



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

## Effects of obstruent voicing on vowel F0

**Citation for published version:**

Kirby, JP & Ladd, R 2016, 'Effects of obstruent voicing on vowel F0: Evidence from "true voicing" languages', *The Journal of the Acoustical Society of America*, vol. 140, no. 4, pp. 2400-2411. <https://doi.org/10.1121/1.4962445>

**Digital Object Identifier (DOI):**

[10.1121/1.4962445](https://doi.org/10.1121/1.4962445)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Publisher's PDF, also known as Version of record

**Published In:**

The Journal of the Acoustical Society of America

**Publisher Rights Statement:**

Copyright 2016 Acoustical Society of America. This article may be downloaded for personal use only. Any other use requires prior permission of the author and the Acoustical Society of America.

The Journal of the Acoustical Society of America 140, 2400 (2016); doi: <http://dx.doi.org/10.1121/1.4962445>

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# Effects of obstruent voicing on vowel $F_0$ : Evidence from “true voicing” languages<sup>a)</sup>

James P. Kirby<sup>b)</sup> and D. Robert Ladd

School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Dugald Stewart Building,  
3 Charles Street, Edinburgh EH8 9AD, Scotland

(Received 14 September 2015; revised 12 August 2016; accepted 22 August 2016; published online  
7 October 2016)

This study investigates consonant-related  $F_0$  perturbations (“ $CF_0$ ”) in French and Italian by comparing the effects of voiced and voiceless obstruents on  $F_0$  to those of voiced sonorants. The voiceless obstruents /p f/ in both languages are found to have  $F_0$ -raising properties similar to American English voiceless obstruents, while  $F_0$  following the (pre)voiced obstruents /b v/ in French and Italian patterns together with /m/, again similar to English [Hanson (2009). *J. Acoust. Soc. Am.* **125**(1), 425–441]. In both languages,  $F_0$  is significantly depressed, relative to sonorants, during the closure for voiced obstruents, but cannot be differentiated from sonorants following the release of oral constriction. These findings are taken as support for a model on which  $F_0$  perturbations are fundamentally the result of laryngeal maneuvers initiated to sustain or inhibit phonation, regardless of other language-particular aspects of phonetic realization.

© 2016 Acoustical Society of America. [<http://dx.doi.org/10.1121/1.4962445>]

[LK]

Pages: 2400–2411

## I. INTRODUCTION

It has been known for well over half a century that speech fundamental frequency ( $F_0$ ) can be affected by intrinsic properties (or “microprosody”) of both vowels and consonants. In particular, high vowels are known to have higher intrinsic  $F_0$  than low vowels; similarly,  $F_0$  following voiceless obstruents tends to be significantly higher than  $F_0$  following voiced obstruents (House and Fairbanks, 1953). Following Kingston (2007), we will refer to these intrinsic  $F_0$  perturbations as  $VF_0$  and  $CF_0$ , respectively, but in this paper our focus will be on  $CF_0$ .

Although the basic facts about  $CF_0$  have been established for some time, proposed explanations of its cause are not always compatible, and make sometimes contradictory predictions. In particular, while there is now a good deal of evidence that a voiceless obstruent can raise  $F_0$  at the onset of a following vowel, it is less clear what effect, if any, voiced obstruents have on  $F_0$ . In this paper, we study  $CF_0$  in two languages with prevoiced stops, French and Italian, considering the time course of  $F_0$  during both the consonant and the following vowel. We show that  $F_0$  is lowered during the closure phase of voiced obstruents and raised following release of voiceless obstruents, suggesting that both effects result from laryngeal maneuvers to facilitate or inhibit phonation, respectively. This, in turn, suggests that obstruent-related perturbations of  $F_0$  arise primarily because of articulatory, rather than perceptual, contingencies related to the production of obstruents.

## A. Background

### 1. Previous work on $CF_0$

In what seems to have been the first systematic study of the “secondary” acoustic characteristics of vowels (i.e., duration, fundamental frequency, and intensity), House and Fairbanks (1953) measured the properties of six American English vowel phonemes flanked by identical consonants that differed in voicing, place and manner of articulation in isolated nonsense sequences like [hə'pip], [hə'mam], and [hə'dud]. They reported that “the [mean] fundamental frequencies of vowels in voiceless environments are invariably higher than those in voiced environments” (1953, p. 109). They also reported a small study showing that the acoustic differences are most pronounced at the beginning of the vowel.

A few years later, in a rather larger study motivated by questions related to intonation rather than the acoustics of vowels, Lehiste and Peterson (1961) refined and extended House and Fairbanks’s findings. They confirmed House and Fairbanks’s preliminary conclusion that  $F_0$  perturbations at vowel onset are primarily due to the voicing of the consonant that precedes the vowel, not the whole consonantal context, and that voicing gives rise to a different  $F_0$  trajectory across the vowel of test words spoken in carrier sentences. Especially important here is Lehiste and Peterson’s observation that the overall trajectory of  $F_0$  is primarily determined by linguistic specifications of intonation, and that  $CF_0$  effects, whatever their underlying cause, must be defined relative to the course that  $F_0$  would have taken if the conditions that give rise to the perturbations were not present. This idea was subsequently developed in work by Lea (1973), Kohler (1982, 1984, 1985), Silverman (1987), and Hanson (2009), and in what follows, we take it for granted that  $CF_0$  effects can best be

<sup>a)</sup>Portions of this work were previously presented at the 14th Conference on Laboratory Phonology, Tokyo, July 2014.

<sup>b)</sup>Electronic mail: j.kirby@ed.ac.uk

understood as a deflection from an underlying linguistic intention (e.g., a lexical tone or an intonation contour).

In the subsequent decades, a large number of studies have replicated and extended the finding that voiceless obstruents raise  $F_0$  of the following vowels, in English as well as in other languages (for some reviews of the literature see Hanson, 2009; Chen, 2011). The question then naturally arises what the source of such an effect might be. An early and still influential account was that of Halle and Stevens (1971), who proposed that voiceless segments canonically involve a stiffening of the vocal folds (Halle and Stevens 1971, pp. 203–204). They suggest that while the primary purpose of this gesture is to inhibit phonation, the increase in vocal fold length and horizontal tension would also predict a rise in the frequency of any glottal vibration that does subsequently occur, as for example, at the onset of a following vowel. This “horizontal-tension” hypothesis later received support from the electromyographic study of Löfqvist *et al.* (1989), who demonstrated that a peak in cricothyroid (CT) activity occurs during the midpoint of closure for voiceless stops in both Dutch (where voiceless stops are canonically unaspirated) and American English (where they are canonically aspirated). On the basis of the relative timing of this gesture, Löfqvist *et al.* argued that its primary function is related to inhibition of vocal fold vibration, rather than to pitch control. Researchers have also suggested that intrinsic aerodynamic properties of the stop release may contribute to  $CF_0$  (Kohler, 1985; Francis *et al.*, 2006).

However, there is also reason to believe that  $F_0$  could be lowered in the vicinity of voiced stops. Ewan (1976) and Hombert *et al.* (1979) advanced the notion that  $CF_0$  is primarily a matter of vertical tension of the vocal cords, the result of articulations designed to enlarge the supraglottal cavity and facilitate voicing, such as expansion of the pharyngeal cavity or lowering of the larynx (Bell-Berti, 1975; Erickson *et al.*, 1982; Westbury, 1983). This “vertical-tension” account predicts that  $F_0$  should be lowered following voiced stops, rather than raised following voiceless stops (Hombert *et al.*, 1979, p. 45). Although Hombert and colleagues did not specify precisely how these articulations would affect vertical vocal fold tension, Honda *et al.* (1999) presented evidence that, at least in low  $F_0$  ranges, a downward vertical movement of the hyoid-larynx complex rotates the cricoid cartilage along the forward-inclined cervical vertebrae, thereby shortening the vocal folds and decreasing their horizontal (rather than vertical) tension, resulting in lower  $F_0$ .

$F_0$  lowering is also predicted by the auditory enhancement account of Kingston and Diehl (1994), who argued that certain articulations, such as those to sustain voicing and lower pitch, tend to covary because the resulting combination of acoustic effects integrate perceptually. In particular, they assert that “ $F_0$  is uniformly depressed next to [+voice] stops, regardless of how the [voice] contrast is otherwise realized” (1994, p. 432). There is also evidence that the lowered  $F_0$  following voiced obstruents (relative to their voiceless counterparts) may become phonologized, as in Shanghai Chinese (Chen, 2011) or Xhosa (Jessen and Roux, 2002), which further suggests an active gesture to explicitly lower pitch in this context.

## 2. Establishing a reference level

An important contribution to the understanding of  $CF_0$  effects was made by Hanson (2009), who compared  $CF_0$  of American English voiced and voiceless stops and fricatives. Hanson treated the time course of  $F_0$  following a nasal /m/ as a reference condition for interpreting the details of  $F_0$  after an obstruent onset.<sup>1</sup> Hanson’s motivation for using nasals as a reference stemmed from the observation that they should have little, if any, effect on the intonationally specified pitch target, due to the facts that (a) nasal sounds do not require any articulatory adjustments to the supralaryngeal cavity or state of the glottis in order for vocal fold vibration to be maintained and (b) the nasal cavity itself offers little resistance to airflow and thus does not condition the kind of decrease in pressure drop across the glottis that would perturb pitch (Ohala, 1975; Hombert *et al.*, 1979). With this approach, Hanson was able to provide clear evidence that, at least in English, it is more appropriate to speak of  $F_0$  raising following [–voice] consonants rather than  $F_0$  lowering following [+voice] consonants. Her results demonstrate that the time course of  $F_0$  after voiced obstruents is qualitatively identical to that after nasals, whereas  $F_0$  at voice onset is higher following voiceless obstruents, and may not converge with the nasal comparison contour for as much as 100 ms. Hanson also showed that the degree to which  $CF_0$  is perturbed in this environment is a function of intonational context.<sup>2</sup> She coached her speakers to produce test words at different pitch levels, from relatively high to relatively low in their speaking range. The summary just given of her findings about voiced and voiceless stops applies most clearly in high  $F_0$  contexts; Hanson also found that post-release  $F_0$  perturbations were present, but much smaller, in neutral and low  $F_0$  environments.

## 3. The relevance of “true voicing” languages

On the face of it, Hanson’s findings seem problematic for at least some theories that predict  $F_0$  lowering following phonologically voiced obstruents, especially those which predict that  $F_0$  is lowered after voiced obstruents relative to sonorants (e.g., Hombert *et al.*, 1979). However, it may also be that an  $F_0$  lowering effect is evading detection, given that voicing during the closure of English [+voice] stops /b d g/ is allophonic and speaker- and dialect-specific (Flege, 1982; Jacewicz *et al.*, 2009). As such, a lowering effect on  $F_0$  may be more easily observed in so-called “true voicing” languages, such as French or Italian, in which phonologically voiced obstruents are typically characterized by vigorous voicing throughout the closure (Benguerel *et al.*, 1978; Vaggies *et al.*, 1978). That is, it may be that the explicitly depressed  $F_0$  predicted by the vertical-tension and auditory-enhancement hypotheses, though not found after English lenis stops, will nevertheless appear after, or indeed during, the fully voiced stops of a language like French or Italian.

A second reason for investigating  $CF_0$  in “true voicing” languages is that the voiceless stops in such languages are typically characterized by relatively short-lag voice onset time (VOT; Lisker and Abramson, 1964), and are thus similar to a possible allophonic surface realization of English

[+voice] stops. This is of interest because past work provides some reason to think that it is the phonological, rather than phonetic, status of stops in a language that triggers local  $F_0$  raising (or lowering) in the following vowel. Both [Ohde \(1984\)](#) and [Hanson \(2009\)](#) found  $F_0$  following unaspirated allophones of English voiceless stops in /s+stop/ clusters to be similar to that following syllable-initial aspirated allophones, and in any event different from that found after syllable-initial [+voice] stops. That is, even though English syllable-initial voiced stops and the voiceless stops in /s+stop/ clusters are both in some sense “voiceless unaspirated,” they induce different  $F_0$  effects on the following vowel<sup>3</sup> [Fig. 1(a)]. Furthermore, in a recent study comparing English stops with those of Spanish, another “true voicing” language, [Dmitrieva et al. \(2015\)](#) found that differences in  $CF_0$  are predictable only from the phonological specification of the stop, and not from the phonetic realization of VOT: English [+voice] stops that were realized with (canonical) short-lag VOT had  $CF_0$  effects indistinguishable from those realized with (less canonical) prevoicing, whereas Spanish [+voice] stops, which are canonically prevoiced, had different  $CF_0$  effects compared to [–voice] stops, which are canonically produced with short-lag VOT.

## B. The present study

The present study applies Hanson’s methodological approach to two languages with unambiguously prevoiced stops, French and Italian. Although there are a number of studies which report  $CF_0$  in languages with phonetically prevoiced stops ([Hombert, 1976](#); [Löfqvist et al., 1989](#); [Chen, 2011](#); [Dmitrieva et al., 2015](#)), none of these studies include a baseline by which we might assess the direction in which  $F_0$  is being perturbed relative to the global intonation target. Another important aspect in which we extend earlier studies is in studying the time course of  $F_0$  during the closure phase of voiced consonants, as well as during the following vowel. Our overall aim is to contribute to the understanding of possible sources of  $CF_0$  and, more generally, of the laryngeal adjustments that characterize voicing in different languages.

If languages like French and Italian employ articulatory strategies to (a) facilitate pharyngeal expansion and maintain the transglottal pressure differential necessary to promote voicing during the stop closure and/or (b) produce lowered

$F_0$  because it contributes to a cluster of perceptually integrated auditory properties, then in a “true voicing” language we might expect to find the opposite pattern to that reported by Hanson: that is, it might be that  $F_0$  is lowered relative to the (presumably unperturbed) nasal baseline after voiced stops but is unaffected after voiceless (unaspirated) stops [Fig. 1(b)]. A second possibility is that, despite the substantial phonetic differences in the realization of voicing, “true voicing” languages may present a picture similar to what Hanson found for English: namely, that voiceless stops induce raised  $F_0$  in the following vowel and (pre)voiced stops match the  $F_0$  pattern found with nasals [Fig. 1(a)]. Such a pattern would be consistent with the finding of Hanson, Ohde, and Dmitrieva *et al.* that English [–voice] stops are associated with raised  $F_0$  regardless of allophonic differences in VOT, as well as with the results of Hanson and Dmitrieva *et al.* that  $F_0$  following English [+voice] stops is unaffected by whether or not voicing is present during the closure. Finally, if both mechanisms are at work, one might find a raising effect following [–voice] obstruents but a lowering effect following [+voice] obstruents, with nasals somewhere in the middle [Fig. 1(c)]. Note that, because [+voice] obstruents are expected to be consistently prevoiced in French and Italian, we can also investigate the time course of  $F_0$  during the closure phase of the obstruent (which may not always be possible in a language like English).

## II. METHODS

### A. Speech materials

We used an approach very similar to that of [Hanson \(2009\)](#) to study  $CF_0$  in French and Italian. Our methods differ from hers primarily in that we used natural meaningful sentences with real words as test words rather than a single carrier sentence with non-words, and we controlled intonation by designing ordinary sentences that could most naturally be read only with the desired intonation pattern. We did this in order to avoid any potentially unnatural effects of detailed coaching and monitoring.

As noted earlier, several studies have shown that the magnitude of intrinsic  $F_0$  perturbations may be attenuated depending on prosodic context ([Ladd and Silverman, 1984](#); [Kingston, 2007](#); [Hanson, 2009](#)). To investigate this possibility, we embedded our test words (see the [Appendix](#)) into

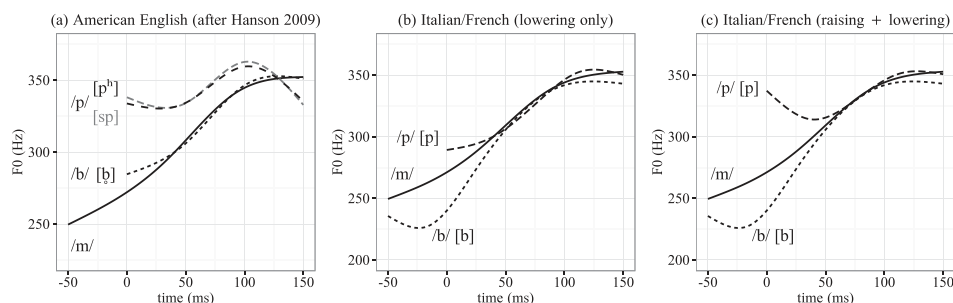


FIG. 1. Schematic comparisons of (a) time course of  $F_0$  production in a high intonational context in American English (after [Hanson, 2009](#)); (b) hypothetical pattern in a “true voicing” language where [+voice] obstruents lower  $F_0$ ; (c) hypothetical pattern in a “true voicing” language where [+voice] obstruents lower  $F_0$  and [–voice] obstruents raise  $F_0$ . Note that panels (b) and (c), unlike panel (a), plot the predicted  $F_0$  trajectory during the voiced stop closure as well as during the following vowel.



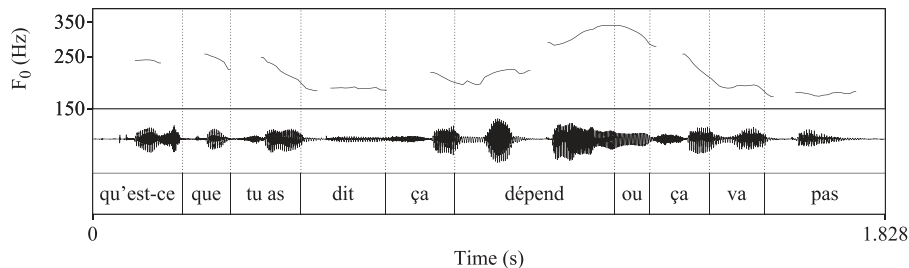


FIG. 2.  $F_0$  trace (log scale) and waveform of the utterance *Qu'est-ce que tu as dit, « ça dépend » ou « ça va pas »?* "What did you say, « it depends » or « it won't work »?" (test item *dépend*, target /p/, high intonation context), female French speaker. Audio example included in the supplementary material (footnote 4).

both of the variable positions in a series of alternative questions of the form *Did you see X or Y?*, *Who did you see, X or Y?*, etc. (see Figs. 2–5). Spoken with the natural intonation contour for an alternative question, these sentences put the test word X in a high-pitch context and the test word Y in a low-pitch context. The questions were all different, tailored to the test words, and aimed to be as natural and colloquial as possible; however, all were alternative questions intended to elicit the intonation contour in which test word X is spoken at high pitch and test word Y at low pitch. In most of the test sentences, only one of the two words in the X and Y positions was a real test word, with the other chosen to ensure that X and Y were pragmatically natural alternatives. We used two versions of all sentences, with X and Y in both orders, so as to record all test words in both positions; the two versions of each test sentence were always widely separated in the recording order. Examples of our test sentences are given in Figs. 2–5.

All test words contained one of the five consonants /p b f v m/ in the onset of the final syllable (in French) or of the stressed syllable (in Italian). Insofar as possible the test consonant was preceded by a vowel and followed by a low or low mid vowel (/a a ɔ ε œ/ in French, /a/ in Italian). Occasionally the test consonant was preceded by a sonorant consonant (e.g., French *halva*, Italian *palma*) rather than a vowel. In both languages, it was necessary to use both open and closed test syllables, though the closed syllables were nearly always closed by a sonorant /l, j, n, m, r, ʁ/.

## B. Participants

Sixteen speakers of Standard French and Standard Italian were recruited for this study. The French speakers (nine female, one male, ages 18–23 years) were all visiting undergraduates at the University of Edinburgh, and were recorded at the University of Edinburgh Phonetics Laboratory. The Italian speakers (four female, two male, ages 20–29 years) were exchange students at the Ludwig Maximilian University (LMU) in Munich, and recorded at the LMU Institute of Phonetics and Speech Processing.

Participants had varying degrees of foreign language competency, but in all cases high enough to facilitate study in a foreign country (the United Kingdom or Germany, respectively). All participants reported normal hearing and no history of speech or language deficit and were paid a small sum for their participation.

## C. Recording procedure

Recordings were made in a sound-treated recording booth at the University of Edinburgh or Ludwig Maximilian University, Munich. Before the recording, participants were given brief oral instructions (in English or German, as appropriate), provided basic demographic information (age, place of birth, other languages spoken) and signed informed consent forms. The intended intonation pattern was then illustrated in French or Italian by a fluent L2 speaker, who emphasized that the sentences were intended to be everyday language and should be read as naturally and colloquially as possible. With these instructions, and without close coaching or monitoring, the speakers all easily produced the intended intonation pattern (as judged by the fluent L2 speaker) on all or nearly all sentences.

Speakers were seated approximately 20 cm from an omnidirectional microphone, with sentence prompts appearing on a computer screen (for the Italian participants) or on sheets of paper printed and placed on a music stand in front of the speaker (for the French participants). All recordings were made direct to disk at a 44.1 kHz sampling rate using the BAS SpeechRecorder software package, version 2.12.10 (Draxler and Jänsch, 2004).

The speakers all read the same list of sentences for their language. After the experimenter's introduction, participants read 10–20 warm-up sentences to insure optimum recording levels and to allow them to become comfortable with the intonation patterns. The French participants then read a total of 58 items in 116 unique test sentences: 12 sentences each with targets /b f v/ and 11 sentences each with targets /p m/ in two different intonation conditions (high and low). Italian participants read a total of 67 items in 134 test sentences: 12

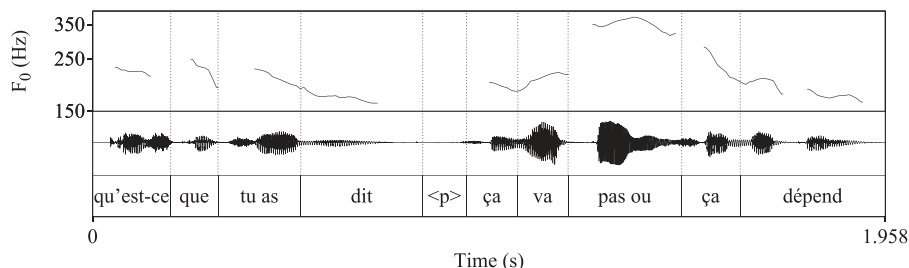


FIG. 3.  $F_0$  trace (log scale) and waveform of the utterance *Qu'est-ce que tu as dit, « ça va pas » ou « ça dépend »?* "What did you say, « it won't work » or « it depends »?" (test item *dépend*, target /p/, low intonation context), female French speaker. (p) represents a pause (silence interval). Audio example included in the supplementary material (footnote 4).

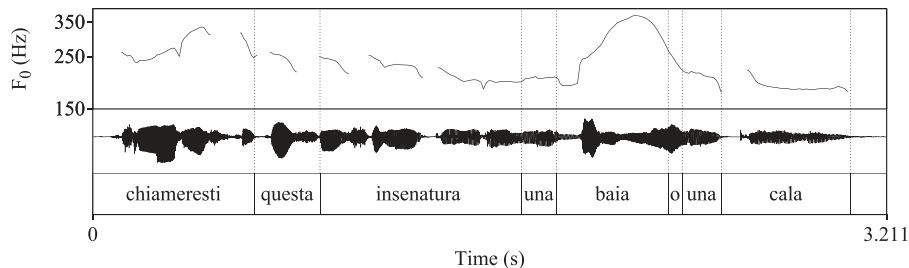


FIG. 4.  $F_0$  trace (log scale) and waveform of the utterance *Chiameresti questa insenatura una baia o una cala?* “Would you call this inlet a bay or a cove?” (test item *baia*, target /b/, high intonation context), female Italian speaker. Audio example included in the supplementary material (footnote 4).

with target /b/, 13 each with targets /p/ and /f/, 14 with target /m/, and 15 with target /v/, again in two different intonation conditions. Examples of the intonation contours produced are given in Figs. 2–5.<sup>4</sup>

#### D. Acoustic analysis

The segments of interest (consonant and following vowel) were manually labeled by the first author using Praat 5.4.03 (Boersma and Weenink, 2015) based on the periodicity in the acoustic waveform, supplemented by spectrographic analysis where appropriate. Four primary acoustic landmarks were annotated: the onset and offset of stop closure (or frication, as appropriate), the onset of voicing, and the duration of the following vowel. For voiced obstruents, onset of voicing was usually coextensive with the onset of oral constriction; in the event of voicing cessation, either during the closure or immediately following the burst, the post-release re-establishment of periodic voicing was also noted. For this data set, we defined the onset of voicing as the onset of the first periodic pattern in the acoustic waveform (see supplementary material<sup>4</sup> for examples). Vowel offset was defined as the last pitch cycle before a significant drop in amplitude when preceding a stop, or the last pitch cycle before significant frication noise when preceding a fricative.

Following segmentation, acoustic measurements were taken using Praat. Pitch analysis was performed using the autocorrelation method of Boersma (1993), with a Gaussian analysis window of 80 ms, a 5 ms frame duration, a pitch floor of 50 Hz, and a pitch ceiling of 500 Hz. Each pitch object was examined visually and checked by hand to correct any instances of pitch halving (or, less commonly, doubling). The resulting  $F_0$  contours were then sampled at 29 equidistant points in each of the closed and open phases (equivalent to once every 5–7 ms). Prior to further analysis, by-speaker raw  $F_0$  measurements were standardized using a z-score transform to facilitate comparison of total degree of pitch change across subjects and tokens.

In addition to  $F_0$ , we took a number of durational measures. For voiceless stops, we measured VOT (Lisker and Abramson 1964), defined here as the duration of the period from stop release to the onset of periodic voicing. Voiceless stops were never prevoiced in our corpus. For stops, we also measured closure duration, and for fricatives, frication noise duration, as both measures are known to vary as a function of voicing in connected speech (e.g., Crystal and House, 1988). Finally, we recorded the duration of the vowel following the onset to use as a proxy for speech rate. Spectrographic examples of each of these segment types can be found in the supplementary material.<sup>4</sup>

### III. RESULTS

All results were analyzed using linear mixed-effect models implemented in the R package *lme4* (Bates et al., 2015), with *lsmeans* (Lenth and Hervé, 2014) used for *post hoc* and pairwise comparison between the predicted marginal means.

#### A. $F_0$

The basic result is presented in graphic form in Fig. 6, which shows the mean standardized  $F_0$  contours over the CV sequence by language, manner, and voicing for both high intonation (test word in X position) and low intonation (Y position) contexts. Focusing first on the open (post-release) phase, visual inspection of the contours for the high context shows clearly that the  $F_0$  pattern following voiced obstruents is similar to that of nasals, while that following voiceless obstruents starts markedly higher and gradually converges with the nasal and voiced-obstruent  $F_0$  contours (more quickly for Italian than for French). A similar pattern is visible in the low intonation context, especially in Italian, though the differences are very small. Any influence of manner (stop vs fricative) on immediate post-release  $F_0$  appears minimal. While some individual differences were observed, the shape and magnitude of the contours was quite similar

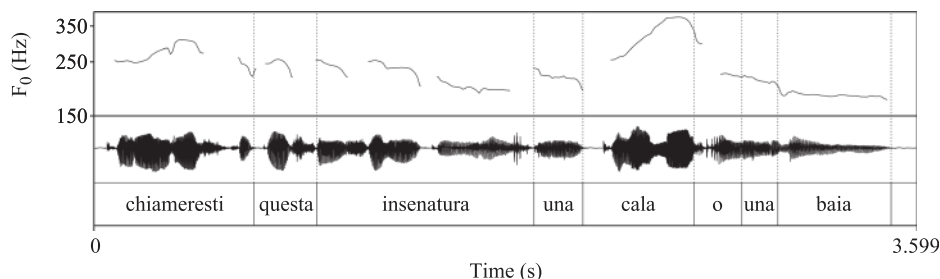


FIG. 5.  $F_0$  trace (log scale) and waveform of the utterance *Chiameresti questa insenatura una cala o una baia?* “Would you call inlet this a cove or a bay?” (test item *baia*, target /b/, low intonation context), female Italian speaker. Audio example included in the supplementary material (footnote 4).

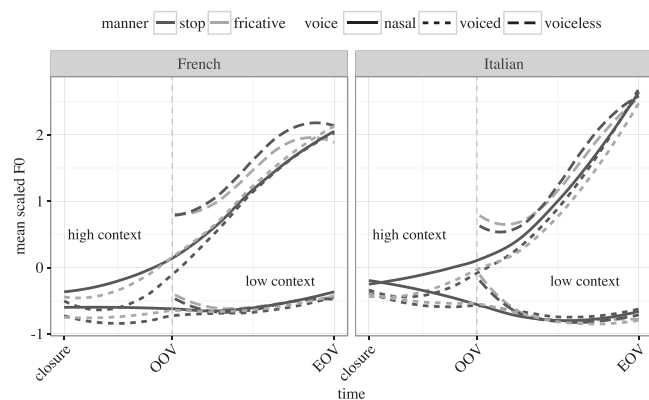


FIG. 6. Average time-normalized standardized pitch contours (loess smoothing) from time of oral closure through end of vowel (EOV) for high and low intonation contexts by language, manner, and voicing. Dashed line indicates onset of post-release voicing, i.e., the first period of the vowel (OOV).

across different speakers, suggesting that the pooled results shown in Fig. 6 are a reasonably accurate representation of our sample (see the supplementary material for individual plots for all participants).<sup>4</sup> Our results thus most closely match the hypothetical pattern sketched in Fig. 1(c) (Sec. 1B).

Support for this interpretation of Fig. 6 is provided by a series of multilevel regression models, each predicting  $F_0$  from SEGMENT TYPE (voiced stop, voiced fricative, voiceless stop, voiceless fricative, or nasal), INTONATION CONTEXT (high or low), and their interaction at one of five temporal landmarks (onset of closure, midpoint of closure, onset of vowel, midpoint of vowel, end of vowel). Separate models were fit for each language. Together with random intercepts for SUBJECT and ITEM, both models included uncorrelated by-subject random slopes for both INTONATION CONTEXT and SEGMENT TYPE. These were the maximally specified models that converged, and were judged to be superior to less fully specified models in a series of likelihood-ratio tests. Tables of predicted marginal means for all post hoc pairwise comparisons can be found in the supplementary material.<sup>4</sup>

### 1. French

At the onset of closure,  $F_0$  of voiced stops was slightly reduced compared to that of nasals in the high but not low context. At the midpoint of closure in the high context,  $F_0$  of both voiced stops and voiced fricatives was lower than that of nasals, and stops were also lower than fricatives at this point. In the low context, voiced stops and fricatives were both lower than nasals, but could not be statistically distinguished from one another.

At the onset of the vowel,  $F_0$  following /p f/ is higher than following /m/;  $F_0$  following /b/ is lower than following /v/, but neither can be statistically distinguished from /m/, although the predicted marginal mean of /b/ is still  $\sim 9$  Hz lower than that of /m/ at this point. In the low context, voiceless stops and fricatives are raised relative to their voiceless counterparts, but neither pair is statistically differentiable from the nasal. In the high context, the situation is much the same at vowel midpoint:  $F_0$  following voiceless stops and fricatives is higher than following nasals, but

none of the voiced segments can be statistically distinguished. At this timepoint,  $F_0$  can no longer be distinguished by segment type in the low series, and at the end of vowel, it is statistically indistinguishable in both intonational contexts.

### 2. Italian

At the point of closure,  $F_0$  is statistically similar for all segment types in both intonation contexts. At the midpoint of closure,  $F_0$  of /b/ and /v/ was lower than the baseline in the high intonation context, while in the low intonation contexts, /b/, but not /v/, was lower than /m/. Voiced stops and fricatives could not be distinguished from one another in either setting.

At the onset of the vowel, voiceless obstruents raise  $F_0$  compared to voiced obstruents and nasals, which cannot be distinguished from one another. These patterns are the same in both intonation contexts. At vowel midpoint, /p/ is still raised relative to /m/ in the high context, but /f/ cannot be distinguished from either /m/ or /p/; in the low context  $F_0$  is the same for all segment types. By vowel offset, no segments can be distinguished from /m/ on the basis of  $F_0$ .

## B. Durational properties

To assess the extent to which VOT, closure duration, and/or frication duration co-vary with  $CF_0$ , we first describe the distribution of these properties in our data.

### 1. Stop VOT

Figure 7 shows the distribution of (positive) VOT in both languages by intonation context, with means and standard deviations given by language and intonation context in Table I. With a few exceptions (two instances of *ballo* and one of *cubana* produced by three different Italian participants), [+voice] stops in both languages were consistently voiced throughout the closure; we removed the three exceptional items from the subset distribution considered here before further analysis. Our VOT model included the two-way interaction of LANGUAGE (French, Italian) and INTONATION CONTEXT (high, low), with random intercepts for subjects and items and by-subject<sup>5</sup> random slopes for INTONATION CONTEXT. The addition of following VOWEL DURATION (as a correlate of speech rate) resulted in an improved model fit ( $p < 0.001$ ,  $df = 1$ ), but the estimated effect size was quite small ( $\beta = 0.05$ ,  $SE = 0.01$ ,  $t = 4.2$ ). French voiceless stops showed slightly greater variation (and slightly higher mean VOTs) than those of Italian (7 ms,  $SE = 3.1$ ,  $df = 19.79$ ,  $t = 2.38$ ,  $p = 0.03$ ). Our Italian findings are roughly comparable to those of Vaggies *et al.* (1978) and Esposito (2002), while the French data indicate a slightly longer voicing lag compared to some previous reports (Caramazza and Yeni-Komshian, 1974), although this may be an effect of the following vowel quality (Nearey and Rochet, 1994). Most germane for present purposes, it is clear that the VOT distributions differ significantly from those of languages like English (see Lisker and Abramson, 1964, and much subsequent work).

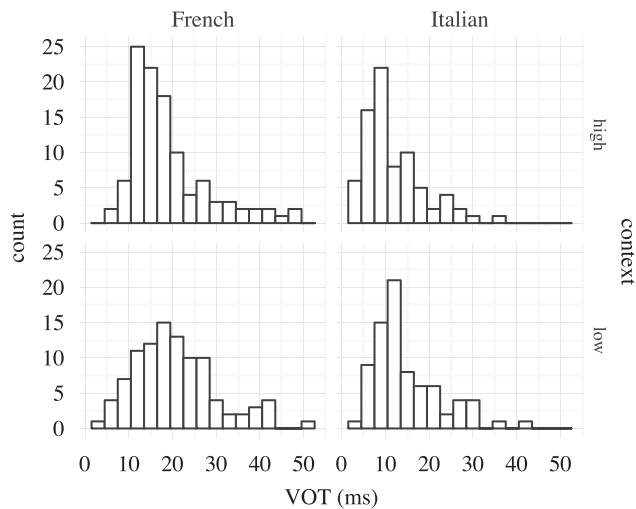


FIG. 7. Distribution of VOT values by language and intonation context.

## 2. Stop closure duration

Summary statistics for closure duration are provided in Table II and presented graphically in Fig. 8. Our model for stop closure duration included a three-way interaction of LANGUAGE (French, Italian), VOICING (voiced, voiceless, nasal), and INTONATION CONTEXT (high, low), with random intercepts for subjects and items and by-subject random slopes for repeated measures VOICING, INTONATION CONTEXT, and their interaction. Addition of a VOWEL DURATION predictor, as a surrogate for speech rate, significantly improved model fit in a likelihood-ratio test ( $p = 0.04$ ,  $df = 1$ ). No significant differences in closure duration were found within each language and intonation context, e.g., the distribution of closure duration for French /p/ in high context is the same as for /b/ in that context.

Within each language and segment type, small but significant differences between predicted marginal means in high versus low intonation context were observed for nasals in French [6 ms, standard error (SE) = 2.2,  $df = 25.40$ ,  $t = 2.7$ ,  $p = 0.01$ ], and for both voiced (10 ms, SE = 2.65,  $df = 20.51$ ,  $t = 3.92$ ,  $p < 0.001$ ) and voiceless (7 ms, SE = 2.9,  $df = 17.75$ ,  $t = 2.4$ ,  $p = 0.03$ ) stops in Italian. In all cases, closure durations were longer in the high compared to the low intonation context. Within language and intonation context, significant differences by segment type were found only between Italian /b/ and /m/ in low intonation environments, with /m/ closures here lasting longer on average (18 ms, SE = 5.23,  $df = 80.14$ ,  $t = 3.4$ ,  $p < 0.01$ ). Between languages, differences in closure duration were not significant for any segment types within a intonation context.

Our findings agree with those of Esposito (2002) who also did not observe voicing-related differences in closure

TABLE I. Mean (standard deviation) for VOT (in ms) by language and intonation context.

Intonation context	Language	
	French	Italian
high	19 (9)	12 (7)
low	20 (10)	15 (8)

TABLE II. Mean (standard deviation) closure durations (in ms) by segment, language and intonation context.

Segment	Intonation context	Language	
		French	Italian
/b/	high	65 (14)	70 (26)
	low	62 (13)	59 (24)
/m/	high	73 (17)	77 (18)
	low	67 (15)	77 (24)
/p/	high	70 (20)	77 (18)
	low	66 (19)	71 (16)

duration between /p/ and /b/ in Italian, and are consistent with Abdelli-Beruh (2004), who found French voiceless stops to have longer closure durations than corresponding voiced stops, a trend that is also present in our data.

## 3. Frication duration

Figure 9 shows the distribution of frication duration in our data for voiced and voiceless fricatives in both languages by intonation context, with means and standard deviations given in Table III. All voiced fricatives in both languages were consistently and robustly voiced throughout the closure period, and we did not observe any instances of spontaneous voicing of the voiceless fricatives. As with closure duration, we modeled frication duration from the three-way interaction of LANGUAGE (French, Italian), VOICING (voiced, voiceless) and INTONATION CONTEXT (high, low), together with a predictor VOWEL LENGTH, random intercepts for subjects and items, and by-subject random slopes for repeated measures VOICING, INTONATION CONTEXT, and their interaction. As seen in Fig. 9, frication duration was significantly longer for /f/ compared to /v/ in both languages and intonation contexts (mean over both languages: 53 ms, SE = 3.15,  $df = 68.09$ ,  $t = -16.8$ ,  $p < 0.001$ ). Frication durations were slightly

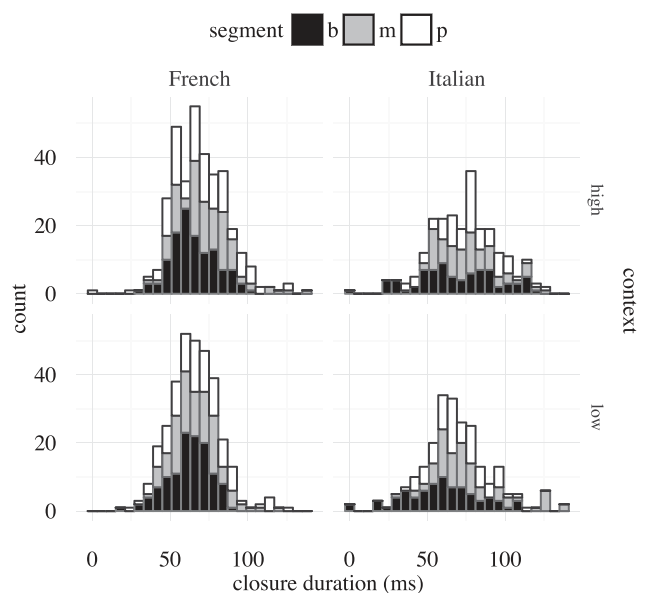


FIG. 8. Distribution of closure durations by segment, language and intonation context.



longer for voiced fricatives in high compared to low intonation contexts in both French (6 ms, SE = 1.85, df = 29.85,  $t = 3.2$ ,  $p < 0.01$ ) and Italian (5 ms, SE = 2.19, df = 20.93,  $t = 2.2$ ,  $p < 0.05$ ). Frication durations for French voiced fricatives were longer than those of Italian in both high (15 ms, SE = 6.1, df = 46.31,  $t = 2.4$ ,  $p < 0.05$ ) and low (14 ms, SE = 5.5, df = 65.06,  $t = 2.5$ ,  $p < 0.05$ ) contexts. Descriptively, Italian voiced fricatives were produced with slightly less variation than those of French. The general pattern observed here (significant temporal overlap but voiceless longer than voiced) is consistent with previous acoustic studies of frication duration (Crystal and House, 1988; Jongman *et al.*, 2000).

### C. VOT-F0 covariance

In an acoustic study of  $CF_0$  based on French and Italian word lists, Kirby and Ladd (2015) found a positive correlation between the duration of voicing lead and post-release onset  $F_0$ : longer voicing leads were associated with lower onset  $F_0$  in both languages. This effect was not replicated in the current study. While there is a very weak negative correlation between closure duration and onset  $F_0$  for French voiced stops in the high intonation context ( $r = -0.08$ ), the effect is in the opposite direction in the low context ( $r = 0.2$ ); in Italian, the pattern is reversed ( $r_{\text{high}} = 0.22$ ,  $r_{\text{low}} = -0.06$ ; see the supplementary material<sup>4</sup> for covariance plots). This is likely due to the fact that, in the present study, [+voice] stops are in intervocalic and often intramorphemic position: because voicing had been initiated prior to the stop closure, it may be that speakers were not controlling articulations related to  $CF_0$  in the same way as when voicing is initiated in utterance-initial position, as in the word-list condition of Kirby and Ladd (2015).<sup>6</sup>

Although there is an overall moderate positive correlation between scaled  $F_0$  and VOT for French voiceless stops in the high intonation context ( $r = 0.39$ ), inspection of this

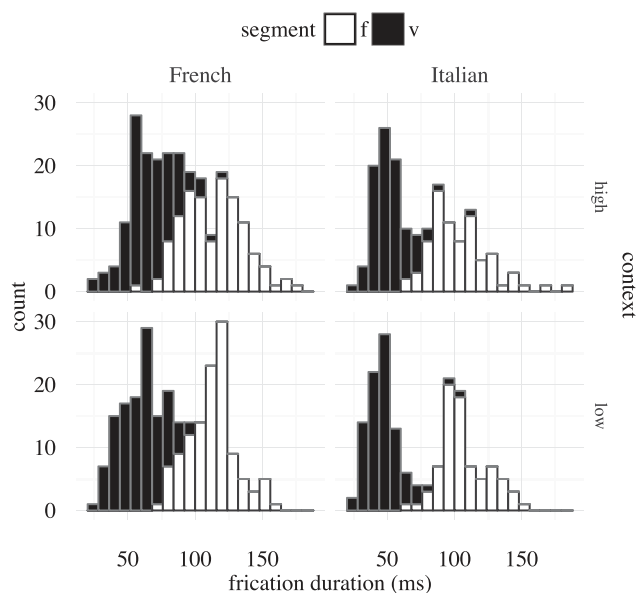


FIG. 9. Distribution of frication duration by segment, language, and intonation context.

TABLE III. Mean (standard deviation) frication duration (in ms) by segment, language and intonation context.

Voicing	Intonation context	Language	
		French	Italian
/v/	high	66 (17)	52 (13)
	low	59 (16)	48 (13)
/f/	high	112 (23)	104 (22)
	low	113 (18)	107 (18)

relationship on a by-speaker basis (not shown here) suggests considerable variation, with some speakers having a negative correlation or no correlation at all. In the low context, no overall correlation was observed ( $r = 0$ ). For Italian, correlations were weakly positive in both contexts ( $r = 0.16/0.17$ ), but once again there is considerable individual variation in the direction of the correlation (cf. Dmitrieva *et al.*, 2015, who report within-category VOT and onset  $F_0$  to be uncorrelated in Spanish). Correlations between noise duration and  $F_0$  at vowel onset for fricatives were similarly weak or absent ( $r < \pm 0.3$  in all cases).

## IV. DISCUSSION

### A. General discussion

Despite the considerable phonetic difference in what it means to be “voiced” (or “voiceless”) in English, French, and Italian, it is clear that in all three languages, voiceless or “fortis” obstruents raise  $F_0$  in the following vowel, regardless of the temporal duration of voicing lag. This effect was attenuated, but not completely obliterated, in the low intonation context (Hanson, 2009), and is consistent with previous studies of  $CF_0$  in “true voicing” languages (Hombert, 1976; Löfqvist *et al.*, 1989; Dmitrieva *et al.*, 2015). Moreover,  $F_0$  immediately following the release of voiced/lenis obstruents is statistically indistinguishable from  $F_0$  following nasals. Thus, in French and Italian as well as in English, post-release  $F_0$  perturbations appear to be primarily a result of laryngeal adjustments that raise, rather than lower pitch (Halle and Stevens, 1971; Löfqvist *et al.*, 1989; Hanson, 2009), although the magnitude of this effect is clearly dependent on the global intonation target (Lea, 1973; Kohler, 1982; Silverman, 1986; Hanson, 2009).

It is also clear from our data that there is pronounced  $F_0$  lowering during the closure phase for voiced obstruents (stops and fricatives) relative to nasals in both languages and in both intonation contexts (although, as with post-release raising, the effect is attenuated in the low intonation context). However, no such lowering was observed during the oral closure for nasals. This supports the use of sonorants as a baseline when studying whether  $F_0$  is perturbed from a global intonation target (Löfqvist *et al.*, 1995; Hanson, 2009), and reinforces the expectation that the linguistic  $F_0$  target of an utterance should be perturbed minimally, if at all, by the production of (voiced) sonorants, for which maintaining transglottal airflow necessary for continued voicing should not require any active or passive vocal tract enlargement (Ohala, 1975) or otherwise require articulatory

adjustments of the glottis or supralaryngeal cavity (see Sec. IV B. below).

However, these findings are accompanied by two caveats. First, as shown in Sec. III A, there is a difference between the  $F_0$  lowering effects of voiced stops and voiced fricatives in French, but not in Italian. In Italian, that is, the lowering effect of voiced obstruents does not seem to depend on the degree of stricture, whereas in French, voiced stops lower  $F_0$  more than voiced fricatives do, at least in the high-pitch environment. Is there any reason to expect such a difference? It is known that the production of voiced fricatives is aerodynamically challenging (Stevens, 1971; Ohala, 1997), involving both glottal adduction as well as maintenance of sufficient airflow to generate frication noise. If the  $F_0$  lowering effect of voiced obstruents is due to active laryngeal adjustments to reduce supraglottal pressure ( $P_O$ ) in order to maintain phonation, this difference suggests that in Italian, stops and fricatives have comparable effects on  $P_O$ , whereas French fricatives have a weaker effect than stops. This may be the case if the production of voiced fricatives in Italian involves greater obstruction of oral airflow (e.g., via a narrower constriction or greater peak closing velocity) than is the case in French. To our knowledge there are no relevant instrumental studies that would confirm or refute this speculation, but it seems very likely that such differences between languages exist. In this connection we note that the second author has had considerable experience analyzing  $F_0$  contours in both English and Greek, and reports impressionistically that Greek voiced fricatives are much more likely than English ones to be fully voiced and to have a substantial lowering effect on  $F_0$ .

Second, while we do observe post-release  $F_0$  raising in high-pitch focus contexts following voiceless (short-lag) stops of French and Italian, the overall magnitude and duration of this raising differed across the two languages, and appears to differ somewhat from languages such as English and German. Compared to Italian,  $CF_0$  in French is in general of greater magnitude at voicing onset and persists for longer into the vowel (Sec. III A, Fig. 6), and is more robustly observed across the speakers in our sample (see the supplementary material for by-speaker plots).<sup>4</sup> Furthermore, our data show a clear difference between the trajectory of  $F_0$  following voiceless obstruents in French and Italian: the post-voiceless contour remains raised relative to the sonorant baseline for much longer in French than in Italian. It is possible that this reflects some fundamental difference in the extent or manner of voicing inhibition, but we suggest that the explanation may lie instead with the prosodic difference between the test syllables in the two languages (for discussion of such prosodic differences see, e.g., Ladd, 2008, pp. 55–61). In French, because intonational pitch accents always associate with the edges of phrases (Jun and Fougeron, 2002), the local  $F_0$  peak after the first constituent of the alternative question (i.e., at the edge of “X” in “X or Y?”) is aligned with the very end of the test syllable (see Figs. 2 and 3). In Italian, on the other hand, the intonational pitch accent associates with the lexically stressed syllable, which was on the penultimate syllable in all our test words. In other words, the later timing of the  $F_0$  peak in French may provide a more favorable pitch environment in which to observe the effect of  $CF_0$ . One possibility is

that this may result from an interaction between the timing of the gestures programmed to produce the intonation target and those associated with suppression of voicing for production of the onset. This idea could be tested by repeating the Italian portion of the present experiment using oxytone test words (*parole tronche*, words with lexical stress on the final syllable), though there is the practical difficulty that such words are not very common. In any case these are questions for further research.

## B. Implications for theories of $CF_0$

In our French and Italian data, as in much previous work on other languages, we find that  $F_0$  is clearly raised, relative to nasals, following the release of voiceless obstruents, and can continue to differ significantly by over 20 Hz from post-sonorant  $F_0$  at vowel midpoint (Sec. III A 1). This finding supports previous proposals that post-release  $CF_0$  effects result from an active gesture to inhibit phonation, such as stiffening of the vocal folds and/or engagement of the cricothyroid (CT) musculature (Halle and Stevens, 1971; Löfqvist *et al.*, 1989). If contraction of the CT is generally employed as a means to inhibit phonation, this would help explain why languages like Italian and French pattern together with English with respect to their  $CF_0$  behavior. Moreover, if the effects of CT relaxation on vocal fold tension take longer to dissipate than do those of CT contraction, a decline in  $F_0$  should lag behind a decline in visible CT activity, despite peak CT activity being temporally coextensive with the midpoint of the consonantal closure (Löfqvist *et al.*, 1989; Sawashima *et al.*, 1982).<sup>7</sup> During the closure phase of voiced obstruents, on the other hand, we observe clear  $F_0$  lowering, as would be expected if speakers employ one of several previously documented strategies to reduce  $P_O$ , such as larynx lowering, velopharyngeal venting, or engagement of the levator palatine and sternohyoid muscles to expand the pharyngeal cavity (Bell-Berti, 1975; Erickson *et al.*, 1982; Westbury, 1983; Solé, 2011). In other words,  $CF_0$  may be explained as a side-effect of the successful production of an obstruent, with the direction of the perturbation (raising or lowering) involved depending on (a) the particular constellation of gestures associated with the phonetic expression of the voicing contrast in a given language, dialect, or idiolect, as well as (b) the point in the speech stream where  $F_0$  is considered.

We emphasize, however, that even if  $CF_0$  arises in the first instance due to articulatory contingencies, awareness of perceptual benefits of certain cue combinations may lead speakers to exaggerate these adjustments for the listener’s benefit: as noted by Hanson (2009, pp. 436–438), there is no necessary contradiction between the kind of evidence presented here and a more general notion of phonologically or auditory-driven enhancement consistent with Kingston and Diehl’s notion of “phonetic knowledge” (on this point see also Hoole and Honda, 2011). Indeed, as has been frequently observed, the very fact that  $CF_0$  perturbations are seemingly crucial components of tonogenetic processes entails that they must at some point come under speaker control. Our contention is simply that, at least in the first instance,

perceptual considerations may not be necessary to account for intrinsic  $F_0$  effects, at least in some languages.

If post-release  $F_0$  raising is the result of gestures intended to suppress phonation, then we might also expect  $F_0$  to be raised following voiceless *sonorants* in languages with such segments. Maddieson (1984) has shown that in Burmese, a language that contrasts voiced with voiceless nasal and lateral sonorants,  $F_0$  is higher following the voiceless than the voiced sonorant series (see also Ohala, 1975). Although Maddieson does not compare with data from obstruents, he points out that this is consistent with much of the historical data on tonogenesis, in which voiced obstruents and sonorants pattern together in terms of their tonal reflexes, as do voiceless obstruents and (voiceless) sonorants. This patterning is predicted if the production of voiceless sonorants in these languages involves an active laryngeal gesture to inhibit phonation similar to that of voiceless obstruents.

### C. Implications for laryngeal phonology

It has at times been suggested that, in languages with two-way laryngeal contrasts, the primary or “active” phonetic indicator of (word-initial) obstruents will be either aspiration or closure voicing, but not both (e.g., Keating, 1984; Kohler, 1984). For example, Keating (1984) argues that in a language like English, where the fortis member of the opposition is marked by aspiration, then the lenis member does not require active voicing. In a language like French, the reverse should hold: if the lenis series involves active gestures to maintain voicing during closure, the fortis series will not be characterized by any particular enhancement. Languages like English may then be described as having an active devoicing gesture, associated with the fortis series only, while languages like French will have an active specification for vocal fold vibration associated with the lenis series, with the fortis series accordingly underspecified. A similar idea is made formally explicit in so-called “laryngeal realism” literature (e.g., Beckman *et al.*, 2013). However, the present results suggest that in languages like French, while the lenis series is characterized by a gesture to support voicing, so too is the fortis series characterized by a gesture to support *devoicing*, i.e., something like [stiff vocal folds]. It is not clear to us in what sense one of these gestures is more essential than the other; rather, they simply reflect the empirical reality of language-specific differences in the implementation of voicing (cf. Kingston, 2007, p. 172). Additional evidence in support of this view comes from Ní Chasaide and Gobl (1993), who found language-specific differences in voice source parameters on the vowel following voiced and voiceless stops in a number of European languages even though these differences would not normally be considered phonologically distinctive. Such findings are problematic only if one insists that phonological contrasts need to be in some sense “economical” or otherwise non-redundant (see Beckman *et al.*, 2011, for some discussion).

On the basis of their findings that Spanish and English short-lag stops were indistinguishable in terms of VOT, but well separated in terms of onset  $F_0$ , Dmitrieva *et al.* (2015) argued that “the phonological status of the consonant may

carry more weight in determining the onset  $F_0$  patterns than do its phonetic properties, such as the presence or absence of laryngeal voicing” (p. 91). Similar arguments have previously been advanced by Keating (1984, p. 294) and Kingston (2007, p. 173). While it is clear that onset  $F_0$  behavior cannot be predicted directly from VOT, we believe that the ultimate determinant of  $CF_0$  patterns is fundamentally phonetic in nature, as argued in Sec. IV B above: the reason that the voiceless stops of French, Spanish, and English pattern together with respect to post-release  $F_0$  raising is not because they share a phonological feature, but because they share a gesture aimed to inhibit phonation (see also Goldstein and Browman, 1986). The fact that Dmitrieva *et al.* found English lenis stops to have the same (non-)effect on onset  $F_0$  regardless of whether or not they were phonated during closure is consistent with this category not having an associated phonation-inhibiting gesture in this language; and in general, we should not expect that the presence or absence of such a gesture is predictable from the surface duration of voicing lag (or lead). In a language like French, the perceptual importance of robust prevoicing may mean that speakers are more likely to take action to insure voicing is sustained throughout the closure, leading to depressed  $F_0$  during the closure (and possibly to some residual  $F_0$  lowering immediately following the release, although we did not find statistical support for this in our data). In English, on the other hand, if voicing during the closure is less perceptually critical, we would not expect the same degree of  $F_0$  lowering either during the closure or in the following vowel (at least, for those speakers/dialects/utterances where [+voice] stops are actually produced with vocal fold vibration during the closure: Flege, 1982; Hanson, 2009).

Taken together, studies of  $CF_0$  serve as an important reminder that cross-linguistic variation in “voicing” is not restricted to the relative timing of laryngeal and supralaryngeal gestures and the resultant differences in VOT; states of the glottis (e.g., spread vs constricted, stiff vs slack) can vary as well, even if VOT alone appears to be acoustically sufficient to distinguish the voicing contrasts in any given language. Indeed, listeners appear to take  $CF_0$  information into account even when the phonological contrast is unambiguously signaled by VOT differences (Whalen *et al.*, 1990; Whalen *et al.*, 1993). These findings support the contention that it is inappropriate to represent obstruent contrasts along a single-dimension acoustic-phonetic continuum, either of timing (e.g., VOT: Lisker and Abramson, 1964; Keating, 1984) or glottal constriction (e.g., a continuum from open to closed: Kim, 1970; Gordon and Ladefoged, 2001). Instead, it should be kept in mind that the articulatory mechanisms underlying the production of laryngeal contrasts of all kinds are considerably more complex (Halle and Stevens, 1971; Kingston and Diehl, 1994; Edmondson and Esling, 2006; Keyser and Stevens, 2006; Hanson, 2009).

### V. CONCLUSIONS

The present findings provide evidence for two types of  $CF_0$  effects. The first—raised  $F_0$  following the release of a voiceless consonant—can be understood as the result of laryngeal adjustments to *inhibit* phonation, while the second—depressed  $F_0$  during the closure phase of a voiced



obstruent—is consistent with known laryngeal adjustments which *sustain* phonation. In both cases, however, the *F0* perturbations accompany, and are thus predicted by, articulatory maneuvers otherwise required for the successful implementation of voicing. While this in no way precludes subsequent enhancement of *CF0* for perceptual purposes, it supports the position that the source of the effect is fundamentally due to articulatory, rather than perceptual, contingency (Halle and Stevens, 1971; Kohler, 1985; Löfqvist *et al.*, 1989; Hanson, 2009; Hoole and Honda, 2011).

## ACKNOWLEDGMENTS

Thanks to the Institut für Phonetik und Sprachverarbeitung at Ludwig-Maximilians-Universität München for their hospitality, the use of their recording facilities, and helpful comments at an early stage of this project. We would also like to thank Marc Brunelle, Helen Hanson, Phil Hoole, John Kingston, and Alice Turk, as well as associate editor Laura Koenig and an anonymous reviewer, for their thoughtful comments on an earlier version of this work. We remain solely responsible for errors of fact or interpretation. This project was supported in part by an award from the University of Edinburgh Moray Endowment.

## APPENDIX: WORDLISTS

Table IV gives wordlists for French and Italian.

TABLE IV. Wordlists.

/b/	/p/	/v/	/f/	/m/
1. French				
bâbord	copains	aval	Beaufort	amant
bain	dépend	avant	confins	amers
Bâle	frappant	baklava	effort	Dumas
banc	lapin	halva	éléphant	grammaire
bas	opale	Laval	enfant	Léman
cubain	pain	L'Avare	enfin	main
Liban	papa	levain	faim	mâle
Robert	part	Levant	faïnes	malle
Robin	pas	quatre-vingts	fin	mer
ruban	sapin	suisant	fort	romans
tabac	vapeur	ver	phare	roumain
tribord		vin	sofa	
2. Italian				
Albano	Campari	L'Avana	Alfano	amante
bagno	capanna	avanti	elefante	amari
baia	incapace	cavallo	Fabio	armadio
baita	marzapane	Giovanni	false	Mafia
ballo	opale	lavagna	falso	maglia
banca	paglia	Novara	fame	mago
Bari	pala	novanta	fango	mano
cubana	palma	rivali	fante	marce
imbarco	panca	savana	farfalle	mare
Obama	pane	Silvana	faro	marmo
rubato	papa	vado	fave	marzo
timballo	pari	valle	rifare	rimane
	riparo	valzer	Stefania	romano
		vanga		romanzi
		vaso		

<sup>1</sup>The early studies by House and Fairbanks and Lehiste and Peterson included both obstruents and sonorants, but much of the experimental work of the 1970s and 1980s focused only on obstruents. Hombert seems to have recognized the importance of establishing some sort of basis of comparison, and anticipated Hanson's idea of comparing obstruents to nasals (Hombert, 1978; Hombert *et al.*, 1979, p. 45). However, these studies were fairly preliminary, and the idea does not seem to have taken up in most subsequent research.

<sup>2</sup>See also Kingston (2007), who studied *CF0* effects in different intonational contexts in the speech of four American English speakers. Control of intonation is important, as earlier studies (e.g., Ladd and Silverman, 1984) have shown that the magnitude of intrinsic *F0* perturbations may be attenuated in some prosodic contexts, a finding confirmed by Hanson's and Kingston's more recent studies, and in the present paper.

<sup>3</sup>Kingston and Diehl (1994, pp. 433–434) present data showing that *F0* following /s+stop/ clusters is in general intermediate between those of voiced and voiceless singletons, a fact they attribute to the neutralization of the [voice] contrast in this context.

<sup>4</sup>See supplementary material at <http://dx.doi.org/10.1121/1.4962445> for audio examples, additional plots, predicted marginal means, and R code to reproduce our figures and analyses.

<sup>5</sup>The addition of by-item random slopes for CONTEXT was not justified by the data, and by-item random slopes for VOICE do not seem to make sense given that each item only ever occurs with one level of VOICE.

<sup>6</sup>We thank John Kingston for suggesting this possible explanation.

<sup>7</sup>It is also possible that sudden post-release activation of the vocalis muscle, which Löfqvist *et al.* (1984) and Hutter (1985) show to be suppressed during voiceless closures in Swedish and Danish, respectively, may also contribute to raised *F0* (Hoole and Honda, 2011, fn. 9).

Abdelli-Beruh, N. B. (2004). "The stop voicing contrast in French sentences: Contextual sensitivity of vowel duration, closure duration, voice onset time, stop release and closure voicing." *Phonetica* 61, 201–219.

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). "Fitting linear mixed-effects models using lme4." *J. Stat. Software* 67(1), 1–48.

Beckman, J., Helgason, P., McMurray, B., and Ringen, C. (2011). "Rate effects on Swedish VOT: Evidence for phonological overspecification." *J. Phon.* 39, 39–49.

Beckman, J., Jessen, M., and Ringen, C. (2013). "Empirical evidence for laryngeal features: Aspirating vs true voice languages." *J. Linguist.* 49, 259–284.

Bell-Berti, F. (1975). "Control of pharyngeal cavity size for English voiced and voiceless stops." *J. Acoust. Soc. Am.* 57, 456–461.

Benguerel, A.-P., Hirose, H., Sawashima, M., and Ushijima, T. (1978). "Laryngeal control in French stop production: A fiberscopic, acoustic and electromyographic study." *Folia Phoniatr. Logop.* 30, 175–198.

Boersma, P. (1993). "Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound." *Proc. Inst. Phon. Sci. Univ. Amsterdam* 17, 97–110.

Boersma, P., and Weenink, D. (2015). "Praat: Doing phonetics by computer (version 5.4.08) [computer program]." <http://www.praat.org/> (Last viewed December 19, 2014).

Caramazza, A., and Yeni-Komshian, G. H. (1974). "Voice onset time in two French dialects." *J. Phon.* 2, 239–245.

Chen, Y. (2011). "How does phonology guide phonetics in segment–f0 interaction?" *J. Phon.* 39, 612–625.

Crystal, T. H., and House, A. S. (1988). "Segmental durations in connected-speech signals: Current results." *J. Acoust. Soc. Am.* 83, 1553–1573.

Dmitrieva, O., Llanos, F., Shultz, A. A., and Francis, A. L. (2015). "Phonological status, not voice onset time, determines the acoustic realization of onset f0 as a secondary voicing cue in Spanish and English." *J. Phon.* 49, 77–95.

Draxler, C., and Jänsch, K. (2004). "SpeechRecorder—a universal platform independent multi-channel audio recording software." in *Proceedings of LREC*, Lisbon, Portugal, pp. 559–562.

Edmondson, J. A., and Esling, J. H. (2006). "The valves of the throat and their functioning in tone, vocal register and stress: Laryngoscopic case studies." *Phonology* 23, 157–191.

Erickson, D., Baer, T., and Harris, K. S. (1982). "The role of the strap muscles in pitch lowering." *Haskins Lab. Status Rep. Speech Res.* SR-70, 275–284.

Esposito, A. (2002). "On vowel height and consonantal voicing effects: Data from Italian." *Phonetica* 59, 197–231.



- Ewan, W. G. (1976). "Laryngeal behavior in speech," Ph.D. dissertation, University of California, Berkeley, CA.
- Flege, J. E. (1982). "Laryngeal timing and phonation onset in utterance-initial English stops," *J. Phon.* **10**, 177–192.
- Francis, A. L., Ciocca, V., Wong, V. K. M., and Chan, J. K. L. (2006). "Is fundamental frequency a cue to aspiration in initial stops?," *J. Acoust. Soc. Am.* **120**, 2884–2895.
- Goldstein, L. M., and Browman, C. P. (1986). "Representation of voicing using articulatory gestures," *J. Phon.* **14**, 339–342.
- Gordon, M., and Ladefoged, P. (2001). "Phonation types: A cross-linguistic overview," *J. Phon.* **29**, 383–406.
- Halle, M., and Stevens, K. N. (1971). "A note on laryngeal features," *MIT Q. Prog. Rep.* **101**, 198–212.
- Hanson, H. M. (2009). "Effects of obstruent consonants on fundamental frequency at vowel onset in English," *J. Acoust. Soc. Am.* **125**, 425–441.
- Hombert, J.-M. (1976). "The effect of aspiration on the fundamental frequency of the following vowel," in *Proceedings of the 2nd Annual Meeting of the Berkeley Linguistics Society* (Berkeley Linguistics Society, Berkeley, CA), pp. 212–219.
- Hombert, J.-M. (1978). "Consonant types, vowel quality, and tone," in *Tone: A Linguistic Survey*, edited by V. Fromkin (Academic Press, San Diego, CA), pp. 77–111.
- Hombert, J.-M., Ohala, J. J., and Ewan, W. G. (1979). "Phonetic explanations for the development of tones," *Language* **55**, 37–58.
- Honda, K., Hirai, H., Masaki, S., and Shimada, Y. (1999). "Role of vertical larynx movement and cervical lordosis in F0 control," *Lang. Speech* **42**, 401–411.
- Hoole, P., and Honda, K. (2011). "Automaticity vs feature-enhancement in the control of segmental F0," in *Where do Phonological Features Come From?: Cognitive, Physical and Developmental Bases of Distinctive Speech Categories*, edited by N. Clements and R. Ridouane (John Benjamins, Amsterdam, the Netherlands), pp. 133–171.
- House, A. S., and Fairbanks, G. (1953). "The influence of consonant environment upon the secondary acoustic characteristics of vowels," *J. Acoust. Soc. Am.* **25**, 105–113.
- Hutters, B. (1985). "Vocal fold adjustments in aspirated and unaspirated stops in Danish," *Phonetica* **42**, 1–24.
- Jacewicz, E., Fox, R. A., and Lyle, S. (2009). "Variation in stop consonant voicing in two regional varieties of American English," *J. Int. Phon. Assoc.* **39**, 313–334.
- Jessen, M., and Roux, J. C. (2002). "Voice quality differences associated with stops and clicks in Xhosa," *J. Phon.* **30**, 1–52.
- Jongman, A., Wayland, R., and Wong, S. (2000). "Acoustic characteristics of English fricatives," *J. Acoust. Soc. Am.* **108**, 1252–1263.
- Jun, S.-A., and Fougeron, C. (2002). "Realizations of accentual phrase in French intonation," *Probus* **14**, 147–172.
- Keating, P. A. (1984). "Phonetic and phonological representation of stop consonant voicing," *Language* **60**, 286–319.
- Keyser, S. J., and Stevens, K. N. (2006). "Enhancement and overlap in the speech chain," *Language* **82**, 33–63.
- Kim, C.-W. (1970). "A theory of aspiration," *Phonetica* **21**, 107–116.
- Kingston, J. (2007). "Segmental influences on F0: Automatic or controlled?," in *Tones and Tunes, Volume 2: Experimental Studies in Word and Sentence Prosody*, edited by C. Gussenhoven and T. Riad (Mouton de Gruyter, Berlin, Germany), pp. 171–201.
- Kingston, J., and Diehl, R. L. (1994). "Phonetic knowledge," *Language* **70**, 419–454.
- Kirby, J., and Ladd, D. R. (2015). "Stop voicing and f0 perturbations: Evidence from French and Italian," in *Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow, Scotland.
- Kohler, K. J. (1982). "F0 in the production of fortis and lenis plosives," *Phonetica* **39**, 199–218.
- Kohler, K. J. (1984). "Phonetic explanation in phonology: The feature fortis/lenis," *Phonetica* **41**, 150–174.
- Kohler, K. J. (1985). "F0 in the perception of lenis and fortis plosives," *J. Acoust. Soc. Am.* **78**, 21–32.
- Ladd, D. R. (2008). *Intonational Phonology* (Cambridge University Press, Cambridge, UK).
- Ladd, D. R., and Silverman, K. E. A. (1984). "Vowel intrinsic pitch in connected speech," *Phonetica* **41**, 31–40.
- Lea, W. A. (1973). "Segmental and suprasegmental influences on fundamental frequency contours," in *Consonant Types and Tone*, edited by L. M. Hyman (University of Southern California, Los Angeles, CA), pp. 15–70.
- Lehiste, I., and Peterson, G. E. (1961). "Some basic considerations in the analysis of intonation," *J. Acoust. Soc. Am.* **33**, 419–425.
- Lenth, R., and Hervé, M. (2014). "lsmmeans: Least-squares means (version 2.13)," <http://CRAN.R-project.org/package=lsmmeans> (Last viewed April 13, 2015).
- Lisker, L., and Abramson, A. S. (1964). "A cross-language study of voicing in initial stops: Acoustical measurements," *Word* **20**, 384–422.
- Löfqvist, A., Baer, T., McGarr, N. S., and Story, R. S. (1989). "The cricothyroid muscle in voicing control," *J. Acoust. Soc. Am.* **85**, 1314–1321.
- Löfqvist, A., Koenig, L. L., and McGowan, R. S. (1995). "Vocal tract aerodynamics in /aCa/ utterances: Measurements," *Speech Commun.* **16**, 49–66.
- Löfqvist, A., McGarr, N. S., and Honda, K. (1984). "Laryngeal muscles and articulatory control," *J. Acoust. Soc. Am.* **76**, 951–954.
- Maddieson, I. (1984). "The effects on F0 of a voicing distinction in sonorants and their implications for a theory of tonogenesis," *J. Phon.* **12**, 9–15.
- Nearey, T. M., and Rochet, B. L. (1994). "Effects of place of articulation and vowel context on VOT production and perception for French and English stops," *J. Int. Phon. Assoc.* **24**, 1–18.
- Ní Chasaide, A., and Gobl, C. (1993). "Contextual variation of the vowel voice source as a function of adjacent consonants," *Lang. Speech* **36**, 303–330.
- Ohala, J. J. (1975). "Phonetic explanations for nasal sound patterns," in *Nasalfest: Papers From a Symposium on Nasals and Nasalization*, edited by C. A. Ferguson, L. M. Hyman, and J. J. Ohala (Language Universals Project, Stanford, CA), pp. 289–316.
- Ohala, J. J. (1997). "Aerodynamics of phonology," in *Proceedings of the 4th Seoul International Conference on Linguistics*, Seoul, South Korea, pp. 92–97.
- Ohde, R. N. (1984). "Fundamental frequency as an acoustic correlate of stop consonant voicing," *J. Acoust. Soc. Am.* **75**, 224–230.
- Sawashima, M., Hirose, H., Yoshioka, H., and Kiritani, S. (1982). "Interaction between articulatory movements and vocal pitch control in Japanese word accent," *Phonetica* **39**, 188–198.
- Silverman, K. E. A. (1986). "F0 segmental cues depend on intonation: The case of the rise after voiced stops," *Phonetica* **43**, 76–91.
- Silverman, K. E. A. (1987). "The structure and processing of fundamental frequency contours," Ph.D. dissertation, University of Cambridge, Cambridge, UK.
- Solé, M.-J. (2011). "Articulatory adjustments in initial voiced stops in Spanish, French and English," in *Proceedings of the 17th International Congress of Phonetic Sciences*, Hong Kong, pp. 1878–1881.
- Stevens, K. N. (1971). "Airflow and turbulence noise for fricative and stop consonants: Static considerations," *J. Acoust. Soc. Am.* **50**, 1180–1192.
- Vaggés, K., Ferrero, F. E., Magno-Caldognetto, E., and Lavagnoli, C. (1978). "Some acoustic characteristics of Italian consonants," *J. Ital. Linguist.* **3**, 68–85.
- Westbury, J. R. (1983). "Enlargement of the supraglottal cavity and its relation to stop consonant voicing," *J. Acoust. Soc. Am.* **73**, 1322–1336.
- Whalen, D. H., Abramson, A. S., Lisker, L., and Mody, M. (1990). "Gradient effects of fundamental frequency on stop consonant voicing judgments," *Phonetica* **47**, 36–49.
- Whalen, D. H., Lisker, L., Abramson, A. S., and Mody, M. (1993). "F0 gives voicing information even with unambiguous voice onset times," *J. Acoust. Soc. Am.* **93**, 2152–2159.