Whistled phoneme categorization: the effect of vowel space range

Anaïs Tran Ngoc[1,2], Julien Meyer[2], Fanny Meunier[1]
[1]Université Côte d'Azur, CNRS, France
[2]Université Grenoble Alpes, CNRS, France

https://doi.org/10.36505/ExLing-2023/14/0027/000621

Abstract

We explore whistled vowel categorization by untrained listeners, focusing specifically on the impact of the different vocalic frequency ranges of two whistlers (for the vowels /i/, /e/, /a/, /o/) and the effect of training on performance. In the experiment, we included stimuli that show inter-individual and intra-individual variations of production. In the analyses we looked at the whistler identity effect and at the learning effect through the experiment for the studied vowels. The results showed an effect of the whistler, where the larger vocalic range led to improved categorization, and highlighted the robustness of the vowel recognition hierarchy. There was no general learning effect, albeit for one vowel and for the whistler with a narrower vocalic range. This study provides insight into one's representation of the vowel space in non-tonal languages.

Keywords: whistled speech; speech perception; vowels; categorization; vocalic range

Introduction

Whistled speech is a natural speech form used for long distance communication. To do so, it transposes spoken speech into whistles produced in the front oral cavity of the mouth. In non-tonal languages, vowels are emitted at different whistled pitch levels depending on spoken vowel qualities (Busnel and Classe, 1976). For example, in Spanish, whistled /i/ has the highest mean values of pitch, /e/ is lower, /a/ is even lower, and /o/ even more so (Meyer, 2008). While whistled speech is not directly understood by naive listeners - i.e. listeners who never heard it before - previous studies have proved that they categorize whistled vowels much better than chance (Meyer et al, 2017, Tran Ngoc et al, 2020). In the present experiment we used whistled speech as a tool to investigate perceptual processes in language processing, more specifically to test the impact of production variations in the Vowel Space Range.

Methods

We ran a behavioral experiment in which we asked 44 naive participants (French language natives) to categorize whistled vowel stimuli. We focused on four whistled vowels: /i, e, a, o/ whistled by two different whistled Spanish teachers in the Canary Islands: whistler A had a more restricted vocalic frequency range, and whistler B had a wider range (see Tran Ngoc et al, 2020 for details). Stimuli were extracted from the stable whistled vowel nuclei of the second vowel of CVCV words (such as /cada/, /nata/…) following various consonants to introduce variations (/d/, /k/, /g/, /t/). The experiment was structured in 3 parts. In part 1, participants listened to 48 whistled vowels (12 versions of each vowel type). Part 2 was a short training session with feedback (16 stimuli) produced by the same whistler as in part 1. Part 3 was similar to part 1. Stimuli were presented in a random order in each part. Four versions of the experiment were built, called AA, BB, AB, BA, according to whether productions of whistler(s) A and/or B were presented in parts 1 and 3.

Results

We took into account the answers given in part 1 and part 3 by each participant. We find that overall, the 44 participants obtained 53.55% (SD=12.99) of correct responses out of the 2112 answers given; well over chance, at 25%. We compared different conditions of the experiment by running several Generalized Linear Mixed Model analyses, described below. When convenient, the post hoc tests (all with Bonferroni corrections) are summarized by the symbols > or =, respectively indicating a significant difference or no difference.

Comparison between AA, BB, AB and AB conditions

We observed that participants who heard version AB obtained 52.46% (SD=9.96%), BA obtained 55.01% (SD=12.91), AA obtained 46.49% (SD=12.70), and BB obtained 60.23% (SD=13.81).

In a first analysis, we looked at the effect of having either only one whistler or two throughout parts 1 and 3. Taking into account the whole set of data, we ran a GLMM on Correct Answers only, with Part (1, 3), Whistler identity (A, B) and the Number of whistlers heard (1 –for AA and BB- or 2 –for AB and BA) as fixed factors and Participants as a random factor.

We find a significant main effect of Whistler identity (X2(3,N=44)=5.9505, p=.01) as well as a significant interaction between Whistler identity and Number of whistlers (X2(3, N=44)=6.8105, p<.01). The post hoc tests revealed a difference between whistler A and whistler B only in the comparison between the versions AA and BB where the same whistler was heard in both parts (p<.05).

Correct answers for the AA/BB experiment

Considering only the 22 participants who heard the versions with only one whistler (either AA or BB), we applied a GLMM on Correct Answer, with Part (1, 3), Vowel (/i, e, a, o/) and Whistler identity (A, B) as fixed factors and Participants as a random factor. We observed significant differences between the vowels, (X2(3, N=22)=247.48, p<.001) (where /i>o>a=e/) and a significant effect of Whistler identity (X2(3, N=22)=6.10, p=.014) showing that *whistler B's productions give rise to much better performances than whistler A (60.23% vs. 46.49%).*

A significant interaction Vowel*Part (X2(3, N=22)=21.62, p< .001) revealed that no vowel showed a significantly better performance in one part compared to the other, though there were differences in vowel recognition hierarchies between parts: /i>o(=a)>e(=a)/in part 1; and/i>o=e>a/in part 3.

Finally, the significant interaction Whistler*Vowel (X2(3, N=22)=7.99, p< .05) showed that for whistler A the hierarchy was /i>o(=e)>a(=e)/; whereas for whistler B /i>o>e=a/ (see % of correct responses in Figure 1). *These results suggest a stable hierarchy between /i>o>a/, with /e/ being less stable and suggesting that, through the experiment with the productions of whistler A, /e/ is better categorised.*

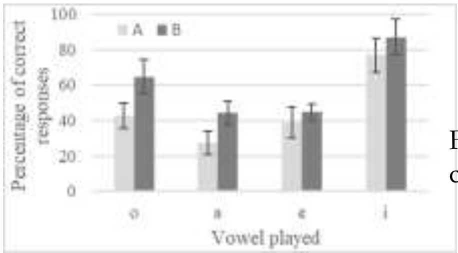

Figure 1. Correct whistled vowel categorization per whistler (in %).

Correct answers analyzed according to Whistler (A,B) for all lists

To gain statistical power and considering we didn't observe an effect of part according to whistler in the previous analysis, we took into account the whole data set (44 participants) and applied the same type of GLMM analysis, thus looking at the global picture. We found a significant main effect of Vowel (X2(3, N=44)=515.02, p<.001) as well as three significant interactions between Whistler and Vowel (X2(3, N=44)=32.36, p<.001), between Part and Vowel (X2 (3, N=44)=7.93, p<.05), and a double interaction Whistler*Vowel*Part (X2 (3, N=44)=11.04, p<.05).

Post-hoc analyses showed that for whistler A (see Figure 2), the only vowel showing a better performance in part 3 than in part 1 is /e/ (p<.01). The hierarchy found in part 1 is /i>o>a=e/; whereas in part 3 it is /i>o=e>a/. Post-hoc analyses also revealed that for whistler B no vowel showed a significant difference in performance between parts 1 and 3 and the hierarchy is similar in both parts: /i/>/o/>/e, a/.

Moreover, the only specific comparisons that reach significance comparing the two whistlers are for /e/ and /i/, both in part 3, (respectively p<.01 and p<.05), in which /i/ is better recognized by listeners while hearing whistler B and /e/ is better recognized while hearing whistler A.

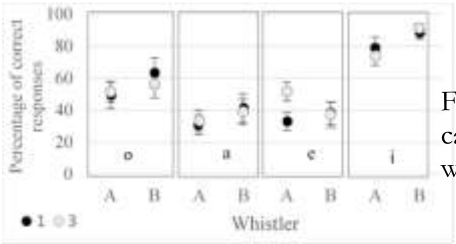

Figure 2. Correct whistled vowel categorization (per vowel, part, whistler).

Discussion and conclusions

This experiment first shows that the range of the vowel space used by different whistlers affects vowel categorization. The whistler with the wider vocalic frequency range gave rise to the best categorization rates, in line with the literature showing that hyper-articulation improves speech processing and that expanded vowel space benefits listeners, both natives and L2 speakers, in silence or in noise (Kagantharan et al. 2022). Moreover, learning through the experiment appeared restricted to only one vowel for the whistler with the narrower frequency range. Interestingly, this vowel /e/ is the one of the experiment that has the least formant convergence in the spoken form, which could explain less stability in recognition (Chistovitch & Lublinskaya 1979). Overall, the results highlight the robustness of the whistled vowel recognition hierarchy previously observed, and a certain stability in the speech perception process when faced with inter-talker variability.

Acknowledgements

We thank the whistlers for their contribution and the participants who volunteered.

References

Busnel, R-G., Classe, A. 1976. Whistled languages. Berlin Heidelberg, Springer.

Chistovitch, L.A., Lublinskaya, V.V. 1979. The center of gravity effect in vowel spectra and critical distance between the formants: psychoacoustical study of the perception of vowel like stimuli, Hear. Res. 1, 185–195.

Kangatharan, J., Uther, M., Gobet, F. 2022. The effect of hyperarticulation on speech comprehension under adverse listening conditions, Psychol. Res. 86, 1535–1546.

Meyer, J. 2008. Acoustic Strategy and Typology of Whistled Languages; Phonetic Comparison and Perceptual Cues of Whistled Vowels. JIPA 38, 69-94.

Meyer, J., Dentel, L., Meunier, F. 2017. Categorization of Natural Whistled Vowels by Naive Listeners of Different Language Background. Front. Psychol. 8, 25.

Tran Ngoc, A., Meyer, J., Meunier, F. 2020. Whistled vowel identification by French listeners. In Interspeech 2020 Proc. of the 21th Annual Conference of the International Speech Communication Association, 1605-1609, Shanghai, China.