

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-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* [wɔ:mpθ], *youngster* [ˈjʌŋkstɚ]. 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, $\Delta P_{\text{glot}} = P_{\text{subglottal}} - P_{\text{oral}}$ is thought to 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 consonants than voiced ones (Halle and Stevens 1971). There is some instrumental data in support of (b) (Löfqvist et al. 1989). We sought to further test these two hypotheses.

2.1.1. Method

We attempted to measure the effect on the post-consonantal F0 perturbation when we artificially perturbed the ΔP_{glot} (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 molars; 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

method was used with two speakers; the vacuum method with three. Audio was recorded and the F0 extracted by the software PCQuirer.

2.1.2 Results and Discussion

Both techniques gave substantially the same results. Fig. 1 shows the F0 contours on the test word *sigh* [saɪ] with the pharyngeal pressure vented or not with the tube inserted via the buccal sulcus (averages of 10 tokens in both conditions).

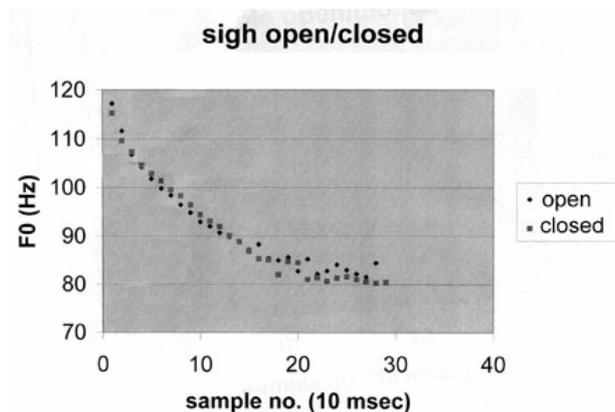


Fig. 1 F0 on the vowel following the [s] of [saɪ] when the pharyngeal pressure was vented or not by a tube inserted into the oral cavity via the buccal sulcus. Subject: JJO.

No measurable difference was found in the F0 contours on the target syllables under the two conditions, using either method of reducing pharyngeal pressure. Based on these results, one could conclude that the hypothesis that the F0 perturbations are due to aerodynamic factors is found wanting. The hypothesis that they are due to differences in vocal cord tension is still viable. However, these methods did not directly test this latter hypothesis and it remains to evaluate it against the hypothesis that the F0 contours are phonological, i.e., purposeful, even in non-tonal languages (J. Kingston, personal communication, 2004).

2.2. Long-term F0 Declination

There is substantial literature documenting that voice F0 drops gradually from the beginning to the end of spoken sentences or phrases (Bolinger 1964; Cohen & 't Hart 1967; Maeda, 1974, 1976). This trend has been variously named "downdrift" or "declination" (the term that will be used in this paper). However, there are also claims that F0 declination has lower incidence or is absent entirely from unscripted speech (Umeda 1982); indeed most of the studies supporting its existence have used scripted speech. One problem with research in this area is that there seems as yet to be no universally-agreed-upon metric for quantitatively defining and measuring declination.

There are two principal hypotheses that propose that declination is an automatic consequence of physiological constraints. Maeda (1974, 1976) suggests it is due to the collapsing rib cage creating a downward pull on the larynx via

the sternum. Strik & Boves (1995), among others, maintain that it is the diminishing air supply in the lungs which leads to a decrease in subglottal air pressure (P_s) (strictly speaking, the pressure transglottal difference between subglottal and atmospheric air P_{atmos} , $\Delta P_{glot} = P_s - P_{atmos}$), the magnitude of which can influence voice F0. An alternative hypothesis is that F0 declination is a purposeful refinement that can be imposed on speech by speakers who have the advantage of being able to plan their speech. Presumably this would be for the sake of the listener – perhaps to allow for easier parsing.

We sought to test at least one of the hypotheses that declination is automatic, namely, Maeda's. We did this by seeing if declination appears in *ingressive* speech, i.e., where the lungs are being filled with air as the utterance progresses.

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 F_0 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 ΔP_{glot} . Although in ingressive speech the lung volume is progressively augmented, not diminished, there is still a progressive decrease in ΔP_{glot} , the difference between P_s and P_{atmos} , 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

- Bolinger, D. L. 1964. Intonation as a universal. *Proc. 9th Int. Congr. Linguists, Cambridge 1962*. The Hague: Mouton. 833-844.
- Cohen, J. & 't Hart. 1967. On the anatomy of intonation. *Lingua*. 19: 177-192.
- Crandall, I. B. 1925. The sounds of speech. *Bell System Technical J* 4.586-626.
- Edkins, J. 1888. *The evolution of the Chinese language*. London. Trubner
- Halle, M. and K. N. Stevens. 1971. "A note on laryngeal features," *Quarterly Progress Report, MIT Res. Lab. of Electronics*, vol. 101, pp. 198-213.
- Kingston, J. personal communication, 2004.

Löfqvist, A., T. Baer, N. S. McGarr, and R. S. Story. 1989. The cricothyroid muscle in voicing control, *J. Acoust. Soc. Am.* 85.1314-1321.

Maeda, S. 1974. A characterization of fundamental frequency contours of speech. *Speech Communication*. 16: 193-210.

Maeda, S. 1976. *A characterization of American English Intonation*. Ph.D. thesis, MIT, Cambridge, MA.

Ohala, J. J. 1997. Emergent stops. *Proc. Seoul International Conference on Linguistics*. Seoul: Linguistic Society of Korea. 84-91.

Ohala, J. J. & Sprouse, R. 2003. Effects on speech of introducing aerodynamic perturbations. Paper presented at the 15h International Congress of Phonetic Sciences, 3-9 August 2003.

Strik, H. & L. Boves. 1995. Downtrend in F_0 and P_{sb} . *Journal of Phonetics*. 23: 203-220.

Umeda, N. 1982. "F₀ declination" is situation dependent. *Journal of Phonetics*. 10: 279-290.