Obstruent voicing effects on F0, but without voicing: Phonetic correlates of Swiss German lenis, fortis, and aspirated stops

D. Robert Ladd [a,*], Stephan Schmid [b]

[a] University of Edinburgh, United Kingdom
[b] University of Zurich, Switzerland

ABSTRACT

It is well known that what are commonly called voicing contrasts in many languages are accompanied by effects on the fundamental frequency (F0) of a following vowel: roughly, F0 is higher after 'voiceless' and lower after 'voiced' obstruents. This is true regardless of how the voicing contrast is manifested in differences of voice onset time (VOT). Such effects potentially provide a window on the nature of voicing itself, but our knowledge is based primarily on typical European two-way voicing contrasts. Here we present a detailed study of voice onset time (VOT), closure duration, and obstruent F0 effects in Zurich Swiss German. The native two-way contrast in oral stops (often termed fortis/lenis) is unusual in being signalled not by VOT – both types are unaspirated – but primarily by closure duration. We confirm studies showing that this distinction is indeed based on duration, and we show for the first time that both types are accompanied by F0 effects that are typical of voiceless obstruents in other languages. In addition, Swiss German has a smallish set of words conventionally pronounced with voiceless aspirated stops. We investigate the VOT and F0 effects of these marginally contrastive aspirated stops, showing that they do exhibit long VOT and are accompanied by a different pattern of F0 effects that is much more variable than that found with fortis and lenis stops. Our findings support the view that the phonetic basis of voicing and related distinctions involves complex interactions of timing and articulatory gestures that cannot always be characterised in terms of a simple VOT continuum from 'voiced' to 'voiceless aspirated' or a simple phonological dichotomy between 'aspirating' and 'true voicing' languages.

1. Introduction

This special issue of *Journal of Phonetics* is inspired by the runaway success of the concept of voice onset time (VOT) in the description of voicing contrasts in the world's languages. VOT was first proposed as a phonetic dimension by Lisker and Abramson (1964), who observed that in a sample of eleven languages with different two-way, three-way, and four-way stop contrasts, there were striking regularities in the relative timing of onset of phonation and release of the word-initial stop closure. Specifically, across languages, regardless of the number of contrastive categories, there were three clusters of values for VOT: a substantial (60–100 ms) voicing lead, a very short (0–20 ms) voicing lag, and a rather longer (50–80 ms) lag corresponding to 'aspiration'. On the basis of these findings, Lisker and Abramson argued that the single dimension of

VOT (or more precisely, relative laryngeal timing; see Abramson, 1977) captured something essential about the phonetic basis of voicing contrasts.

The fact that VOT data from a range of languages seemed to fall into three distinct clusters, which was unexpected, has justifiably influenced thinking in both phonetics and phonology ever since. The most prominent current phonological echo of Lisker and Abramson's work is the theory of 'laryngeal realism' (e.g. Beckman, Jessen, & Ringen, 2013), which proposes a universal feature [voice] and a universal feature [spread glottis]; proponents of this idea argue that languages with two-way stop distinctions are generally either 'true voicing' languages (based on [voice] and typically contrasting fully voiced stops with short-lag voiceless unaspirated ones) or 'aspirating' languages (based on [spread glottis] and typically contrasting short-lag stops with voiceless aspirated ones). However, the apparent separation of clusters of 'unaspirated' and 'aspirated' voiceless stops along the VOT dimension may be to some extent an artifact of their limited sample. The larger study by

Cho and Ladefoged (1999) suggests that there is a more or less continuous range of typical VOT values up to about 100 ms, and shows that at least a few languages have values as high as 140 ms; the relevance of this larger linguistic dataset for universalist phonological claims like 'laryngeal realism' has not yet been explored. More generally, despite the unquestioned usefulness of the VOT concept (reviewed at length by Abramson and Whalen, 2017), the time seems ripe for a reconsideration of cases that fit awkwardly into the Lisker-Abramson scheme.

Lisker and Abramson's own sample presents certain problems for their reductive approach to the phonetic data, particularly in languages with more than a two-way laryngeal contrast. They acknowledged that the distinction between the 'voiced unaspirated' and 'voiced aspirated' stops of the Indo-Aryan languages (many of which show a four-way laryngeal contrast) requires an additional phonetic dimension, and that the difference between 'tense' and 'lax' stops in Korean (which has three laryngeally distinct stop types) cannot be expressed in terms of VOT alone. It seems at least possible that systems with more than a two-way laryngeal contrast are likely to make use of other phonetic properties besides relative laryngeal timing to ensure the robustness of their lexical contrasts. With this in mind, the purpose of this paper is to report a large laboratory study of the unusual set of laryngeal contrasts in Zurich Swiss German.

Swiss German (*Schwyzertüütsch*) is a rather heterogeneous group of German dialects in the Alemannic dialect continuum that extends from Alsace through southwestern Germany and Switzerland into westernmost Austria. In German-speaking Switzerland, unlike Austria and Germany, the status of Alemannic dialects is affected by the existence of stable diglossia (Ferguson, 1959), i.e. the systematic use of distinct 'high' and 'low' language varieties in the same speech community, namely Standard German ('high') and a Swiss German dialect ('low'). While Standard German is used for most written purposes, and as a spoken language in some formal contexts (education, law, much broadcasting), most people – of all social classes – speak a Swiss German dialect in their daily lives, and this is what gives Swiss German a certain sociolinguistic unity despite its dialectal diversity. There are many widely shared phonetic features across the different Swiss German dialects, especially in their consonant systems; although our study is specifically based on speakers from the Zurich area, we believe that the instrumental findings reported here will apply quite generally.

Swiss German is generally described as having a 'fortis/lenis' or 'tense/lax' contrast rather than a voicing contrast. The phonetic sketch of Zurich Swiss German presented by Fleischer and Schmid (2006) takes for granted that the fortis/lenis distinction is pervasive in the obstruent inventory, corresponding to distinctions based on voicing in many other West Germanic varieties. In keeping with a long-standing tradition in the use of the IPA, Fleischer and Schmid transcribe lenis obstruents using the IPA symbols for voiced obstruents together with the devoicing diacritic, distinguishing for example between /p/ and /b̥/ or /t/ and /d̥/. Typical word-medial minimal pairs are /ˈhuːpə/ 'honk (a car horn)' vs. /ˈhuːb̥ə/ 'bonnet, hood of a car' and /ˈlɒtə/ 'lath' vs. /ˈlɒd̥ə/ 'store, shop'. As the examples just given show, the word-medial distinction is found after

both long and short vowels, which are phonologically distinct (Schmid, 2004). The contrast is also lexically exploited both word-finally and word-initially: cf. /ĝrɒːt/ 'ridge' vs. /ĝrɒːd̥/ 'degree', and /tɒː/ 'done' ∼ /d̥ɒː/ 'here'. Phonetic evidence reviewed in the next section makes clear (and our study confirms) that the fortis and lenis stops do not differ in VOT and that both are unaspirated; the contrast is manifested most obviously by greater closure duration.

However, we extend the investigation of the acoustic correlates of this distinction into new phonetic territory in two ways. First, we investigate the patterns of F0 perturbation on the vowel following fortis and lenis stops, looking for differences of the sort that are commonly associated with 'voicing' distinctions in many languages (House & Fairbanks, 1953; Kingston & Diehl, 1994). Previous work on the Swiss fortis-lenis distinction, discussed in Section 2.2, has investigated a number of its articulatory and acoustic aspects, but never, to our knowledge, its effect on F0. Second, in addition to the native Swiss German fortis and lenis stops, we investigate the phonetics of words (mainly but by no means exclusively loanwords) that are normally pronounced with aspirated stops. While the phonological status of the aspirated stops is debatable, the presence of three distinct phonetic categories (lenis, fortis, aspirated) allows us to compare not only fortis with lenis, but fortis (unaspirated) with aspirated as well.

2. Background

2.1. 'Fortis' and 'lenis'

Given the background just sketched, it comes as no surprise that the history of the terms 'fortis' and 'lenis' is closely connected with the study of Swiss German. The terms appear in Sievers (1876), referring to a contrast between two homorganic stops realized through a difference in intensity and duration. Sievers stated that the terms had been suggested by his student Winteler (1876), who had just finished his PhD thesis on a Swiss German dialect, though Braun (1988) notes that they had already been used previously by other scholars. In any case, fortis/lenis (or the ostensibly equivalent pair tense/lax) have often been used since then in dialectological and phonetic studies of Swiss German (e.g. Heusler, 1888; Dieth, 1950; Willi, 1996). Jakobson and Halle (1964), quoting Winteler, explicitly mentioned the 'Swiss German consonantal pattern' as an example of their proposed universal feature [±tense], which was intended to cover duration and height differences between vowels and duration and intensity differences between obstruents.

Beyond Swiss German, the dichotomy fortis vs. lenis has been used in descriptions of numerous languages, referring to various different mechanisms of speech production both at the supraglottal and the glottal level. For instance, in Kohler's (1984) attempt to establish a typology of consonantal contrasts relying on this single universal feature, fortis vs. lenis is conceived of as a 'power feature' which interweaves the dimensions of timing and tension. In this view, fortis and lenis can surface in a variety of different phonetic manifestations. For instance, lenis stops can be 'passively voiced' (as in German) but also 'actively voiced' (as in Italian); fortis stops can appear as geminates (as in Finnish) or can involve tension of the vocal