Effects of obstruent voicing on vowel F0: implications for laryngeal realism

Abstract

It is sometimes argued that languages with two-way laryngeal contrasts can be classified according to whether one series is realized canonically with voicing lead or the other with voicing lag. In languages of the first type, such as French, the phonologically relevant features is argued to be [voice], while in languages of the second type, such as German, the relevant feature is argued to be [spread glottis]. A crucial assumption of this position is that the presence of certain contextually stable phonetic cues, namely voicing lead or lag, can be used to diagnose the which feature is phonologically active.

In this paper, we present data on obstruent-intrinsic F0 perturbations (CF0) in two [voice] languages, French and Italian. Voiceless obstruents in both languages are found to raise F0, while F0 following (pre)voiced obstruents patterns together with sonorants, similar to the voiceless unaspirated stops of [spread glottis] languages like German and English. The contextual stability of this cue implies that an active devoicing gesture is common to languages of both the [voice] and [spread glottis] types, and undermines the idea that a strict binary dichotomy between true voicing and aspirating languages can be reliably inferred based on properties of the surface phonetics.

Keywords: laryngeal realism : voicing : CF0 : phonetic grounding : privativity

1. Introduction: the problem with [±voice]

It has long been observed that a distinctive feature system with a single, universal [voice] feature appears inadequate to describe the typological variability found in systems of laryngeal contrast in the world's languages. One aspect of this inadequacy has been much discussed in the recent literature: in some languages the phonetic property that is somehow 'available' to the phonology seems to be voicelessness, while in others, the relevant property

appears to be voicing. To take one example among many, Iverson and Salmons (1995: 381–2) cite differences of voicing assimilation in obstruent+obstruent sequences in German and Dutch: German shows evidence of assimilation to voiceless stops and does not allow anticipatory assimilation to voicing (1), whereas Dutch exhibits cross-syllable assimilation of voiceless stops to a following voiced obstruent (2).

(1) German: *fragte* /fʀag + tə/ > [fʀaktə] 'ask.3SG.PAST', but *undenkbar* [ʊndɛŋ**k**bar] 'unthinkable', not *[ʊndɛŋgbar]

(2) a. Dutch: *zaken* /zak + ɛn/ > [zakɛn] 'pockets' but *zakdoek* /zak + duk/ > [za**g**duk] 'handkerchief"

b. *ik* [ɪk] 'I' but *ik ben* /ɪk + bɛn/ > [ɪ**g**bɛn] 'I am'

c. *kook* [kok] 'cook', *boek* [buk] 'book', but *kookboek* [ko**g**buk] 'cookbook'

The actual phonetic facts are acknowledged to be far more complex and varied than such transcription-based examples suggest (see e.g. Slis 1986 for careful experimental phonetic work on Dutch), and there are complications that arise from a general tendency to devoicing in coda position in both languages, but it does appear that there is a genuine difference in voicing assimilation behaviour between Dutch and German. This is only one of many cross-linguistic differences that go unexplained in any descriptive system, such as that proposed in the *Sound Pattern of English* (Chomsky and Halle 1968), that treats all two-way laryngeal contrasts in terms of a single feature [±voice].

One proposed solution to such issues is a descriptive system with two separate features that reflect two different laryngeal gestures (Avery and Idsardi 2001; Beckman, Jessen and Ringen 2013; Honeybone 2005; Iverson and Salmons 1995; Jessen and Ringen 2002). This idea has antecedents as far back as Sievers 1876 (and more recently Jakobson 1949); one recent manifestation of this idea sometimes goes by the name of LARYNGEAL REALISM, a term coined by Honeybone (2002). On the basis of phonetic, phonological, and

diachronic evidence, laryngeal realism draws a distinction between 'true voicing' languages, such as French and Dutch, and 'aspirating' languages, such as German and English. The precise details vary somewhat from author to author, but true voicing languages are generally said to specify the feature [voice], contrasting *fully voiced* (lenis) stops with *voiceless unaspirated* (fortis) stops, while aspirating languages specify [spread glottis], and distinguish *aspirated* (fortis) stops from stops that are crucially *unaspirated* (lenis) and may sometimes be 'passively voiced' (see Table 1).

Table 1: Comparison of SPE-style phonological representation of laryngeal contrast with the laryngeal realist proposal.

	Phonetic	**Orthographic**	**SPE**	**Laryngeal realism**
true voicing	[p t k]	*p t k*	[-voice]	[Ø]
	[b d g]	*b d g*	[+voice]	[voice]
aspirating	[p^h t^h k^h]	*p t k*	[-voice]	[spread glottis]
	[p t k]/[b̥ d̥ g̊]	*b d g*	[+voice]	[Ø]

At this point it becomes potentially important to distinguish between (at least) two strands of laryngeal realism, which differ in terms of what types of evidence is deemed relevant. For some researchers, laryngeal realism is based on two fundamental assumptions about phonological features, namely, that features are both PRIVATIVE and PHONETICALLY GROUNDED. Privative feature values, familiar from theories of underspecification in phonology, are taken to be categorically either present or absent in the phonological representation of any given segment (Lombardi 1991; Steriade 1995). At the same time, grounding features in phonetic substance creates a more transparent mapping between the distinctive feature specification and the phonetics (Archangeli and Pulleyblank 1994; Boersma 1998; Donegan and Stampe 1979; Hayes, Kirchner and Steriade 2004), which has

the attractive property of helping to explain cross-linguistic similarity in the phonological behavior of segments among unrelated languages.

However, there are also self-professed 'laryngeal realists' such as Iverson and Salmons (1995, 2006) or Iosad (2017), who would appear to adopt the notion of featural privativity without demanding transparent phonetic grounding of features. For these researchers, the more decisive evidence is that of phonological processes, such as neutralization and voicing assimilation. For example, the apparent asymmetry between languages which only show assimilation to fortis obstruents versus those which only show assimilation to lenis obstruents is taken as strong evidence of some kind of difference in featural specification.

For concreteness, we focus our critique primarily on the theoretical variant of the first type, specifically that pursued by Beckman, Ringen and colleagues (Beckman, Jessen and Ringen 2013; Jessen and Ringen 2002; Petrova, Plapp, Ringen and Szentgyörgyi 2006; Ringen and Kulikov 2012, *inter alia*), because it makes a strong and explicit claim that derives both from the assumption of privativity and the relevance of phonetic behavior. Specifically, the prediction is that phonetic voicing will be phonetically 'active' in true voicing languages and 'passive' in aspirating languages:

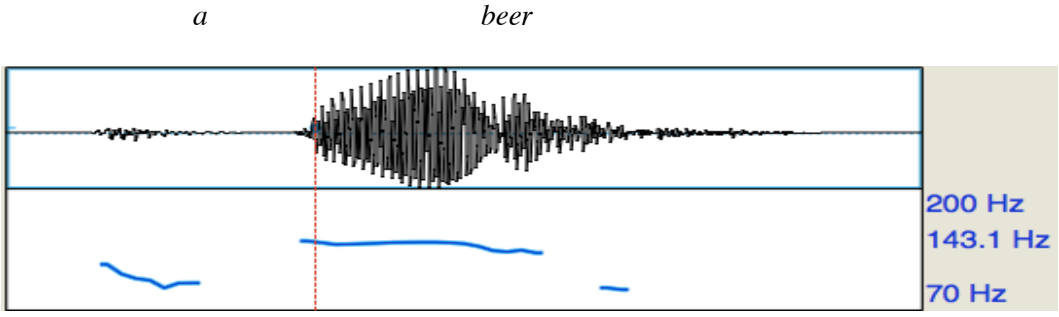
> "…the claim is that speakers [of aspirating languages] are **not actively aiming** to voice the intervocalic/intersonorant lenis stops, just as they are **not actively aiming** to voice the word/utterance-initial lenis stops…" (Beckman, Jessen and Ringen 2013: 261; emphasis added)

A corollary of this claim would seem to be that speakers of true voicing languages do not actively 'aim for voicelessness' in their unmarked obstruent series, and that in such languages, categorical phonetic processes associated with voicelessness should be absent.

In this paper, we show that there is empirical evidence which is inconsistent with this corollary. First, we consider data on *CF0 effects*—systematic differences in F0 accompanying voiced and voiceless consonants—and their implications for the laryngeal realist position. Based on these data, we argue that voicelessness involves an active gesture even in languages argued to be true voicing languages, and suggest that this undermines the superficial attraction of the privativity assumption. We also show that a variety of other phonological evidence for the putative typological distinction is at best equivocal. This leads us to conclude that even phonetic behavior which appears categorical may have an at best opaque relationship to phonological representations.

2. Effects of voicing specification on vowel F0

It has been known at least since House and Fairbanks (1953) that differences of consonant voicing are associated with 'microprosodic' effects on fundamental frequency, which we refer to as *CF0* (after Kingston 2007). House and Fairbanks found that F0 in vowels adjacent to voiceless consonants is higher than F0 in vowels adjacent to voiced consonants; these observations have been corroborated many times since then. An example of the CF0 effect in American English is illustrated in Figure 1.



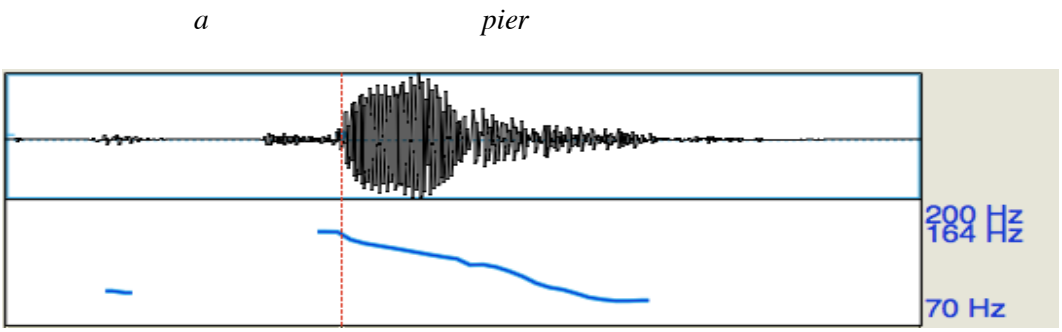


Figure 1: Illustrations of F0 contours accompanying American English productions of *a beer* and *a pier*, perceived as intonationally equivalent. F0 at vowel onset is approximately 20 Hz higher following /p/ than /b/, while F0 at offset is roughly comparable.

Hanson (2009) studied CF0 in American English using a set of materials carefully controlled for segmental and intonational context. Crucially, in addition to voiced and voiceless stops, Hanson's materials included nasals, which are not predicted to affect F0[1]; she also included /sC/ onset clusters, which allowed her to compare CF0 effects in different structural (phonological) contexts. Her main findings are illustrated schematically in Figure 2. American English lenis stops /b d g/, which in her materials were never realized with closure voicing in absolute initial position, did not significantly perturb F0 away from the nasal baseline. By contrast, fortis stops and (and fricatives) always had an F0-raising effect, regardless of whether they were initial or part of an /sC/ cluster). These findings appear to show that the *phonological* context somehow determines whether or not a CF0 effect is present. This finding is consistent with the strong prediction of laryngeal realism, namely that, in aspirating languages like English, a gesture to support voicelessness is active, even when (as in /sC/ clusters) 'aspiration' in the sense of long-lag VOT is superficially absent. Specifically, the acoustic effect of local F0 raising is what we might expect from a gesture to inhibit voicing, such as stiffening of the vocal folds and/or engagement of the cricothyroid

[1] This follows from the observation that nasals do not entail any specific articulatory adjustments in order to maintain vocal fold vibration and the fact that the lack of resistance offered by the nasal cavity does not condition the type of decrease in transglottal pressure that would be expected to perturb pitch (Ohala, 1975; Hombert et al., 1979).

musculature, either as an aerodynamic (Hombert and Ladefoged 1977; Kohler 1982) or an articulatory consequence (Halle and Stevens 1971; Löfqvist, Baer, McGarr and Story 1989). The precise physical mechanisms are less important for laryngeal realism than (1) the presence of an active gesture, and (2) the fact that the putatively unmarked obstruents have no effect on F0, suggesting that an active gesture is absent.

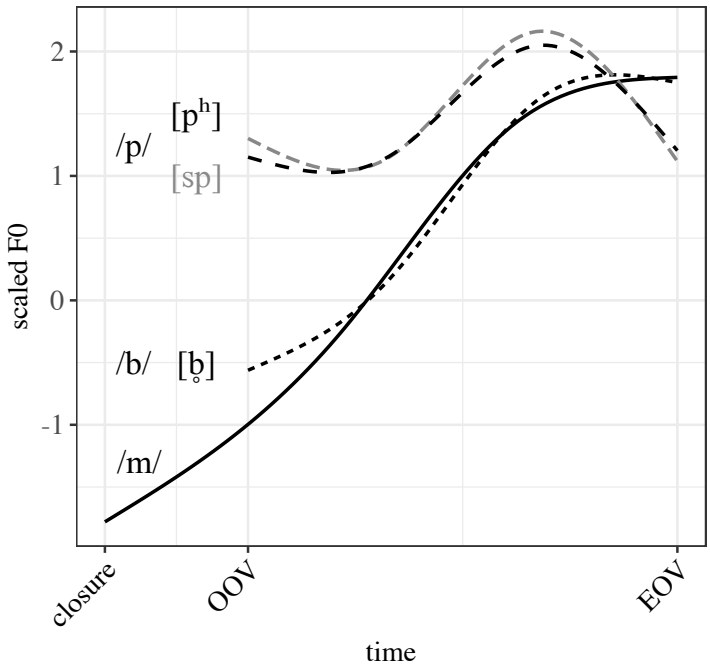


Figure 2: Schematic comparisons of time course of F0 production in a high intonational context in American English (after Hanson 2009). Adapted from Kirby and Ladd (2016: 2402).

However, this line of reasoning gives rise to the question of what happens in true voicing languages. If (as suggested by the laryngeal realist literature) true voicing languages are characterised by an active feature [voice] manifested phonetically by an active articulatory gesture to support voicing, then in such languages we might expect to find lower F0 accompanying voiced stops, while the effect of unmarked ('voiceless') obstruents should be

negligible. This hypothesis, illustrated in Figure 3, was tested experimentally by Kirby and Ladd (2016).

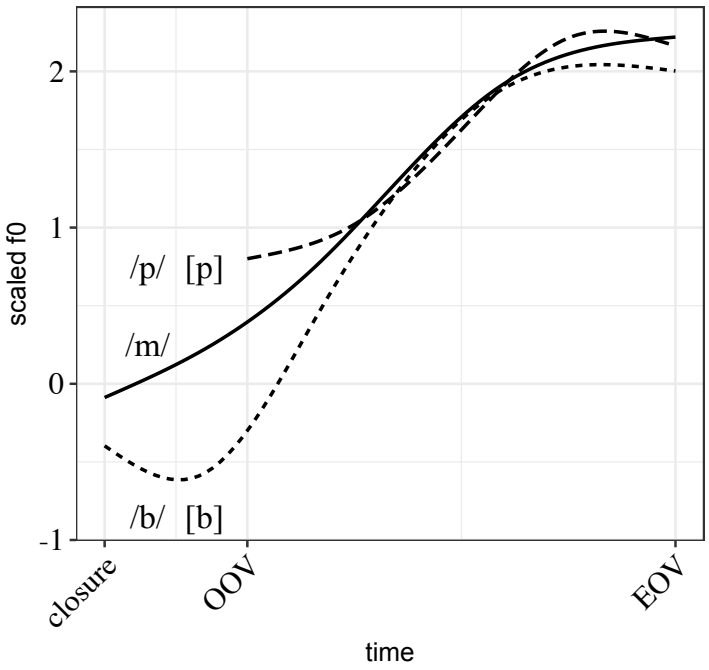


Figure 3: Hypothetical comparisons of time course of F0 production in the language where [voice] obstruents lower F0 during the closure phase. Adapted from Kirby and Ladd (2016: 2402).

Kirby and Ladd conducted experiments similar to Hanson's in two 'true voicing' languages, French and Italian. Their main results are reproduced in Figure 4, where it can be seen that F0 is indeed lowered during the closure phase of phonologically voiced obstruents. This F0 lowering is precisely what we might expect if gestures to support voicing are actively controlled in these languages. Maintenance of voicing during an obstruent closure is known to be aerodynamically challenging (Ohala 1983): in order to have vocal fold vibration, a constant transglottal pressure differential needs to be maintained, but if the cavity above the glottis is closed off, then very quickly the pressure on both sides of the glottis equalizes, and phonation ceases. As a result, many languages take action to help maintain voicing during fully voiced obstruents, such as larynx lowering, velic leakage, and pharyngeal expansion

(Bell-Berti 1975; Erickson, Baer and Harris 1982; Solé 2018; Westbury 1983). Lowering F0 may also serve to reinforce other acoustic properties associated with (phonetic) voicing (Kingston and Diehl 1994; Kingston, Diehl, Kirk and Castleman 2008). Kirby and Ladd's results clearly support the expectation that there is a phonetic gesture in French and Italian corresponding to the phonological feature [voice].

However, in addition to the lowering effect of closure voicing on F0, Kirby and Ladd found that both French and Italian also exhibit ***raised*** F0 after phonologically voiceless obstruents, exactly as Hanson found for American English[2]. Similar results are reported in a multi-language corpus study by Sonderegger et al. (2017), who found raised F0 following fortis stops in a diverse range of languages including Croatian, Swedish, Russian, Spanish, French, Turkish, and Korean. That is, phonetically voiceless obstruents appear to induce F0 raising in many (or even most) languages, independent of any other aspects of how the voicing contrast is realized.

The apparent generality of the link between F0 raising and voicelessness suggests that it results from an active gesture to inhibit phonation, which is present irrespective of whether [voice] or [spread glottis] is treated as the active feature in the language's system of laryngeal contrasts. This would seem to contradicts the assumption that 'active' phonetic gestures are predictive of privative phonological feature values. More generally, it makes trouble for the assumption that only a single feature – either [voice] or [spread glottis], but not both – is phonetically active in two-way systems of laryngeal contrast. A similar conclusion was reached by Beckman and colleagues in their study of Central Swedish (Beckman, Helgason, McMurray and Ringen 2011), which appears to require both active [voice] and [spread glottis] features despite having just a two-way laryngeal contrast. We suspect that this state of affairs is the rule rather than exception.

[2] A reviewer poses the question of whether the magnitude of the F0 raising following French and Italian lenis stops is in fact comparable to that of American English fortis stops. At least for French, the magnitude of the effect appears to be roughly comparable, although in the Kirby and Ladd (2016) data it does not persist throughout the vowel to the extent typically observed in English, likely due to language-particular differences in the timing of the intonational F0 peak.

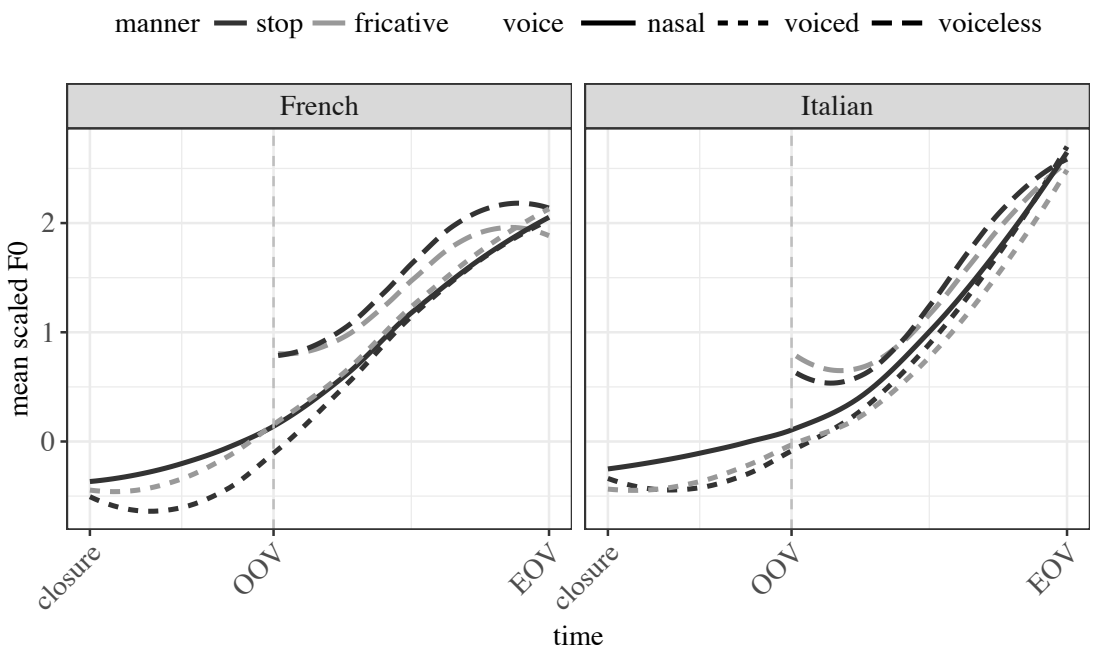


Figure 4: Average time-normalized standardized F0 contours from time of oral closure through end of vowel (EOV) for stops, fricatives, and nasals in French and Italian. Dashed vertical line indicates onset of post-release voicing, i.e., the first period of the vowel (OOV). Adapted from Kirby and Ladd (2016: 2405).

3. Evidence from phonology: regressive voicing assimilation

There are other reasons to be sceptical of the claim that [voice] is the only feature 'available' to the phonology in true voicing languages. A number of authors describe processes in true voicing languages in which voiceless obstruents appear to be phonologically active. Rubach (1996) argues that voiced fricatives may devoice following voiceless obstruents in Polish. Wetzels and Mascaró (2001) discuss instances of regressive voicing assimilation in French, Romanian, and Serbo-Croatian. Iosad (2017) argues that analysis of final neutralization and initial consonant mutations in Bothoa Breton are simplified if the voiceless (fortis) series is taken to be phonologically active. And Bennett and Rose (2017)

describe a dissimilation process in the Kordofanian language Moro, whereby voiceless stops and affricates become voiced before voiceless obstruents. Despite the fact that Moro is otherwise a prototypical 'true voicing' language, the authors argue that this process is best understood as a process of dissimilation to [-voice].

It is not difficult to find additional examples of this type. Here, we briefly discuss two instances of regressive voicing assimilation, a process which is suggested to covary with the presence of prevoiced obstruents (Kohler 1984; Wissing and Roux 1995) and thus to be a potential diagnostic of a [voice] language (Iverson and Salmons 1995; van Rooy and Wissing 2001; but cf. Ringen and Helgason 2004).

Our first example is the Italian prefix *s-*, a semi-productive affix meaning something like 'un-' or 'de-'. This prefix surfaces with two phonetic variants, [s] and [z], as shown in (3).

(3) a. with [s-]: [s]*taccare* 'detach', [s]*fasciare* 'destroy', [s]*comodo* 'uncomfortable'

b. with [z-] before obstruents: [z]*bagliare* 'make a mistake', [z]*ganciare* 'unhook, release', [z]*venire* 'faint', but

c. also before sonorants[3]: [z]*radicare* 'uproot', [z]*nodare* 'untie'

If the underlying form of the morpheme is assumed to be /s/, there is a problem generating the phonetic form [z] before the stem-initial sonorants /r/ and /n/ in *sradicare* and *snodare*: according to theories on which phonological feature specification are truly privative, sonorants would not bear a specification for [voice], voicing not being contrastive for this class of segments. However, in order to trigger the assimilation of the /s-/ prefix, it would be necessary to assume that such sonorants are (perhaps redundantly) specified with a [voice] feature, a problematic proposal given that pre-sonorant voicing differs systematically from regressive voicing assimilation triggered by obstruents (see Strycharczuk 2012 for detailed

[3] Note that *s-* never occurs before a vowel, so it is not possible to use this environment to diagnose its underlying representation.

arguments and discussion). The alternative would be to accept that the underlying form is in fact /z/, and that examples of type (3a) are the result of a gesture associated with the stem-initial obstruents /t/, /f/, and /k/ which induces the prefix to surface as voiceless. This seems to us to be the more reasonable account, especially given the evidence reviewed above that Italian voiceless unaspirated stops are indeed produced with a laryngeal gesture to inhibit voicing.

Our second example comes from regressive voicing assimilation in French. Traditional descriptions of this process generally state that in sequences of obstruents, voicing assimilation is regressive – the voicing specification of the second obstruent is anticipated in the first – and that voiced and voiceless consonants are equally affected. Grevisse (1993: §36b), a traditional reference grammar, says that in a sequence of two consonants differing in voicing 'la première s'assimile à la seconde' (the first is assimilated to the second); Valdman et al. (1964: 254ff.), a pronunciation teaching book for foreign learners of French, says that 'the first consonant is influenced by the second'. Indeed, Valdman et al. go on to provide exercises giving equal time to assimilation to both voiced and voiceless obstruents: 'Practice devoicing of voiced consonants ... in the middle of the following utterances', then 'Now practice voicing of voiceless consonants'. The first practice list includes phrases like those in (4a); the second gives examples like those in (4b):

(4) a. *coup de pied* /kudpje/ 'kick (noun)' > [kutpje]

quinze francs /kɛ̃zfrɑ̃/ '15 francs' > [kɛ̃sfrɑ̃]

b. *tasse de thé* /tasdəte/ 'cup of tea' > [tazdəte]

petite gamine /ptitgamin/ 'little child (fem.)' > [ptidgamin]

Such descriptions also often make two further points: first, that the assimilation is not necessarily complete, and second, that there is a difference between assimilation to voicing

and assimilation to voicelessness. Both points are neatly summarised in Valdman et al.'s succinct statement: 'If the first consonant is voiced and the second voiceless, the first becomes devoiced; if the first is voiceless and the second voiced, the first become partially voiced.' Grevisse's account is more extensive. With respect to the completeness of the assimilation, his statement that the first consonant assimilates to the second, quoted above, is qualified: he goes on to say '...mais du point de vue de la sonorité seulement: elle garde sa force articulatoire' (but only from the point of view of voicing; [the consonant] retains its force of articulation[4]). Grevisse further says that complete assimilation 'est généralement considérée comme incorrecte', but from this judgement he explicitly excludes /b/ and /d/ immediately followed by a voiceless consonant, which are correctly pronounced [p] and [t]; his example is *obtenir* [ɔptənir] 'obtain'. Other sources (e.g. Grundstrom 1983: 63 and many dictionaries) further note the existence of an exception to this latter generalisation, namely the verb *subsister* 'subsist', whose prescriptively condoned pronunciation is [sybziste].

None of the foregoing is consistent with the expectations of an analysis in terms of a single active feature [voice]. First, assimilation is triggered by both voiced and supposedly unmarked 'voiceless' consonants; second, to the extent that there is an asymmetry, it appears that the voiceless consonants are more effective triggers than voiced consonants; and third, everyone agrees that the assimilation is frequently incomplete and that there are lexical exceptions. These traditional statements are backed up by more recent experimental phonetic evidence. For example, Snoeren, Hallé and Segui (2006) found that 27% of underlyingly voiced word-final stops surface as voiceless in the appropriate assimilation contexts. In a related finding, Hallé and Adda-Decker (2011) show that French voicing assimilation is by no means a categorical substitution of one sound for another, but a complex matter of partial assimilation in both directions. It is difficult to see how these facts can be reconciled with the

[4] The notion 'force of articulation' refers to an earlier passage (§32a) presumably based on the notions fortis and lenis: 'Les consonnes sourdes sont dites aussi *fortes*, parce qu'elles exigent un effort plus considérable que les consonnes sonores, dites aussi *faibles* ou *douces*' (Voiceless consonants are also called *strong*, because they require greater effort than voiced consonants, which are also called *weak* or *soft*).

idea that only the feature [voice] is active in French phonology (see also Wetzels and Mascaró 2001).

4. Implications

Beckman et al. (2013) develop a descriptive framework for dealing with phonetic manifestations of features that, according to the theory, are supposed to be phonologically inactive. Although they do not consider CF0 effects (but see Jessen 2001), they discuss at some length the fact that fortis stops in 'true voicing' languages routinely do *not* undergo passive voicing, which is unexpected if these stops are laryngeally unmarked (2013: 277-280). Their response to this objection is to introduce a level of representation based on 'phonetic distinctive features' between the phonetics and the phonology. This involves two additional stipulations:

(5) a. Every segment receives a numerical specification for all active features;
b. Phonetic processes cannot apply to segments which bear a numerical feature specification.

On this account, (unmarked) fortis stops in true voicing languages would be transformed into something like [1VOICE] prior to phonetic spell-out (vs. something like [9VOICE] for [voice] stops in the same language) while (unmarked) lenis stops in aspirating languages will receive a specification of something like [1SPREAD] (vs. something like [9SPREAD] for [spread glottis] stops in an aspirating language). Stipulation (5b) is then invoked to block the application of passive voicing of lenis stops in true voicing languages.

We see two main issues with this proposal. First, it is not clear how it improves upon a simple, accurate description of the empirical facts. This is seen most clearly in Beckman et

al.'s treatment of the difference between German and Icelandic (2013: 280-281). At issue is the fact that (underspecified) intervocalic lenis stops in German are passively voiced, while the same stops in Icelandic, which is nominally also a [spread glottis] language, are not. The proposed solution is to assign different numerical specifications under stipulation (5a) (e.g. [5SPREAD] for the Icelandic lenis stops vs. [1SPREAD] for the German ones) and to modify (5b) to state that phonetic processes can apply only when the numerical specification is below some threshold. It is not clear how one would go about determining such a threshold short of observing the phonetic behaviour of obstruents in all the relevant environments in a given language, at which point phonetic behaviour is being invoked to motivate the very features which are meant to predict it. The heart of the matter is that, at least in some aspects of their realizations, German /p/ is not like Icelandic /p/, which is in turn not like French /p/. Such observations are of course nothing new (Chao 1934; Trubetzkoy 1958; Docherty 1992), but if we cannot predict whether /p/ in a given language will be [1SPREAD] or [5SPREAD], then we have lost whatever typological advantage privativity might have provided.

Related to this, there is considerable empirical evidence which runs counter to the assertion that intervocalic voicing processes are absent from 'true voicing' languages. The examples of Beckman et al. (2013: 278) and Jansen (2004: 48) notwithstanding, closure voicing of intervocalic fortis stops has been documented for Italian (Hualde and Nadeu 2011), Polish (Keating 1980), and Spanish (Hualde, Simonet and Nadeu 2011). If this type of voicing differs in kind or frequency compared with intervocalic voicing of fortis stops in languages such as German or English, this is again an empirical fact demanding explanation, but on the face of it the existence of these processes runs counter to the predictions of Beckman et al.'s proposal, as well as to a strong typological asymmetry between languages which only ever have assimilation to one class of obstruent or the other.

The second issue relates to CF0 effects. A key component of Beckman et al.'s proposal is that the presence of prevoicing on obstruents implicates the feature [voice], just as aspiration implicates [spread glottis] (2013: 277). By extension, then, the presence of consistently raised CF0 should implicate a feature associated with voicelessness. It clearly

cannot be the case that CF0 is a redundant property of underspecified segments, since it appears consistently with segments specified for [spread glottis] in aspirating languages. Yet neither can it be a function of the [spread glottis] specification, even a gradient numerical function, since true voicing languages will not have any numerical specification for (inactive) [spread glottis] at all. The only other option is to admit that the (underspecified, laryngeally 'inactive') fortis stops of [voice] languages *do* have an associated gesture some kind, and that this same gesture may also be associated with voiceless aspirated stops in languages such as German and English.[5]

We think this is the only interpretation consistent with the phonetic facts, and that it fatally undermines the idea that a strict binary dichotomy between true voicing and aspirating languages can be inferred based on properties of the surface phonetics. While the [voice]/[spread glottis] distinction correctly hints at the phonetic complexities which give rise to differences between languages, it is too coarse-grained to accurately reflect either the phonetic reality or the cross-linguistic typology.

On this last point, we would simply call attention to several more counter-examples to the idea that languages with two-way laryngeal contrasts can be neatly classified as belonging to either the true voicing or aspirating types. Without even considering complex cases such as Danish, it is not difficult to find examples of two-way languages which are not well-described by this dichotomy:

- Although Javanese contrasts two series of plosives, neither series is produced with vocal fold vibration during the closure phase, and both are produced with a similar short-lag VOT (Fagan 1988);

[5] A more extreme version of this might be to propose that voicelessness is *always* phonologically active, regardless of language type. For an example of this approach as pursued in the Onset Prominence framework see Schwartz (2017); Schwartz and Arndt (2018).

- Central Swedish contrasts prevoiced with long-lag aspirated stops, does not show evidence of regressive voicing assimilation (Ringen and Helgason 2004), and evidence from speech-rate effects suggests that both [spread glottis] and [voice] are 'active' in this language (Beckman, Helgason, McMurray and Ringen 2011);

- In Tokyo Japanese, phonologically voiceless stops are only lightly aspirated, and phonologically voiced stops frequently undergo devoicing in utterance-initial position, but without any accompanying change in their onset F0 behavior (Gao and Arai 2018).

Far from being exceptional, we suspect that such systems are actually quite common. Given the multidimensional and highly redundant acoustic space in which laryngeal contrasts are realized, however, it is unsurprising that different languages might simply 'choose' to amplify or privilege different acoustic parameters or combinations thereof (Jacewicz, Fox and Lyle 2009).

In our view, a fundamental problem in the way many phonologists think about cases like these is the persistent belief that some phonetic properties are somehow more 'phonological' while others are 'purely phonetic'. Many researchers appear to regard CF0 as somehow fundamentally different from other quasi-controlled effects, such as closure voicing or VOT. For example, in discussing the similarity of CF0 effects between English and Spanish, Dmitrieva et al. (2015) write that:

> "...the phonological status of the consonant may carry more weight in determining the onset f0 patterns than do its phonetic properties, such as the presence or absence of laryngeal voicing." (Dmitrieva 2015: 91)

In the same way, Jessen (1998, 2001) distinguishes between 'basic' vs. 'non-basic' correlates, 'basic' correlates being those with 'high contextual stability':

> "*Contextual stability* means that in many contexts in a language (often the clear majority) the relevant distinction is expressed by the correlate whose status is classified as basic" (Jessen 2001: 243).

Jessen (2001: 249-252) explicitly discusses CF0 as an example of a 'non-basic' correlate which nonetheless 'can be phonologized to a level of importance otherwise limited to the basic correlates' (252). While we think it is uncontroversial that the CF0 correlate is frequently phonologized, it is not clear to us in what sense the division of correlates into 'basic' and 'non-basic' is predictive rather than post-hoc, rather like the division of languages into [voice] and [spread glottis] categories. In our view, both fail to predict the substantive and consistent differences—and similarities—in the phonetic implementation of laryngeal contrasts.[6]

This issue was at the heart of Kingston and Diehl's landmark paper (1994) on phonetic knowledge, which draws a distinction between 'controlled' and 'automatic' phonetics. Although the term 'controlled' seems to suggest volition or action on the part of the speaker (and hence perhaps available for use in phonology), this is expressly *not* what the authors intend. Instead, a phonetic property is deemed to be 'controlled' if it is effectively arbitrary, in the sense that the speaker, or the speech community, might have settled on some other set of covarying phonetic properties to target. 'Automatic' phonetics, on the other hand, refers to the inevitable acoustic consequences of executing some particular gesture or set of gestures, based on physical properties of the vocal tract, the biomechanical particulars of how

[6] One possible metric for distinguishing basic from non-basic correlates, advocated by Jessen (2001, 2004), is perceptual saliency. While CF0 effects are clearly perceptible to listeners (Whalen, Abramson, Lisker and Mody 1990; Whalen, Lisker, Abramson and Mody 1993), the extent to which they are used as a cue to voicing in connected speech has not been established (Abramson and Whalen 2017).

the vocal folds oscillate, principles of aerodynamics, and so forth. The production of a particular canonical VOT would therefore be an example of controlled phonetics, because there exists a wide range of possible values that speakers or languages might aim to produce (Cho and Ladefoged 1999).

Kingston and Diehl make a strong case that CF0 effects (among other things) are the outcome of controlled phonetics. However, it would appear that many, if not all, of the phonetic correlates of voicing are similarly controlled. Consider:

- VOT is the acoustic (phonetic) effect of a particular (hence controlled, arbitrary, 'phonological') glottal aperture target, timed relative to the onset of voicing.
- Closure voicing is the acoustic (phonetic) effect of a particular (hence controlled, arbitrary, 'phonological') set of gestures to ensure a sufficient transglottal pressure differential during an obstruent closure;
- CF0 is the acoustic (phonetic) effect of a particular (hence controlled, arbitrary, 'phonological') laryngeal tension setting, timed relative to a constriction in the oral cavity.

We may assume that the same phonetic laws govern the acoustic realisations which follow from the implementation of these articulations. In this sense, they may be deemed automatic. At the same time, however, they are also clearly language-specific, in the sense that a particular choice of timing of peak glottal aperture does not entail a particular laryngeal tension setting. Spanish speakers produce short-lag stops accompanied by some laryngeal manoeuvre which increases vocal-fold tension; English speakers do not. It seems curious to us to argue that one consistent and perceptible language-specific effect is 'just' phonetic, while another is worthy of phonological reification. If VOT, CF0 and the presence or absence of laryngeal voicing are all controlled in Kingston and Diehl's sense, they must all be specified as part of the representation of contrasts.

5. Summary and conclusions

In this paper, we have discussed the relevance of obstruent-intrinsic F0 perturbations for privative theories of laryngeal features. In languages of both the French/Italian and English/German types, phonologically voiceless/fortis obstruents raise F0 in the following vowel, regardless of any other aspects of their phonetic realization. We have argued that this shows that even putatively unmarked members of a laryngeal contrast can also have active, canonical gestures associated with them. We therefore suggest that even predictable phonetic effects should not be taken as reliable indicators of the nature of phonological features; and phonologically active features, e.g. those which define a natural class for the purposes of a process or lexical generalization, may not necessarily, or even usually, be reliably signaled by particular phonetic properties.

While we believe that laryngeal realism is correct to highlight the typological diversity of voicing systems, a two-feature typology seems insufficient to predict consistent effects such as CF0 in French or Italian (§2), nor is it clear that the phonetic predictions it makes about the behaviour of supposedly underspecified segments in 'true voicing' languages stands up to empirical scrutiny (§4). We conclude, with Keating (1984), that any such attempts to provide ever more transparent grounding of phonological features will inevitably run into problems, for predicating a feature set on the basis of phonetic accuracy 'will require ever more additional features as new articulatory mechanisms are discovered' (289).

As we see it, the outstanding research challenge is to determine how such findings contribute to a more adequate understanding of the relation between phonological abstractions and phonetic detail, as well as where 'ungrounded' phonological features might come from (Mielke 2008). While it is by no means the only approach, we believe there is merit in a bottom-up perspective, and that renewed attention to the phonetic details of how

voicing is and can be implemented will ultimately improve our understanding of laryngeal typology.

Acknowledgments

This paper was originally presented at the 47th Poznań Linguistic Meeting as part of the thematic session '*Beyond VOT – searching for realism in laryngeal phonology*', convened by Geoff Schwartz and Eugeniusz Cyran. Thanks to the audience, as well as P. Honeybone, P. Iosad, P. Strycharczuk and an anonymous reviewer, for many insightful comments and suggestions.

References

Abramson, A. S. and D. H. Whalen. 2017. "Voice Onset Time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions". *Journal of Phonetics* 63. 75–86.

Archangeli, D. and D. Pulleyblank. 1994. *Grounded Phonology*. Cambridge, MA: The MIT Press.

Avery, P. and W. J. Idsardi. 2001. "Laryngeal dimensions, completion and enhancement". In: Hall, T. A. and U. Kleinhanz (eds.), *Distinctive feature theory*. Berlin: Walter de Gruyter. 41–70.

Beckman, J., P. Helgason, B. McMurray and C. Ringen. 2011. "Rate effects on Swedish VOT: Evidence for phonological overspecification". *Journal of Phonetics* 39(1). 39–49.

Beckman, J., M. Jessen and C. Ringen. 2013. "Empirical evidence for laryngeal features: Aspirating vs. true voice languages". *Journal of Linguistics* 49(02). 259–284.

Bell-Berti, F. 1975. "Control of pharyngeal cavity size for English voiced and voiceless stops". *The Journal of the Acoustical Society of America* 57(2). 456–461.

Bennett, W. G. and S. Rose. 2017. "Moro voicelessness dissimilation and binary [voice]". *Phonology* 34(3). 473–505.

Boersma, P. 1998. *Functional phonology*. PhD dissertation. University of Amsterdam.

Chao, Y. R. 1934. "The non-uniqueness of phonemic solutions of phonetic systems". *Bulletin of the Institute of History and Philology* 4. 363–397.

Cho, T. and P. Ladefoged. 1999. "Variations and universals in VOT: evidence from 18 languages". *Journal of Phonetics* 27. 207–229.

Chomsky, N. and M. Halle. 1968. *The sound pattern of English*. Cambridge, MA: The MIT Press.

Docherty, G. J. 1992. *The timing of voicing in British English obstruents*. Berlin: Walter de Gruyter.

Donegan, P. J. and D. Stampe. 1979. “The study of Natural Phonology”. In: Dinnsen, D. (ed.), *Current approaches to phonological theory*. Bloomington: Indiana University Press. 126–173.

Erickson, D., T. Baer and K. S. Harris. 1982. “The role of the strap muscles in pitch lowering”. *Haskins Laboratories Status Report on Speech Research* SR-70. 275–284.

Fagan, J. L. 1988. “Javanese intervocalic stop phonemes: The light/heavy distinction”. In: McGinn, R. (ed.), *Studies in Austronesian Linguistics*. Athens, OH: Ohio University Center for International Studies. 173–202.

Gao, J. and T. Arai. 2018. “F0 perturbation in a ‘pitch-accent’ language”. In: *TAL2018, Sixth International Symposium on Tonal Aspects of Languages*. 18 June 2018 ISCA. 56–60.

Grevisse, M. 1993. *Len bon usage: Grammaire française*. 13e ed. Goosse, A. (ed.), . Paris: Editions Duculot.

Halle, M. and K. N. Stevens. 1971. “A note on laryngeal features”. *MIT Quarterly Progress Report* 101. 198–212.

Hallé, P. and M. Adda-Decker. 2011. “Voice assimilation in French obstruents: A gradient or a categorical process?”. In: Goldsmith, J., E. Hume, and L. Wetzels (eds.), *Tones and features: Phonetic and phonological perspectives*. Berlin: Mouton de Gruyter. 149–175.

Hanson, H. M. 2009. “Effects of obstruent consonants on fundamental frequency at vowel onset in English”. *The Journal of the Acoustical Society of America* 125(1). 425–441.

Hayes, B., R. Kirchner, and D. Steriade (eds.), . 2004. *Phonetically based phonology*. Cambridge: Cambridge University Press.

Hombert, J.-M. and P. Ladefoged. 1977. “The effect of aspiration on the fundamental frequency of the following vowel”. *UCLA Working Papers in Phonetics* 36. 33–40.

Honeybone, P. 2002. *Germanic Obstruent Lenition: some mutual implications of theoretical and historical phonology*. PhD dissertation. Newcastle upon Tyne: University of Newcastle upon Tyne.

Honeybone, P. 2005. “Diachronic evidence in segmental phonology: the case of obstruent laryngeal specifications”. In: van Oostendorp, M. and J. van de Weijer (eds.), *The internal organization of phonological segments*. Berlin: Mouton de Gruyter. 319–354.

Hoole, P. and K. Honda. 2011. “Automaticity vs. feature-enhancement in the control of segmental F0”. In: Clements, N. and R. Ridouane (eds.), *Where do phonological features come from?: Cognitive, physical and developmental bases of distinctive speech categories*. Language Faculty and Beyond (LFAB): Internal and External Variation in Linguistics. Amsterdam: John Benjamins. 133–171.

House, A. S. and G. Fairbanks. 1953. “The influence of consonant environment upon the secondary acoustic characteristics of vowels”. *The Journal of the Acoustical Society of America* 25(1). 105–113.

Hualde, J. I. and M. Nadeu. 2011. “Lenition and phonemic overlap in Rome Italian”. *Phonetica* 68(4). 215–242.

Hualde, J. I., M. Simonet and M. Nadeu. 2011. “Consonant lenition and phonological recategorization”. *Laboratory Phonology* 2(2). .

Iosad, P. 2017. *A substance-free framework for phonology: an analysis of the Breton dialect of Bothoa*. Edinburgh: Edinburgh University Press.

Iverson, G. K. and J. C. Salmons. 1995. “Aspiration and laryngeal representation in Germanic”. *Phonology* 12(03). 369–396.

Iverson, G. K. and J. C. Salmons. 2006. “On the typology of final laryngeal neutralization: Evolutionary Phonology and laryngeal realism”. *Theoretical Linguistics* 32(2). 205–216.

Jacewicz, E., R. A. Fox and S. Lyle. 2009. “Variation in stop consonant voicing in two regional varieties of American English”. *Journal of the International Phonetic Association* 39(03). 313–334.

Jakobson, R. 1949. “On the identification of phonemic entities”. *Travaux du Cercle Linguistique de Copenhague* V. 205–213.

Jansen, W. 2004. *Laryngeal contrast and phonetic voicing: A laboratory phonology approach to English, Hungarian, and Dutch*. [University Library Groningen][Host].

Jessen, M. 2001. “Phonetic implementation of the distinctive auditory features [voice] and [tense] in stop consonants”. In: Hall, T. A. and U. Kleinhanz (eds.), *Distinctive feature theory*. Berlin: Walter de Gruyter. 237–294.

Jessen, M. 2004. “Instability in the production and perception of intervocalic closure voicing as a cue to bdg vs. ptk in German”. *Folia Linguistica* 38(1–2). 27–42.

Jessen, M. and C. Ringen. 2002. “Laryngeal features in German”. *Phonology* 19(2). 189–218.

Keating, P. A. 1980. *A phonetic study of a voicing contrast in Polish*. PhD dissertation. Providence, RI: Brown University.

Keating, P. A. 1984. “Phonetic and phonological representation of stop consonant voicing”. *Language* 60(2). 286–319.

Kingston, J. 2007. “Segmental influences on F0: Automatic or controlled?”. In: Gussenhoven, C. and T. Riad (eds.), *Tones and tunes, volume 2: Experimental studies in word and sentence prosody*. Phonology and Phonetics. Berlin: Mouton de Gruyter. 171–201.

Kingston, J. and R. L. Diehl. 1994. “Phonetic knowledge”. *Language* 70(3). 419–454.

Kingston, J., R. L. Diehl, C. J. Kirk and W. A. Castleman. 2008. “On the internal perceptual structure of distinctive features: The [voice] contrast”. *Journal of Phonetics* 36(1). 28–54.

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.

Löfqvist, A., T. Baer, N. S. McGarr and R. S. Story. 1989. “The cricothyroid muscle in voicing control”. *The Journal of the Acoustical Society of America* 85. 1314–1321.

Lombardi, L. 1991. *Laryngeal features and laryngeal neutralization*. PhD dissertation. University of Massachusetts, Amherst.

Lombardi, L. 1996. “Postlexical rules and the status of privative features”. *Phonology* 13(1). 1–38.

Mielke, J. 2008. *The emergence of distinctive features*. Oxford: Oxford University Press.

Ohala, J. J. 1983. “The origin of sound patterns in vocal tract constraints”. In: MacNeilage, P. F. (ed.), *The production of speech*. New York: Springer. 189–216.

Petrova, O., R. Plapp, C. Ringen and S. Szentgyörgyi. 2006. “Voice and aspiration: Evidence from Russian, Hungarian, German, Swedish, and Turkish”. *The Linguistic Review* 23(1). 1–35.

Ringen, C. and P. Helgason. 2004. “Distinctive [voice] does not imply regressive assimilation: evidence from Swedish”. *International Journal of English Studies* 4(2). 53–71.

Ringen, C. and V. Kulikov. 2012. “Voicing in Russian stops: cross-linguistic implications”. *Journal of Slavic Linguistics* 20(2). 269–286.

van Rooy, B. and D. Wissing. 2001. “Distinctive [voice] implies regressive voicing assimilation”. In: Hall, T. A. (ed.), *Distinctive feature theory*. Berlin: Mouton de Gruyter. 293–334.

Rubach, J. 1996. “Nonsyllabic analysis of voice assimilation in Polish”. *Linguistic Inquiry* 27(1). 69–110.

Schwartz, G. 2017. “Formalizing modulation and the emergence of phonological heads”. *Glossa: a journal of general linguistics* 2(1). .

Schwartz, G. and D. Arndt. 2018. “Laryngeal Realism vs. Modulation Theory – evidence from VOT discrimination in Polish”. *Language Sciences* 69. 98–112.

Sievers, E. 1876. *Grundzüge der Lautphysiologie zur Einführung in das Studium der Lautlehre der indogermanischen Sprachen*. Leipzig: Breitkopf & Härtel.

Simpson, A. P. 1999. “Fundamental problems in comparative phonetics and phonology: does UPSID help to solve them?”. In: *Proceedings of the 14th International Congress of Phonetic Sciences*. 1999 349–352.

Slis, I. H. 1986. “Assimilation of voice in Dutch as a function of stress, word boundaries, and sex of speaker and listener”. *Journal of Phonetics* 14. 311–326.

Snoeren, N., P. A. Hallé and J. Segui. 2006. “A voice for the voiceless: Production and perception of assimilated stops in French”. *Journal of Phonetics* 34(2). 241–268.

Solé, M.-J. 2018. “Articulatory adjustments in initial voiced stops in Spanish, French and English”. *Journal of Phonetics* 66. 217–241.

Sonderegger, M., M. McAuliffe and H.-Y. Bang. 2017. *Segmental influences on F0: cross-linguistic and interspeaker variability of phonetic precursors*.

Steriade, D. 1995. “Underspecification and markedness”. In: Goldsmith, J. (ed.), *The handbook of phonological theory*. Oxford: Basil Blackwell. 114–174.

Strycharczuk, P. 2012. *Phonetics-phonology interactions in pre-sonorant voicing*. PhD dissertation. University of Manchester.

Trubetzkoy, N. S. 1958. *Grundzüge der phonologie*. Göttingen: Vandenhoeck & Ruprecht.

Valdman, A., R. J. Salazar and M. A. Charbonneaux. 1964. *A drillbook of French pronunciation*. New York: Harper & Row.

Westbury, J. R. 1983. “Enlargement of the supraglottal cavity and its relation to stop consonant voicing”. *The Journal of the Acoustical Society of America* 73(4). 1322–1336.

Wetzels, W. L. and J. Mascaró. 2001. “The typology of voicing and devoicing”. *Language* 77(2). 207–244.

Whalen, D. H., A. S. Abramson, L. Lisker and M. Mody. 1990. “Gradient effects of fundamental frequency on stop consonant voicing judgments”. *Phonetica* 47. 36–49.

Whalen, D. H., L. Lisker, A. S. Abramson and M. Mody. 1993. “F0 gives voicing information even with unambiguous voice onset times”. *The Journal of the Acoustical Society of America* 93(4). 2152–2159.

Wissing, D. and J. Roux. 1995. “The interrelationship between VOT and voice assimilation”. In: *Proceedings of the 13th International Congress of the Phonetic Sciences*. 1995 Stockholm. 50–53.