The Impact of Allophony versus Contrast on Speech Perception[*]

Amanda Boomershine,[a] Kathleen Currie Hall,[b] Elizabeth Hume,[b] Keith Johnson[c]

[a]University of North Carolina, Wilmington
[b]The Ohio State University
[c]University of California, Berkeley

1. Introduction

The perceptual consequences of phonological contrast have long been of interest to phonologists and phoneticians. In Trubetzkoy's well-known *Grundzüge der Phonologie* (1939: 78), for example, he speculates that an opposition between speech sounds that is always contrastive in a given language will be perceived more clearly than an opposition that is neutralizable in some context. Furthermore, even within the category of neutralizable oppositions, he predicts that perception will fluctuate depending on factors such as context. There are three important assumptions that underlie Trubetzkoy's speculations. First, that one's native language experience influences the ability to perceive speech sounds. Second, that the phonological relation holding between sounds in a language has an impact on a listener's perception of those sounds. And third, that it is not simply the presence versus the absence of phonological contrast that is relevant to perceiving a sound. Rather, Trubetzkoy pinpoints different categories, or degrees, of contrast and suggests that each may have a particular consequence for speech perception.

Trubetzkoy's first and second assumptions, that one's native language experience – particularly the phonological relations between sounds – influences the ability to perceive speech sounds, are now well established in the literature. For example, studies in second language learning have found that listeners are more adept at perceiving sounds of their native language than those of a second language acquired later in life, e.g., Polka and Werker (1994), Strange (1995), Dupoux et al. (1997), Best et al (1998), Francis and Nusbaum (2002). Familiar illustrations include the perception of English /l/ and /r/ by Japanese listeners and that of Hindi dental and retroflex stops by American English listeners. Since the liquids /l/ and /r/ are non-contrastive in Japanese, Japanese listeners have difficulty distinguishing between them, even though they are fully

contrastive in English (Goto 1971; MacKain et al. 1981). For similar reasons, perceiving a distinction between the Hindi stops is more challenging for English speakers than it is for Hindi speakers (Werker et al. 1981; Pruitt et al. 1998). The conclusion that can be drawn from these and other studies is that while listeners have little difficulty distinguishing between contrastive native sounds, they are less successful when it comes to non-native sounds that do not serve a contrastive function in their own language.

Less is known, however, concerning Trubetzkoy's third assumption, especially as it relates to the potential impact of phonological relations other than contrast on speech perception. It is this last point that we are especially concerned with in this paper.

As noted above, it is well established that while listeners have no difficulty distinguishing between native sounds that are contrastive, they are less successful when it comes to sounds that do not occur in their own language. Furthermore, there is evidence suggesting that it is not simply the presence versus the absence of phonemic contrast that is relevant to perceiving a sound. Partial contrast, where an otherwise contrastive pair of elements is neutralized in some context, has also been shown to influence perception, as Trubetzkoy predicted. For example, drawing on perception data on Mandarin tone (Huang 2001), Hume and Johnson (2003) conclude that not only is perceptual distinctiveness a function of phonological contrast, but that *partial* contrast reduces perceptual distinctiveness for native listeners. Thus, contrast seems to be more nuanced than is often assumed in the speech perception literature.

This finding then raises the question as to whether other phonological relations also shape perception. Consider non-contrastiveness. As noted above, listeners typically have greater difficulty distinguishing between sounds that do not occur in their own language, and are thus non-contrastive, than they do with native sounds that are contrastive. In addition to this typical notion of non-contrastiveness, sounds that do in fact co-occur within a single language can also be in a non-contrastive relation, such as when they are allophones of the same phoneme. While two sounds with an allophonic distribution both occur in a speaker's phonetic inventory, they never effect a change in meaning. In English, for example, the phones [d] and [ɾ] can be considered allophones of a single phoneme, /d/, with [ɾ] occurring intervocalically when the first vowel is stressed, e.g. [rayɾɪŋ] "riding," and [d] occurring elsewhere, e.g. [rayd] "ride." Crucially, however, substituting [d] for [ɾ] in "riding" has no effect on the meaning of the word.

Given the lack of contrast between a pair of allophones, we would expect them to be perceived as less distinct than a pair of contrastive sounds, all else being equal. Theories of speech perception generally predict this result (see, e.g. Lahiri 1999; Gaskell and Marslen-Wilson 2001), although the means by which they do so vary in their predictions for the perception of other pairs of sounds in the language. There is also some experimental support for the idea that allophony plays a role in speech perception (e.g., Dupoux et al. 1997; Harnsberger 2001; Johnson 2004), though its precise influence on perception has not been directly tested. For example, Harnsberger's (2001) results from an AXB classification task point to a near merger in the perception, by Malayalam listeners, of allophonically-related dental and alveolar nasal consonants. These coronal nasals are in complementary distribution in the language, with the dental occurring morpheme-initially and the alveolar occurring both morpheme-finally and intervocalically (Mohanan and Mohanan 1984). Contrastive nasals such as bilabial [m] versus velar [ŋ], on the other hand, showed greater perceptual separation in Harnsberger's study. Findings such as these suggest that the simple presence of a sound in an inventory is not the only source of information concerning the relative perception of that sound. The sound's phonological relatedness to other sounds in the inventory must also be taken into consideration.

This paper explores the impact of contrast versus allophony on the perception of speech sounds in a series of four experiments contrasting the behavior of Spanish-speaking and English-speaking listeners, and considers how these empirical results should be integrated into a theory of speech perception. In addition to the basic finding that models of speech perception are in fact correct in their prediction that phonemic contrasts are more perceptually distinct than allophonic contrasts, the results of experiments like the ones presented here can be used to differentiate models of speech perception based on the mechanisms by which this more basic finding is predicted in the different models, as will be discussed. In section 6, we consider the effectiveness of two different models in accounting for the results: a phonological inferencing model (e.g. Gaskell and Marslen-Wilson 1998) and an exemplar model (e.g. Goldinger 1992, 1996; Palmeri et al. 1993; Johnson 1997a, b, 2004; Coleman 2002; Pierrehumbert 2003; Hawkins 2003). To anticipate our conclusion, both models are successful in predicting our findings relating to allophony versus phonemic contrast. Only the exemplar model, however, is able to account for the full range of results obtained in this study.

The experiments presented in this paper make use the fact that English and Spanish place similar sounds, namely [d], [ð], and [ɾ], in very different positions in the linguistic system of contrasts. As illustrated in (1), the phones [d] and [ɾ] are allophones of a single phoneme in English while

[d] and [ð] are contrastive ([do] "dough" versus [ðo] "though"). Conversely in Spanish, [d] and [ð] are allophones of a single phoneme (de [ð]onde 'from where', [d]onde 'where'), while [d] and [ɾ] are separate phonemes. Note, however, that [d] and [ɾ] are never lexically contrastive in Spanish since the sounds do not appear in the same context: [ɾ] occurs in medial position and [d] in initial position.

(1) Phonological grouping of [ð], [d], and [ɾ] in English and Spanish. Sounds within parentheses pattern as allophones of a single phoneme, and are contrastive with sounds outside parentheses.

English   [ð]   ([d]   [ɾ])
Spanish   ([ð]   [d])   [ɾ]

In general, we expect that when sounds are contrastive in a language, listeners will be more attuned to the phonetic contrast between these sounds and thus judge them to be more different from each other than sounds that are in a non-contrastive relationship within a given language.

While the pairs [ð]/[d] and [d]/[ɾ] display different phonological relations in Spanish and English, the pair [ɾ]/[ð] patterns similarly in terms of phonological representation. In each language, these sounds are associated with different phonemes, but one sound of the pair is in an allophonic relationship with a different sound that is also present in the inventory of the language, as shown in (2).

(2) Surface and phonemic correspondences of [ð] and [ɾ] in Spanish and English
    (a) Spanish: surface contrast [ð] - [ɾ] corresponds to phonemic contrast /d/ - /ɾ/
    (b) English: surface contrast [ð] - [ɾ] corresponds to phonemic contrast /ð/ - /d/

The patterning of the two sounds [ɾ]/[ð] is also similar in that in both languages the distinction between the phones signals lexical, or surface, distinctions, as (3) illustrates.

(3) Surface contrast of [ɾ] and [ð]
    (a) English
            [lɛðɹ̩] "leather"   [lɛɾɹ̩] "letter"
            [mʌðɹ̩] "mother"   [mʌɾɹ̩] "mutter"
    (b) Spanish
            [kaða] cada "each"   [kaɾa] cara "face"

To summarize, the phonological relations of each of the three pairs of sounds are given in (4). The first pair, [d]/[ɾ], is contrastive in Spanish and allophonic in English. In neither language does this pair display a surface contrast. The pair [d]/[ð], on the other hand, displays contrast at the phonemic and surface levels in English while in Spanish, it is allophonic and thus contrasts on neither level. Finally, the phonological relations of the pair [ɾ]/[ð] are the same in both languages, being contrastive both at surface and phonemic levels.

(4) Summary of phonological relations among [d], [ð], and [ɾ] in English and Spanish

Pair:	[d] - [ɾ]		[d] - [ð]		[ɾ] - [ð]	
Language:	English	Spanish	English	Spanish	English	Spanish
phonemic (underlying) contrast	-	+	+	-	+	+
surface contrast	-	-	+	-	+	+

Given the similar patterning of the latter pair across the two languages, we would expect that the perceived difference between intervocalic [ð] and [ɾ] to be about the same for both Spanish and English listeners. On the other hand, given the allophonic/contrastive differences with the remaining two pairs, we would expect the pairs to pattern differently in the two languages. Specifically, contrastive pairs should show greater perceptual separation than the allophonic pairs.

To explore the perception of the contrastive and allophonic relations among [d, ɾ, ð] in the two languages, we used two experimental paradigms, intending to differentiate processing that might emphasize surface contrast from processing at a more phonemic level. To capture phonological processing, listeners were asked to rate the perceived difference between the sounds, forcing them to categorize each sound and then compare it to a second categorized sound. To capture surface phonetic processing, listeners were asked to make speeded AX discrimination judgments; such tasks are generally assumed in the literature to access a more purely auditory level of discriminability (see, e.g., Fox 1984; Strange and Dittman 1984; Werker and Logan 1985). Because the pattern of contrasts at surface and phonemic levels differs for the [d]/[ɾ] comparison, we expected that if one task taps surface contrast effects while the other taps phonemic contrast then we might see differing patterns of response with the two paradigms. It will be seen in the following sections, however, that these predictions regarding paradigm differences were not borne out.

1.1. Structure of the paper

Section 2 describes an experiment in which Spanish-speaking and English-speaking listeners were asked to rate the perceived similarities of pairs of non-identical stimuli: [d]/[ɾ], [d]/[ð], and [ɾ]/[ð]. Because the phonologies of Spanish and English group these sounds differently (see (1) above) and because the rating task is an off-line judgment task, we expected to see a strong effect of native language background on the listeners' similarity ratings. Section 3 presents results from a speeded discrimination study using the same stimuli that were used in experiment 1. We expected to find in this experiment a much smaller effect of native language on perceptual distance because the speeded discrimination task is a much more on-line task which may tap earlier "phonetic" processing (Werker and Logan 1985). Surprisingly, Spanish-speaking and English-speaking listeners differed in this experiment just as they differed in the rating task using these stimuli. Sections 4 and 5 present rating and speeded discrimination experiments that are identical to experiments 1 and 2 in every regard, except that in these experiments the stimuli were produced by speakers of Greek, who in their native language make all of the contrasts tested in the experiments (whereas the speakers for experiments 1 and 2 were English-speaking linguists). Finally, the differences and similarities between the two sets of experiments, as well as the implications of the experiments for theories of speech perception, are presented in section 6.

2. Experiment 1: Rating [d], [ɾ], [ð] pairs

2.1. Methods

2.1.1. Stimuli

Materials consisted of two tokens of each of the following VCV sequences: [ada], [aɾa], [aða], [idi], [iɾi], [iði], [udu], [uɾu], and [uðu]. The tokens were produced by two American English speaking trained phoneticians, one male and one female. The speakers recorded multiple examples of the stimuli using a head-mounted microphone in a soundproof booth. The speakers attempted to produce equal stress on the first and second syllables. In order to control the amplitude across tokens and speakers, the peak amplitude was equated for each of the tokens. The two best recordings for each VCV sequence were used as stimuli in the studies. These materials were used as stimuli in both experiment 1 and experiment 2.

2.1.2. Participants

One group of native Spanish speakers and one group of native American English speakers participated in the experiment. The native Spanish speakers (N = 10, 3 men, 7 women) were students or friends of students at The Ohio State University, and were from a variety of Spanish-

speaking countries, including Mexico, Colombia, Spain, Argentina, Puerto Rico, and Peru. They were paid a small sum for participating in the experiment. The native English speakers (N = 18, 8 men, 10 women) were undergraduate students at The Ohio State University enrolled in introductory linguistics courses who participated in the experiment for partial fulfillment of a course requirement. They were screened in a post-test questionnaire and only subjects who had no Spanish speaking experience were included in this experiment. The native English speaking participants thus had a mean self-rating of their Spanish ability of 0 on a scale from 0-7, where a score of 7 is equivalent to native competency, and a score of 0 is equivalent to no experience in that language. The native Spanish speaking participants had a mean self-rating of their English ability of 5 on a scale from 0-7. None of the speakers reported any history of speech or hearing disorders.

It should be noted that all of the native Spanish-speaking participants in the experiments reported here had an advanced level of English (i.e. they were bilingual). They were, however, run in a Spanish setting (the experimenter spoke to them in Spanish and the post-experiment questionnaire was presented in Spanish), so we believe that their English abilities had a minimal influence on their perception (see e.g. Marian and Spivey 2003 for a discussion of how the language of the experimental setting affects participant performance). We are currently running experiments on monolingual native Spanish speakers, and we expect to find that the monolingual Spanish speakers pattern very similarly to the Spanish speakers with a high degree of English. If anything, we expect that the inclusion of Spanish speakers with some knowledge of English in our experiments would bias the results against finding a difference between the perception of phonemic and allophonic pairs across languages; foreshadowing the results, the fact that such a difference was found is further indication that these Spanish speakers were operating in a Spanish mode.

Furthermore, while some of the native English-speaking participants did have knowledge of another foreign language (e.g. French, German, Japanese, etc.), none had familiarity with any language where the phones [d], [ð], and [ɾ] are in a fully contrastive relationship, such as Greek. Also, their mean self-rated ability in any foreign language was at a very low level, and such a superficial acquaintance with a second language does not seem to affect perception to any significant degree (see Boomershine et al. 2004).

2.1.3. Procedure

In this similarity rating task, participants were told that they would hear a pair of sounds and be asked to rate how similar those sounds were on a scale of 1 - 5, where 1 was 'very similar' and 5

was 'very different.' The participants were each seated at a computer that was connected to a 5-button response box, with up to four participants taking part in the study at a time. The participants listened to the stimuli through headphones, and then judged the similarity of the sounds using the button box. The pairs were presented in a different random order for each participant, using E-Prime software (v. 1.1; Psychological Software Tools, Pittsburgh, PA). The listeners heard pairs of stimuli, separated by one second of silence, such as [ada] <1 sec silence> [aɾa]. The talker and vowel context was the same for every pair so that the only difference in each pair was the consonant. The stimuli presented in each pair were always physically different tokens, even when they were both examples of a single sound (e.g. [ada] . . . [ada]). The participants were given four practice trials, and then the opportunity to ask questions before proceeding to the four test blocks (360 test trials total). They received no feedback in this experiment.

2.2. Results

To analyze the rating task results, the rating scores for each speaker were normalized to compensate for differences in use of the 5-point scale (e.g. avoiding use of the endpoints, etc.). The scores were normalized using a standard z-score transformation, such that each participant's scores were centered around 0, with scores above zero indicating "more different" and scores below zero indicating "more similar." The normalized results with their 95% confidence intervals are shown in Figure 1.

Insert Figure 1 about here

A repeated measures analysis of variance showed that there was a main effect of pair (F[2, 52] = 31.621, $p < 0.05$). That is, regardless of native language, the pairs were not all rated the same. There was also a significant pair by group interaction effect (F[2,52] = 22.174, $p < 0.05$), meaning that a participant's response to a given pair was dependent on the language group he was in. As shown in the figure, Spanish speakers found the pair [d]/[ɾ] (which is phonemically contrastive in Spanish but allophonic in English) more different than did the English speakers. Subsequent planned comparison independent samples t-tests showed that this difference was significant (t(26) = 3.29, $p < 0.05$). Furthermore, English speakers found the pair [d]/[ð] (which is phonemically contrastive in English but allophonic in Spanish) more different than did the Spanish speakers (t(26) = 4.902, $p < 0.05$). The pair [ɾ]/[ð], however, was rated the same by both

8

Spanish and English speakers (t < 1); this pair is composed of allophones of different phonemes in each language.

2.3. Discussion

The results from experiment 1 provide strong evidence that allophonic relationships influence the perceived distance between sounds at a phonological level of processing. As expected from the fact that [d] and [ð] are allophones of the same phoneme in Spanish, but are separate phonemes in English, Spanish-speaking listeners rated pairs of stimuli contrasting [ð] and [d] as being much more similar sounding than did the American English listeners. Parallel to this, as expected from the fact that [d] and [ɾ] are in an allophonic relationship in English while phonemic in Spanish, English-speaking listeners rated [d]/[ɾ] pairs as being more similar than did Spanish-speaking listeners. There was no significant difference in the ratings by both groups of listeners of the pair [ɾ]/[ð], which are allophones of different phonemes, an expected result given the similarity in the phonological relations of the pair in the two languages.

The results also indicate that on average, listeners rated [d]/[ɾ] pairs as more similar to each other than the [d]/[ð] pairs, and we hypothesize that this is due to the raw auditory discriminability of these particular tokens. Experiment 3 returns to this question, but first we turn to experiment 2 which uses a "phonetic" listening task that might be sensitive to patterns of surface contrast.

3. Experiment 2: Discriminating [d], [ɾ], [ð] pairs

In experiment 1, it was found that the native language of a listener had a strong impact on the listener's judgments of phonetic sound similarity. Given that the similarity rating task invites the listener to use metalinguistic knowledge and ponder the sounds during each trial, it is perhaps not surprising that the language difference was observed. Experiment 2 tests the same contrasts, with speakers of Spanish and English again, but this time using a discrimination task that is intended to require much more "phonetic" or "psychoacoustic" listening, as Werker and Logan (1985) found. Because the patterns of contrast among [d], [ð], and [ɾ] in Spanish and English differ depending on whether we are focusing on surface phonetic contrast or on phonemic category-level contrast, we sought to test in this experiment whether the surface pattern of contrast would influence listeners' responses in a lower-level listening task.

It should be noted that there is some evidence that even in a speeded discrimination task, which should tap a much lower level of processing than similarity rating, listeners' responses are influenced by linguistic experience. Huang (2001, 2004) observed that Mandarin listeners responded with relatively longer reaction times in a speeded discrimination task (as compared

with English-speaking listeners) when the sounds they were asked to discriminate were lexically related to each other. Specifically, the phonological neutralization of the dipping and rising tones of Mandarin resulted in longer reaction times for discriminations pairing these tones. English listeners did not show any effect of the Mandarin tone neutralization pattern. Interestingly, Huang found this effect of lexical/phonological contrast in a speeded discrimination task, which is generally assumed to be less prone to such language-specific effects. What Huang did not show is whether the linguistic experience reflected in her experiments relates to surface contrast or phonemic contrast. This experiment addresses this issue.

3.1. Methods

3.1.1. Stimuli

The stimuli that were used for experiment 1 were also used in this experiment.

3.1.2. Participants

The participants in this experiment were drawn from the same pool as those in experiment 1. The native Spanish speakers (N = 13, 3 men, 10 women) self-rated their ability in English at a mean value of 5.7; the native English speakers (N = 17, 3 men, 14 women) had no reported knowledge of Spanish. None of the speakers reported any history of speech or hearing disorders.

3.1.3. Procedure

In this discrimination task, the participants were told that they would hear a pair of sounds and be asked to judge whether the sounds were identical or different. "Identical" meant physically the same token (e.g. the same token of [ada] twice), while "different" meant either a different token of the same stimulus or two completely different stimuli (e.g. [ada] – [ada] where the two were not the same production, or [ada] – [aɾa], etc.). As with the rating task of experiment 1, the participants were seated at a computer connected to a 5-button response box. The participants were asked to indicate whether the pair of sounds they heard was physically identical or different by pressing button 1 on the response box if they were the same tokens and button 5 if they were different tokens. The stimuli, presented with E-Prime software, were again separated by one second of silence. Participants were given four practice trials before completing the three test blocks (288 test trials). After responding to each stimulus pair, the participants were given feedback as to the accuracy of their response, their average percent correct overall, and their response time (ms). This feedback was used to encourage both heightened accuracy and shorter response times.

3.2. Results

The average results and 95% confidence intervals for the "different" pairs from the discrimination task, shown in Figure 2, are very similar to those from the rating task. This figure shows normalized reaction times. Reaction time for these "different" pairs is taken to be a measure of perceptual distance, where slower reaction times indicate a smaller distance (see for example, Takane and Sergent, 1983); hence "more different" is at the bottom of the graph and "more similar" is at the top. As with the rating scores of experiment 1, we normalized the data in this experiment using a z-score transformation to correct for individual differences in overall reaction time. Consistent with the results from experiment 1, there was a main effect of pair ($F[2,56] = 22.162$, $p < 0.05$), indicating that some pairs were harder to discriminate than others, regardless of the native language of the listener.

Insert Figure 2 about here

There was also a significant pair by group interaction effect ($F[2, 56] = 3.876$, $p < 0.05$), indicating again that the pattern of pair reaction times differed depending on which group the listener was in – i.e., that native language influenced discrimination reaction time. Recall that slower reaction times are associated with more difficult discrimination and therefore with higher similarity. As predicted by the rating task, Spanish listeners were faster at discriminating the pair [d]/[ɾ], which is phonemic in Spanish, than were English listeners for whom [d]/[ɾ] are allophonically related. In subsequent planned comparison independent samples t-tests, this difference was found to be significant ($t(28) = 2.373$, $p < 0.05$), indicating that [d]/[ɾ] is perceived as less similar by the Spanish listeners. Not surprisingly, the English listeners were faster than the Spanish listeners at discriminating the pair [d]/[ð] ($t(28) = 2.823$, $p < 0.05$), given that these sounds have a phonemic relation in English but an allophonic one in Spanish. Finally, for the pair [ɾ]/[ð], the difference in reaction times of the two groups was not statistically significant ($t < 1$).

3.3. Discussion

The results from the discrimination task in experiment 2 are strikingly similar to those from the rating task in experiment 1. Again, there is strong evidence that allophony influences the perceived distance between sounds. As we found in the first experiment, a pair of sounds that is phonemic in one language (e.g. Spanish [d]/[ɾ]; English [d]/[ð]) was judged to be less similar than in the language where it is allophonic. Further, the native language of the listener did not

impact the judgment of [ɾ]/[ð]. The pair [d]/[ɾ] is of particular interest here because it does not contrast on the surface in Spanish, just as it doesn't in English. We expected that this lack of surface contrast might make it pattern more like the English [d]/[ɾ] pair. However, even in this discrimination task, Spanish listeners found [d]/[ɾ] to be more different than English listeners did. For Spanish listeners, these two sounds are allophones of different phonemes, so evidently this more abstract level of contrast influences perception even in this on-line reaction-time experiment.

With respect to task, the results from experiment 2 support the findings of Huang (2001, 2004) where cross-linguistic speech perception differences were found using a discrimination task. As noted above, it is commonly assumed in the L2 perception literature that "phonetic" listening tasks, such as discrimination, may obscure cross-linguistic speech perception differences (Werker and Logan, 1985; Huang, 2001 and 2004). The observation that the phonological relations of the pairs in each language impacted the discrimination of the sounds in experiment 2 thus provides further evidence that language-specific influences that emerge in an off-line task can also be observed in an on-line task.

One concern regarding experiments 1 and 2 is that the stimuli were produced by English speakers (linguists trained to be able to produce IPA symbols, but native English speakers nonetheless), and we were comparing responses of English-speaking listeners with those of Spanish-speaking listeners. In a post-test questionnaire the majority of the Spanish-speaking listeners identified the stimuli as having been produced by English speakers, presumably because the coronals were pronounced with an alveolar place of articulation rather than with the dental place of articulation used in the pronunciation of coronals in Spanish. As a result, the stimuli may have been less natural for Spanish listeners than they were for English listeners. (Interestingly though, the majority of the native English speakers did not identify the stimuli as English.) To address this concern, we conducted two further experiments identical to the first two, except that the stimuli were produced by Greek speakers, as opposed to American English speakers. Discussion of these experiments follows.

4. Experiment 3: Rating Greek [d], [ɾ], [ð] pairs

Experiments 3 and 4 replicate experiments 1 and 2 in almost every detail. The listeners were drawn from the same populations and the tasks were the same as in the first two experiments. The only difference was that new speech tokens were used in experiments 3 and 4. We were

interested to know whether the evidence for a role of phonemic contrast in speech perception could be replicated in an experiment with new stimuli.

An additional test inherent in these last two experiments has to do with two separable factors in speech perception. Experiments 3 and 4 manipulate one of these factors and hold the other constant, allowing us to examine the former's effect on speech perception. The first factor is the raw auditory/phonetic contrast between sounds. Thus, although [l] and [m], for example, are just as phonemically different from each other as are [p] and [m], we expect that listeners would rate the [p]/[m] contrast as more different than they would the [l]/[m] contrast because the auditory contrast between [p] and [m] is greater than that between [l] and [m]. The second factor is a language-specific mechanism of some sort that responds to speech in a way that is appropriate for, or trained by, the speech sounds and phonological patterns of a particular language. This factor operates the same way across the four experiments; that is, there are no changes in the linguistic identities of the stimuli (still intervocalic [ɾ], [ð], and [d]), and there are no changes in the characteristics of the populations of listeners being tested (though the actual participants were different in all four experiments). By rerunning experiments 1 and 2 with a new set of stimuli produced by speakers of a different language, we expect that the first factor, raw phonetic/auditory discriminability, of the stimuli may change. Comparing the results of experiments 1 and 2 with those of experiments 3 and 4 may thus help us identify aspects of the listeners' response patterns that are affected by the linguistic system of contrast, and pull these apart from aspects of the data that may be due solely to phonetic properties of the particular stimuli used in the test.

4.1. Methods

4.1.1. Stimuli

New stimuli were prepared for this experiment and for experiment 4. Materials consisted of two tokens of the same VCV sequences that were used in experiments 1 and 2: [ada], [aɾa], [aða], [idi], [iɾi], [iði], [udu], [uɾu], and [uðu]. Multiple tokens of these were produced and recorded by two native speakers of Greek, one male and one female, using a head-mounted microphone in a soundproof booth. Greek speakers were chosen because all three of the test phones, [d], [ɾ], and [ð], are contrastive in Greek and are produced naturally in intervocalic position. The speakers attempted to produce equal stress on the first and second syllables. In order to control the amplitude across tokens and speakers, the peak amplitude was equated for each of the tokens. The two best recordings for each VCV sequence were used as stimuli in the studies. These materials were used as stimuli in both experiment 3 and experiment 4.

4.1.2. Participants

Again, participants were drawn from the same pools as experiments 1 and 2. The native Spanish speakers (N = 7, 2 men, 5 women) had a mean self-rating of their English ability of 5.5. The native English speakers (N = 10, 3 men, 7 women) had a mean self-rating of their Spanish ability of 1.6.[1] None of the participants reported any history of speech or hearing disorders.

4.1.3. Procedure

The similarity rating procedure that was used in experiment 1 was also used in this experiment. Participants heard pairs of physically different stimuli and responded with a rating score from 1 (very similar) to 5 (very different).

4.2. Results

The results of experiment 3 are shown in figure 3; as in the graph from experiment 1, "more similar" is at the bottom of the graph and "more different" is at the top, and the means are plotted along with their 95% confidence intervals. These results were analyzed in the same way as those of experiment 1, reported in section 2.2, using a repeated measures analysis of variance on z-score normalized rating scores. There was not a significant main effect of pair ($F[2,30] =$ 2.389, $p > 0.05$). However, as in experiments 1 and 2, there was a significant pair by group interaction effect ($F[2,30] = 20.289$, $p < 0.05$). Subsequent planned comparison independent samples t-tests show that the English listeners rated the pair [d]/[ɾ] (which is allophonic in English) as more similar than did the Spanish speakers (for whom the pair is phonemic) ($t(15) =$ 4.652, $p < 0.05$). Similarly, Spanish listeners rated the pair [d]/[ð] (which is allophonic in Spanish) as more similar than did English listeners (for whom the pair is phonemic) ($t(15) =$ 5.162, $p < 0.05$). Finally, there was no significant difference between the two language groups in the rating of [ð]/[ɾ] ($t < 1$).

Insert Figure 3 about here

---

[1] Note that in this experiment, some of the English-speaking subjects in this experiment did in fact have some exposure to Spanish, unlike those in experiments 1 and 2. We included these participants because, in an experiment not reported on here (see Boomershine et al. 2004, 2005), we found no significant difference in responses to these stimuli by native English speakers who had anywhere from no Spanish experience to an intermediate level with a self-rating of 4.5 on a scale from 0-7. As is reported in that study, only native English speakers who are advanced Spanish speakers (with a self rating greater than 5) begin to approach the perceptual characteristics of the native Spanish speakers; the native English speakers with an advanced level of Spanish patterned almost identically to the native Spanish speakers in the discrimination task and in between the native English speakers with little or no experience in Spanish and native Spanish speakers in the rating task.

4.3. Discussion

The results from experiment 3 also provide evidence that allophonic relationships influence the perceived distance between sounds in phonological processing. The allophonic pairs for English listeners ([d] and [ɾ]) and for Spanish listeners ([d] and [ð]) were both rated as being more similar than the non-allophonic pairs. These results are very similar to those for experiment 1, which used the same task but involved different stimuli. One interesting difference between experiments 1 and 3 is that in experiment 1, the native Spanish speakers thought that the [ɾ]/[ð] distinction was the most salient, while in experiment 3, they found the [d]/[ɾ] distinction most salient. This is most likely due to the change in the raw perceptibility of the stimuli; in experiment 1, the stimuli were produced by native English speakers who perhaps did not make a particularly clear distinction between [d] and [ɾ], which are allophonic in English, but in experiment 3, the stimuli were produced by native Greek speakers, who do make a distinction between [d] and [ɾ] in production. The Spanish listeners found these stimuli, therefore, more perceptually distinct than those produced by English speakers.

5. Experiment 4: Discriminating Greek [d], [ɾ], [ð] pairs

5.1. Methods

5.1.1. Stimuli

The stimuli that were used for experiment 3 were also used in this experiment.

5.1.2. Participants

The participants in this experiment were drawn from the same pools as the other three experiments. The native Spanish speakers (N = 7, 4 men, 3 women) had a mean self-rating of their English ability of 6.1; the native English speakers (N = 11, 5 men, 6 women) had a mean self-rating of their Spanish ability of 1.18. None of the speakers reported any history of speech or hearing disorders.

5.1.3. Procedure

The speeded discrimination procedure that was used in experiment 2 was also used in this experiment. Participants heard pairs of stimuli and responded "same" if the stimuli were the same tokens of the same type of stimuli (e.g. the same token of [ada] twice) and "different" if they were not (either different types, e.g. [ada] – [aɾa], or different tokens, e.g. [ada] – [ada] where the two were not the same production). Participants were told after each pair whether they were correct or incorrect and were given their response time and their overall average percent correct, in order to encourage fast, accurate responses.

5.2. Results

The results and 95% confidence intervals for experiment 4 are shown in figure 4; as in figure 2, "more similar" is at the top of the graph and "more different" at the bottom. As with the reaction time data of experiment 2, reaction time is taken as a measure of perceptual distance, and each listener's reaction times are z-score normalized to remove individual differences in overall speed of responding.

```
--------------------------------
Insert Figure 4 about here
--------------------------------
```

In this experiment, there was not a significant effect of pair (F = 1.122). There was, however, a significant pair by group interaction (F[2,32] = 5.939, p < 0.05). Subsequent planned comparisons independent samples t-tests showed that as with experiment 3, the English listeners found [d]/[ɾ] more similar than the Spanish listeners did, though this difference was not quite significant in this particular experiment (t = 1.156). As in the previous experiments, too, the Spanish listeners found [d]/[ð] to be significantly more similar than the English listeners did (t(16) = 2.538, p < 0.05). Interestingly, there was also a trend in this experiment toward a difference between the two groups for the pair [ð]/[ɾ]; unlike all three of the other experiments, where the two groups had responded to this pair in the same way, in this experiment, the English listeners found [ð] and [ɾ] to be much more similar than the Spanish listeners did, though as with [d] and [ɾ], this difference was not quite significant (t(16) = 1.664).

5.3. Discussion

The results from experiment 4 again confirm our hypotheses about the role of allophony as opposed to phonemic contrast in perception: each pair was found to be less perceptually distinct by listeners for whom the pair is allophonic than by listeners for whom it is phonemic. The lack of significance between the two groups in the discrimination of [d]/[ɾ] may again be due to the raw auditory discriminability of the stimuli in this experiment as opposed to experiment 2, which used the same task but English-produced stimuli. That is, in experiment 4, perhaps the native English listeners found the Greek [d]/[ɾ] to be more distinct than the English [d]/[ɾ] of experiment 2 because the Greek [d] and Greek [ɾ] are inherently more different. The difference between the English and Greek stimuli might also explain why there was a (non-significant) tendency for Spanish speakers to find [ɾ]/[ð] more distinct than the English speakers did in experiment 4; if the Greek stimuli are acoustically more like Spanish phones, then perhaps the Spanish listeners simply had an easier time perceiving the difference than did the English

16

listeners. Further experimentation on the raw phonetic discriminability of all of these sounds needs to be carried out to confirm these conjectures. Importantly, however, the Spanish speakers still found the [d]/[ɾ] pair to be more distinct than did the English speakers, while the English speakers found the [d]/[ð] pair to be more distinct than did the Spanish speakers.

6. General Discussion and Conclusion

6.1. Discussion

In summary, all four experiments showed a similar pattern. Across languages, speakers of a language in which a particular pair of sounds is contrastive at a phonemic level perceive that pair as being more perceptually distinct than do speakers of a language in which the pair is not phonemically contrastive. In each of the experiments, the English speakers found [d]/[ð], which is a phonemically contrastive pair in English but allophonic in Spanish, to be more perceptually distinct than the Spanish speakers did. Similarly, the Spanish speakers found [d]/[ɾ], which is phonemically contrastive in Spanish but allophonic in English, to be more perceptually distinct than the English speakers did. The pair [ð]/[ɾ] had about the same level of perceptual distinctiveness in the two languages; recall that in each language one sound of the pair is in an allophonic relationship with a different sound that is also present in the inventory of the other language.

This pattern of results is interesting because while [d] and [ð] are both phonemically and surface contrastive in English, [d] and [ɾ] are only phonemically contrastive in Spanish since they do not contrast in any surface minimal pair (see (4)). It is not too surprising to find that the phonemic level of contrast was related to listeners' ratings of sound similarity in experiments 1 and 3, given that the rating task used in these two experiments encourages a degree of off-line contemplation of the sounds. However, the AX discrimination experiments reported here (experiments 2 and 4) used a "phonetic" listening task that was designed to tap an earlier level of processing in order to see possible effects of the presence or absence of surface contrast. In fact the AX speeded discrimination task is a common psychoacoustic task that is generally assumed to show phonetic responses, but here it apparently does not: the results of experiments 2 and 4 closely matched those of 1 and 3. This leads us to wonder if it would be possible in any listening task to see "phonetic" responding independent of phonological structure.[2]

---

[2] It should be noted that there have also been claims that shorter inter-stimulus intervals and a lower degree of uncertainty in the task may reduce language-specific effects (Polka 1991; Fox 1984). It would be interesting to see if these effects, however, can ever actually be eliminated from processing.

Of course, this is not to say that phonetic characteristics do not matter. One of the most noticeable differences between experiments 1 and 2 on the one hand and experiments 3 and 4 on the other was that the contrast between [d] and [ɾ] seemed to be much more salient in the second set of experiments, for both English and Spanish listeners. Because the only thing that changed between the two sets of experiments was the specific acoustic stimuli being used, we assume that this change in experimental materials created the difference in results; that is, the differences between the two sets of experiments (1 and 2 on the one hand, and 3 and 4 on the other) was due to the raw phonetic differences between the stimuli, *not* to differences in phonological patterning. The similarities between the two sets of results, on the other hand, are strongly tied to the phonological systems of the native languages of the listeners. Evidently the Greek [d] and [ɾ] tokens were more distinct from each other than were the American English [d] and [ɾ] tokens. Given the lack of a [d]/[ɾ] contrast in English and the presence of such a contrast in Greek, it makes sense to believe that the Greek speakers would be better at keeping them separate in production, which would then transfer over to a better ability by listeners to differentiate them.

It is also interesting to note that in all of the experiments, there was a tendency for English listeners to perceive [d]/[ð] as more distinct than [ɾ]/[ð] despite the observation that there is no apparent representational difference in English between the two pairs; the sounds in each pair are contrastive at phonemic as well as surface levels, as shown in (5) (repeated from (4)).

(5)   English [d]/[ð] versus [ɾ]/[ð]

	d/ð	ɾ/ð
Phonemic contrast	/d/ - /ð/	/d/ - /ð/
Surface contrast	[d] - [ð]	[ɾ] - [ð]

It may be that this tendency is simply a result of the raw overall auditory qualities of the sounds in question, an issue that must be explored by further research. It is also possible, however, that the difference is due to the fact that /d/ and /ð/ are each phonemes of English in a traditional analysis, while [ɾ] is simply an allophone of /d/. Although this difference is not indicated by the representations given in (4) or (5), perhaps the notion of contrast is even more finely nuanced than we have shown here. Again, we leave this question to later research.

In sum, the data presented in this paper suggest that phonemic contrast strongly influences speech perception, and that surface phonetic detail influences perceptual discrimination judgments. These results are important in that any model of speech perception must account for

them, making sure that the phonemic level of representation is kept distinct from the allophonic level, with the phonemic level resulting in more distinct perceptual contrasts than the allophonic level. There are multiple perceptual models that achieve or could achieve this result; we outline two of them below: a phonological inferencing model and a lexical processing model.

6.2. Modeling the role of allophony and contrast in speech perception

In Gaskell and Marslen-Wilson's (1998) phonological inferencing model of speech perception, the acoustic signal is perceived in terms of the phonological representation that produced it. For instance, suppose that in Spanish there is a lenition rule that changes an underlying stop /d/ into the fricative [ð], and that the word "donde" 'where' has the abstract lexical representation /donde/. With these assumptions, this type of model predicts that the lenition rule is "undone" during the perception of "de [ð]onde" 'from where,' producing a formal (phonetic/phonological) representation that matches the lexical representation. Thus, the prediction is that in a language with an allophonic relation between [d] and [ð], the acoustic signal of "de [ð]onde" is perceived exactly like that of "[d]onde." The difference between Spanish and English is that English has no such rule, so that the perception of [ð]onde would not be subject to the undoing of such a rule, and there would be a distinction between the signals [d]onde and [ð]onde. Hence English speakers are correctly predicted to find [d] and [ð] more distinct than Spanish speakers. Note that the same argument can be made, with the role of the languages reversed, for the relation between [d] and [ɾ] – English listeners undo a flapping rule and perceive [d] and [ɾ] as the same, while Spanish listeners have no such rule and perceive [d] and [ɾ] as distinct. In either case, the distinction between [ð] and [ɾ] would be correctly predicted to pattern similarly in the two languages, because in each language, each of these phones is mapped to a different phonological representation.

Interestingly, this model also predicts that the difference between different realizations of [d] will be indistinct from [ð] in Spanish or [ɾ] in English, as each sound is immediately linked to its underlying phonological representation. This prediction was indirectly tested in the rating experiments (experiments 1 and 3) by the comparison of ratings of [d]/[ð] to [d]/[d] and [ð]/[ð] pairs in Spanish and the comparison of ratings of [d]/[ɾ] to [d]/[d] and [ɾ]/[ɾ] pairs in English. In both sets of comparisons, it was found that the pair containing two different articulatory realizations of the same phoneme (e.g. [d]/[ð] or [d]/[ɾ]) was rated as significantly more different than the pairs containing the same articulatory realizations (e.g. [d]/[d], [ð]/[ð], or [ɾ]/[ɾ]). This result was found in experiment 1 with the English-produced stimuli, in which planned comparison paired samples t-tests showed that for Spanish listeners, the difference between

[d]/[ð] and [d]/[d] was significant [t(9) = -7.45, p < 0.05], as was the difference between [d]/[ð] and [ð]/[ð] [t(9) = 9.403, p < 0.05]. Similarly, for English listeners the difference between [d]/[ɾ] and [d]/[d] was significant [t(17) = -7.324, p < 0.05], as was the difference between [d]/[ɾ] and [ɾ]/[ɾ] [t(17) = 7.558, p < 0.05]. The same pattern was found in experiment 3 with the Greek-produced stimuli, where for Spanish listeners, [d]/[ð] versus [d]/[d] was significantly different [t(6) = 12.304, p < 0.05], as was [d]/[ð] versus [ð]/[ð] [t(6) = 11.072, p < 0.05]. For the English listeners in experiment 3, the comparison of [d]/[ɾ] versus [d]/[d] was significantly different [t(9) = 12.613, p < 0.05], as was [d]/[ɾ] versus [ɾ]/[ɾ] [t(9) = 12.260, p < 0.05]. While the inferencing model can thus account for the differences found across languages in the comparison of allophonic versus phonemic pairs, it is not powerful enough to correctly predict perceptual differences for the different types of "allophones of the same phoneme" found within a single language.

In a lexical processing model, on the other hand, both types of results are predicted. In this approach, differences between phonological representations come at the lexical level, once listeners have tried to access words themselves, rather than being a property of the signal-to-representation mapping.

One type of lexical processing model is an exemplar model (see, e.g., Goldinger 1992, 1996; Palmeri et al. 1993; Johnson 1997a, b, 2004; Coleman 2002; Pierrehumbert 2003; Hawkins 2003). In an exemplar model, grammar is an emergent property over stored exemplars of all utterances heard. Word recognition is achieved by matching an incoming acoustic signal with the most similar stored representation of the signal, as defined by the amount of activation of the various stored representations. Hence an incoming [d] will activate stored examples of [d] more than it will activate stored examples of, say, [z], and so it will be recognized as [d]. Allophonic relations in this kind of model are represented by high co-activation (Johnson, in press). For example, in Spanish, an incoming [d] will activate both [d] and [ð] because there are words that variably contain each different pronunciation. In English, on the other hand, [d] will activate [d] and [ɾ], but not [ð]. High rates of co-activation will make two sounds less perceptually distinct; the results of the experiments here would therefore be correctly predicted. Further, as with the phonological inferencing model, [ð] and [ɾ] are correctly predicted to pattern similarly across the two languages; in this case, because they are not activated by the same incoming signals. That is, in English, [ð] is activated only by an incoming [ð], and [ɾ] is activated by an incoming [d], [t], or [ɾ]. Thus, the signals that activate [ð] and [ɾ] do not overlap. Similarly for Spanish, [ð] is

activated by an incoming [d] or [ð], while [ɾ] is activated by an incoming [ɾ]; the activation signals are again non-overlapping.

An exemplar model also predicts that even though an incoming [d] in Spanish will activate both [d] and [ð], as will an incoming [ð], the perception of a [d]/[d] pair will differ from that of a [d]/[ð] pair. This result comes about for a number of reasons. First, the acoustic representation of an incoming signal is not completely removed; every utterance is stored with its acoustic representation intact, and similarity between signals is calculated over these acoustic representations. Second, the words that are activated by an incoming signal will depend on this similarity matching. Consequently, the words activated by an incoming [d] might be somewhat different than those activated by an incoming [ð], and words that are activated by both signals may be activated to a greater or lesser extent by one than the other. This use of the acoustic representation of the signal in activating words in the lexicon allows such a model to predict both the difference in the perception of phonemic and allophonic pairs across languages as well as the difference in the perception of pairs of allophones of the same phoneme within a language.

In summary, while both phonological inferencing and exemplar models are able to correctly predict the differing influences of allophony versus phonemic contrast on perception (a phonemic relationship is perceptually more distinct than an allophonic relationship, regardless of the actual identity of the sounds in question), only the exemplar model is successful in accounting for the differences in perception that one finds within a language between pairs of the "same" allophone of one phoneme (e.g. [d]/[d]) and "different" allophones of one phoneme (e.g. [d]/[ð]).

Returning to the speculations of Trubetzkoy, we see this study as providing further evidence for his claim that the particular phonological relation holding between sounds in a language has an impact on a listener's perception of those sounds. Our direct test of allophony versus contrast points to the need to include both in the inventory of phonological relations shown to influence perception. The inventory thus includes phonemic contrast, partial contrast due to phonological neutralization, the non-contrastive relation of allophony, as well as non-contrastiveness due to the absence of one or more of the sounds in a language's sound system. The extent to which the two types of non-contrastiveness differ with regards to their impact on speech perception remains an open question and one that must be addressed in future research.

References

Best, Catherine   (1994) The emergence of native-language phonological influences in

infants: A perceptual assimilation model. In Judith Goodman and Howard Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167-224). Cambridge, MA: MIT Press.

Best, Catherine, McRoberts, Gerald, and Sithole, Nomathemba (1988) Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 345-360.

Boomershine, Amanda, Hall, Kathleen Currie, Hume, Elizabeth, and Johnson, Keith (2004, 30 October) *The influence of contrast vs. allophony on perception: The case of Spanish and English.* Paper presented at the Mid-Continental Workshop on Phonology, Northwestern University.

Boomershine, Amanda, Hall, Kathleen Currie, Hume, Elizabeth, and Johnson, Keith (2005, 10 November) *The influence of language experience on speech perception: The case of Spanish and English.* Paper presented at The Hispanic Linguistics Symposium, Penn State.

Coleman, John (2002) Phonetic representations in the mental lexicon. In Jacques Durand and Bernard Laks (Eds.), *Phonetics, phonology, and cognition* (pp. 96-130). Oxford: Oxford University Press.

Dupoux, Emmanuel, Pallier, Christophe, Sebastian, Nuria, and Mehler, Jacques (1997) A destressing 'deafness' in French? *Journal of Memory and Language, 36*, 406-421.

Fox, Robert (1984) Effect of lexical status on phonetic categorization. *Journal of Experimental Psychology: Human perception and performance*, *10*(2), 526-540.

Francis, Alexander, and Nusbaum, Howard (2002) Selective attention and the acquisition of new phonetic categories. *Journal of Experimental Psychology: Human Perception and Performance, 28*(2), 349-366.

Gaskell, M. Gareth, and Marslen-Wilson, William (1998) Mechanisms of phonological inference in speech perception. *Journal of Experimental Psychology: Human Perception and Performance, 24*(2), 380-396.

Gaskell, M. Gareth, and Marslen-Wilson, William (2001) Lexical ambiguity resolution and spoken word recognition: Bridging the gap. *Journal of Memory and Language, 44*(3), 325-349.

Goldinger, Stephen (1996) Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(5), 1166-1183.

Goto, Hiromu (1971) Auditory perception by normal Japanese adults of the sounds "l" and "r." *Neuropsychologia, 9*, 317-323.

Harnsberger, James (2001) The perception of Malayalam nasal consonants by Marathi, Punjabi, Tamil, Oriya, Bengali, and American English listeners: A multidimensional scaling analysis. *Journal of Phonetics, 29*, 303-327.

Hawkins, Sarah (2003) Roles and representations of systematic fine phonetic detail in speech understanding. *Journal of Phonetics, 31*(3-4), 373-405.

Huang, Tsan (2001) The interplay of perception and phonology in tone 3 sandhi in Chinese Putonghua. In Elizabeth Hume and Keith Johnson (Eds.), *Studies on the interplay of speech perception and phonology* (Vol. 55, pp. 23-42). Columbus, OH: Ohio State University Working Papers in Linguistics.

Huang, Tsan (2004) *Language-specificity in auditory perception of Chinese tones.* Unpublished PhD dissertation, The Ohio State University, Columbus, OH.

Hume, Elizabeth, and Johnson, Keith (2001) A model of the interplay of speech perception and phonology. In E. Hume and K. Johnson (Eds.), *The role of speech perception in phonology* (pp. 3-26). San Diego: Academic Press.

Hume, Elizabeth, and Johnson, Keith (2003) The impact of partial phonological contrast on speech perception. *Proceedings of the Fifteenth International Congress of Phonetic Sciences* (pp. 2385-2388).

Johnson, Keith (1997a) Speech perception without speaker normalization. In Keith Johnson and John Mullennix (Eds.), *Talker variability in speech processing* (pp. 145-165). San Diego: Academic Press.

Johnson, Keith (1997b) The auditory/perceptual basis for speech segmentation. In Kim Ainsworth-Darnell and Mariapaola D'Imperio (Eds.), *Papers from the linguistics laboratory* (Vol. 50, pp. 101-113). Columbus, OH: Ohio State University Working Papers in Linguistics.

Johnson, Keith (2004) Cross-linguistic perceptual differences emerge from the lexicon. In Augustine Agwuele, Willis Warren, and Sang-Hoon Park (Eds.), *Proceedings of the 2003 Texas Linguistics Society Conference: Coarticulation in speech production and perception* (pp. 26-41). Sommerville, MA: Cascadilla Press.

Johnson, Keith (in press) Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics.*

Lahiri, Aditi   (1999) Speech recognition with phonological features. In *Proceedings of the XIVth International Congress of Phonetic Sciences* (pp. 715-718). San Francisco.

MacKain, Kristine, Best, Catherine, and Strange, Winifred (1981) Categorical   perception   of English /r/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics, 2*, 369-390.

Marian, Viorica, and Spivey, Michael   (2003) Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition, 6*(2), 97-115.

Mohanan, Karuvannur, and Mohanan, Tara (1984) Lexical phonology of the consonant system in Malayalam. *Linguistic Inquiry, 15*(4), 575-602.

Palmeri, Thomas, Goldinger, Stephen, and Pisoni, David   (1993) Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*(2), 309-328.

Pierrehumbert, Janet (2003) Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech, 46*, 115-154.

Polka, Linda   (1991) Cross-language speech perception in adults: Phonemic, phonetic, and acoustic contributions. *Journal of the Acoustical Society of America, 89*, 2961-2977.

Polka, Linda, and Werker, Janet   (1994) Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 421-435.

Pruitt, John, Akahane-Yamada, Reiko, and Strange, Winifred   (1998) Perceptual assimilation of Hindi dental and retroflex consonants by native speakers of Japanese and English. *Proceedings of the Joing Meeting of ASA/16th ICA*.

Strange, Winifred   (1995) Cross-language studies of speech perception: A historical review. In W. Strange (Ed.), *Speech perception and linguistic experience* (pp. 3-45). Baltimore: York Press.

Strange, Winifred, and Dittman, Sibylla   (1984) Effects of discrimination training on the perception of /r-l/ by Japanese adults learning English. *Perception and Psychophysics, 36*(2), 131-145.

Takane, Yoshio and Sergent, Justine (1983) Multidimensional scaling models for reaction times and same different judgments. *Psychometrika, 48*, 393-423.

Werker, Janet, Gilbert, John, Humphrey, Keith, and Tees, Richard (1981) Developmental aspects of cross-language speech perception. *Child Development, 52*, 349-353.

Werker, Janet, and Logan, John   (1985) Cross-language evidence for three factors in speech perception. *Perception and Psychophysics, 37*, 35-44.

Figure Captions

Figure 1. Results of experiment 1. Normalized similarity rating of [d], [ð], and [ɾ] by Spanish-speaking and English-speaking listeners.

Figure 2. Results of experiment 2. Normalized reaction times for speeded discrimination [d], [ð], and [ɾ] by Spanish-speaking and English-speaking listeners.

Figure 3. Results of experiment 3. Normalized similarity rating of [d], [ð], and [ɾ] by Spanish-speaking and English-speaking listeners. Stimuli produced by Greek speakers.

Figure 4. Results of experiment 4. Normalized reaction times for speeded discrimination [d], [ð], and [ɾ] by Spanish-speaking and English-speaking listeners. Stimuli produced by Greek speakers.
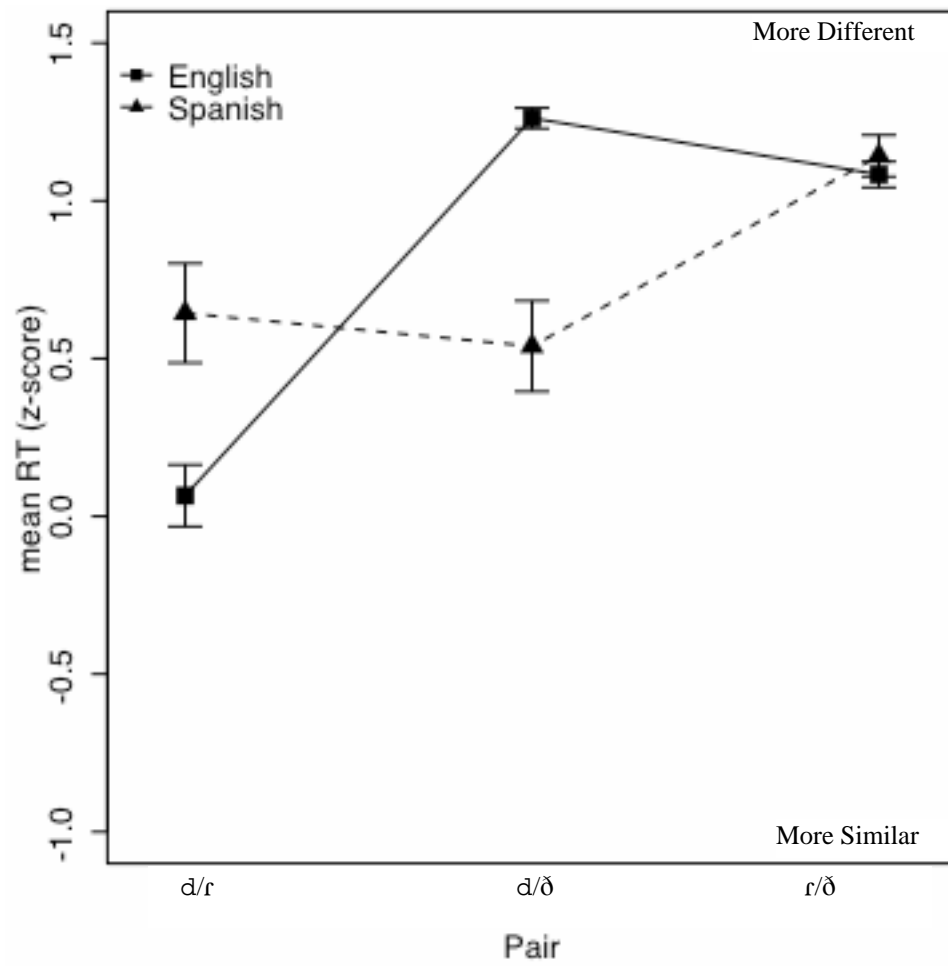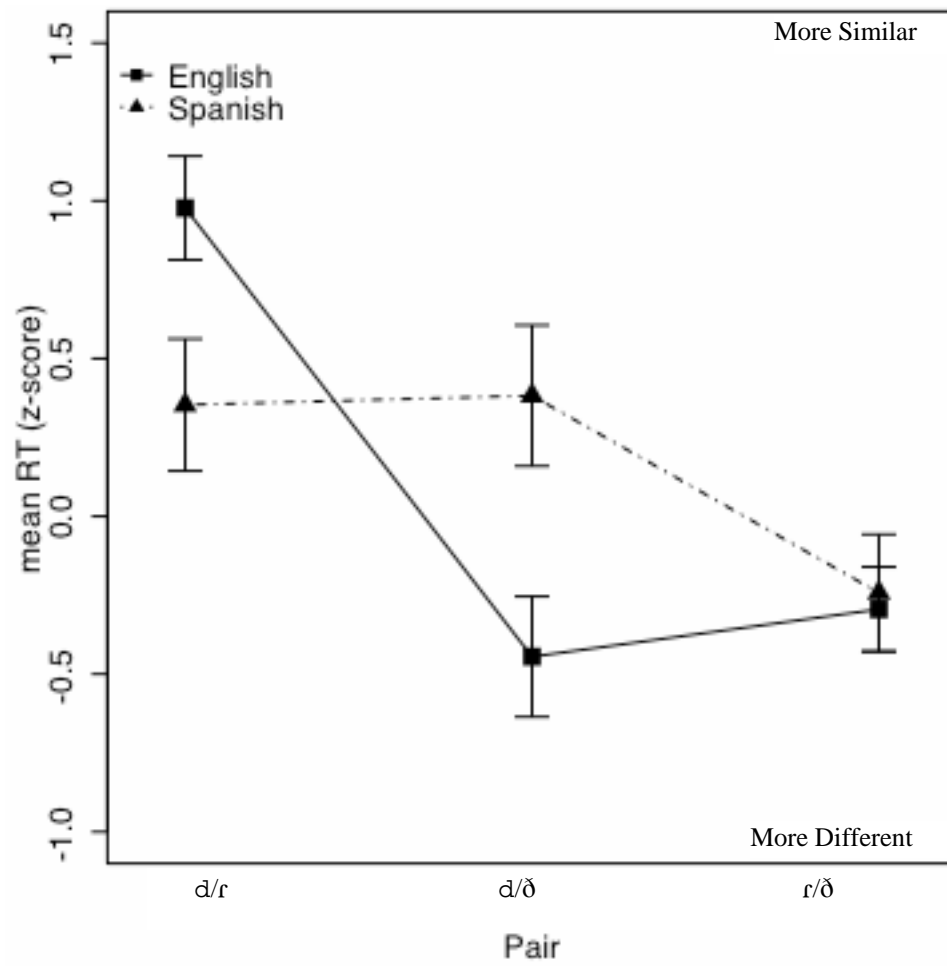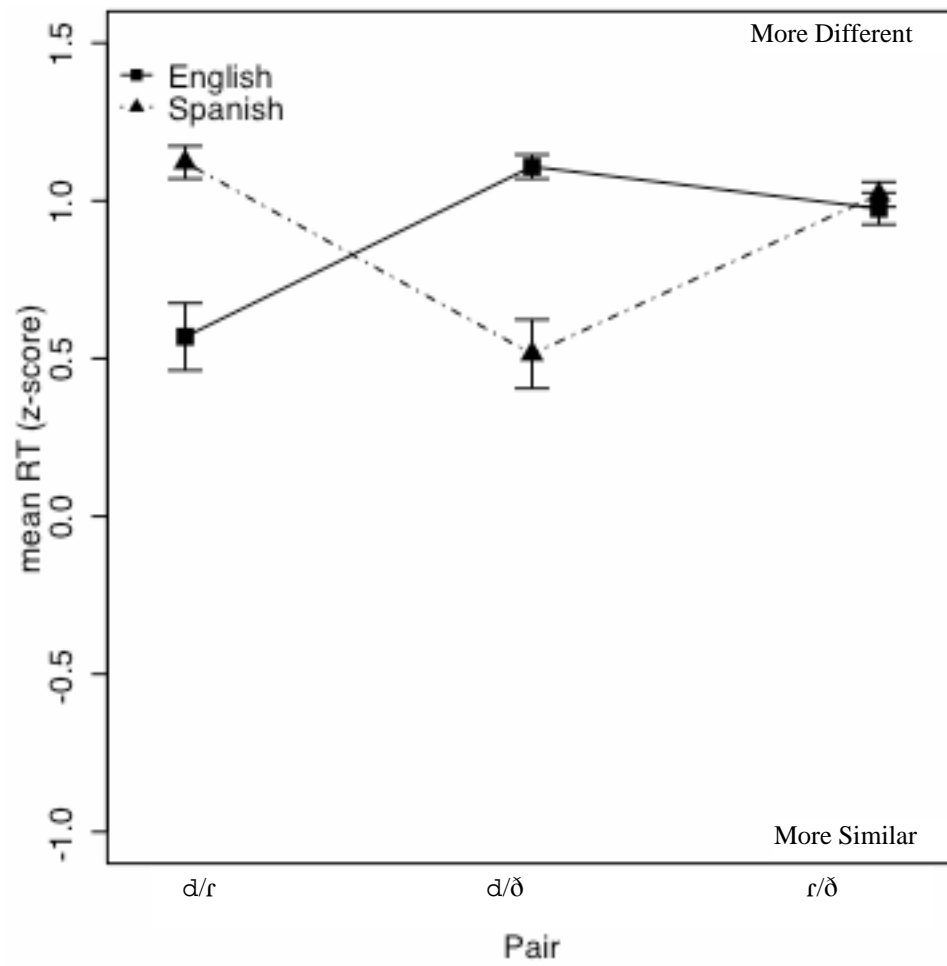
Figure 1:

Figure 1:

Figure 2:

Figure 3:

Figure 4: