Use of Segmental and Prosodic Cues in Adult Language Discrimination: Effects of Language Experience and Speaking Style

Crystal Chow
University of Toronto Mississauga

1 Introduction

In a multilingual environment, listeners are tasked with recognising language-specific cues to help them discriminate one language from another. By isolating various aspects of the speech signal – in particular, segments and prosody – the use of and reliance on these cues during a language discrimination task can be analysed and interpreted with respect to understanding adult language processing. We explore the effects of language experience and speaking style on listeners' abilities to accurately differentiate Mandarin Chinese from Malay.

1.1 *Background* Many acoustic cues exist in the speech signal for listeners to make use of when processing language. Lexical information (i.e. the words of a language) and segmental information (i.e. the phonemic inventory of a language) are two of the more prominent cues that can be used by adult listeners. Specifically, the adult listener's familiarity with the words and sounds of their own language (or languages) can facilitate the process of language discrimination. Such information is important for the categorisation of input utterances into the respective languages to which they belong.

The lexical and segmental cues that are used by adults are not available to infants who have not yet acquired the words and sounds of their first language(s). Prior work has shown that, aside from these obvious cues, languages also differ more subtly in their rhythmic properties (Grabe & Low, 2002) and in their tonal properties (Maddieson, 2013). Researchers have used this knowledge of language-specific suprasegmental differences to prove that infants are able to tell languages apart, even without segmental information.

For example, Nazzi et al. (1998) showed that French-speaking newborns were able to discriminate between two rhythmically different languages, English and Japanese. Their stimuli consisted of sentences that had been low-pass filtered to degrade segmental information while preserving prosody. Results from their study indicate that languages – even those that are unfamiliar to the participants (e.g. discrimination between English and Japanese sentences by French newborns) – can be differentiated from each other by prosody only, as long as they are rhythmically different. In another infant language discrimination experiment, Molnar et al. (2013) used low-pass filtered stimuli to find that 3.5-month-old monolingual Basque, monolingual Spanish, and bilingual Basque-Spanish infants were able to discriminate between the two rhythmically similar languages. Without words and segments in the stimuli, other prosodic cues (such as linguistic rhythm) could still be perceived; thus, the abilities of the monolingual infants to discriminate these languages could be attributed to a reliance on native language familiarity with the prosodic cues of one of the two test languages. However, the bilingual infants' discrimination implies that they must have been making use of other suprasegmental differences (besides linguistic rhythm, which is said to be similar between Basque and Spanish) to complete the task. Overall, the abilities of newborns and infants to successfully discriminate languages when hearing low-pass filtered stimuli suggests that, aside from lexical and segmental information, there must be other (suprasegmental) cues in the acoustic signal that listeners can use to differentiate languages.

Adult language discrimination has been studied to a lesser degree. In one such study, Ramus and Mehler (1999) tested French adult listeners on their abilities to discriminate between English and Japanese. They used a speech resynthesis technique to delexicalize the utterances by replacing all phonemes with pseudo-randomly selected phonemes (e.g. the English phrase "the truck" was resynthesised into /satlat/). This modification of the signal resulted in four manipulation conditions for the stimuli, in each of which three types of suprasegmental cues – segments, tone, and rhythm – were preserved or degraded. The French-speaking adults performed well in

* I am grateful to Jessamyn Schertz for her guidance and support. I would also like to thank Gauri Chaudhari, Sarah Khan, Anna Lyashenko, Nur Sakinah Nor Kamal, and Hui Zhong Zhu for their help in data collection and annotation. This work was supported by an Undergraduate Research Grant at the University of Toronto Mississauga.

three of the four conditions: in one condition when the segments, tone, and rhythm were all preserved, in another condition when the tone and rhythm were preserved but the segmental information was degraded, and in a third condition when only the rhythm was preserved but both segmental and tonal information were degraded. However, they were not able to do the task in a fourth condition when only the tone was preserved but both segmental and rhythmic information were degraded. Without segmental information, the listeners were able to differentiate the two languages in the rhythm-only condition but not in the tone-only condition; this implies that rhythmic information alone is sufficient for French listeners to use as a cue when discriminating between English and Japanese. To investigate a potential effect of participant language background on accuracy in language discrimination, Ramus and Mehler recruited a group of native English speakers (i.e. speakers who were familiar with one of the two test languages) to do the task in the same tone-only condition that the French speakers had not been able to do. They found that the English speakers performed well in this condition, which suggests that listeners might make more use of a certain cue if they have had experience with (e.g. as a speaker or learner of) a language. In other words, the English speakers' native language familiarity with the intonational properties of English could have been the reason for their reliance on the tonal cue, which did not prove to be useful for the French speakers.

1.2 *Research questions and hypotheses* The implications of infant language discrimination on language acquisition and development have been researched extensively, but fewer studies have investigated the implications of adult language discrimination on cue use in language processing. We conducted an adult language discrimination experiment to investigate listeners' use of different aspects of the acoustic signal when differentiating between two languages, Mandarin Chinese and Malay. To assess the dependence of accuracy in language discrimination on segmental, tonal, and rhythmic cues, we proposed the following research questions: (1) Which acoustic signals – segments, tone, rhythm – do adult listeners depend on when differentiating between languages? (2) How is the reliance on cues affected by individual language backgrounds? and (3) How is the reliance on cues affected by different speech styles (careful versus casual)?

We presented listeners with Mandarin and Malay sentences that had been modified to preserve or degrade segmental or tonal information. By comparing the listeners' performance across different experimental conditions, we were able to assess the extent of reliance on segmental, tonal, and rhythmic cues. Following the existing literature on both infant and adult language discrimination, we hypothesised that the listeners in this experiment will perform best when all information (segments, tone, and rhythm) is present, but that they will still be able to perform above chance with only suprasegmental cues. Specifically, the linguistic rhythm cue alone should be sufficient for successful language discrimination, even for listeners with no exposure to either test language (Ramus & Mehler, 1999).

Furthermore, we explored the effect of language experience (knowing one of the test languages) on the extent of reliance on cues. Previous studies have shown that native language familiarity can be a factor influencing performance on the task (e.g. the English speakers in Ramus and Mehler's (1999) English-Japanese language discrimination experiment performed better in the tone-only condition than the French speakers). Thus, we predicted that listeners with prior exposure to one of the two test languages (i.e. first-language (L1) speakers of Mandarin and of Malay) will perform better overall in the task than listeners with no exposure to either of the test languages (i.e. L1 speakers of English). Based on the findings of Ramus and Mehler (1999), we also predicted that the tonal cue will only be used by the L1 Mandarin and L1 Malay speakers, who are familiar with one of our test languages. In particular, we expected the Mandarin-speaking listeners to show the most reliance on the tonal cue in comparison to the other two listener groups. This could be attributed to the fact that tone is a contrastive feature in Mandarin, but not in either Malay or English. Generally, speakers of tonal languages may weight the tonal cue more heavily, since it has an important, phonemic role in speech processing.

Lastly, we examined any differences in accuracy based on speaking style (conversational versus read speech). Prior work has focused on a single speech style for the experimental stimuli: either casual speech (e.g. Bosch & Sebastian-Galles, 1997) or careful speech (e.g. Ramus & Mehler, 1999). Since conversational speech is heard more often in the linguistic environment, we expected the participants to perform better in the task when they were presented with conversational speech than with read speech.

2 Methodology

An AX discrimination experiment was run on PsychoPy, where the test languages were Mandarin Chinese and Malay. Listeners heard pairs of stimuli in various conditions and were asked to decide whether the two sentences that they heard belonged to the same or different languages.

2.1 *Recording procedure* Four native speakers of Mandarin (two male, two female) and four native speakers of Malay (two male, two female) were recruited to record the stimuli for the experiment. Each speaker filled out a consent form, a compensation form, and a language background form, and they were all paid for their contributions. They were paired by gender within the same language group and were spoken to by an experimenter who was fluent in the relevant language (i.e. Mandarin or Malay). In each recording session, the speakers were asked to read materials and converse in their languages. First, they read an excerpt from respective translations of "The Little Prince" (5 minutes). Then, they had a 15-minute unstructured conversation with the other participant, where they were given some speaking prompts (translated into both languages) in case they were unable to think of topics on their own. Lastly, they repeated the same excerpt from "The Little Prince" (5 minutes). Prior to beginning the recording procedure, the participants were given some time to look over and familiarise themselves with the reading passages.

2.2 *Annotation* Segmentation and annotation of the recordings were carried out using TextGrids on Praat. Recordings of the conversation and one of the reading passages were segmented into prosodically-delineated phrases, which were separated by pauses. Boundaries were placed around each phrase, which was given a number from 1 to *n*. Long pauses and apparent hesitations were omitted from the segmentations. Laughter, overlapped speech, and any code-switching (with languages other than the target language) were omitted from the segmentations of the conversations. For the conversations, another tier was created on the TextGrid to concatenate shorter phrases into longer ones, with the following criteria: if two consecutive phrases were less than 2 seconds each and were separated by a pause less than 0.5 seconds long, they were combined into a longer phrase. The recordings were segmented by annotators who were not familiar with the target language.

2.3 *Manipulation conditions* All stimuli were manipulated with four filtering conditions: (1) natural – preserve segments, tone, rhythm; (2) low-pass filtered – preserve tone and rhythm, degrade segmental information; (3) f0-flattened – preserve segments and rhythm, degrade tonal information; and (4) filtered-flattened – preserve rhythm, degrade both segmental and tonal information. Manipulations were done using Praat, using the PSOLA algorithm with the standard settings and a 75 Hz/500 Hz floor/ceiling. Short, prosodically-delineated phrases between a duration range of 2.5 to 6 seconds were taken from the read and conversation recordings.

This duration range (2.5 to 6 seconds) was chosen as the narrowest range of duration (in order to minimise duration-based variability across stimuli) for which enough stimuli could be found. The cut-off frequency for low-pass filtering was 450 Hz and was chosen as the highest frequency for which there was no audible segmental information. The f0 values for the flattened conditions differed by speakers and genders and were chosen in the following way: for each speaker, the average pitch was taken from two randomly chosen phrases; then, another average value was taken from the averages for the speakers of the same gender within a language group. The intensity was normalised to 55 Hz for the natural and f0-flattened conditions, and to 75 Hz for the other two filtered conditions (since low-pass filtered sounds are less loud than non-filtered sounds of the same intensity).

2.4 *Participants* L1 English speakers (n=12) with no prior exposure to Mandarin or Malay, L1 Mandarin speakers (n=12) with no prior exposure to Malay, and L1 Malay speakers (n=12) who do not speak Mandarin were recruited to complete an AX discrimination experiment. Each listener filled out a consent form, a compensation form, and a language background form, and they were all paid for their contributions. Upon hearing each pair of phrases within the experiment, the participants were asked to decide whether the utterances that they heard came from the same or different languages. For the "same" and "different" responses, they were instructed to press the corresponding labelled buttons on the keyboard.

2.5 *Procedure* The actual experiment was preceded by a practice set with three pairs of English and Russian sentences, which were recorded by a native English speaker and a bilingual Russian-English speaker. Short, prosodically-delineated phrases were taken from recordings of the speakers, who read excerpts of "The Little Prince" in the respective translations.

Listeners heard a total of 128 trials (excluding the practice set) and were given feedback on their accuracy after every trial throughout the experiment. Feedback consisted of: (1) whether they got the correct answer on the previous trial, and (2) overall accuracy for the block. For each of the read and conversational speech styles, there was an initial training block (with phrases in the natural condition) followed by three experimental blocks (with phrases in each of the three filtering conditions: filtered, flattened, filtered-flattened); this resulted in a total of 8 blocks ([1 training + 3 experimental] * 2 styles). The participant number determined which of the 12 possible orders (6 experimental block orders * 2 style orders) were presented during the experiment.

The design was balanced such that each listener heard each pair of phrases (and thus, each stimulus) once; each stimulus was heard in a different filtering condition among all of the orders. All participants heard a randomised order of the same set of trials throughout the entire experiment (i.e. the pairing of two phrases was fixed, while the order of these pairings was randomised). The training block consisted of the same sets of trials for all participants, but the order was randomised. Among the three experimental blocks, each of the remaining stimuli (that were not heard during the training block) were heard in a different filtering condition by the groups. Each block contained an equal number of “same” and “different” stimuli, and the order of the ‘A’ and ‘X’ stimuli were randomised by trial.

2.6 *Analysis* Data were analysed using a logistic mixed-effects regression model in R, using the *lme4* package. The response variable of analysis was accuracy in language discrimination (0=incorrect, 1=correct). Predictor variables were language group (English, Malay, Mandarin), filtering condition (natural, (f0-)flattened, filtered, and filtered-flattened), and speech style (read, conversational). Language group was Helmert-coded, such that two comparisons were tested: (1) English vs. Malay/Mandarin, and (2) Malay vs. Mandarin. Filtering condition was forward difference coded, testing the following three comparisons: (1) natural vs. flattened, (2) flattened vs. filtered, and (3) filtered vs. filtered-flattened. Speech style was simple-coded, with Read as the reference level. An alpha-level of 0.05 was used to determine significance of main effects and interactions. When appropriate, follow-up tests were done using the *phia* package in R.

3 Results

Graphs of the results are shown in Figures 1-4, and the statistical results are given in Table 1.

	β	*SE*	*z*	*p*
Intercept	2.07640	0.10197	20.363	<.001 ***
Lang1 (English vs. *Mandarin/Malay*)	-1.62566	0.17410	-9.338	<.001 ***
Lang2 (Malay vs. *Mandarin*)	-0.12755	0.26009	-0.490	0.6239
Cond1 (Natural vs. *Flat*)	0.17678	0.30604	0.578	0.5635
Cond2 (Flat vs. *Filtered*)	2.00938	0.24302	8.268	<.001 ***
Cond3 (Filtered vs. *Filtered-Flattened*)	1.00422	0.10632	9.445	<.001 ***
Style (Read vs. *Conversational*)	0.10767	0.17261	0.624	0.5328
Lang1 * Cond1	0.05686	0.46337	0.123	0.9023
Lang2 * Cond1	-1.54241	0.83627	-1.844	0.0651 .
Lang1 * Cond2	-1.84033	0.41919	-4.390	<.001 ***
Lang2 * Cond2	1.01768	0.66554	1.529	0.1262
Lang1 * Cond3	-1.05882	0.21546	-4.914	<.001 ***
Lang2 * Cond3	-0.48354	0.26756	-1.807	0.0707 .
Lang1 * Style	-0.03039	0.27976	-0.109	0.9135
Lang2 * Style	-0.34752	0.45883	-0.757	0.4488
Cond1 * Style	0.34194	0.58402	0.585	0.5582
Cond2 * Style	-0.34286	0.42402	-0.809	0.4188
Cond3 * Style	0.41863	0.19513	2.145	0.0319 *

Table 1: Statistical results from a logistic mixed-effects regression model predicting accuracy from language group, filtering condition, and speech style. Only main effects and two-way interactions are shown. No three-way interactions were significant. Reference levels are shown in italics. The comparison Lang1 compares English to Mandarin/Malay, while the comparison Lang2 compares Malay to Mandarin. Significant or trending results are shaded.

3.1 *Effect of language group* Overall, across both speech styles (read, conversational) and all four filtering conditions (natural, f0-flattened, low-pass filtered, filtered-flattened), all listeners performed above chance (accuracy>50%). Figure 1 shows the effect of language group (L1 English, L1 Malay, L1 Mandarin) on accuracy in language discrimination (in percentages). The L1 English speakers were less accurate overall than either the L1 Malay or L1 Mandarin group. There was no significant difference in overall accuracy between the L1 Malay and L1 Mandarin speakers.

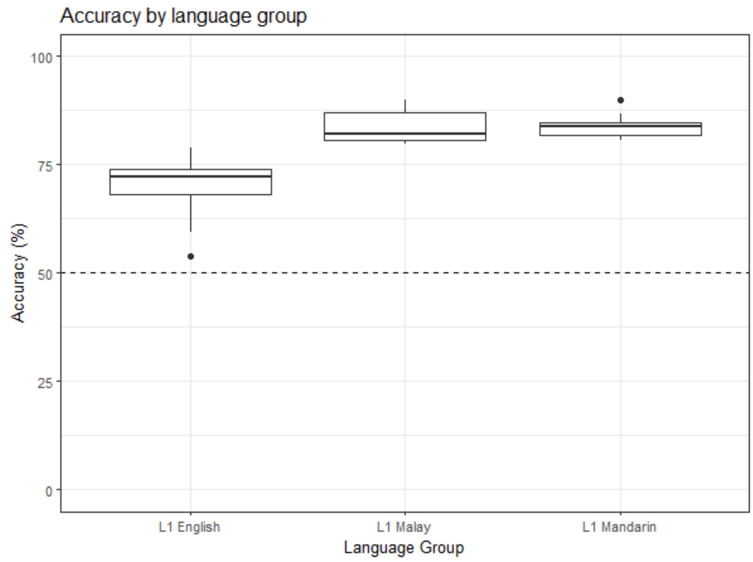

Figure 1: Overall accuracy (shown in percentages) by the three language groups.

3.2 *Effect of filtering condition* Figure 2 shows the effect of the four filtering conditions on accuracy in language discrimination, split by language group. In the natural condition (when all cues were present), listeners performed well above chance. Degrading the tonal information in the flattened condition had no effect on their accuracy – there was no significant difference in performance between the natural and flattened conditions. However, the accuracies of all three listener groups decreased in the filtered condition, when the segmental information was degraded. The difference between the flattened and filtered conditions was significant for all language groups, but was significantly bigger for the L1 Malay and L1 Mandarin groups than for the L1 English group. Accuracies decreased further in the filtered-flattened condition, when both segmental and tonal information were degraded. Follow-up tests showed that there was a trending but non-significant difference in the English speakers' performance between the filtered and filtered-flattened conditions (χ^2=3.3, p = .07), but this difference was significant for the other two language groups (Malay: χ^2=45.9, p<.001; Mandarin: χ^2=67.5, p<.001). Furthermore, the difference was larger for the L1 Mandarin than for the L1 Malay group, though the language group by condition interaction estimating this difference was only trending (p=0.07).

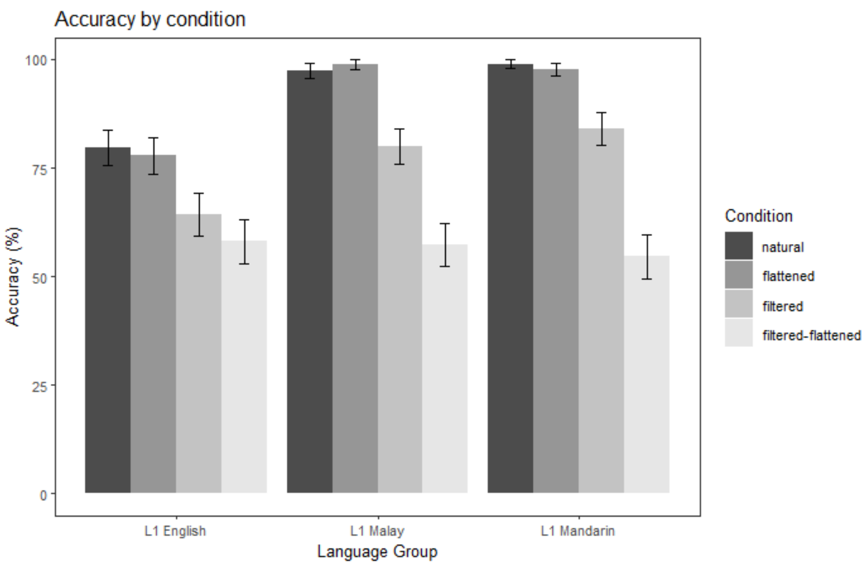

Figure 2: Accuracy by filtering condition.

3.3 *Effect of speaking style* The plot in Figure 3 shows the effect of conversational versus read speech on accuracy in language discrimination. There was no significant difference in overall performance between the two speech styles.

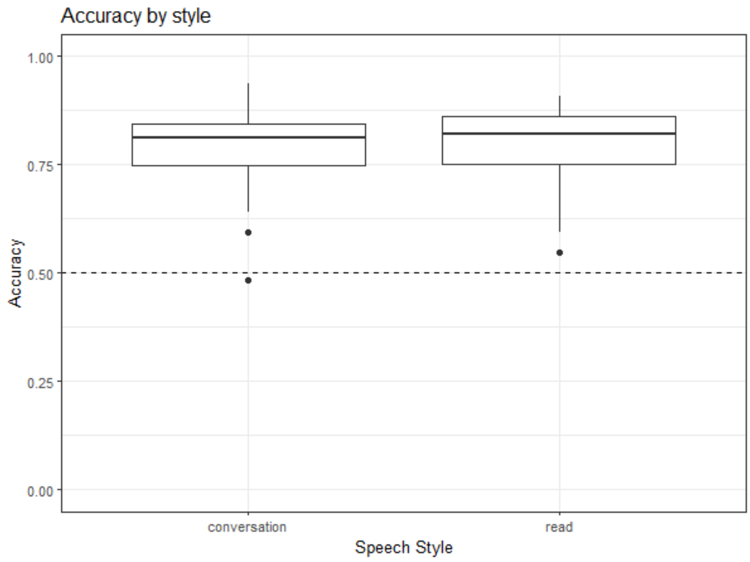

Figure 3: Accuracy by speech style.

For all three language groups, there was a significant interaction between speech style and the filtered versus filtered-flattened conditions (Figure 4). The difference between the two conditions was larger for read speech than for conversational speech.

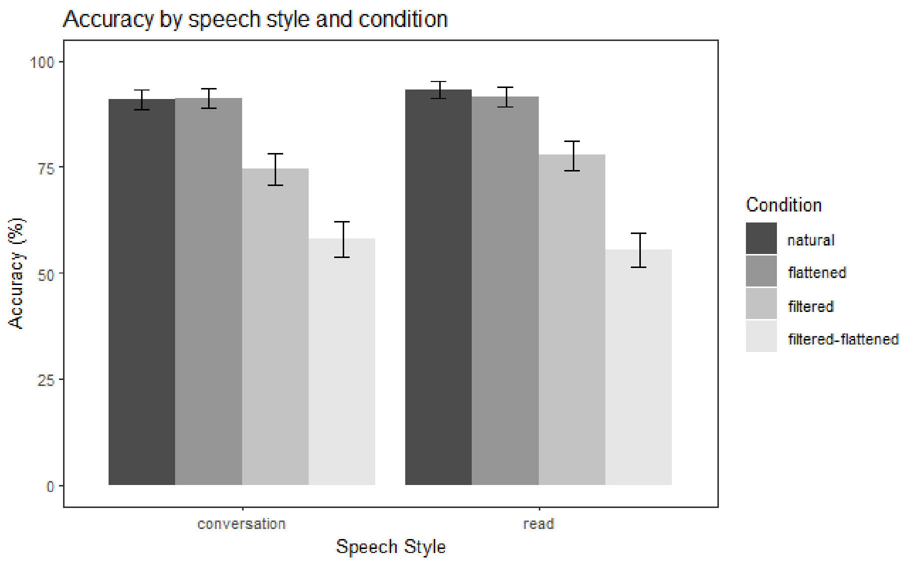

Figure 4: Accuracy by speech style and filtering condition.

4 Discussion

We predicted that speakers who are familiar with one of the two test languages will perform better overall in the task than speakers who are not familiar with either test language. Preliminary results were consistent with this hypothesis: Figure 1 shows that the L1 Malay and L1 Mandarin speakers performed better overall in the task than the L1 English speakers. This indicates an effect of language experience on the language discrimination task: speakers who have native language familiarity will be better at differentiating between the two languages than speakers who do not have familiarity with either test language. Native speakers receive regular exposure to the segmental and suprasegmental properties of their language, so it is expected that they are better able to process and make use of those cues.

Although the L1 Malay and L1 Mandarin speakers performed better than the L1 English speakers overall, the patterns in accuracy across the four filtering conditions were similar across all three language groups (Figure 2). Listeners performed well in the natural condition (when segments, tone, and rhythm were all present), and their performance was unaffected when tonal information was degraded in the flattened condition. This suggests that the listeners in these three language groups do not need to make use of the tonal cue as long as they have access to segmental information. Furthermore, we predicted that listeners will be able to perform above chance with only suprasegmental cues (specifically, with only the rhythmic cue). While the listeners performed above chance in the filtered condition (when tone and rhythm were preserved but segmental information was degraded), they were only at chance in the filtered-flattened condition (when only the rhythmic cue was present). In the absence of segmental information, it seems that tonal information can help listeners in language discrimination.

The L1 Malay and L1 Mandarin groups showed a significantly bigger difference between the flattened and filtered conditions than the L1 English group, indicating that they benefit more from the presence of segmental information. This is expected, since the Malay and Mandarin speakers have native speaker knowledge of the words and phonemes of their languages. The English speakers, who do not understand the Malay and Mandarin words of the stimuli and who are not familiar with the phonemic inventory of either language, seem to be making use of the segmental information as well, but not as much as either the Malay or the Mandarin groups.

Examining the difference in accuracies between the filtered and filtered-flattened conditions provides information on the listeners' reliance on the tonal cue. Since the segmental information has been degraded in both of these conditions, the only difference is the presence of pitch in the filtered condition and the absence of pitch in the filtered-flattened condition. For the L1 English group, the lack of a significant difference between these two conditions suggests that pitch (tonal information) does not benefit the English speakers when differentiating between these two languages: the presence or absence of pitch does not seem to have an effect on their performance. In contrast, the significant difference between these two conditions for both the L1 Malay and L1 Mandarin groups implies that the Malay and Mandarin speakers benefit from the tonal cue, unlike the English speakers. Since they have constant exposure to the intonation patterns of their own languages, it is expected that they are able to extract more information from the cue than the English speakers.

It is interesting to note that the difference between the filtered and filtered-flattened conditions was larger ($p=0.07$) for the L1 Mandarin group than for the L1 Malay group. In the absence of segmental information, this implies that the Mandarin speakers use more tonal information than the Malay speakers, which is consistent with our hypothesis. We suggested that this could be attributed to the fact that Mandarin is a tonal language and Malay is not, so the contrastive use of tone might prompt Mandarin speakers to rely on the tonal cue more. Our study is balanced such that we have speakers of Mandarin and of Malay, who are equally familiar with one of the two test languages (i.e. either Mandarin or Malay). In other words, the Mandarin speakers are just as familiar with the properties of Mandarin (such as pitch) as the Malay speakers are familiar with the properties of Malay. We would expect the Mandarin and Malay speakers to show similar accuracies in the task and therefore a similar reliance on cues. However, the fact that the segmental information has been degraded in both of these conditions seems to imply that the Mandarin speakers are showing an overall heavier reliance on pitch – even in non-lexical contexts – than the Malay speakers. It should be noted that this difference was only trending ($p=0.07$) from our preliminary results and should be interpreted with caution. Further research is required to explore the language-specific processing techniques used by native speakers.

We predicted better performance for the conversational versus the read speech blocks, since conversational speech is used and heard more often than read speech. Our preliminary results in Figure 3 show that there was no significant effect of speech style on accuracy in language discrimination. This indicates that listeners do not differ in their use of cues when they are processing conversational versus read speech.

An unpredicted interaction of speech style and filtering condition was found: the difference in accuracy between the filtered and filtered-flattened conditions was larger for read speech than for conversational speech (Figure 4). This did not interact with language group and was equally true for all three groups. We can interpret this finding as a greater use of pitch when processing read versus conversational speech. Perhaps a greater range of pitch is used in careful speech (as in a reading passage), so listeners might be able to make more use of pitch as a cue. Further research is required to analyse this interaction and interpretation in depth.

5 Conclusion

This study has assessed the dependence of accuracy in language discrimination on the presence of segmental, tonal, and rhythmic cues. Through an AX experiment, we tested discrimination of Mandarin Chinese and Malay sentences by L1 English, L1 Mandarin, and L1 Malay adult listeners. Our results suggest that performance is unaffected by loss of pitch as long as segmental information is present: listeners performed equally well in the natural condition (segmental, tonal, and rhythmic information preserved) and in the flattened condition (segmental and rhythmic information preserved, tonal information degraded). All three groups of listeners only need to rely on the tonal cue in the absence of segmental information; this is particularly true for the Mandarin and Malay groups. Moreover, accuracies in the filtered-flattened condition show that linguistic rhythm alone does not seem to be sufficient for discrimination between Mandarin and Malay.

We also examined the effects of language experience (knowing one of the test languages) and speech style (conversational versus read speech) on the extent of reliance on different cues in the acoustic signal. Preliminary findings indicate that listeners with prior exposure to one of the test languages (i.e. the L1 Mandarin and L1 Malay speakers) performed better than listeners with no exposure to either test language (i.e. the L1 English speakers). While the conversational versus read speech styles seemed to have no effect on overall accuracy, it was found that listeners rely more heavily on the pitch cue when processing read versus conversational speech.

There are a couple of limitations to our study that can be addressed. In our design, we used low-pass filtering at a cut-off frequency of 450 Hz to create an experimental condition in which the segmental information was degraded. Although this technique allowed us to degrade the segmental information significantly, there could have been segmental artefacts remaining at frequencies above the cut-off. One way to mitigate this in future studies would be to use a different technique for manipulating the stimuli, such as the speech resynthesis technique proposed by Ramus and Mehler (1999). Their technique would allow for a more controlled means of eliminating language-specific lexical and syntactic information. Another limitation is that all of the L1 Malay participants have received indirect exposure to Mandarin in their environments (e.g. in media), despite not being able to speak the language. Even though the Malay listeners did not understand the lexical information in the Mandarin sentences, their varying amounts of indirect exposure may have influenced their abilities to process and recognise certain cues in Mandarin.

Future adaptations of this study could take into consideration some additional factors in order to further interpret the findings. The participants' response times during the AX experiment could be recorded and analysed with respect to their accuracies. Since the response time can be reflective of how information is processed up to the moment of decision, it can allow for some interpretation of correct and incorrect responses. Longer response times (i.e. >1 second) may be typical of more difficult tasks and may also indicate a change in decision (Ratcliff & Rouder, 1998). Another variable that could be included in a future adaptation of this experiment may be a fourth language group, consisting of speakers of a tone language that is not Mandarin. In our study, we found that the Mandarin speakers seemed to be making more use of pitch (even non-lexical pitch) than the Malay speakers. To investigate this further, the performance of a fourth group of tonal language speakers may provide some insight into how the tonal cue is used and processed, and whether reliance on this cue could be attributed to language-specific phonological patterns.

Preliminary results from this Mandarin-Malay language discrimination experiment have revealed that accuracy is improved with prior exposure to a test language but is unaffected by differences in speaking style. Listeners are able to perform significantly above chance as long as they have access to segmental cues, and they differ in their reliance on the tonal cue based on their language backgrounds. Our language discrimination experiment contributes to the existing research on acoustic cue use and language processing techniques, in particular for the case of adult bilingual speakers. Future research may attempt to explore any effects of language background on listeners' tendencies to weight certain acoustic cues more than others.

References

Bosch, Laura & Sebastián-Gallés, Núria. (1997). Native-language recognition abilities in 4-month-old infants from monolingual and bilingual environments. *Cognition, 65*(1). 33-69.

Grabe, Esther & Low, Ee Ling. (2002). Durational variability in speech and the rhythm class hypothesis. *Papers in Laboratory Phonology 7,* 515-546.

Maddieson, Ian. (2013). Tone. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online.* Leipzig: Max Planck Institute for Evolutionary Anthropology. Retrieved from http://wals.info/chapter/13

Molnar, Monika; Gervain, Judit; & Carreiras, Manuel. (2013). Within-rhythm class native language discrimination abilities of Basque-Spanish monolingual and bilingual infants at 3.5 months of age. *Infancy*, *19*(3), 326-337.

Nazzi, Thierry; Bertoncini, Josiane; & Mehler, Jacques. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance, 24*(3), 756-766.

Ramus, Franck & Mehler, Jacques. (1999). Language identification with suprasegmental cues: A study based on speech resynthesis. *The Journal of the Acoustical Society of America, 105*(1), 512-521.

Ratcliff, Roger & Rouder, Jeffrey N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, *9*(5), 347-356.