



BROWN

LAUREL VS. YANNY: FORMANTS BEHAVING BADLY

CHELSEA SANKER CHELSEA_SANKER@BROWN.EDU
26TH MANCHESTER PHONOLOGY MEETING

Introduction

- Listeners are remarkably accurate in perceiving phonological categories across speakers, though context and acclimation to a speaker improve accuracy (Ladefoged & Broadbent 1957, Mullennix, Pisoni, & Martin 1989)

- The recent viral clip heard variously as the startlingly dissimilar *laurel* and *yanny* (Matsakis 2018) provides an auditory illusion based on ambiguous formant structure

F1 and F2 in the clip are low intensity and close in frequency; some listeners interpret them as a single formant (*yanny*)

Other listeners hear them as distinct formants (*laurel*)

- Listeners hear one item or the other, with no possibility for anything in between

- Manipulations of the clip demonstrate the importance of the relative intensity of formants in allowing listeners to disambiguate (Katz, Corum, & Huang 2018)

Notable Behaviors

- The two perceived forms are phonologically extremely distinct; ambiguity is not due to sounds that are at category boundaries

While some listeners are able to switch between hearing the two forms, perception is categorical. The two possibilities are based on different frames of perception of the formant information

- The two forms sound like different voices of different timbre

Some accounts describe *yanny* as 'higher' and *laurel* as 'lower'

The *laurel* interpretation has a formant structure reflecting a long vocal tract, while *yanny* has structure that would be produced by a much shorter vocal tract

- Differences in clarity and naturalness

Laurel sounds like a normal human voice, with no variation in the particular sounds listeners hear, as it reflects the natural formant structure of a real speaker

Yanny sounds unnatural, because it doesn't follow prototypical phonological patterns; many listeners disagree about the central sonorant, as it doesn't align well with any English consonant

- Manipulating intensity of different frequencies changes perception

Increasing low-frequency intensity increases perception of *laurel*, as it provides more information to help separate the formants, in addition to being more typical for merged formants

Increasing high-frequency intensity increases perception of *yanny* (and variants), because the relative intensity of F1 and F2 is even lower, making them harder to perceive as two formants

The Recording

- The clip originally comes from a male speaker saying *laurel*. Figure 1 illustrates this version; Praat readily identifies both F1 and F2.

- The widely discussed version of the clip comes from a recording of it being played through one listener's speakers. In addition to more noise, low frequencies are notably attenuated in this version, as illustrated in Figure 2. Automatic formant tracking, marked here, captures the *yanny* pattern.

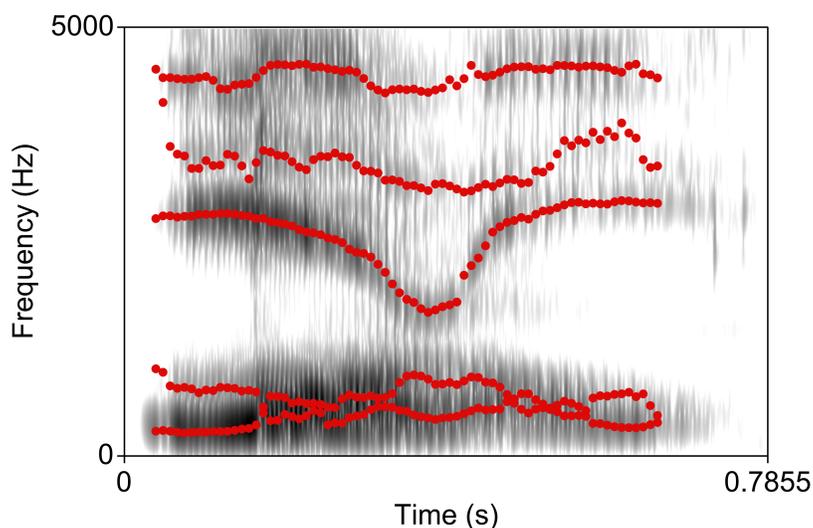


Figure 1: Original recording, *laurel* [tɔɹɪ]

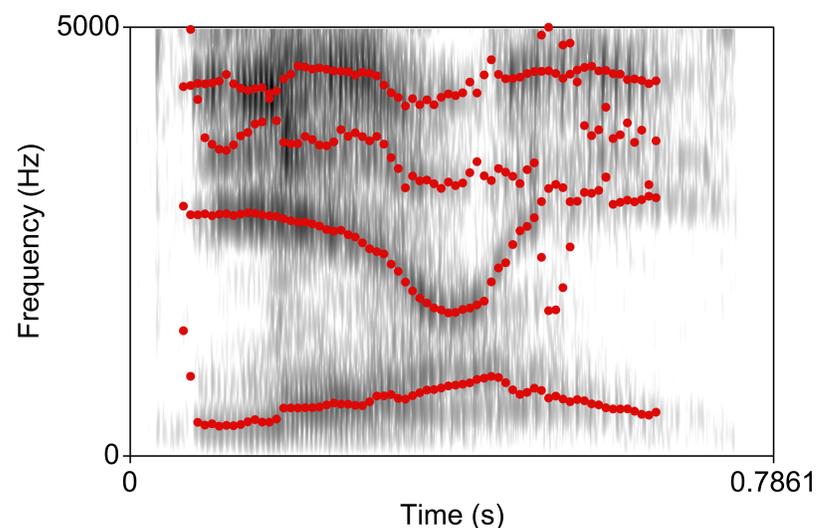


Figure 2: The ambiguous clip, *laurel* [tɔɹɪ] OR *yanny* [jɛni] (and variants)

Acoustic-Phonological Mapping

- The production of *laurel* has close low frequency F1 and F2 throughout; intensity at these low frequencies in the ambiguous clip is weak

- Interpreting F1 and F2 as separate preserves the original formant pattern, sounding like a normal male voice saying *laurel*

- Interpreting F1 and F2 as a single formant produces a structure in which the original F3 is interpreted as F2, sounding like an unnatural form of something like *yanny*

The low F1+F2 and moderate F3 of [t] and [ɹ] instead become a low F1 and high F2 for [j]/[i]

The low F3 of [ɹ] instead becomes a moderate F2, consistent with an alveolar constriction but not typical of any English consonant; the transition resembles a common prenasal realization of /æ/

Individual Variation

What influences the form which different listeners hear?

- The speakers the clip is played on: in particular, the strength of low frequencies

- Differences within listeners perhaps include:
Attention to detail (cf. Yu 2013), e.g. separation of close formants and relative intensity of individual vs. merged formants

Dialect/native language(s), e.g. how low the expected F2 of /o/ is, depending on fronting

Exposure to different voices, setting expectations about formant structures

- More demographic data and information about listeners' other perception behaviors may help clarify the differences

References

Katz, J., Corum, J., & Huang, J. (2018, May 16). We made a tool so you can hear both yanny and laurel. *New York Times*.
Ladefoged, P., & Broadbent, D. E. (1957). Information conveyed by vowels. *The Journal of the acoustical society of America*, 29(1), 98–104.
Matsakis, L. (2018, May 16). The true history of 'yanny' and 'laurel'. *WIRED*.
Mullennix, J. W., Pisoni, D. B., & Martin, C. S. (1989). Some effects of talker variability on spoken word recognition. *The Journal of the Acoustical Society of America*, 85(1), 365–378.
Yu, A. (2013). Individual differences in socio-cognitive processing and the actuation of sound change. In A. Yu (Ed.), *Origins of Sound Change: Approaches to Phonologization* (201–227). Oxford University Press, Oxford, England.