Towards a quantized representation of phonological stop contrasts

Puggaard-Rode, R.; Botma, E.D.; Grijzenhout, J.; Breit, F.; Veer, M. van 't; Oostendorp, M. van

12

Towards a quantized representation of phonological stop contrasts

Rasmus Puggaard-Rode, Bert Botma, and Janet Grijzenhout

12.1

The cross-linguistic comparison of phonetic and phonological patterns has progressed tremendously in the past century and a half, due in part to the development of the International Phonetic Alphabet (IPA). Apart from a few symbols pertaining to prosody, the bulk of the IPA provides articulatory information about segments. The individual symbols (with accompanying diacritics) can essentially be thought of as bundles of articulatory properties: [b] is an abbreviation for a voiced bilabial plosive, [ɾ̥̃] is a devoiced nasalized alveolar tap, etc. The idea of representing segments in this way is also central to the influential model of generative phonology in Chomsky and Halle (1968), where phonological representations consist of linear sequences of segments, and the segments themselves of bundles of articulatory features.

Several approaches have since downplayed the role of the segment, or have even rejected segments altogether. Autosegmental Phonology (Goldsmith 1976; 1990) offers a non-linear approach in which features can behave independently from segments. In Articulatory Phonology (e.g. Browman and Goldstein 1986; 1989; 1992), where the basic units of phonology are articulatory gestures, segments emerge as the result of gestural constellations, but have no formal status in the grammar. The role of segments is explicitly rejected in strong versions of Exemplar Theory (e.g. Bybee 2001), where phonological patterns are assumed to emerge from detailed episodic representations of encountered words. A similar claim is made by Silverman (2017), who maintains that learners do not decompose words into segment-sized units, unless there is evidence to do so based on alternations.

Other approaches assume that segments have a more fine-grained structure than is traditionally assumed. One such approach is Q-theory (the Q is short for 'quantized'), which, in its original form, proposes a tripartite division of segments

Rasmus Puggaard-Rode, Bert Botma, and Janet Grijzenhout, *Towards a quantized representation of phonological stop contrasts*. In: *Primitives of Phonological Structure*. Edited by Florian Breit, Bert Botma, Marijn van 't Veer, and Marc van Oostendorp, Oxford University Press. © Rasmus Puggaard-Rode, Bert Botma, and Janet Grijzenhout (2023).
DOI: 10.1093/oso/9780198791126.003.0012

and assumes that these 'subsegments' (or 'q-positions') form the primary units of phonological structure (Shih and Inkelas 2014; 2019a; 2019b; Inkelas and Shih 2016; 2017).[1] This makes Q-theory well-equipped to represent segments that are not internally stable, such as affricates, prenasalized stops, and less commonly attested sounds like circumoralized nasals (e.g. [^{b}m^{b}]). Indeed, Q-theory has thus far been mostly applied to cross-linguistically uncommon segments with contour characteristics. This begs the question whether the model is also suitable for representing more commonly attested sounds. We address this question here by exploring how Q-theory—more specifically, a version of Q-theory called 'Q-CV'—can be applied to a segment type that is extremely common in languages of the world, viz. stops.

We believe that the quantized approach to stops that we develop in this chapter offers two general advantages: (1) contrasts involving differences in relative timing can be represented as such, and (2) contrasts involving differences in quantity can be represented in terms of the number of q-positions. Our approach makes a number of ancillary assumptions regarding segmental structure, the most important of which is an additional layer of structure called the 'CV-level'. This level is inspired by the representation of major class features in Dependency Phonology (Anderson and Ewen 1987). The CV-level constrains the order in which q-positions can occur and, by replacing the traditional major class features, reduces the number of distinctive features needed to represent segmental contrasts. We refer to the combined architecture of quantized segments with the CV-level as 'Q-CV'.

The main challenge facing a quantized approach to segmental structure is the risk of overgeneration, in terms both of the number and of the kind of q-positions that can be present in a segment. Although we suggest some ways in which overgeneration can be curtailed, a detailed examination of this issue is beyond the scope of this chapter. In what follows, our main aim is to explore how stop contrasts can be modelled in Q-CV, and to highlight some of the advantages that this approach offers.

The chapter is structured as follows. Section 12.2 outlines the main tenets of Q-theory as well as our proposed extension of the framework. Section 12.3 provides a brief typological overview of laryngeal contrasts in stops, and section 12.4 examines how these contrasts are represented in our Q-CV model. Section 12.5 shows that Q-CV offers an account of sonorant devoicing which is arguably superior to previous analyses of this phenomenon. Section 12.6 discusses some of the challenges facing Q-CV. Section 12.7 concludes.

[1] A precursor of this approach is Aperture Theory (Steriade 1993), where the internal structure of segments consists of closure and release positions.

12.2 The Q-CV model

12.2.1 Q-theory

Q-theory is a theory of segmental structure laid out in, among others, Shih and Inkelas (2014; 2019a; 2019b) and Inkelas and Shih (2016; 2017). The model offers a quantized view of phonology by splitting the traditional segment ('Q') into distinct subsegments ('q'). This creates the potential to represent a variety of contour segments. The original assumption was that segments have exactly three q-positions. More recent work allows for greater flexibility in the number of q-positions, in order to model such aspects as duration and articulatory strength (Garvin et al. 2018; 2020; Schwarz et al. 2019). We make the same assumption here.

Some examples of Q-theoretic representations are given in (1), from Inkelas and Shih (2016). The q-positions, represented here as subscripted IPA symbols, should be interpreted as abbreviations of feature bundles. The segments in (1a) are internally stable, since their q-positions contain identical features. The contour segments in (1b), on the other hand, involve intrasegmental feature changes for tone and nasality.

(1) a. [a] = [a_1 a_2 a_3]
[k] = [k_1 k_2 k_3]
b. [â] = [$á_1$ $á_2$ $à_3$]
[^{n}d] = [n_1 d_2 d_3]

We will assume that plain stops have the general structure in (1a), while certain types of laryngeally modified stops have contour-like properties, similar to the segments in (1b).

Inkelas and Shih (2016) note that a tripartite structure for stops is intuitively appealing, since this reflects the approach, hold, and release phase of stop articulations. However, since it is unclear whether the approach phase of stops has any phonological relevance, and since the release phase seems to be relevant only for certain types of stops, we prefer to assign a more abstract interpretation to q-positions. We view q-positions primarily as timing positions which mediate between subsegmental and segmental structure, much as skeletal positions mediate between segmental and syllabic structure (e.g. Clements and Keyser 1983).

12.2.2 The CV-level

We assume that q-positions are anchored by a level of subsegmental root nodes that we call the CV-level. This level consists of C- and V-components, as used in Dependency Phonology (Anderson and Ewen 1987) and related approaches

(see e.g. van der Hulst 2020). We follow these approaches in assuming that (combinations of) C and V specify major segment classes; they replace the traditional major class features [consonantal], [sonorant], and [continuant]. Anderson and Ewen (1987: 151) define C as a component of 'periodic energy reduction' and V as 'relatively periodic'. When occurring in isolation, C and V identify the two extremes of the sonority scale, viz. plain voiceless stops and vowels. This is also what we assume; see (2a,b). Combinations of C and V identify segments that are intermediate to these extremes. Unlike in Dependency Phonology, where C and V can enter into various types of head–dependency relations, our Q-CV model assumes just two combinations of C and V, given in (2b,c).

(2) a. C voiceless stop b. C_V voiceless fricative c. V_C sonorant d. V vowel

(2b) identifies voiceless fricatives. Their structure involves a combination of C and V in which C is more prominent than V, i.e. C_V. The resulting sound is stop-like but with egressive airflow, which is turbulent due to the prominent constriction. (2c) identifies sonorant consonants. These have a structure in which V is more prominent than C, i.e. V_C. Sonorant consonants are like vowels in having non-turbulent airflow, but the presence of C implies that they have a more prominent constriction (albeit one that allows for unobstructed airflow through the oral or nasal cavity).

The inclusion of the CV-level has a number of advantages. One is that it provides a built-in sonority hierarchy, and as such goes some way towards predicting core syllable structure. The CV-level also reduces the number of features. This includes not just the major class features, but also the laryngeal features [constricted glottis] and [spread glottis]. We assume that a placeless C-component is interpreted as involving blocked airflow at the glottis, i.e. as [ʔ], while a placeless C_V-component is interpreted as having turbulent airflow through an open glottis, i.e. as [h]. In this respect, our approach is similar to the interpretation of [ʔ] and [h] as 'degenerate segments' in models of Feature Geometry (e.g. McCarthy 1988).

The only laryngeal feature which remains in our approach is therefore [voice]. The difference between C- and V-dominance has implications for the compatibility of this feature. Vocal-cord vibration requires a sufficient transglottal pressure differential. This is difficult to maintain if little or no air can escape through the oral or nasal cavity (e.g. Ohala and Riordan 1979; Westbury 1983). We interpret this to mean that C-dominant q-positions (i.e. obstruents) are voiceless by default. Such positions can be voiced, e.g. by actively expanding the supraglottal cavity or by loosening the constriction, but this is the marked option and requires association of [voice]. V-dominant q-positions (i.e. sonorants) have an unobstructed airflow. This guarantees a sufficient transglottal pressure drop, so that their natural state is to be voiced. Observe in this respect that so-called 'voiceless sonorants' are produced with turbulent airflow, suggesting that they are fricatives (Ohala and

Ohala 1993) and therefore contain C_V-positions as part of their internal structure. We will see an example of this in our analysis of sonorant devoicing, in section 12.5.

12.3 Laryngeal contrasts in stops

In their simplest form, stops consist of an oral occlusion whose manner is cued by a period of silence and whose place is cued mainly by formant transitions of surrounding vowels and the release burst. Stops can be modified in a number of ways, providing contrasts based on, among other things, the duration of the closure, the presence of an aspirated or affricated release, the presence of voicing, and the relative timing of oral and laryngeal gestures. Ladefoged (1973: 80) lists the following types, which differ from each other in terms of phonation and/or airstream mechanism:

(3) *Stop types* (Ladefoged 1973)

laryngealized (creaky voiced)
voiced implosive
voiced plosive
voiceless lenis plosive
voiceless plosive
murmured (breathy voiced)
aspirated plosive
voiceless fortis plosive
voiceless ejective
glottal stop plus plosive
glottalic ingressive

Henton et al. (1992) add to this list a number of stops involving pre-aspiration, affrication, nasalization, nasal or lateral release, and differences in length. The phonetic picture becomes even more varied when we consider that the stops in (3) are not realized identically across different languages (see e.g. Kingston 1985; Ladefoged and Maddieson 1996; Cho and Ladefoged 1999).

Phonologically, the picture is much less varied, in that many of the stop types in (3) are not known to contrast within a single language. For example, there do not appear to be any languages in which laryngealized stops contrast with ejectives or with sequences of glottal stop plus plosive, suggesting that these form a single phonological category. More generally, Kehrein and Golston (2004) have shown that the relative order of oral and laryngeal articulations is never contrastive within the same syllabic position (e.g. within onsets), which substantially reduces the number of possible contrasts.

The representation of laryngeal contrasts in stops continues to be a matter of debate. Approaches differ, among other things, in the features which are assumed. For example, Halle and Stevens (1971) propose to distinguish the stop types in (3) with four binary features: [spread glottis], [constricted glottis], [stiff vocal cords], and [slack vocal cords]. Lombardi (1995) employs a set of three privative features: [voice], [glottalized], and [aspirated]. Another point of contention is the phonetic exponence of the features. For example, Kingston and Diehl (1994) argue that the feature [voice] is not necessarily cued by closure voicing, but rather by the relative timing of voicing onset, the relative levels of F_0 and F_1 after the burst, and the strength of the burst. Keating (1984) assigns a more abstract interpretation to [voice]. In her approach, [voice] distinguishes any two-way laryngeal contrast, the specific implementation taking place at the level of phonetic representation. The most radical position is in this respect a fully substance-free approach in which phonological contrasts are expressed in purely formal terms (for discussion of this position see Hall, Chapter 5 this volume). This type of approach treats all two-way laryngeal contrasts in a unified fashion, but it requires an elaborate spell-out procedure to translate the phonological structure into language-specific phonetic forms.

Alternatively, it has been proposed that the language-specific phonetic correlates of stop contrasts are reflected in the choice of phonological feature—an approach which goes by the name of 'laryngeal realism' (e.g. Honeybone 2005; Iverson and Salmons 1995). In this approach a distinction is made between voicing languages and aspiration languages. The former contrast plain stops with a series of stops with long-lead voicing, which are specified for a privative feature [voice]. The latter contrast plain stops with a series of stops with long-lag voicing, which are specified for [spread glottis].

The Q-CV analysis that we develop below is similar to this approach to the extent that it employs distinct representations for contrastively voiced and aspirated stops. However, rather than use features, we will explore the idea that aspirated stops have an additional q-position. More generally, we will argue that such a representation is appropriate for a range of stop types which are traditionally labelled 'fortis', a category which also includes glottalized stops and geminate stops.

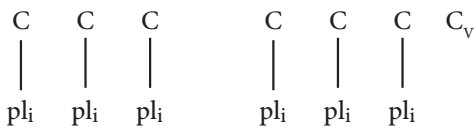
12.4 Stop contrasts in Q-CV

Having outlined the relevant background, we consider now how laryngeal contrasts in stops can be represented in our Q-CV approach. As we will see, most of these contrasts can be derived by (a) varying the number of q-positions, (b) combining different types of q-positions, and (c) varying the association of place features. This is shown in section 12.4.1. Section 12.4.2 considers representations involving the feature [voice].

12.4.1 Quantity-based contrasts in stops

We begin by considering languages with post-aspirated stops, such as English, German, and Danish. In pretonic position, the stop contrast in these languages is based primarily on a temporal distinction in the release phase of the stops. Plain stops are realized with short-lag voicing; we assume that these have the unmarked tripartite structure as in (4a). Post-aspirated stops are realized with long-lag voicing, i.e. with a longer voiceless release phase, suggesting that in these stops the release phase is phonologically relevant. We model this in terms of an additional q-position which takes the form of a placeless C_V node, as in (4b). In Example 4, 'pl' represents a place feature; co-indexing indicates that q-positions in a segment have the same place specification.[2]

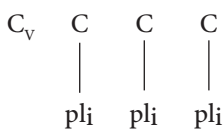
(4) a. plain stop b. postaspirated stop



The structure in (4b) is consistent with the observation that post-aspirated stops tend to have a longer overall duration than plain stops (Fischer-Jørgensen 1954; Pohl and Grijzenhout 2010). It also allows a natural interpretation of the relation between post-aspirated and affricated stops, which we consider below.

Some languages, e.g. Icelandic, Scottish Gaelic, and Saami, have pre-aspirated stops (Henton et al. 1992). These typically occur in post-tonic position, where they are allophones of post-aspirated stops. Reasons of pattern congruity suggest that pre-aspirated stops also have an extra C_V-position, this time preceding the closure positions, as in (5).

(5) preaspirated stop



The structure in (5) is appropriate for the pre-aspirated stops of Icelandic. Thráinsson (1978) analyses these as surface realizations of aspirated geminate stops. This is supported by alternations in inflectional forms, such as those in (6).

(6) mætti /mæt^h-t^{h}ɪ/ ['mæhtɪ] 'meet-INF'
feitt /feit^h-t^h/ ['feiht] 'fat-NEUT.SG'

[2] This is done for ease of exposition only. We do not take a position on whether structures of the kind in (4) contain individual place features for each q-position, or whether they share a single, multiply linked place feature. (The latter is clearly the more restrictive option.)

Thráinsson extends this analysis to pre-aspirated stops in underived contexts, e.g. *nátt* ['nauht] 'night', *epli* ['ɛhplɪ] 'apple'. However, our concern here is not so much with the distribution of pre-aspirated stops, but with how pre-aspiration interacts with a process of metrical lengthening, which lengthens vowels in stressed syllables. The effect of this process can be seen in the forms in (7a). The stressed vowels in these forms surface as long in open syllables, before single word-final consonants, and also before clusters which form possible onset sequences. Lengthening does not apply in the forms in (7b), where the clusters following the vowel do not form possible onset sequences. Notice that such sequences include pre-aspirated stops.

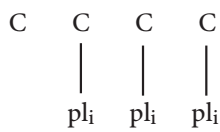
(7)	a.	tala	['tʰaːla]	'talk-INF'	b.	lamb	['lamp]	'lamb'
		vekja	['vɛːca]	'wake up-INF'		ólga	['oulka]	'foam-INF'
		ís	['iːs]	'ice'		nátt	['nauht]	'night'
		sötr	['søːtʰr]	'slurping'		epli	['ɛhplɪ]	'apple'

The fact that a pre-aspirated stop occurs with a preceding short vowel leads Thráinsson (1978) to conclude that the pre-aspirated portion forms a separate segment, i.e. [h]. Thus, in a form like ['ɛhplɪ], the oral portion of the stop occupies the onset of the unstressed syllable, while [h] is syllabified as the coda of the preceding stressed syllable, thereby blocking lengthening of the vowel.

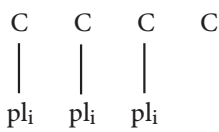
In our Q-CV analysis, there is no need to make the counterintuitive assumption that the two parts of the pre-aspirated stop surface as separate segments. Instead, we propose that the q-positions of the stop receive a heterosyllabic parse, in such a way that the C_V position is grouped with the stressed syllable, while the following C-positions form the onset of the following unstressed syllable. The issue of heterosyllabic parsing in Q-theory is explored in more detail in Garvin et al. (2018).

Turning now to glottalized stops, these also involve a temporal difference between oral and laryngeal gestures, which in our approach are represented at the CV-level. Pre-glottalized stops occur as allophones of post-aspirated stops in some varieties of English, such as RP and Tyneside English (Docherty et al. 1997). In these segments the glottal closure precedes the supralaryngeal closure, as in (8a). In post-glottalized stops, the glottal closure is retained after the oral occlusion has been released. Such stops are typically realized as ejectives (Henton et al. 1992); they have the structure in (8b):

(8) a. preglottalized stop



b. postglottalized stop (ejective)



There are several parallels between the patterns in Icelandic and English. For one, while preglottalization is found allophonically in e.g. RP and Tyneside English, other varieties have pre-aspirated stops in this context (for an overview, see Hejná and Jespersen 2019; Hejná et al. 2021). Another parallel is that pre-glottalized stops in English occur with preceding shortened vowels, a process called 'pre-fortis clipping'; compare *beat* [b̥iʔ͡t] vs *bead* [b̥iːd̥], and *bit* [b̥ɪʔ͡t] vs *bid* [b̥ɪːd̥]. This might suggest a trade-off, whereby the longer duration of the stop is compensated for by a shorter duration of the vowel. In Q-theory, this scenario would imply that syllables (or, more specifically, syllable rhymes) impose conditions on the number of permissible q-positions.

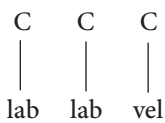
In varieties like Tyneside English, pre-glottalized stops are also found in foot-medial context, as in the forms in (9) (Carr and Montreuil 2013: 207).

(9) ['hɐʔ͡pi] 'happy'
['pʰɹɪʔ͡ti] 'pretty'
['mɪʔ͡ki] 'Mickey'

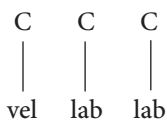
This suggests a further parallel between pre-aspirated and pre-glottalized stops; the latter can also be analysed as receiving a heterosyllabic parse, with the initial placeless C-position forming part of the stressed syllable.

Although we will have little to say about them, clicks can also be viewed as resulting from the relative timing of distinct releases, in this case of a velar closure and a closure at a more forward location in the mouth. For example, the bilabial click [ʘ] is produced by first releasing the bilabial closure while retaining the velar closure, as in (10a). This is crucially different from how labial-velar stops like [k͡p] are released; in the latter, the closures overlap to a much larger extent than in clicks, but the velar closure is released before the labial closure (see e.g. Ladefoged and Maddieson 1996). This suggests that labial-velar stops have the representation in (10b).

(10) a. bilabial click



b. labial-velar stop

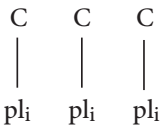


One interesting property of labial-velar stops is that /g͡b/ often occurs in the absence of /k͡p/ (Cahill 1999). This may suggest that labial-velars function as 'hypervoiced' stops; we discuss such stops below, in section 12.4.2.

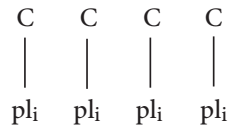
Some languages, e.g. Swiss German, employ a stop contrast based on the duration of the closure itself, i.e. between singleton and geminate stops (Kraehenmann 2001). The former are plain stops (11a); the latter are represented by scaling up the

number of q-positions associated with the articulatory closure target, as in (11b) (see also Garvin et al. 2018).[3]

(11) a. singleton stop



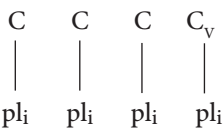
b. geminate stop



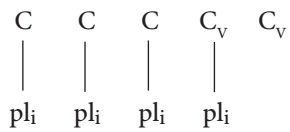
As is the case for pre-aspirated and pre-glottalized stops, the phonetic realization of geminate stops depends on the phonological context. Swiss German maintains a length contrast in the initial, medial, and final position of words, but the contrast in word-initial position can be realized (and perceived) only in phrasal context, i.e. when the first q-position can be syllabified in the coda position of a preceding syllable. Word-finally, the contrast is neutralized if the preceding nucleus is branching (Kraehenmann 2001: 114), which suggests, once again, that there is an upper limit to the number of q-positions in the syllable rhyme.

The representation of affricates is similar to that of post-aspirated stops. Both have a final C_V-position, but in affricates this is linked to a place feature, as in (12a). The place specification of the C_V-position is identical to that of the preceding C-positions, given that affricates involve homorganic friction. The structural similarity between affricates and post-aspirated stops accounts for why, historically, the former often develop from the latter (e.g. Hock 1991). In our approach, this development would involve copying the place specification of the C-positions onto the empty C_V, a process which may be motivated by a dispreference for empty q-positions (and by a dispreference for contour structures; see section 12.6). Some languages, e.g. Standard Mandarin Chinese, have aspirated affricates (Duanmu 2007). We assume that these have an additional, placeless C_V node, as in (12b).

(12) a. plain affricate



b. aspirated affricate



The 'ultra-long' structure of aspirated affricates is consistent with their greater duration, at least in Standard Mandarin Chinese (Liu and Jongman 2013; Puggaard 2020). Garvin et al. (2018) similarly suggest 'ultra-long' representations with five q-positions for the Hungarian geminate affricates [ts: tʃ:]. Such segments are clearly highly marked. We do not know of any stops, or indeed any other segment types, which require more than five q-positions in their representation.

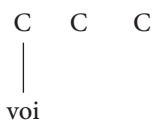
[3] Garvin et al. also consider the possibility of scaling down the number of q-positions, which they propose is appropriate for such segments as taps and flaps.

Summarizing, we have proposed that a number of stop types—aspirated, glottalized, geminate, and affricate—differ from plain voiceless stops in that the former have an additional q-position. In Q-CV, any two-way contrast of this type therefore involves a difference in quantity. Notice also that aspirated, glottalized, and geminate stops are traditionally classified as 'fortis'. This suggests, therefore, that fortis segments share the property of having more q-positions.

12.4.2 Voiced stops

The contrasts discussed in section 12.4.1 can all be modelled in structural terms, based on differences in the number and in the type of q-positions. However, voicing does not readily lend itself to such an approach. For this reason, we suggest that voicing contrasts are represented by means of the feature [voice]. We assume that stops with modal voicing have [voice] associated to the first q-position only, as in (13).

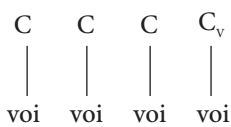
(13) (pre-)voiced stop



(13) reflects the idea that the initial q-position is the articulatory target for voicing, even if vocal cord vibration typically continues throughout the closure phase, though with decreasing intensity. This is because the complete closure which characterizes stops is not amenable to voicing; it causes the intra-oral air pressure to rise, which inhibits vocal cord vibration (e.g. Westbury and Keating 1986).

In addition to modally voiced stops, a number of other voiced stop types must be considered. The first are voiced aspirated (or breathy voiced) stops. These tend to have vocal-cord vibration throughout both the closure and the lengthy release, suggesting that they have the structure in (14), with [voice] attached to each of the four q-positions.[4]

(14) breathy voiced stop

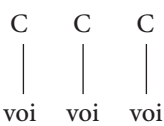


Our interpretation of breathy voiced stops is similar to that in feature-based accounts, where they are represented in terms of the combination [voice] and [spread glottis]. See e.g. Lombardi (1995).

[4] The q-positions in (14) are also specified for place, of course; here we focus on voicing.

Feature-based approaches use the combination of [voice] and [constricted glottis] to represent implosive stops, making these sounds the voiced counterparts of ejective stops. There are some languages, e.g. Hausa, where ejectives and implosives pattern together in root co-occurrence restrictions, suggesting that both must be specified as glottalized (Mackenzie 2009). However, such cases seem to be the exception rather than the norm. Cross-linguistic evidence shows that implosive stops are only rarely produced with glottal constriction (Ladefoged 1971). Phonetically, the only consistent property of implosives appears to be a lowered larynx. However, larynx lowering is also frequently observed in modally voiced stops as a mechanism to enlarge the supraglottal cavity, which counteracts the rise in intra-oral air pressure. Clements and Osu (2002) argue that the absence of air pressure build-up is a correlate of 'non-explosive' stops, which they represent as [–obstruent,–sonorant]. According to this account, implosives can be regarded as a type of non-explosive stop. Another type of non-explosive stops are the voiced lenis stops which occur in some languages of West Africa, e.g. Ikwere and Cama. Like implosives, these lack a strong release burst, and they pattern as non-obstruents in the phonology (Clements and Osu 2002; Botma and Smith 2006). We suggest that non-explosives are phonologically 'hypervoiced'; such stops have the feature [voice] attached to all q-positions, as in (15).

(15) hypervoiced stop



Associating [voice] to each of the q-positions will ensure that voicing is not attenuated during the stop's closure phase. This leads us to expect that the realization of hypervoiced stops will involve articulatory strategies which counteract the intra-oral rise in air pressure. This may involve different mechanisms for expanding the supra-glottal cavity, such as larynx lowering, as in implosives, but also tongue root advancement and velic lowering (Westbury 1983). We suggested above that labial-velars are another stop type which may be analysed as hypervoiced. Recall from (10) that in labial-velars the velar closure is released first. This may be viewed as a strategy to expand the supra-glottal cavity, which for velar stops is small; the rapidly increasing air pressure above the glottis makes such stops particularly ill-equipped for voicing. Releasing the velar closure into a bilabial closure results in a sudden and significant cavity expansion, thus making it easier to maintain vocal cord vibration throughout the closure phase.

Finally, notice that the stop types which we classify as hypervoiced do not display the markedness relation that we observe in stops with modal voicing. Languages with modally voiced stops also always have plain voiceless stops—an observation which suggests that modally voiced stops are marked. We do not find this markedness relation in hypervoiced stops. We saw already that voiced

/g͡b/ frequently occurs without its voiceless counterpart, while voiceless implosives and non-explosive lenis stops are extremely marginal. In this respect, hypervoiced stops pattern like sonorants, in that voicing in these segments is unmarked. In other respects, hypervoiced stops pattern as obstruents, in that they can be neither tone-bearing nor syllabic; this is reflected in their specification at the CV-level, which consists of C-positions.

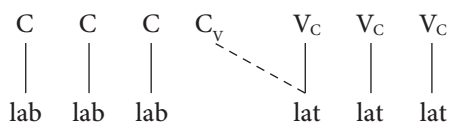
12.5 Sonorant devoicing

Our focus in this chapter is primarily typological, in that we explore how stop contrasts in the world's languages can be represented in Q-CV. However, our approach should also be capable of representing how segments interact. A detailed investigation of the dynamic aspects of Q-CV is beyond the scope of this chapter. We limit ourselves to showcasing how our approach handles one cross-linguistically common pattern of laryngeal assimilation—that of devoicing of sonorants by aspirated stops. Iverson and Salmons (1995) maintain that sonorant devoicing provides support for the feature [spread glottis]. It is therefore interesting to see how the process is analysed in Q-CV, which lacks this feature.

One language with sonorant devoicing is English. The process is typically described as affecting /j w ɹ l/ when these follow an aspirated stop in the same syllable (see e.g. Ladefoged and Johnson 2015: 113). We focus here on devoicing of /l/. A phonetic study by Tsuchida et al. (2000), which simultaneously recorded the sound and glottal activity of speakers, shows that the process is always partial, and that devoicing after voiceless fricatives is practically negligible (for the latter point, see also Fischer-Jørgensen 1954).

We propose that devoicing involves spreading of the laterality of the /l/, rather than (as is commonly assumed) spreading of the laryngeal features of the stop. This is shown in (16) for the initial consonant cluster in a word like *plan*. Notice in (16) that the final V_C-position of stop is empty, and so provides a target for spreading of [lateral]:

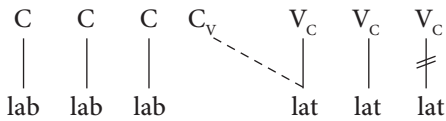
(16)



(16) appropriately reflects the partial nature of the devoicing process. In this respect, our analysis departs from earlier ones in which devoicing is treated as categorical (see e.g. Iverson and Salmons 1995; Ringen 1999).

Tsuchida et al. (2000) in fact find that roughly two-thirds of /l/ is devoiced after an aspirated stop. This suggests that devoicing also leads to the loss of one of the V_C-positions of /l/, as shown in (17).

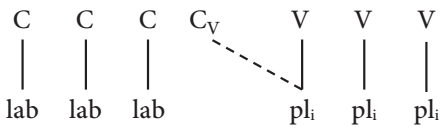
(17)



We suspect that the loss of the V_C-position may be tolerated because voiceless laterals have a quite distinct spectral profile (for this, see Maddieson and Emmorey 1984).

It is interesting to compare the realization of a cluster like /pl/ with a /p^hV/ sequence, in a word like *pan*. Here the place features of the vowel [æ] also spread to the final C_V-position of the stop. Throughout most of the aspiration phase, the tongue is already in position for [æ]; this is represented in (18) by the leftward spreading of the vowel's place features to the stop's release. However, the absence of modal voicing means that the overtones of [ḁ] are not very salient. We therefore propose that, unlike for [l̥], the voiced portion of the vowel is not significantly shortened.

(18)



Our interpretation of the English data ties in with the findings of Juul et al. (2019) for Danish. They observe that sonorant devoicing in Danish is more prominent after liquids than after nasals, in the order ʁ > l > n.[5] This is consistent with devoiced [ʁ̥] having salient acoustic cues, which [n̥] lacks. (Notice that English does not have devoicing of nasals, since historical stop–nasal clusters in words like *knife* have lost the stop.)

Finally, we suggest that the negligible devoicing of laterals after voiceless fricatives, in both English and Danish, does not result from spreading but from laryngeal mistiming. The supraglottal air pressure in voiceless fricatives is too high to maintain voicing. When the fricative is released, it will therefore take a couple of milliseconds to achieve a transglottal pressure differential that is sufficiently high to allow vocal cord vibration. The resulting voicing lag corresponds to the short positive voice onset time which accompanies voiceless unaspirated stops, and to the brief voiceless portion which follows the release of voiceless fricatives.[6] In view of this, we conclude that sonorant devoicing is phonologically relevant only in the context of preceding aspirated stops.

The main insight of the analysis that we have presented here is that devoicing of a sonorant by a preceding aspirated stop does not result from spreading

[5] The Danish rhotic is variably described as a voiced fricative or approximant, and as either velar, uvular, or supra-pharyngeal (Sobkowiak 2018). We transcribe it as /ʁ/.

[6] Fischer-Jørgensen (1954) in fact describes Danish /s/ as having a short period of 'aspiration'.

of a laryngeal feature from the stop to the sonorant, but from spreading of sonorant features into the voiceless stop release. If the sonorant features are compatible with voicelessness, as is the case for /l/, then the voiced portion of the sonorant will be correspondingly shorter. (We might view this as 'compensatory shortening', in order to retain the preferred number of three q-positions.) If, on the other hand, the sonorant features are incompatible with voicelessness, as is the case for vowels, the voiced portion of the sonorant is realized in full, and no q-positions are removed.

One question raised by our analysis is whether it can be extended to devoicing of a sonorant by a *following* aspirated stop. This phenomenon is found in some dialects of Icelandic, e.g. in forms like *mjólka* ['mjoul̥ka] 'milk' (see Thráinsson 1978). There is a parallel here with pre-aspirated stops, in which, as we have seen, the aspiration phase is also realized as part of the stressed syllable; see (7b). This may suggest that in Icelandic the distribution of the C_V-component of an aspirated stop is prosodically conditioned; it follows the stop's C-positions in the onset position of stressed syllables (where we find post-aspiration), but it precedes the stop's C-positions in the onset position of unstressed syllables (where we find pre-aspiration, with the aspiration phase syllabified in the coda of the preceding stressed syllable). We leave this kind of interaction between subsegmental and syllabic positions for further research.

12.6 Discussion

The preceding discussion shows that the Q-CV model is capable of representing a wide range of stop contrasts. Indeed, with the quantized representations that we use, the challenge is not so much to represent the segment types that are found in languages of the world, but rather to restrict the model from generating segment types that are not attested. A mechanism is therefore needed to limit the generative power of the model. (This is true for Q-CV, and also for Q-theory in general.) Below, we will offer some brief and necessarily speculative thoughts on how the range of possible contrasts can be curtailed. Since the model that we are proposing is new, a proper formalization of our ideas will have to be explored in future work.

We envisage that the inclusion of a CV-level makes it possible to derive a number of restrictions on the segment-internal order of q-positions. As was noted in section 12.2.2, the CV-level provides a basic sonority hierarchy. As such, the CV-level can be used to place restrictions on sequences of segments and, within segments, on sequences of q-positions. Many segments which are internally 'complex' are indeed structured in such a way that the V-dominant portion follows the C-dominant portion. This is the case for post-aspirated stops and affricates, for example, and also for consonants with secondary articulations. (We have

not considered these, but a reasonable assumption is that they contain a final V-position.) However, since there are also complex segments in which the V-dominant portion precedes the C-dominant portion, such as prenasalized stops, this issue requires further research.

As we noted in section 12.1, Q-theory has thus far been mainly concerned with the representation of typologically uncommon contour segments. This is not surprising, of course, since the model was developed for precisely this purpose. We suspect that contour segments are uncommon because intrasegmental q-positions prefer to be identical.[7] Segments which are internally stable are presumably favoured because they make it is easier for the listener to chunk the speech stream into meaningful units, and for the language-acquiring child to build underlying representations.

But if this is correct, then why do languages have contour segments at all? One reason is that such segments can emerge as the result of phonetic dispersion and contrast maximization. Consider the observation that there are languages, e.g. Maxakalí and Southern Barasano, in which voiced stops are prenasalized, a phenomenon that Wetzels and Nevins (2018) call 'nasal venting'. Prenasalization enhances the contrast between voiced and voiceless stops; prenasalized stops are more clearly distinct from voiceless stops, so that they will be correctly perceived as voiced at a higher rate than 'regular' voiced stops. This type of enhancement thus involves strengthening of a phonetic cue, which may eventually be promoted to a phonological contrast (see Wedel 2006 for a way to model this kind of development). Wetzels and Nevins note that nasal venting is especially common in languages which lack an underlying contrast between voiced stops and nasals. This is not surprising from the perspective of contrast maximization; in such languages, partial nasalization of stops will not lead to increased perceptual confusion with nasals.

Another example of enhancement leading to the loss of internal stability can be seen in the diachronic development of the laryngeal stop contrast in Germanic. Iverson and Salmons (2003) argue that, at some point in the history of Germanic, plain voiceless stops were enhanced with aspiration, thereby increasing the contrast with voiced stops. This enhancement was eventually followed by the 'phonological marginalization' of voicing (Iverson and Salmons 2003: 63), leading to present-day systems of the kind found in English. In Q-CV, enhancement with aspiration involves adding a C_V-position to the structure of plain stops, making them contour-like.

Contour segments can also emerge to guarantee the internal stability of neighbouring segments. An example of this is the phenomenon of 'nasal shielding', in which a nasal consonant is realized as an nasal–oral contour when adjacent to an oral vowel, e.g. /ma/ is realized as [m^{b}a]. Wetzels and Nevins observe that nasal

[7] However, notice that coronal affricates such as /ʦ ʧ/ are common. We will consider these shortly.

shielding is especially frequent in languages which contrast oral and nasal vowels (examples are Kaingang and Nɨkak). The function of shielding is then to protect oral vowels from becoming nasalized, at the cost of losing the nasal's internal stability.

The examples considered above all involve the loss of internal stability at the edges of segments. The edges are also the part of segments which are primarily affected by co-articulation. An example of this is the devoicing of sonorants by aspirated stops, which we examined in section 12.5. We saw there that this process leads to the creation of segmental contours; see e.g. (16). Some processes in fact involve both co-articulation and enhancement. An example is the cross-linguistically common development in which velar stops become coronal affricates, perhaps the most common type of contour segment (e.g. Old English *cinn* [kɪnː] > Modern English *chin* [ʧɪn]; see Campbell 2013: 35). The palato-alveolar place of coronal affricates is the result of assimilation to a front vowel, while their affrication likely serves to enhance the contrast with alveolar stops.

Overgeneration is also an issue when it comes to the number of q-positions in a segment. As was noted, the original assumption in Q-theory was that segments consist of exactly three q-positions, while later work assumes that the number of q-positions in a segment is flexible. For example, Garvin et al. (2018) analyse taps and flaps as having only two q-positions, and excrescent vowels as having no more than one. Garvin et al. further propose that geminate stops have four q-positions, while geminate affricates have five. We have assumed much the same flexibility here. In Q-CV, 'fortis' stops and affricates have four q-positions, while highly marked stop types such as aspirated affricates have five; see (12b). We do not know of any segment types which would require more than five q-positions—but this does not mean, of course, that five is the principled upper bound on the number of q-positions that segments can have. Given this, we might speculate that the unmarked number of q-positions for a segment is three, in line with the original Q-theory proposal. Languages can expand on this number, but they can do so only in single, incremental steps, so that a segment with, say, five q-positions is possible only if all options with four q-positions have been exhausted. The likely result is that a segment with eight q-positions, while theoretically possible, is never encountered in practice.

12.7 Conclusion

This chapter has outlined the main tenets of Q-CV, a new approach to segment-internal structure that is based on Q-theory. Our motivation to develop the Q-CV model was to explore whether the quantized representations of Q-theory can also be fruitfully applied to common segment types and common phonological contrasts, such as those found in stop systems. We have seen that our model makes

it possible to represent a wide range of stop types by varying the number of q-positions and the relative timing of oral and laryngeal articulations, without recourse to traditional major class features, and to the laryngeal features [spread glottis] and [constricted glottis]. A distinct advantage of Q-CV is that it can represent the phonetically gradient nature of sonorant devoicing. We have shown that a quantized analysis of this process is more in line with the phonetic facts than an account in which segments are viewed as 'monolithic' entities.

Since Q-CV is a new model, many of the issues raised by our approach are in need of further research. It seems to us that the greatest challenge facing Q-CV—and indeed any model which uses quantized representations—is that it is unrestrictive. Q-theory is capable of representing a wide range of segmental contrasts, but it does so at the cost of predicting many segment types that are unattested. The inclusion of a CV-level goes some way towards mitigating this problem, since it helps to restrict possible orderings of q-positions. Further restrictions are necessary to restrict the number of intrasegmental feature specifications, and also the number of the q-positions themselves. We leave the formalization of these restrictions for future work.