Prosody and Phonology

John J. Ohala* Alexandra Dunn†, & Ronald Sprouse*

*Department of Linguistics, University of California, Berkeley 94720-2650
ohala@socrates.berkeley.edu

†Department of Linguistics
University of California, Santa Cruz
pipper_dunski@mac.com

Abstract

We address the problem of how to differentiate between phonetically- and phonologically-caused aspects of prosody — those that arise from purely physical phonetic factors and are not reflected in the mental lexicon — vs. phonologically-maintained aspects — those that have a psychological component, i.e., arise from the representation in the mental lexicon. Two case studies are reported: the first focusing on the F0 perturbation caused by pre-vocalic voiced and voiceless consonants, and the second, F0 declination in utterances. Differentiating between these two distinct sources of contextual variation may involve testing the influence of posited phonetic causes one-by-one.

1. Introduction

In the life history of most common sound patterns found in unrelated languages, i.e., those deemed to be ‘universal’ at least in tendency, the first two stages are (a) phonetically-predictable variation due to articulation or perception and (b) the fossilization of these phonetic variants — what is called ‘phonologization’ — due, we would maintain, to listener-speakers making them purposeful (having a psychological representation). For example, there is often an emergent (or ‘epenthetic’) stop between a nasal and a following oral obstruent: warmth [wɔrmθ], youngest ['jʌŋstə]. This arises for phonetic reasons, essentially anticipatory assimilation during the latter portion of the nasal of the velic elevation required of the following obstruent (Ohala 1997). Such an unintended phonetic event, arising out of the transition between two intended sounds, may confuse the listener and it may be taken as purposeful, i.e., phonologized. Once phonologized, this emergent stop is produced purposely — that is, is present in the mental lexical representation of the listener. Historically a number of words show the phonologization of what must originally have been a stop emerging phonetically as described above: e.g., English dempster from deem + ster (a term for ‘judge’), Thompson < Thom + son. The accepted spellings with a ‘p’ would suggest that the emergent stop is now phonologized.

Sound patterns in language prosody may also be either phonetic or phonological. The question addressed here is how to differentiate the two stages. We consider two case studies: (a) the fundamental frequency (F0) perturbation on vowels following voiced and voiceless consonants and (b) F0 declination, i.e., the slow drop in F0 from beginning to end of phonological phrases. To anticipate the answer to the question: it is not easy to do such differentiation of stages, especially when the contextual factors (as in the case of the nasal + oral obstruent in the example cited above) remain. What we discuss here is the classic method of tracing effects to their posited causes: systematically varying the supposed causes and seeing whether the effect is still manifested.

2. Two Case Studies

2.1. F0 perturbation by voiced and voiceless consonants.

Philological evidence from Chinese (Edkins 1888) reveal that the development of lexical distinctive tone was influenced by the voicing of the obstruent that originally preceding the tone-bearing vowel. Also early phonetic studies (Crandall 1925) showed that even in non-tone languages like English the onset F0 on vowels was affected by the voicing of the preceding obstruent. In both cases, either the phonetic F0 or the lexically-distinct tone was lower after voiced than after voiceless obstruents. But why does this pattern occur? There have been two dominant hypotheses: (a) the aerodynamic conditions are different at vowel onset after voiced and voiceless obstruents — the pressure drop, \( \Delta P_{glott} = P_{subglottal} - P_{oral} \) is thought be higher after voiceless obstruents than after voiced ones and (b) it is thought that the tension of the vocal cords is higher during voiceless obstruents than after voiced ones. Despite this difference the onset of tone was lower after voiced than after voiceless obstruents. In both cases, the F0 perturbation was found to be more pronounced after voiceless obstruents. The reason for this difference is still unclear and is the focus of this paper.

2.1.1. Method

We attempted to measure the effect on the post-consonantal F0 perturbation when we artificially perturbed the \( \Delta P_{glott} \) (Ohala & Sprouse 2003). This was done in two ways: speakers produced targeted speech samples while the pharyngeal pressure was reduced (a) by a relatively large catheter (inner diameter = 6 mm) inserted into the oral cavity via the buccal sulcus (the space between the teeth, the upper teeth in this case, and the cheek) and bending into the oral cavity behind the upper molar; the external opening was opened and closed by another person at moments unpredictable to the speaker, and (b) by inserting a tube with a smaller diameter (outer diameter = 4 mm) into the pharynx via the nasal passage and connecting the outer end to a vacuum source; in this case, due to the noise and vibration of the onset of the vacuum connection, the timing could not be done without the speaker being aware of it. The buccal sulcus...
2.2.1 Method

Ten speakers, adult males and females, were given a list of seven scripted sentences or paragraphs to study. They uttered in total 13 sentences, some of them isolated sentences and some of them in short paragraphs. These were spoken one time egressively and one time ingressively. Thus there were a total of 260 scripted utterances (13 sentences x 2 airstream mechanisms x 10 speakers).1 (Of the 130 ingressive samples, some were unanalyzable for F0 because of aperiodicity.) The speech was directly digitized on a computer, using Speech Analyzer. Pitch plot information was also computed using the Speech Analyzer software.

Each of the subjects was briefly trained on how to speak ingressively and was then given several minutes to practice before the recording began. Many subjects did not feel confident about speaking ingressively, and had to take several breaths during the articulation of longer sentences, or series of sentences. Most of these breaths were at grammatical pauses, or at the ending of ideas. The difference in F0 before and after the breaths was negligible. In determining the presence or absence of declination and its degree, we used a subjective but generous criterion for detecting F0 declination by visual examination of F0 traces (as is common in the literature on declination).

2.2.2 Results

Whereas 75% of the egressive utterances had some negative F0 slope (declination), only 50% of the measurable ingressive utterances did. But the important point is that in spite of the rib cage increasing in volume, declination could still be found in the ingressive samples. We take this as partial support for our claim that declination is purposely imposed on speech by speakers who have the time and presence of mind to be considerate of their listeners and thus appears most commonly in prepared or scripted speech.

This experiment did not directly test the hypothesis that declination is due to the collapse of the lungs from beginning to end of an utterance because the crucial aspect of this

---

1 Three subjects had prior knowledge of this experiment. The other seven did not. There was no evident difference in the data between these two groups.
hypothesis is the decrease in the $\Delta P_{\text{pl}}$. Although in ingressive speech the lung volume is progressively augmented, not diminished, there is still a progressive decrease in $\Delta P_{\text{pl}}$, the difference between $P_i$ and $P_{\text{mon}}$, which in ingressive speech is negative, but still diminishing from beginning to end of an utterance.

However, there is some suggestion in the data that declination is purposeful and thus phonological, not phonetic, since the rate of declination varies with utterance length. There is a broadly inverse correlation between length of utterance and the degree of (negative) declination. In other words, short utterances have a steeper rate of F0 decline than long utterances. If declination were purely a function of lung volume one might expect much the same rate of F0 declination no matter how long the utterance.

3. Conclusion

Variations in speech prosody, like many other phonetic details, may be caused by purely phonetic factors—those that arise automatically from the constraints of the speech production mechanism—or by phonological factors—those that are part of the mental representation that directs speech production. Differentiating between these two factors is important for an understanding of the speech production process. To do this it may be necessary to systematically evaluate, one by one, all hypotheses attributing the variation to phonetic processes. To the extent that these hypotheses are falsified, one can have increased confidence that they may be phonological. Two cases studies were presented. One evaluated a hypothesis that the F0 perturbation that voiced and voiceless obstruents create on following vowel were caused by aerodynamic factors. This hypothesis was falsified. A second study evaluated the hypothesis that the commonly seen F0 declination in scripted utterances was due to the diminishing volume of the lungs and the consequent downward movement of the larynx via its connections to the sternum. When speakers used ingressive speech there was still a high incidence of F0 declination, though less than with egressive speech. We took this as undermining this hypothesis.

4. References


