = 3.18, \text{MSE} = 10,157.67, p < .05]\), where increased phonological overlap was associated with shorter reaction times (although the relationship was non-linear). Participants responded slower to words with 0-phoneme overlap \((M = 560.43, SD = 239.77)\) than to words with 1-phoneme overlap \((M = 494.47, SD = 166.93), t (25) = 2.158, p < .05, or to words with 3-phoneme overlap \((M = 475.50, SD = 159.12), t (25) = 2.61, p < .05.\) Similarly, reaction times to words with 2-phoneme overlap \((M = 530.66, SD = 182.28)\) were slower than to words with 3-phoneme overlap, \(t (25) = 2.98, p < .01.\)

In the second English phase, no main effect of Phonological Overlap was found. However, planned contrasts showed that reaction times to words with 0-phoneme overlap \((M = 567.82, SD = 261.33)\) were slower than to words with 2-phoneme overlap \((M = 477.61, SD = 164.95), t (25) = 2.73, p < .05.\) No differences in reaction times to words were found between the first English phase \((M = 515.27, SD = 172.36)\) and the second English phase \((M = 516.14, SD = 370.62).\)

In the Russian phase, participants responded more accurately to words with 0-phoneme overlap \((M = .95, SD = .06)\) than to words with 2-phoneme overlap \((M = .89, SD = .14), t (25) = 2.60, p < .05, or to words with 3-phoneme overlap \((M = .90, SD = .12), t (25) = 2.60, p < .05.\) Similarly, they responded more accurately to words with 1-phoneme overlap \((M = .95, SD = .06)\) than to words with 2-phoneme overlap, \(t (25) = 2.29, p < .05, or to words with 3-phoneme overlap, t (25) = 2.62, p < .05\) (See Figure 2).

**Discussion**

The degree of cross-linguistic phonological overlap was found to influence participants' response speed and accuracy. However, different patterns were observed for L1 and L2 processing. In the second language (English), greater cross-linguistic phonological overlap was associated with shorter latencies and greater accuracy of response. The opposite pattern was observed for the first language (Russian), where, in general, phonological overlap with L2 was associated with longer latency rates and decreased accuracy. It is important to note that the patterns of results observed in the present study may not hold for bilinguals.
with a different language-history profile, such as bilinguals who are balanced across both languages, who acquired both languages in parallel, or whose L1/L2 proficiencies differ more drastically.

In both English phases of the present study, words that shared phonology with Russian were identified faster and more accurately than words comprised of unique English phonemes. Moreover, as phonological overlap increased, responses were provided faster and with more accuracy. The observed facilitation of the second language as a function of phonological overlap with the first language is consistent with previous research reporting facilitation during masked priming of non-native words with phonologically similar native words (Brysbaert, Van Dyck, & Van de Poel, 1999). However, although reaction time and accuracy followed the same patterns in both English phases, follow-up pair-wise comparisons suggested that the magnitude of the differences was greater in the first English phase than in the second English phase. That is, while in the first English phase, a graded increase in RT was observed as a function of phonological overlap, in the second English phase, only considerable differences such as between unique phonology and shared phonology affected response latency and accuracy. It is possible that practice effects and increased familiarity with the task in the course of the experiment reduced sensitivity to fine-grained differences in phonological overlap. Alternatively, the discrepancies between the two English phases could be attributed to the change in linguistic context and baseline of activation. In the Jared and Kroll (2001) study greater interference in production of words with French competitors was observed after completion of the French phase. Completing a task in Russian could introduce greater overall facilitation when processing English stimuli and attenuate differences in reaction times to stimuli with various degrees of phonological overlap.

In the Russian phase, response latency and accuracy were also affected by degree of phonological overlap. However, unlike English word recognition, Russian word recognition appeared to be inhibited by increased phonological overlap with L2. The finding that lexical decision was slower for words with 0-phoneme overlap than for words with 1-phoneme overlap was inconsistent with the overall pattern. One possible explanation relies on the speed-accuracy trade-off, i.e., while reaction times were slower to stimuli with 0-phoneme overlap than to stimuli with 1-phoneme overlap, accuracy was greater for stimuli with 0-phoneme overlap than for stimuli with 2-phoneme overlap. Another possibility is that the English context of the first phase suppressed access to uniquely Russian phonological information. Reaction time data reflected this suppression, while accuracy data did not, possibly due to greater sensitivity of reaction time to minor changes in linguistic context. This hypothesis is consistent with previous research on the effects of phonological similarity in bilingual naming, where reaction time measures of processing in a non-native language were more sensitive to phonological neighborhood effects (Marian & Blumenfeld, in press). To further test this hypothesis, future research may vary the order of native and non-native language input, so that both are presented with and without prior exposure to the other language. Overall, it appears that lexical decision in L1 is slowed by interference effects from phonologically overlapping L2 phonemes.

The facilitation and interference effects observed in the present study provide evidence for an integrated account of bilingual lexical organization. Cross-linguistically overlapping input activated both languages, regardless of the task-relevant language. In L1, activation of the non-native language phonology delayed or compromised lexical decision, possibly due to competition between viable word-form representations as a result of simultaneous activation of L2. In L2, activation of the native-language phonology aided processing, perhaps due to faster phoneme recognition in L1 as a result of extensive previous use. Alternatively, the facilitation and interference effects could be explained in terms of order of acquisition, with L1 mediating subsequent language learning (Best et al., 2001). L1 phonetic categories acquired early in life are tightly constrained and may compete with similar L2 phonemes, while L2 phonetic representations acquired later in life, are organized into wide categories of sounds and can be co-activated by similar L1 phonemes. Therefore, the observed pattern of results could be explained by the structure of phonological representations in L1 and L2. Lexical decision in L1 was delayed by competition of highly constrained L1 categories and co-activated L2 phonemes, and lexical decision in L2 was facilitated by co-activation of wider L2 categories with similar L1 phonemes.

In sum, native-language words that shared phonology with the second language were processed slower and with less accuracy than words with unique native phonology. However, first-language words that were completely unique in phonological characteristics were also recognized slower, suggesting a possible influence of linguistic-context on first language processing. These differences in the direction and magnitude of the effect were uncovered only because degree of phonological overlap was systematically manipulated across four levels. Such graded manipulation of phonological overlap emerged as a valuable tool for exploring processing in the bilingual language system. Studies of language interaction in bilinguals typically use cognates, homophones, or homographs, which are usually the exception to bilingual linguistic input rather than the rule. Non-cognate, non-homophonic/non-homographic stimuli that are comprised of either overlapping or non-overlapping phonology, such as the words used in the present study, provide a window into the more general system of bilingual organization.

To conclude, both the facilitation and the interference effects observed in the present study support parallel accounts of bilingual language processing and integrated accounts of bilingual lexical organization. However, they also suggest an asymmetry in first and second language
phonological processing in unbalanced bilinguals. This asymmetry may be accounted for by variability in L1 and L2 phonological representations, with additional research needed to explore the nature of these differences.

Acknowledgments
Data collection for this study was supported by an Undergraduate Research Grant from SUNY Binghamton to the first author; data analyses and writing of this paper were supported in part by Grants NICHD 1R03HD046952-01A1 and NSF BCS-0418495 to the second author. We thank Dr. Cynthia Connine, Henrike Blumenfeld, Margarita Kaushanskaya, Avital Rabin, Jay Mittal, and Allison Moy for their help.

References