# Removing the disguise: the matched guise technique and listener awareness

Kyler Laycock1 and Kevin B McGowan2

1The Ohio State University

2The University of Kentucky

# Author Note

# Abstract

*Keywords*:

# Removing the disguise: the matched guise technique and listener awareness

# Introduction

A great deal of attention has been paid in the phonetics and sociophonetics literatures to the perception of the voiceless fricatives [ʃ] and [s] in English. To a first approximation, these fricatives differ in the distance between the point of lingual articulation and the teeth, which give them their characteristic sibilance Shadle ([1991](#ref-shadle1991)). English [s] has a short resonating chamber behind the teeth; it is typically produced by holding the tongue blade near enough to the alveolar ridge to cause turbulent airflow. English [ʃ] has a comparatively larger resonating chamber; it is typically produced with a more posterior, palato-alveolar tongue position and lip rounding both of which serve to reinforce this posteriority. But listeners do not perceive via {++this type of++}{>>KM<<} first approximation{++; we are sensitive to fine phonetic details far beyond these gross, categorical, differences++}{>>KM<<}. Indeed, these two fricatives have been exciting to researchers precisely because of the sensitivity listeners bring to their perception and how that perception interacts with both linguistic and social knowledge.

## Coarticulatory and Social Information Influence [ʃ]-[s] perception

{>>KL: ok, this is the specific section the editors didn’t like I think vis-a-vis “rewriting the introduction to state more strongly why this study is important to sociolinguistics, and not mainly interesting to cognitive linguists”<<} Listeners are sensitive to articulatory mismatches between the fricatives [ʃ]-[s] and neighboring sounds. Whalen ([1984](#ref-whalen1984)) conducted a series of experiments to investigate listeners’ responses to articulatory mismatches in synthetic speech. Overall, the result of these investigations was that subcategorical phonetic mismatches slow phonetic judgments. In onset position, in isolation, or in coda position, misleading coarticulatory information inhibited reaction times. Listeners, Whalen cautions in the conclusion, are sensitive to articulatory patterns that are below the level of conscious awareness and not available to direct experimenter scrutiny. While listeners will readily fill-in missing or ambiguous information, the presence of actively *conflicting* articulatory information is inhibitory.

A commonly used methodology involves the creation of synthetic fricative continua. These continua have endpoints in prototypical examples of [ʃ] and [s] with some number of equal-sized acoustic steps generated, synthesized, or even mixed between these. Somewhere in the middle of such a continuum will be fricative-like noise that is ambiguous as to category membership: not clearly a [ʃ] and not clearly an [s]. paired a continuum from [ʃ] (2.9 kHz) to [s] (4.4 kHz) with synthetic [æ] vowels to form CV pairs. May found that listeners perceived a higher proportion of the fricative continuum as [ʃ] when paired with vowel stimuli from a smaller vocal tract. The logic here is that smaller resonating chambers between the lingual articulation and teeth will have a higher mean frequency than larger resonating chambers. Listeners’ use of apparent vocal tract size in perception reflect their knowledge of this variation ([Munson, 2011](#ref-munson2011)).

Mann and Repp ([1980](#ref-MannRepp1980)) replicated this finding, extending it to natural productions of vowels spoken by a male or female-identified talker. Similar to May’s results with simulated vocal tract size, Mann & Repp found a higher proportion of the fricative continuum was heard as [ʃ] when paired with the speech of the female talker. This early work, like others of the period ([Ohala, 1984](#ref-ohala1984)), theorized size as being a relatively deterministic feature of talker sexual dimorphism. One consequence of this view is that gender-related variation in the speech signal is considered mechanistic, universal, and following from purely physical laws. Vocal tract size is presumably not available for individual performance and so listener knowledge of this variation can be correspondingly simple. Vocal tract size may influence perception, but it does so implicitly, automatically, and below the level of introspective awareness.

{++ Mann and Repp ([1980](#ref-MannRepp1980)) also replicated and extended previous work ([Kunisaki & Fujisaki, 1977](#ref-kunisakifujisaki1977); [Whalen, 1981](#ref-whalen1981)) demonstrating that listeners report hearing more of the synthetic fricative continuum as [s] when followed by a rounded vowel quality such as English [u] than when followed by an unrounded quality such as [i] or [a]. Listeners experience the fricative continuum differently in the presence of anticipatory coarticulation. The presence of nasal coarticulation on a vowel similarly allows listeners to make a lexical decision between words like *bend* and *bed* as soon as that information is present in the acoustic signal ([Beddor et al., 2013](#X461759748dc6d5ad6d078d8d0fc840233f13e27), [2018](#ref-beddorcoetzeestylermcgowanboland2018)). Mann & Repp’s participants in this study experienced auditory evidence of posteriority in the ambiguous portion of the fricative continuum as the presence of coarticulation with a following rounded vowel. As with vocal tract length, above, the behavioral result was a shift in the listeners’ fricative category boundary toward [s]. ++}{>>KM: clarifying and putting this back at least for now.<<}

Strand and Johnson ([1996](#ref-strandJohnson1996)) conducted a pair of experiments investigating the influence of purported gender of a talker on the perception of the [ʃ]-[s] boundary. In their experiment 1, listeners heard a [ʃ]-[s] continuum paired with voices previously normed as prototypical female, non-prototypical female, non-prototypical male, and prototypical male voices. The result replicates Mann and Repp ([1980](#ref-MannRepp1980)) and extends it to show that the influence of a gendered voice correlates with the protypicality of that voice (exp1). They then extend this research to show that presenting listeners with prototypically-gendered videos of their purported talker can, again, shift perceptions of the [ʃ]-[s] such that listeners report hearing a higher proportion of the continuum as [ʃ] when watching a female talker and a higher proportion of [s] when watching a male talker. The AV condition of their experiment 2 is reminiscent of McGurk and MacDonald ([1976](#ref-McGurkMacDonald1976)) and is presented in that context. A striking feature of the McGurk Effect is its automaticity; participants can not choose to perceive the two components of a fused percept independently. It is unclear from Strand and Johnson ([1996](#ref-strandJohnson1996)) and subsequent work whether the perceptual influence of visually-presented social information is implicit and automatic, like vocal tract size, the McGurk effect, etc., or whether the effect disappears when listeners are aware of the guise manipulation.

This is an incomplete sample of the literature on the perception of these fricatives. We hope, however, that the message is clear that even when arriving at a purely linguistic percept, listeners’ judgments depend on a rich constellation of evidence and expectation. Vocal tract size, formant transitions, following vowel quality ([Mann & Repp, 1980](#ref-MannRepp1980)), and coarticulatory cues, along with the acoustic properties of the fricative itself, can all shape how listeners report experiencing that fricative. Rather than relying on a single, invariant, phonetic cue, listeners take the entire fricative and context into account Whalen ([1991](#ref-whalen1991)).

{~~One imagines~> It is conceivable that~~} {>>KM<<} such exquisite sensitivity to the phonetic cues conveying linguistic category membership might restrict language users’ freedom to communicate and perceive social information via the same phonetic signal. This would be the prediction of a phonetic theory in which linguistic information and social information battle for control of the air waves –where listeners must normalize away social variation to recover linguistic information. Instead, with these fricatives, at least, we can observe the opposite. The fricatives /ʃ/ and /s/ often carry social meaning ([Mack & Munson, 2012](#ref-mackMunson2012b); [Podesva & Kajino, 2014](#ref-podesvakajino2014)) with /s/ being “perhaps the most iconic phonetic variable in the field” ([Calder, 2018](#ref-calder2018)). The implication is that the social and linguistic meanings of particular phonetic cues are not in competition with one another.

## Phonetics, Speech Perception, and the Social-Construction of Gender

La Palma is one of the west most islands in the Volcanic Archipelago of the Canary Islands ([Figure 1](#fig-map)).

[Figure 2](#fig-spatial-plot) shows the location of recent Earthquakes on La Palma.

## Data & Methods

## Conclusion

## References

Beddor, P. S., Coetzee, A. W., Styler, W., McGowan, K. B., & Boland, J. E. (2018). The time course of individuals’ perception of coarticulatory information is linked to their production: Implications for sound change. *Language*, *94*(4), 931–968.

Beddor, P. S., McGowan, K. B., Boland, J. E., Coetzee, A. W., & Brasher, A. (2013). The time course of perception of coarticulation. *The Journal of the Acoustical Society of America*, *133*(4), 2350–2366.

Calder, J. (2018). From “gay lisp” to “fierce queen”: The sociophonetics of sexuality’s most iconic variable. In K. Hall & R. Barrett (Eds.), *The oxford handbook of language and sexuality* (pp. 1–23).

Fant, G. (1960). *Acoustic theory of speech production*. Mouton.

Kunisaki, O., & Fujisaki, H. (1977). On the influence of context upon perception of voiceless fricative consonants. *Annual Bulletin*, *11*, 85–91.

Mack, S., & Munson, B. (2012). The association between/s/quality and perceived sexual orientation of men’s voices: Implicit and explicit measures. *Journal of Phonetics*, *40*(1), 198–212.

Mann, V. A., & Repp, B. H. (1980). Influence of vocalic context on perception of the [ʃ]-[s] distinction. *Perception & Psychophysics*, *28*(3), 213–228.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746–748.

Munson, B. (2011). The influence of actual and imputed talker gender on fricative perception, revisited (l). *The Journal of the Acoustical Society of America*, *130*(5), 2631–2634.

Ohala, J. J. (1984). An ethological perspective on common cross-language utilization of F₀ of voice. *Phonetica*, *41*(1), 1–16.

Podesva, R. J., & Kajino, S. (2014). Sociophonetics, gender, and sexuality. *The Handbook of Language, Gender, and Sexuality*, 103–122.

Shadle, C. H. (1991). The effect of geometry on source mechanisms of fricative consonants. *Journal of Phonetics*, *19*(3-4), 409–424.

Strand, E. A. (1999). Uncovering the role of gender stereotypes in speech perception. *Journal of Language and Social Psychology*, *18*(1), 86–100.

Strand, E. A., & Johnson, K. (1996). Gradient and visual speaker normalization in the perception of fricatives. *KONVENS*, 14–26.

Whalen, D. H. (1981). Effects of vocalic formant transitions and vowel quality on the english [s]–[š] boundary. *The Journal of the Acoustical Society of America*, *69*(1), 275–282.

Whalen, D. H. (1984). Subcategorical phonetic mismatches slow phonetic judgments. *Perception & Psychophysics*, *35*, 49–64.

Whalen, D. H. (1991). Perception of the english/s/–//distinction relies on fricative noises and transitions, not on brief spectral slices. *The Journal of the Acoustical Society of America*, *90*(4), 1776–1785.

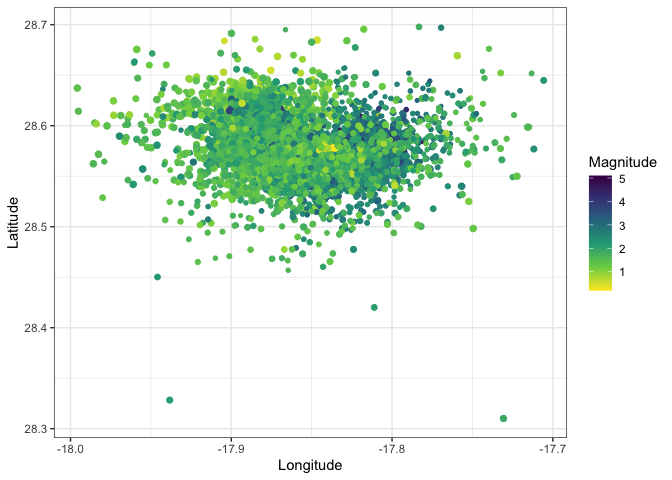
Figure 1

Map of La Palma



Figure 2

Locations of earthquakes on La Palma since 2017



Source: [Explore Earthquakes](https://kbmcgowan.github.io/play-manuscript/notebooks/explore-earthquakes.qmd.html#cell-fig-spatial-plot)