Testing perceptual flexibility in speech through the categorization of whistled Spanish consonants by French speakers

Anaïs Tran Ngoc, Fanny Meunier, Julien Meyer

JASA EXPRESS LETTERS ARTICLE asa.scitation.org/journal/jel

Testing perceptual flexibility in speech through the categorization of whistled Spanish consonants by French speakers

Anaïs Tran Ngoc,[1,a)] Fanny Meunier,[1] and Julien Meyer[2]
[1]Université Côte d'Azur, CNRS, BCL, France
[2]Université Grenoble Alpes, CNRS, GIPSA-Lab, Grenoble, France
anais.tran-ngoc@etu.univ-cotedazur.fr, fanny.meunier@unice.fr; julien.meyer@cnrs.fr

Abstract: Whistled speech is a form of modified speech where, in non-tonal languages, vowels and consonants are augmented and transposed to whistled frequencies, simplifying their timbre. According to previous studies, these transformations maintain some level of vowel recognition for naive listeners. Here, in a behavioral experiment, naive listeners' capacities for the categorization of four whistled consonants (/p/, /k/, /t/, and /s/) were analyzed. Results show patterns of correct responses and confusions that provide new insights into whistled speech perception, highlighting the importance of frequency modulation cues, transposed from phoneme formants, as well as the perceptual flexibility in processing these cues.

[Editor: Charles C. Church] https://doi.org/10.1121/10.0013900

Received: 19 May 2022 **Accepted:** 17 August 2022 **Published Online:** 7 September 2022

1. Introduction

Whistled speech is a naturally modified speech form characterized by the transposition of the spoken signal into whistled frequencies, drastically changing the spoken timbre. Whistled vowels of non-tonal languages are produced at relatively stable frequencies, which depend on the vowel position, the whistling technique, the language, the whistler's oral cavity, and vowel coarticulation with surrounding phonemes. In such languages, the whistled $F0$ codes and simplifies the timbre of modal speech (see supplementary material SuppPub1 for a figure showing the Waveform and Spectrogram of a spoken and whistled Spanish sentence[1]). Typically, /i/ corresponds to the highest whistled frequencies and /o/ to the lowest, while /e/ and /a/ are placed in the middle (with /e/ higher than /a/; Meyer, 2015). Just like in spoken speech, the consonants modify/modulate the vowel frequencies by adding stops and pitch changes as the whistlers "pronounce" the consonants while whistling (cf. Fig. 1). With the added constraint of a rather closed mouth to whistle, this speech transformation generally creates an emphasis on the upper resonant cavities. This explains why whistled frequencies usually reflect frequency shapes of the second or third formants of modal speech ($F2$ most frequently, but also $F3$ for front vowels) (Shadle, 1983; Meyer, 2015). However, in the case of coarticulation between back vowels and velar/uvular consonants, whistled $F0$ also often resembles the $F1$ of modal speech (see, for example, /k/ in Fig. 1, and see Meyer *et al.*, 2019).

We can consider whistled speech akin to other forms of modified speech, such as speech in noise or artificial sine-wave or vocoded speech, where untrained listeners are able to identify and categorize certain aspects, such as phonemes (Blanco *et al.*, 2018). Whistled speech recognition and categorization experiments first started in the 1960–1970s with Bearnese and Turkish whistlers, focusing on word (Busnel, 1970) and CV nonsense syllable recognition between local whistlers (Busnel *et al.*, 1962; Moles, 1970; and see Meyer, 2015 for a reanalysis). In 2005, Rialland ran a behavioral experiment on whistled VCV nonsense utterances identified by a fluent Spanish whistler, obtaining 57% of correct answers with better performance for certain consonants and vowels (Rialland, 2005). More recently, Meyer *et al.* (2019) conducted a syllable recognition experiment (/ta, /da, /ka, /ga/) with Tashlhiyt Berber whistlers. Experiments with participants who were not previously familiar with whistled speech ("naive listeners") only date back to 2005. Such studies included participants with different language backgrounds (Spanish, French, Chinese) who were tested on a whistled vowel recognition paradigm based on Spanish vowels. The results obtained were well over chance for all categories of listeners with striking differences between vowel positions and language background (Meyer, 2008; Meyer *et al.*, 2017). These previous results make us wonder whether such a capacity, or form of perceptual flexibility allowing for phoneme categorization in spite of the reduced phonetic cues, can extend to whistled consonants. We tested French speakers' categorization capacities for whistled

[a)]Also at: Université Grenoble Alpes, CNRS, GIPSA-Lab, Grenoble, France. Author to whom correspondence should be addressed.

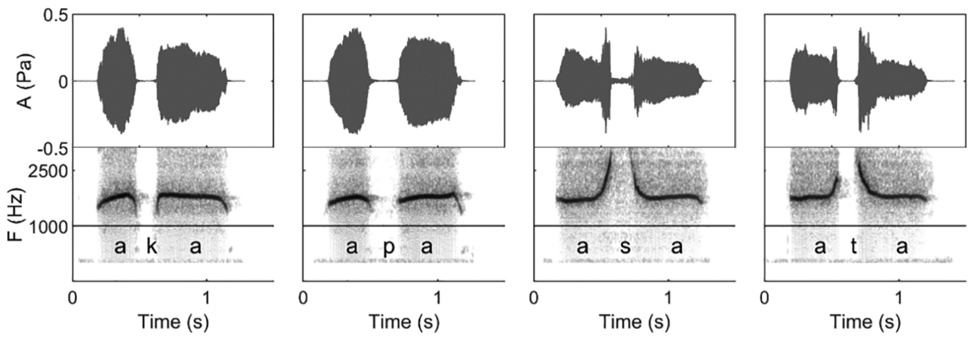

Fig. 1. Waveform and Spectrogram of VCV forms of the experiment [whistled Spanish /aka/ (Mm.1a), /apa/ (Mm.1b), /asa/ (Mm.1c, and /ata/ (Mm.1d)] [see supplementary material SuppPub1[1] (Ref. 1) for a figure showing the Waveform and Spectrogram of the corresponding spoken utterances].

Spanish consonants through a behavioral experiment. The setup of the experiment also allowed us to explore other complementary questions: does the inclusion of a training portion allow for a learning affect? Which factors underlie participants' consonant categorization?

To answer these questions, we constructed our experiment in three parts including a section with whistled consonant categorization without feedback, one section with feedback, and a final section with natural variations of each consonant. This allows us to test whether participants learn to apply consonant models to multiple varieties of each consonant (a method based on a perceptual learning experiment on noise-vocoded speech, see Hervais-Adelman *et al.*, 2008). To understand the cues used for consonant categorization, we will compare the participants' responses with previous interpretations and classifications of the whistled Spanish consonant system. The whistled consonants chosen (/p/, /k/, /s/, and /t/) and recorded in Silbo (the local name for whistled Spanish in the Canary Islands), have previously been grouped in different categories based on acoustic loci, as well as frequency and/or amplitude modulations. Trujillo (2006), for example, proposed four consonant groups and Rialland (2005) proposed eight groups for all whistled Spanish consonants, both using distinctions such as "low," "acute," "continuous," and "interrupted." Both Trujillo and Rialland oppose /p/ and /k/ said to be "low" consonants, to /s/ and /t/ said to be "acute." Rialland also proposes secondary distinctions such as higher loci in /s/ than /t/ and sharper attacks in /t/. However, whistled consonant classifications are in fact more complex as they are influenced by parameters that were not systematically controlled by Trujillo and Rialland. These include whistling technique, the position of the consonant in the word, speech rate, and proficiency of the whistler (Meyer, 2015). Nevertheless, all researchers agree on two clear distinctions among whistled consonants in Spanish: one between consonants with high (/s/ and /t/) or low (/p/ and /k/) whistled loci, and one between semi continuous or continuous (represented by /s/ in our experiment) and interrupted consonants (/k, p, t/). This continuous/interrupted opposition is explained by the amplitude decay either corresponding to a dip (maintaining continuity), applied to /s/, or to a complete interruption (resulting in a silence), which applies to stops. On the other hand, "acute" consonants with high loci systematically correspond to consonants rising after the previous vowel (V1) and falling towards the next vowel (V2) (see /asa/ and /ata/ in Fig. 1), and low loci ("grave consonants") to the reverse (Meyer, 2015). Secondary whistled distinctions, such as those suggested by Rialland (2005), notably the sharper CV attack, also come into play. This is typically the case of /k/ and /t/ (Meyer, 2015), whereas a more progressive CV frequency slope characterizes /p/ (see Fig. 1). In this context, the classification of /s/ is one of the most complex because it emulates the continuous fricative aspect of spoken speech, expressed by an amplitude dip only by expert whistlers (see Díaz, 2017, and Fig. 1). This type of whistled fricative is well termed "semi-continuous" or "continuous" because its acoustic continuity depends on the whistler's proficiency, the speech rate, and the listening distance. In fast speech, whistled fricatives are more clearly continuous because the speed reduces the dip through a more gradual amplitude envelope modulation (Meyer, 2015); however, the low dB level present in the amplitude dip can cause its dynamics to be partly masked by the background noise in increased emitter-receiver distances.

Students learning Silbo from La Gomera Island generally follow recommendations based on Trujillo's groupings, however, those from the Yo Silbo association, the most active Silbo revitalization association in the Canary Islands, assemble the consonants into five pronunciation-based groups using VCV configurations as a didactic basis for the classification. This classification opposes /t, s/ to /p/ and to /k/ (Díaz, 2017), where /p/ is considered low but continuous. This difference may take into account the contrast between the sharp attack for /k/, with a more rapid release, and the softer frequency slope and amplitude modulation which characterize the consonant stop in /p/ (see Fig. 1). Even if it is not highlighted by the groupings presented above, the sharpness of the interruption could also be a defining commonality in /t/ and /k/, in opposition with /s/ and /p/. Moreover, the plosive and fricative consonant opposition is not usually proposed as a grouping characteristic, but skilled whistlers manage to develop it (Díaz, 2017). It is therefore considered as another secondarily developed opposition in the whistlers' community. To sum up, our experiment includes contrastive consonants that will shed light on the importance of these cues for categorization. Though our experiment targets a naturally modified speech form (whistled speech), the speech cues employed reflect more generalized phoneme processes as well as mental representations of phonological cues.

2. Methods

2.1 Stimuli

We chose to test four distinct consonants of spoken Spanish that have either identical or easily learned pronunciation differences in Spanish and French (Molina Mejia, 2007). These include three occlusive or plosive consonants ([p]-bilabial), [t]-dental/alveolar), [k]-velar), and a fricative ([s]-alveolar), followed and preceded by the vowel [a], giving the following VCV forms: [ata], [aka], [asa], and [apa]. The use of a VCV form enables us to take into account variations due to the duration of consonant closure, as well as amplitude and frequency modulations. In addition, using only one vowel reduces the effects of different Consonant–Vowel coarticulations (Meyer, 2015). Four instances of each /aCa/ segment were whistled by the same proficient whistler-teacher of "Silbo" (whistled Spanish of the Canary Islands) and recorded by the last author. The frequencies before and after each consonant closure vary between 1141.9 and 2628.7 Hz, with an average of 1715.86 Hz.

2.2 Procedure and design

The experiment was programmed using PCIbex Farm and took place online from participants' own homes. Before starting the experiment, participants were asked their age, the languages they speak (and their level), as well as if they play any musical instruments. As this experiment was online, they were to indicate whether they used headphones, earbuds, or speakers, and to give the name of the brand. We recruited the participants through various social media networks, excluding participants with self-declared speech/hearing impairments. Before beginning part 1 of the experiment, participants are shown the recordings of the four whistled VCV forms heard in part 1 (in a randomly chosen order) without any indication of their categorization. This allows participants to familiarize themselves (briefly) with the acoustic specificities of whistled signals as well as to adjust the volume to a comfortable listening level. The four /aCa/ recordings presented (one of each consonant, chosen according to the stability of whistled vowel frequencies, see Fig. 1) are used during part 1 without any indication of the consonant heard. The participants then hear these clips in a random order and are asked to respond with either "p," "k," "t," or "s" after each clip. These consonants are attributed to the arrow keys on the keyboard according to the layout of both azerty and qwerty keyboards. Participants see this figure on screen as they listen and respond to the 40 recordings (ten times each consonant) which make up part 1.

Part 2 is a training phase with feedback, using the same whistled audio tracks as part 1. We first present the four different consonants in a random order by playing a spoken version of the VCV segment, followed by the whistled version. An image of the "written consonant" appears on the screen simultaneously. Following this, participants complete a shorter version of the previous test albeit with feedback. Participants hear each clip (each consonant) four times, amounting to 16 total excerpts. Feedback is given after each response: "Bravo" when correct and "Non ce n'était pas la bonne réponse"—"No that was not the correct answer," when false. In part 3 of the experiment, participants hear sound clips and are requested to indicate which consonant was heard. However, in this portion, three additional versions of each consonant are included, amounting to four total variations per consonant. As this applies to all four consonants, 16 recordings are heard, out of which 12 are unfamiliar variations (i.e., not heard in part 1). Each recording is played three times and participants hear a total of 48 stimuli in part 3.

2.3 Participants

This study included 30 adults (21 women, 9 men, mean age: 29.6 years, standard deviation: 8.77) whose first language was French and who did not have any language or hearing impairments. A number of participants had experience in different languages, notably in Spanish. 19 participants indicated having some experience in Spanish, where 8 participants declared being beginners, 8 participants had an intermediate level, and 3 had a confirmed level. Participants gave informed consent before starting the experiment.

3. Results

Our analysis focused on parts 1 and 3, excluding the short training portion (part 2) due to the small sample size. We compared both parts 1 and 3 by taking into account the 40 answers given in part 1 by each participant as well as the 48 answers given in part 3. This gave us 3520 data components. After presenting the results, we analyzed the correct answers and the confusion separately.

When analyzing the results for correct answer percentages and confusion for the task with four possible answers, we find significantly different categorizations for the four consonants [X2(9) = 1850, $p < 0.001$]. Overall, the agreement of the answers with the consonant categories was different from chance and not accidental, being "moderate" according to Cohen's kappa (k) statistics ($k = 0.454$, $p < 0.001$).

3.1 Correct answers

Participants obtained 59.2% of correct answers obtained (well above chance at 25%), i.e., participants categorized the whistled consonants properly, with the results of parts 1 and 3 pooled together. We ran a Generalized Linear Mixed model with Spanish as a second (or third or fourth) language as a Fixed Factor and Participant as a Random effect but found no effect. We ran a global analysis of variance (ANOVA) on participants with repeated measures, that included Consonant type (k, p, s, t) and Part (part 1, part 3) as within factors. We observed that the scores varied significantly depending on the main effect of Consonant type [$F(3, 87) = 16.893$; $p < 0.001$]. Meanwhile, the main effect of Part and the interaction between the two factors were not significant. This suggests that there was no significant increase in performances between Part 1 and Part 3.

Concerning consonant types, "s" and "t" obtained the largest amount of correct answers (respectively, 74.5% and 68.8%), while "k" was intermediate (52.9%) and "p" was the least well-recognized (40%). We also ran *post hoc* multiple comparisons with a Bonferroni correction ($p < 0.05$) revealing that correct "p" categorizations are significantly different from those of "t" and "s" ($p < 0.001$), that "k" correct answers are also significantly different from "s" ($p < 0.001$) and from "t" ($p < 0.02$). This opposes "p" and "k" to "s" and "t" in the following manner:

"t" = "s" > "k" = "p."

3.2 Confusions

Observation of confusion in the incorrect answers allowed us to gain further understanding of the participants' behavior. To look at confusions between consonant types we first ran a non-parametric ANOVA with repeated measures showing that the interaction between the two factors—Played consonant and Answered consonant—was significant ($p < 0.005$). Thus, for each played consonant, we applied a pairwise comparison (Durbin Conover), and the significant differences obtained are presented below and illustrated in Fig. 2.

The image presented shows the proportion of confusion for each consonant played (arrows), as well as the amount of correct answers obtained (by the size of the blue circle of the consonant played). The three different sizes of the smaller consonant bubbles (in black) allow us to illustrate the confusion hierarchies described thanks to pairwise comparisons.

As shown in the image presented in Fig. 2, when /t/ was played, it was mistaken for "s" 12% of the cases (noted t/s), for "p" 9.5% of the cases (t/p), and for "k" 9.7% of the cases (t/k). There are significant differences between the correct answers obtained for /t/ (68.8%, noted t/t) and t/s, t/k, and t/p ($p < 0.001$). These significant differences confirm that the consonant /t/ was mistaken as often for "s," as for "p" and finally for "k."

When the consonant /s/ was played, it was answered as "t" 13.2% of the cases (s/t), as "p" 7% of the cases (s/p), and as "k" 5.3% of the cases (s/k). Here, correct answers obtained for /s/ (74.5%, noted s/s) are significantly different from s/p, s/t, and s/k ($p < 0.001$). There is also a significant difference between s/k and s/t ($p < 0.01$), indicating that /s/ was confused more often with "t" than with "k."

When the consonant /k/ was played, it was confused with "p" 26.4% of the cases (k/p), "t" 15.3% of the cases (k/t), and "s" 5.5% of the cases (k/s). Moreover, the level of correct answers for /k/ (52.9%, noted k/k) was significantly different from k/s, k/t ($p < 0.001$), and k/p ($p < 0.01$). In addition, k/s was significantly different from k/t ($p < 0.001$) and k/p

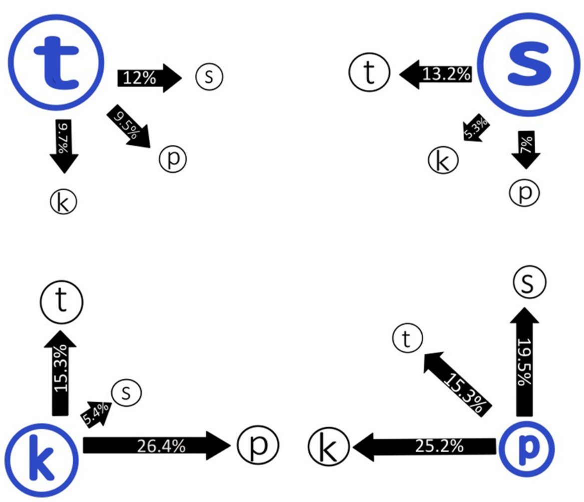

Fig. 2. Schematic representation of the statistical relations in the confusion matrix.

($p < 0.001$). This indicates that /k/ was taken as much for "t" as for "p" (and more than for "s"); however, the percentage of confusion presented above (Fig. 2) suggests that /k/ was most often mistaken for "p," then for "t" and finally for "s."

Finally, when consonant /p/ (40% of correct answers, noted p/p) was played, it was confused with "k" in 25.2% of the cases (p/k), "s" in 19.5% of the cases (p/s), and "t" in 15.3% of the cases (p/t). Moreover, p/p is different from p/t ($p < 0.001$), and from p/s ($p < 0.01$), however, not from p/k. In addition, p/k is different from p/t ($p < 0.05$). This means the /p/ is most often confused with "k," then with "s" and finally with "t."

We can also compare the mirrored confusions (t/p and p/t; t/s and s/t; p/s and s/p; t/k and k/t; and k/s and s/k), where we find a significant difference between p/s and s/p ($p < 0.001$) and a tendency between t/k and k/t ($p = 0.09$). This is not surprising due to the higher levels of consonant recognition for "s" and "t" as opposed to the low levels associated with "k" and "p."

4. Discussion

The overall performance shows that participants recognized the set of consonants well over chance (25%) for every consonant type. General whistled consonant recognition averages at 59.9%, with no significant difference between the first (Part 1) and the last part (Part 3) of the experiment. In addition, there was no impact of experience with Spanish.

Though Part 3 included a greater proportion of new tokens for each consonant (75% of new stimuli), this did not affect the overall performances. This is demonstrated not only by the absence of significant effect between parts but also by the lack of interaction of this factor with the consonant type. One could have expected that the categorization rate would decrease and that the new recordings would not be as well identified as the previous tokens. However, as there was in fact no difference between parts, this could suggest that participants learned consonant categorization and managed to categorize the different variations. One possible explanation is that participants learn from the consonant heard (presented in part 1 and part 2) which could act as an exemplar or instance of the phonological category. It would be interesting to test the exact same stimuli in both parts 1 and 3 to see if a stronger learning effect can be observed.

In addition, our results show that certain consonants are easier to recognize (/s/ or /t/) and others are more difficult to recognize (/p/). The hierarchy derived from the correct answers and confusions shows a preference for the consonants with high loci, or those containing a rising pitch towards these high loci ("s" and "t," see Fig. 1). This could be because the magnitude of the transitional pitch movements is greater than for the low consonants. This could also suggest that pitch movements are easier to identify than changes in articulation, envelope gaps, or interruption duration.

Indeed, correct answers reflect certain preferences, reprising some aspects of previous research. In the hierarchy obtained ("s" = "t" > "p" = "k"), the preference for "s" and "t" corresponds to the opposition between "high frequency modulated whistled consonants" with high loci and whistled consonants with low loci ("k" and "p") (Busnel and Classe, 1976; Trujillo, 2006; Rialland, 2005; Meyer, 2015). The significant difference between the recognition rate of "k" and "p" from those of "s" and "t" suggests that the clear stop which characterizes our /t/, /k/, and /p/ stimuli is not a strong enough characteristic to compete with differences in saliency due to frequency slopes induced by high loci. However, the sharp attack cue, present for "k" and "t," does seem to influence perception, as /t/ is correctly categorized at 68% and is therefore differentiated from /s/.

The relative proportions of confusions reflect similarities in the perception of different consonants. Their comparison enables us to track more closely the phonetic traits to which such similarities may be due. There are three main types of traits coded in whistling: frequency modulations, amplitude envelope modulation, and gap or interruption duration. For example, when looking at the main confusions for each consonant, we note that /s/ is significantly more confused with "t" than with other consonants, reinforcing the interpretation derived from the correct answers obtained, and underlining that the acoustic cues associated with a high locus are key for whistled consonant categorization. Indeed, this is the principal acoustic cue that these consonants share, as /s/ is otherwise semicontinuous with a slower attack than /t/. Such a view is also supported by the high percentage of /t/ mistaken for "s" (even if, due to high variability between listeners, t/s errors are not significantly different from the errors of /t/ for "k" and "p"). Despite the fact that the confusion s/t is preferred over other confusions and the reverse is not the same for t/s, there is no significant asymmetry between s/t and t/s confusions.

The patterns of confusions of the two least well recognized consonants (/k/ and /p/) also highlight interesting phonetic aspects. As we saw earlier, /k/ is answered more as "p" than "t" or "s," and /p/ is answered more as "k," though there is no significant difference between the confusion as p/k and p/s. Interestingly, whistled realizations of /p/ and /t/ both share key common phonetic cues with whistled /k/: /p/, /t/ and /k/ realize a full stop, /t/ and /k/ use a sharp attack, and /p/ and /k/ share the flat frequency shape. The consonant /s/ however shares none of these characteristics with /k/. Moreover, /p/ is answered "k" and "s" at relatively similar proportions (25.12% "k," 19.5% "s," which are not statistically different), while "t" is answered at a significantly lower rate (15.3%, statistically different from p/k). These results may be explained by the fact that a whistled /p/ shares two phonetic traits with /k/ (full stop + flat frequency), one with /s/ (a more gradual attack than /k/ and /t/), and one with /t/—full stop).

Overall, the results strongly confirm the hierarchy found in the correct answers: High loci (frequency shape towards high frequencies) are preferred over other phonetic cues. They also show that when several phonetic cues are shared between two consonants, this augments their probability of confusion. However, the present study does not include enough consonantal types to classify the other key phonetic cues in whistled speech: clear silent gap, sharp/gradual attack. Interestingly, the confusion patterns also underline the relative facility to identify /s/. Does this suggest that continuous sound with pitch change is easiest to identify in extremely modified speech?

All the results highlighted here are confirmed by the asymmetry s/p vs p/s and the tendency t/k vs k/t, as opposed to the symmetries k/p vs p/k and t/s vs s/t. Such asymmetries would be interesting to explore further with more data from the perspective of debates opened by Chang *et al.* (2001).

The relative ease at which /s/ is categorized by naive French listeners also contrasts with the documented difficulty for whistlers to learn to produce it. This asymmetry is all the more interesting as it may have implications for teaching whistled speech in the context of current revitalization of the practice (Díaz, 2017; Meyer, 2021). It also opens the possibility of convergence/divergence in production vs perception during spoken speech acquisition (Moskowitz, 1975).

Finally, the results obtained here for this modified speech form are in line with those previously obtained by studies also dealing with whistled phoneme recognition. (a) Performance levels are coherent with those found by Meyer and colleagues for whistled vowel recognition by untrained listeners (Meyer, 2008; Meyer *et al.*, 2017). (b) This experiment highlighted consonant preferences just as Rialland found for Silbo whistlers (Rialland, 2005). (c) Rates of correct answers + confusion were analyzed similarly to Meyer *et al.* (2019), who also found that /t/ was better recognized than /k/ for traditional whistlers of Tashlhiyt Berber (the other consonants of their test were not tested here).

Overall, with such an approach, we have shown that the naive listener capacity for recognition and categorization found in whistled vowels also applies to whistled consonants, which opens rich experimental possibilities to observe the notion of perceptual flexibility both with non-standard but natural, whistled consonant articulations, and across different language backgrounds.

5. Conclusion

In conclusion, naive French listeners recognize whistled consonants above chance and generally use frequency change to identify the sound correctly, which is coherent with the fact that frequency modulations are the most salient and resilient aspects of the signal with better propagation for long distance communication. These results underline the strong perceptual flexibility present in naive listeners who can successfully identify and attribute these cues to a modified form of speech. This analysis highlights certain phoneme processing methods that could apply to other forms of modified speech, paving the way for more research on whistled speech and processing methods.

Acknowledgments

This work was supported by the doctoral grant of the Université Côte d'Azur attributed to Anaïs Tran Ngoc and the framework of the project GeoTeleLing with the support of the CNRS 80Prime interdisciplinary program. We thank the whistler David for the consonant recordings, and the participants who volunteered their time.

References and links

[1]See supplementary material at https://www.scitation.org/doi/suppl/10.1121/10.0013900 SuppPub1 for a figure showing the Waveform and Spectrogram of a spoken and whistled Spanish sentence, and the Waveform and Spectrogram of the corresponding spoken utterances.

Blanco, N., Meyer, J., Hoen, M., and Meunier, F. (**2018**). "Phoneme resistance and phoneme confusion in noise: Impact of Dyslexia" in *Proceedings of Interspeech 2018*, September 2–6, Hyderabad, India, pp. 2290–2294.

Busnel, R., and Classe, A. (**1976**). *Whistled Languages* (Springer-Verlag, Berlin-Heideleberg).

Busnel, R.-G. (**1970**). "Recherches expérimentales sur la langue sifflée de Kusköy" ("Experimental Research on the Whistling Language of Kusköy"), Rev. Phonétique Appl. **14/15**, 41–57, available at https://eric.ed.gov/?id=EJ041036.

Busnel, R.-G., Moles, A., and Vallencien, B. (**1962**). "Sur l'aspect phonétique d'une languesifflée dans les Pyrénées françaises" ("On the phonetic aspect of whistled speech in the French Pyrenees"), in *Proceedings of the International Congress of Phonetical Science*, September 4–9, Helsinki, Finland, pp. 533–546.

Chang, S., Plauché, M. C., and Ohala, J. J. (**2001**). "Markedness and consonant confusion asymmetries," in *The Role of Speech Perception in Phonology*, edited by E. Hume and K. Johnson (Academic Press, San Diego, CA), pp. 79–101.

Díaz, R. D. (**2017**). *El Lenguaje Silbado en la Isla de El Hierro* (*The Silbo Language on the Island of El Hierro*), 2nd ed. (Tenerife, Le Canarien Ediciones, La Orotava, Spain).

Hervais-Adelman, A., Davis, M., Johnsrude, I., and Carlyon, R. (**2008**). "Perceptual learning of noise vocoded words: Effects of feedback and lexicality," J. Exp. Psychol. Hum. Percept. Perform. **34**(2), 460–474.

Meyer, J. (**2008**). "Acoustic strategy and typology of whistled languages: Phonetic comparison and perceptual cues of whistled vowels," J. Int. Phon. Assoc. **38**(1), 69–94.

Meyer, J. (**2015**). *Whistled Languages, a Worldwide Inquiry on Human Whistled Speech* (Springer, New York).

Meyer, J. (**2021**). "Environmental and linguistic typology of whistled languages," Annu. Rev. Ling. **7**, 493–510.

Meyer, J., Dentel, L., Gerber, S., and Ridouane, R. (**2019**). "A perceptual study of CV syllables in both spoken and whistled speech: A Tashlhiyt Berber perspective," in *Proceedings of Interspeech 2019*, September 15–19, Graz, Austria, pp. 2296–2299.

Meyer, J., Dentel, L., and Meunier, F. (**2017**). "Categorization of natural Spanish whistled vowels by naive listeners of different language background," Front. Psychol. **8**, 25.

Moles, A. (**1970**). "Etude sociolinguistique de la langue sifflée de Kusköy" ("Sociolinguistic Study of the Whistling Language of Kusköy"), Rev. Phonétique App. **14/15**, 78–118, available at https://eric.ed.gov/?id=EJ041037.

Molina Mejia, J. (**2007**). *Diagnostique et Correction Des Erreurs de Prononciation en FLE DesApprenants Hispanophones: Mémoire M1, Direction de Dominique Abry* (*Diagnostic and Error Correction of the Pronunciation of Hispanic Learners of French as a Second Language*) (Grenoble, Université de Stendhal, Gières, France).

Moskowitz, B. A. (**1975**). "The acquisition of Fricatives: a study in phonetics and phonology," in J. Phonetics **3**(3), 141–150.

Rialland, A. (**2005**). "Phonological and phonetic aspects of whistled languages," in *Phonology* (Cambridge University Press, Cambridge, UK), Vol. 22, pp. 237–271.

Shadle, C. H. (**1983**). "Experiments on the acoustics of whistling," Phys. Teach. **21**(3), 148–154.

Trujillo, C. R. (**2006**). *El Silbo Gomero. Nuevo Estudio Fonológico (Edición Bilingüe Español-Inglés) (Academia Canaria de la Lengua El Silbo Gomero: Analisis Linguistico*) [*The Gomeran Whistle. New Phonological Study (Spanish-English Bilingual Edition) (Canarian Academy of Language El 297 Silbo Gomero: Linguistic Analysis*)] (Andres Bello, Santa Cruz de Tenerife, Spain).