PERCEPTUAL SIMILARITY SPACE OF SERBIAN LEXICAL PITCH ACCENTS: EVIDENCE FROM SERBIAN AND FRENCH LISTENERS

Dušan Nikolić, Stephen Winters

University of Calgary
dusan.nikolic1@ucalgary.ca, swinters@ucalgary.ca

ABSTRACT

Serbian lexical pitch accents have been characterized by two related perceptual dimensions: length and pitch [1]. In line with these two dimensions, the present study explored perceptual similarity space of Serbian lexical pitch accents as perceived by Serbian (native) speakers, the speakers of a pitch-accent language, and French (non-native) speakers, the speakers of a language with no word-level prosody. Participants completed an online free classification task which required them to position the given lexical pitch accents close to each other based on their similarity. The multi-dimensional scaling and results revealed length as a clear dimension of perceptual salience of lexical pitch accents, while pitch movement distinctions were often confused. The findings suggest that the perceptual space created by listener groups is largely shaped by listeners' language experience, and that a lack of word-level prosodic categories in one's language does not necessarily induce perceptual "insensitivity" to lexical pitch accents.

Keywords: lexical pitch accents, perceptual space, Serbian, French, free classification task

1. INTRODUCTION

Using a sequence recall task (SRT), Dupoux and colleagues [2, 3] found that French speakers often have difficulty identifying stress contrasts because they fail to encode stress in their phonological representations. This "stress deafness" effect--which we are recasting as "stress insensitivity"--was also investigated by Rahmani and colleagues [4] in a similar study that involved speakers of Japanese, French, Persian, Dutch, and Indonesian. The authors discovered that French and Indonesian listeners exhibit similar insensitivity patterns when perceiving and processing stress contrasts, while no such effect was found for listeners from the rest of the language groups. This combination of results suggests that speakers of *any* language without lexical stress, tone markings or, in fact, any word-level prosodic category fail to perceive the distinctions between stress contrasts. Nikolić and Winters [5] also used the SRT to explore whether listeners might be similarly insensitive to another word-prosodic category – lexical pitch accent. Nikolić and Winters found that both English and Serbian listeners could identify Serbian lexical pitch accent categories at a better than chance rate. This finding is consistent with the observation that speakers of a language which has word-prosodic contrasts can identify other word-prosodic contrasts better than speakers of languages which have no word-prosodic distinction [4].

These studies used the sequence recall task (SRT) to explore the topic of perceptual insensitivity at the word-prosodic level [3, 4, 5]. One of the aims of using this methodological approach is to preclude the listeners in the task from superficially processing and encoding the stimuli at an acoustic level only, which makes it easier for the listeners to detect fine-grained prosodic differences between tokens of different prosodic categories. Instead, the SRT encourages listeners to develop and rely on phonological representations of word-prosodic categories in order to complete the experimental task. With the SRT, listeners are first required to associate arbitrary category labels with example tokens from one of the two word-prosodic categories under study. Listeners then hear a sequence of tokens from different word-prosodic categories and, after a delay, are required to input the sequence of labels associated with those tokens, in the same order. Only entirely correct sequences are then counted as indicative of "non-insensitivity" to the contrast.

Instead of arbitrarily associating labels with categories in this study, we used a free classification task to investigate how listeners might classify tokens of word-prosodic categories based on their similarity, without imposing any labels upon them. Imai and Gardner [6] first developed the free classification task and showed that there were a number of advantages to using it, the main one being that the task did not need to employ experimenter-imposed categories. Interest in the task resurged in linguistic research when Clopper and Pisoni [7] used it to investigate the perceptual similarity of regional dialects in the United States. Listeners in [7] carried out a free classification task in which they were instructed to group regional dialects based on their similarity into as many groups as they heard among the stimuli. Multidimensional scaling analyses of the data showed that gender, geography, and linguistic markedness were the three

determining dimensions of similarity among the dialects. The listeners' performance on the task therefore reflected relevant linguistic and sociolinguistic factors in the perception of regional dialect variation.

Daidone and colleagues [14] used a free classification task to investigate the perceptual space of German vowels and Finnish phonemic length contrasts by L2 learners of these languages. The results showed that the task could be effectively used to describe discriminability of both segmental contrasts (vowels) and suprasegmental contrasts (length), and that the task could be a replicable and reliable method for examining the perception of non-native speech categories.

Clopper [8] described the free classification task by emphasizing that the method could be applied to "examine the perceptual similarity of a range of speech and nonspeech stimulus materials, including linguistic structures such as segments, tones, or intonation contours, and indexical properties such as voices, speaking styles, emotions, dialects, foreign accents, or languages" (p. 575). Apart from [14], no study has yet used this task for examining any of the linguistic purposes described in [8]. With the present study, we addressed this large research gap.

2. THE PRESENT STUDY

In the present study, our broad goal was to use the free classification task to examine the perception of *lexical pitch accents* by listeners of languages both with and without word-prosodic contrasts in their native phonological inventories. To this end, we specifically explored whether native Serbian and French speakers could distinguish between Serbian lexical pitch accents. Serbian is a language that has four lexical pitch accents which combine short and long vowels (S and L) and rising and falling pitch contours (R and F). The lexical pitch accents in Serbian are thus descriptively categorized as short-falling (SF), short-rising (SR), long-falling (LF), and long-rising (LR). Unlike Serbian, French has no word-level prosodic contrasts, and perceptual prominence in French is expressed at the phrasal level only [9]. Based on the previous stress insensitivity findings, we expected Serbian listeners to display greater distances between the lexical pitch accent categories in the perceptual spaces that emerged from the free classification task than the French listeners would. However, since we were applying a new methodology to the study of this task, we also expected the analysis to yield subtler, finer-grained differences in perceptual categorization (for both groups of listeners) that the Sequence Recall Task could not capture.

2.1. Study Design

40 native speakers of Serbian and 60 native speakers of French participated in the study. Both groups of participants completed an online free classification task, in which they freely grouped recordings based on their similarity. The free classification task was adapted from an online rating toolbox built by Donhauser and Klein [10]. The participants listened to 2 groups of 2 nonwords, "rejav" [rejav] and "kavan" [kavan], both of which were patterned on Serbian phonotactic rules. Each word was produced with 4 Serbian lexical pitch accents by one male and one female speaker. 16 nonwords were produced in total (2x4x2). In the free classification task interface, the listeners could hover the computer cursor over distinct icons, spaced around a circle, each of which played a unique recording of a nonword production (Fig. 1). The listeners' task was to place the icons representing each of these words next to each other in the circle, based on similarity. The listeners were instructed to place more similar tokens closer to each other, and more dissimilar tokens further away from each other (Fig. 2).

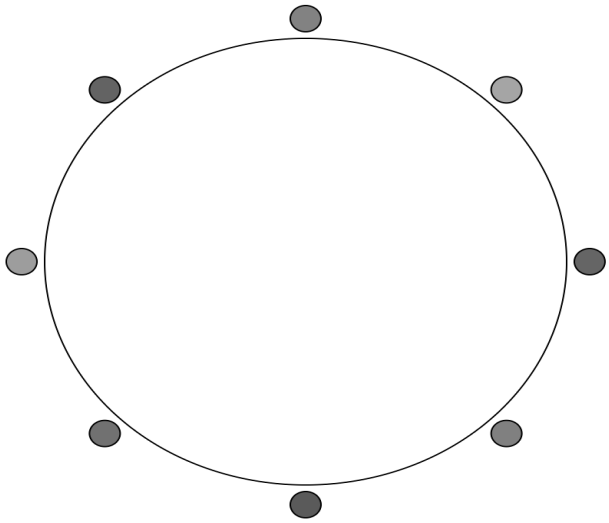

Figure 1: The set-up of the free classification task. Each small circle is a recording. Hovering the cursor over a circle plays the recording.

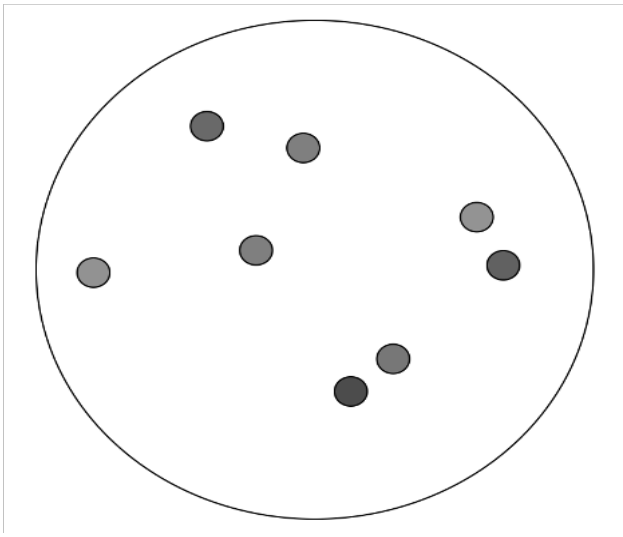

Figure 2: The display of the free classification task after the classification has been completed.

There was no time limit on the task and the participants were instructed *not* to group the words simply based on their segmental structure (i.e., all "kavan" and all "rejav" words) or based on gender (i.e., all male- and all female-produced words next to each other). Since there were sixteen tokens in total, we separated them into distinct groups for two iterations of the task, which we will refer to free classification task I and free classification task II.

2.2. Data Analysis

The perceptual similarity structure of the nonwords was extracted from the free classification data following the guidelines in [10]. Using a python script, we converted the raw coordinates into similarity matrices reflecting the normalized Euclidean distances. The resulting matrices were submitted to multi-dimensional scaling (MDS) to "produce the best fit for the entire distance matrix in a specified number of orthogonal dimensions" ([8], p. 578). All the analyses were performed in RStudio following the procedure outlined in [11].

3. RESULTS

We created an average similarity matrix per listener group (Serbian and French) out of all the received similarity matrices. For Serbian listeners, the most optimal number of dimensions for the free classification task I was 3, while the most appropriate number of dimensions for the free classification task II was 2.

The MDS analysis for the Serbian listeners' performance on the free classification tasks yielded the results displayed in figures 3 and 4. The stress values for the two free classification tasks were 0.177 and 0.24, respectively. The "M" stands for a male-produced item, while the "F" stands for a female-produced item. Thus, the "MRejav_LR" is a male-produced "rejav" nonword in a long-rising lexical pitch accent.

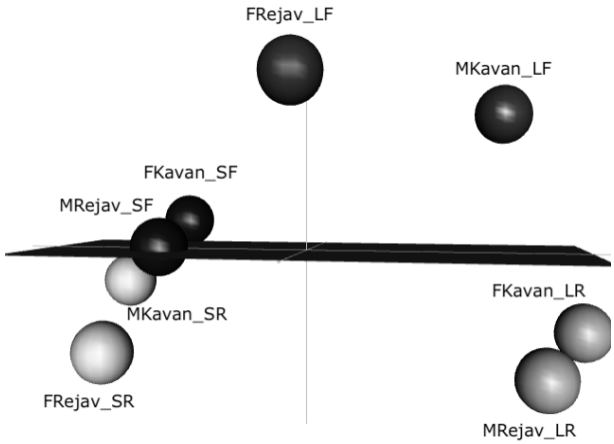

Figure 3: The 3-dimensional multidimensional scaling of Serbian listeners' classification of lexical pitch accents on free classification task I.

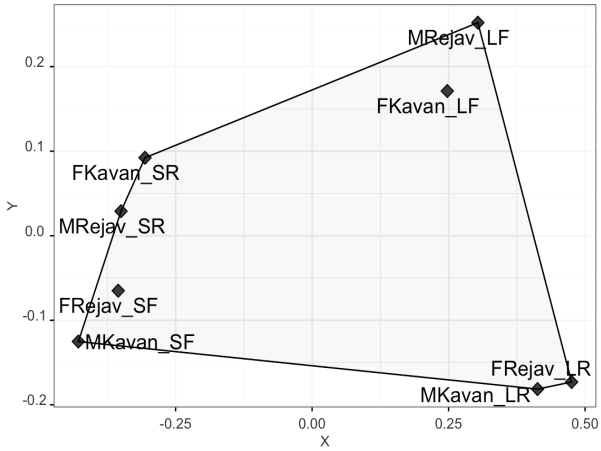

Figure 4: The 2-dimensional multidimensional scaling of Serbian listeners' classification of lexical pitch accents on free classification task II.

An identical MDS analysis was performed on the French listeners' classifications of the lexical pitch accents. Both analyses yielded 3-d models with stress values of 0.18 and 0.06. Figures 5 and 6 below display the perceptual spaces received from these analyses.

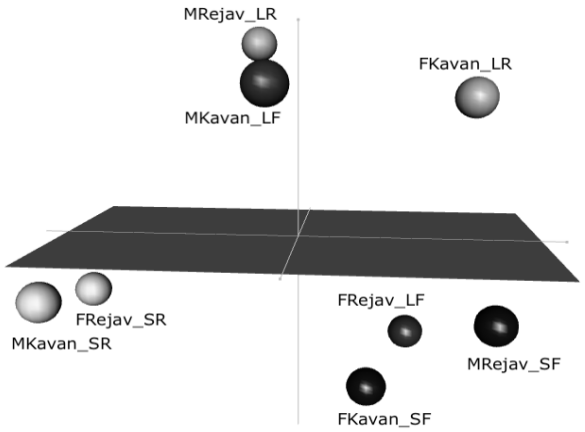

Figure 5: The 3-dimensional multidimensional scaling of French listeners' classification of lexical pitch accents on free classification task I.

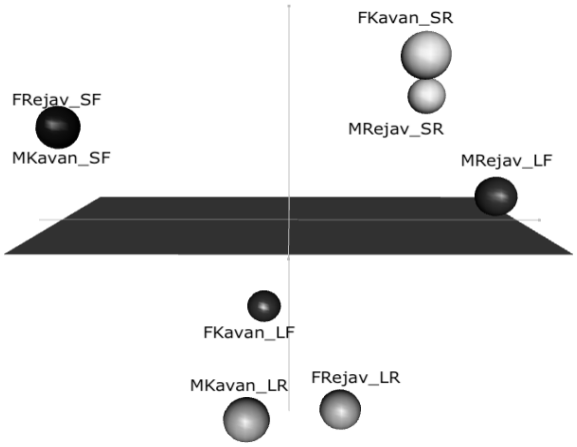

Figure 6: The 3-dimensional multidimensional scaling of French listeners' classification of lexical pitch accents on free classification task II.

4. DISCUSSION

The French listeners clearly distinguish the SF, SR and LR groups, but the classification of the LF tokens is inconsistent. The Serbian listeners distinguish between the LR and the LF tokens, while the SR and SF stimuli form "one group". The Serbian listeners maintain better coherence between tokens of the four existing pitch accent categories than the French listeners do (with the LF category causing the big perceptual grouping problems for the French listeners).

Given all the perceptual support that the task provides, the listeners would likely process the stimuli at an acoustic level. However, the perceptual categories they form are not strictly based on the obvious acoustic distinctions in the stimuli, and they also differ from one language group to the other. There are two main findings of the study. First, language experience and knowledge must be playing some role in shaping those perceptual categories. The classifications formed by both groups consisted of different lexical pitch accent categories. Considering that the stimuli that both groups heard were identical, this distinction is likely reflected in the prosodic system of the listeners' first language. The French listeners could hear some nuanced phonetic differences in the pitch movement as the pitch movement is characteristic of the French post-lexical phrases (e.g., AP, iP, IP). The Serbian listeners could clearly hear the differences in length, and the pitch distinctions in items with long lexical pitch accents. Pitch distinctions were less obvious in items with short vowels, which indicates that the Serbian listeners tend to conflate short-falling and short-rising lexical pitch accents.

Second, methodologically, the free classification task can reveal perceptual sensitivities to word-level prosodic distinctions in listeners of languages with no native word-level prosodic contrasts. That is, unlike what the previous research suggests [4], a lack of word-level prosodic categories does not strictly impose perceptual "insensitivity" to a prosodic category in a task such as the free classification task. Non-native listeners could be processing lexical pitch accent differently from native listeners, i.e., non-native listeners' processing of a word-prosodic category that does not exist in their first language could be phonetic, while native listeners' processing could be phonological (cf. [12, 13]). The future research should attempt to address the processing mechanisms of native and non-native listeners when listening to word-level prosodic categories.

5. CONCLUSION

The study explored the perceptual similarity space of lexical pitch accents by using an adapted and novel version of the free classification task showing more fine-grained phonetic and phonological subtleties of the perceptual sensitivity towards native and non-native prosodic categories. The study has revealed that taking a more nuanced methodological approach to the investigation of prosodic insensitivity may require a revision of our understanding of when and why the phenomenon occurs.

6. REFERENCES

[1] Sredojević, D. 2017. How much do phonetic realisations of Serbian accents actually differ from each other in various dialects?. *Годишњак Филозофског факултета у Новом Саду*, 42(1), 323-337.

[2] Dupoux, E., Pallier, C., Kakehi, K., Mehler, J. 2001. New evidence for prelexical phonological processing in word recognition. *Language and cognitive processes*, 16(5-6), 491-505.

[3] Dupoux, E., Sebastián-Gallés, N., Navarrete, E., Peperkamp, S. 2008. Persistent stress 'deafness': The

case of French learners of Spanish. *Cognition*, 106(2), 682-706.

[4] Rahmani, H., Rietveld, T., Gussenhoven, C. 2015. Stress “deafness” reveals absence of lexical marking of stress or tone in the adult grammar. *PloS one*, 10(12), e0143968.

[5] Nikolić, D., Winters, S. 2022. Are Serbian and English listeners insensitive to lexical pitch accents in Serbian?. *Phonetica*, 79(4), 397-423.

[6] Imai, S., Garner, W. R. 1965. Discriminability and preference for attributes in free and constrained classification. *Journal of Experimental Psychology*, 69(6), 596.

[7] Clopper, C. G., Pisoni, D. B. 2007. Free classification of regional dialects of American English. *Journal of phonetics*, 35(3), 421-438.

[8] Clopper, C. G. 2008. Auditory free classification: Methods and analysis. *Behavior Research Methods*, 40(2), 575-581.

[9] Post, B. 2000. *Tonal and phrasal structures in French intonation* (Vol. 34). The Hague: Thesus.

[10] Donhauser, P. W., Klein, D. 2022. Audio-Tokens: A toolbox for rating, sorting and comparing audio samples in the browser. *Behavior Research Methods*, 1-8.

[11] Levshina, N. 2015. *How to do linguistics with R. Data Exploration and Statistical Analysis*, Amsterdam-Philadelphia.

[12] Chen, J., Best, C. T., & Antoniou, M. 2020. Native phonological and phonetic influences in perceptual assimilation of monosyllabic Thai lexical tones by Mandarin and Vietnamese listeners. *Journal of Phonetics*, 83, 101013.

[13] Post, B., & Alter, K. (2015). Neural correlates of categorical linguistic and gradient paralinguistic intonation. *Proc. 18th ICPhS* Glasgow, 1-5.

[14] Daidone, D., Lidster, R., & Kruger, F. (2023). Free classification as a method for investigating the perception of nonnative sounds. *Studies in Second Language Acquisition,* 1-27.