The Grammar of Coarticulation

Edward Flemming
Department of Linguistics & Philosophy, MIT

1. Introduction

It is well established that coarticulatory patterns are language-specific and therefore must be specified in the grammars of languages (e.g. Beddor, Harnsberger & Lindemann, 2002, Clumeck, 1976, Huffman, 1988, Keating & Cohn, 1988, Magen, 1984, Manuel, 1990, Manuel & Krakow, 1984, Oh, 2002), but there is less consensus on the nature of the grammar of coarticulation. Here we use evidence from the typology of coarticulatory patterns to argue for a model based on weighted constraints. Through two case studies we see that coarticulation shows typological variation that is parallel in many respects to phonological typology: there are common or universal cross-linguistic patterns, such as F_0 transitions between adjacent tones, that point to universal constraints, such as physiological limitations on rates of F_0 change, but language-specific variation in the details of coarticulation, e.g. variation in the timing of F_0 transitions, shows that languages differ in their responses to these constraints. I propose that this is due to interaction between conflicting constraints: constraints on rate of change interact with perceptually motivated constraints requiring the realization of phonetic targets. These constraints can conflict, with conflicts being resolved by constraint prioritization. The relative priority of constraints can differ from language to language resulting in language-specific patterns of coarticulation. This is essentially the same approach to the analysis of typology that has been successful in Optimality Theoretic analyses of phonology (Prince & Smolensky, 2004).

This account of coarticulation has implications for analyses that attempt to explain phonological generalizations in terms of phonologization of coarticulation. Phonologization is hypothesized to be a process by which phonetic phenomena are misinterpreted by listeners as phonological (Hyman, 1976, Ohala, 1992). It has been argued that phonologization can explain typological generalizations about phonology in terms of properties of the phonetic phenomena that provide the inputs to phonologization. For example phonological tone spreading processes almost always involves left-to-right spreading (Hyman & Schuh, 1974, Hyman, 2007). Hyman (2007) offers an account of this generalization based on the idea that phonological tone spreading arises from phonologization of tonal coarticulation. As we will see below, tones predominantly show carryover coarticulation, so according to Hyman's analysis, tone spreading inherits its directional bias from its phonetic source.

This account of directionality in tone spreading is incomplete in that it does not offer any explanation for the predominance of carryover tonal coarticulation. In fact most analyses based on phonologization do not account for the phonetic patterns that are hypothesized to constitute the inputs to phonologization. In a sense this paper attempts to fill this gap by providing explicit analyses of coarticulatory patterns, but the analyses that we arrive at are not consistent with basic assumptions of most theories of phonologization. First, the concept of phonologization depends on a clear cut distinction between phonetics and phonology to give substance to the idea that phonetic effects could be misinterpreted as phonological. For example Ohala (1981, 1992) argues that phonologization arises where language users misinterpret unintended properties of speech that 'are not under the active control of the speaker; they are instead added by the physical constraints of the speech production anatomy or even neuro-anatomy' (Ohala, 1981:179) as intentional, phonological processes. Ohala regards many aspects of coarticulation as falling into the category of unintended properties of speech - coarticulation is described as a ‘distortion’ of the speaker’s intentions (e.g. p.182). This conception of phonologization is appealing because it promises to explain generalizations about grammar in terms of extra-grammatical facts concerning anatomy (cf. Bermúdez-Otero, 2006). We will see that this picture of coarticulation is not plausible, rather coarticulation is part of grammar, and the structure of its grammar is very similar to the structure of the phonological component. So, if phonologization of coarticulation exists, it is a phenomenon in which processes that are derived in one component of grammar are misconstrued as processes that are derived in a similar component of grammar (perhaps even in the same component of grammar).

Furthermore, the conclusion that the grammar of coarticulation is built from constraints on articulatory effort and perceptual recoverability places limitations on the reductionist goals of phonologization-based approaches to phonological typology. It has been argued that phonetically-based constraints can be eliminated from synchronic phonological grammars because the diachronic process of phonologization can account for all influences of phonetic factors on phonological typology (e.g. Blevins, 2004, Barnes, 2006, Hale & Reiss, 2000). But phonologization-based analyses presuppose facts about the nature of coarticulation (and other phonetic processes) that are argued here to follow from universal phonetically-based grammatical constraints. So a process of phonologization cannot provide a basis for eliminating phonetically-based constraints from synchronic grammar.

The picture of coarticulation developed here is consistent with the hypothesis that phonetic and phonological processes are shaped by the same constraints (Flemming, 2001) – we will see that the constraints required for the analysis of coarticulatory patterns are similar to constraints that have been proposed in analyses of unambiguously phonological phenomena, suggesting the possibility that they are in fact the same constraints. We will also see evidence of mutual influence between patterns of coarticulation and phonological patterns, suggesting that these two aspects of grammar are closely integrated.

The analysis of coarticulation will be developed through analyses of two cases of cross-linguistic variation in coarticulatory patterns, one involving tonal coarticulation, and the other coarticulatory fronting of vowels by coronals.

2. Tonal coarticulation and spreading

Our first case study concerns tonal coarticulation and phonological tone spreading. The starting point is the observation, noted above, that both obey similar generalizations concerning directionality: tone spreading is almost always rightwards, while tonal coarticulation is predominantly carryover coarticulation. Hyman & Schuh (1974) and Hyman (2007) observe that rightward spreading of tones (1) is extremely common, whereas leftward spread of tones is very rare (2).

(1) L.H → L.LH   H.L → H.HL
(2) L.H → LH.H   H.L → HL.L

Rightward tone spread is exemplified by Yoruba (3) (Akinlabi & Liberman, 2001).

(3) /rárà/ (H.L) → rárâ (H.HL) ‘elegy’
/àlá/ (L.H) → àlǎ (L.LH) ‘dream’

Laniran & Clements (2003) argue that this rightward tone spreading is not simply a coarticulatory effect because it interacts with a process of vowel deletion. Contour tones surface even when the vowel that carried the conditioning tone is deleted in hiatus. In the example in (4), a rising tone surfaces on the final syllable although there is no preceding low tone on the surface.

(4) /ófɛ́ ìgbá/ → ófɛ́gbǎ 's/he wants a garden egg'

The rightward bias in tone spreading corresponds to a widespread asymmetry in the direction of tonal coarticulation. In general, rightward, or carryover, coarticulation between tones is much stronger than leftward, or anticipatory, coarticulation. Xu (1997) shows that in Mandarin Chinese the transition from one tone to the next generally does not begin until the onset of the second syllable, resulting in substantial coarticulatory variation at the beginning of the second tone, but very little variation at the end of a tone. This is illustrated in figure 1, which shows stylized F_0 contours for disyllabic words, based on Xu (1997). The right panel shows the realization of the level high tone (H) followed by three different tones on the second syllable: high (H), low (L) and falling (F). The vertical dashed line marks the syllable boundary.

The realization of the high tone on the first syllable is very consistent across all three contexts because movement towards the second tone does not begin until just before the onset of the second syllable. The left panel illustrates significant coarticulatory variation in the realization of the H tone depending on the preceding tone. Where the second-syllable H tone is preceded by another H tone, F_0 is high at the offset of the first syllable, and the second H is realized with a high, level F_0 contour, but where the preceding tone is low, the H tone is realized with F_0 rising through the duration of the second syllable – i.e. the bulk of the transition from low to high is realized on the second syllable, with little anticipation of the upcoming tone during the first syllable. As a result, there is substantial coarticulatory variation at the onset of a tone, depending on the F_0 level at the offset of the preceding tone, but little variation in the offset of a tone regardless of the following tone. Consequently a high tone is realized by rising F_0 following a low tone (L-H, left panel), and a low tone tone is realized by falling F_0 following a high tone (H-L, right panel), much as in Yoruba. However, in Mandarin, a simple tone spreading analysis is not appropriate since the phonetically falling tone on the second syllable of an H-L sequence contrasts with the underlying falling tone in the same context (H-F, right panel). The falling tone is realized as a fall with a later onset and lesser extent.

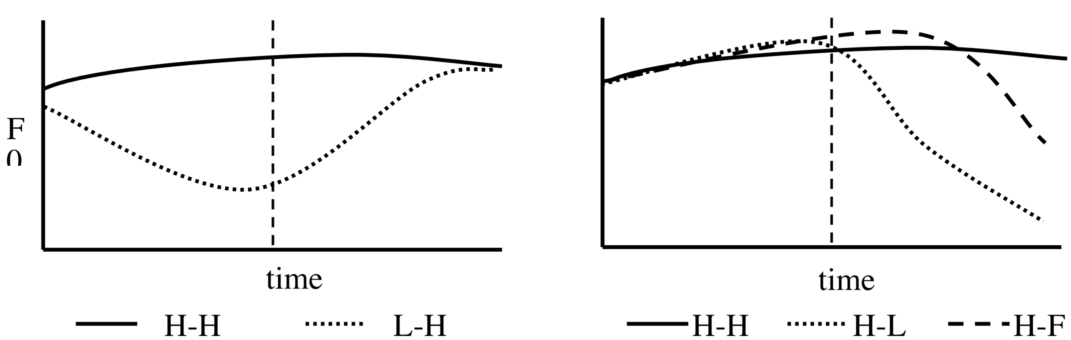


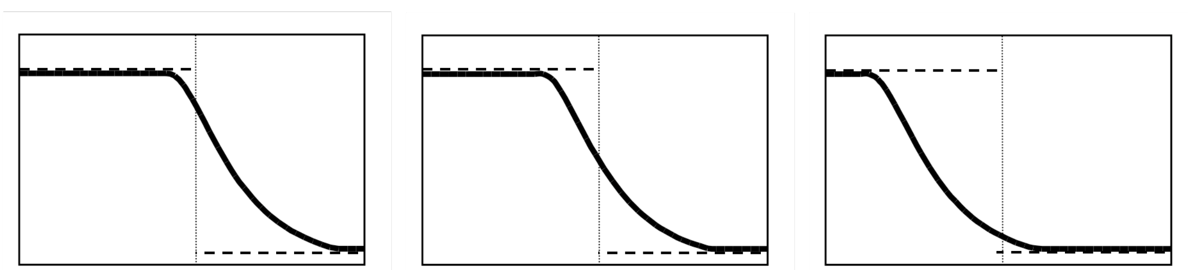
Fig. 1. Stylized F_0 contours for disyllabic mandarin tone sequences, after Xu (1997).

This general pattern of tonal coarticulation is cross-linguistically common (Xu & Liu, 2006), also being found in Cantonese (Li et al, 2004), Thai (Gandour et al, 1994) and Vietnamese (Brunelle, 2003).

This parallel might seem to invite an analysis in terms of phonologization: phonological tone spreading results from phonologization of tonal coarticulation, so tone spreading inherits the directional bias of tonal coarticulation (Hyman, 2007). However, this account leaves the question why there is a rightward bias in tonal coarticulation. We will see that an explanation solely in terms of physical and physiological constraints is not viable and will argue that a more plausible explanation is based on a violable grammatical constraint.

The predominance of carryover coarticulation cannot be a direct consequence of the nature of speech production since the pattern is not universal. A striking exception is Kinyarwanda (Myers, 2003). This language contrasts high and low tones on short vowels, and low, falling and rising tones on long vowels. In a sequence where a low tone is followed by a high tone (either a level high tone or the beginning of a falling tone), the rising transition occupies over half of the low-toned syllable, i.e. there is substantial anticipatory coarticulation. Myers does not quantify the carryover coarticulation after high tones, but it does not appear to be any greater than the anticipatory coarticulation.

This cross-linguistic variation in patterns of tonal coarticulation is not surprising because consideration of the articulatory factors involved in the production of tone sequences suggests that they are unlikely to favor a particular direction of tonal coarticulation. Physiological constraints do limit the rate of F_0 change, so there must be significant coarticulatory transitions between adjacent tones of different levels, but there is no physiological reason why transitions must be realized at the beginnings of syllables. That is, in a two syllables sequence with a high tone on the first syllable and a low tone on the second syllable, a substantial transition duration is physically necessary, but there is no physiological reason why the transition should begin at the onset of the second syllable (fig. 2a) (as in Mandarin Chinese) rather than finishing at the beginning of the second syllable (fig. 2c), or spanning the syllable boundary (fig. 2b).



(a) carryover (b) split (c) anticipatory

Fig. 2. Schematic illustration of possible timing patterns for an F_0 transition from high to low tone. The vertical dotted line marks the syllable boundary and the horizontal dashed lines indicate tone targets.

The predominance of carryover coarticulation is also not a general property of speech production. For example, in English the duration of anticipatory nasalization exceeds carryover nasalization in CVC words (Sefton & Beddor, 2005), and vowel-to-vowel coarticulation shows both anticipatory and carryover effects, with their relative magnitude depending on language and prosodic context (e.g. Beddor et al, 2002, Fowler, 1981).

The analysis of tonal coarticulation proposed here involves interacting constraints: articulatory constraints necessitate significant transition durations between tones at different F_0 levels, as outlined above, but the timing of the transition is determined by constraints on the realization of tone targets. Specifically, the bias to carryover coarticulation is accounted for by a constraint that places greater importance on realizing tone targets during syllable rhymes rather than onsets while the Kinywarwanda pattern is analyzed in terms of a bias to realize High tone targets at the expense of Low tone targets. The different patterns result from different weightings of these constraints.

The non-articulatory constraints pertain to the faithful realization of tonal targets, and so are ultimately concerned with perceptual recoverability of tones. We assume that in general tones specify F_0 targets that span a syllable, as in Xu (1997), so the transitions between tones result in deviations from targets as in figure 2, where the tone targets are indicated by the horizontal dashed lines. Accordingly, the timing of transitions can be influenced by constraints that specify which tone targets, and which parts of those targets, are most important. Specifically, we suggest that the rightward bias arises because it is more important to realize tone targets in the syllable rhyme than in the onset so the deviations from tone targets that arise during transitions between tones are located during syllable onsets rather than syllable rhymes as far as possible. This is achieved by beginning transitions between consecutive tones at the onset of the second syllable, as observed in Mandarin. A realization with exclusively anticipatory coarticulation (fig. 2c) is undesirable because it results in large deviations from the tone target of the first syllable during the rhyme of that syllable. Realizing the transition across the syllable boundary results in less deviation from targets during the syllable rhyme, but it still results in a worse violation of this preference than a transition that begins near the onset of the second syllable.

The syllable rhyme is important for the realization of tone contrasts because it is generally the part of the syllable with highest intensity periodicity, and therefore the place where tone is most perceptible (cf. Zhang, 2004). That is, tone contrasts are primarily realized by pitch, and pitch is better discriminated where the periodic sound contains more harmonics (Fastl & Weinberger, 1981) and where the signal level is higher (e.g. Freyman & Nelson, 1991), so tones should be easiest to identify when they are realized on sonorant sounds – i.e. relatively intense sounds with rich harmonic structure – and vowels provide the best platform for the realization of tone. The importance of the sonorous portion of the rhyme for the realization of tone contrasts has been demonstrated by Zhang (2004), who shows that the duration of the sonorous rhyme is a key factor in licensing contour tones. In Mandarin, the only permissible coda consonants are nasals, so the rhyme is always a good environment for the realization of tones.

Of course onset consonants can be sonorant also, but in some cases onset sonorants have lower intensity than their counterparts in coda, e.g. English liquids [l, ɹ] (Parker, 2008:73), so it is likely that the sonorous rhyme is still the preferred location for realization of tone. However, it is predicted that syllables with obstruent codas should pattern differently from syllables with sonorant codas. Since F_0 cannot be realized clearly on obstruents, the transition to a following tone may as well begin during an obstruent coda, resulting in earlier transitions with obstruent codas than sonorant codas. One piece of evidence that is consistent with this hypothesis is Roengpitya & Ohala's (2001) observation that there is more carryover tonal coarticulation after nasal final words than after stop final words in their study of Thai, but most studies do not directly examine the effect of segmental context on tonal coarticulation, or look at languages that only allow sonorant codas.

2.1 Implementation of the analysis of the carryover bias in tonal coarticulation

The operation of the proposed constraints is illustrated here by a simple but explicit implementation. The articulatory limitations that regulate the duration of F_0 transitions between tones are implemented by adopting the core of Fujisaki's model of F_0 production according to which commands to change F_0 are step functions – abrupt rises and falls in – but they are enacted by an articulatory system that is effectively a critically damped second order system with a fixed time constant (Fujisaki & Hirose, 1984), so the resulting F_0 trajectory involves a smooth transition, as illustrated in figure 2. Specifically, the shape of the transition between level tones is specified by the equation in (5), with the size of the transition being scaled by the magnitude A of the F_0 command[1]. Figure 2 illustrates the transition resulting from a step function from high to low F_o. The time taken to complete transitions is regulated by the time constant $1/\beta$, which is taken to be fixed at 0.05 s ($\beta = 20\ s^{-1}$), following Fujisaki et al (2005). The maximum value of 0.9 flattens out the end of the transition, allowing for the realization of an F_0 plateau at the target value of a tone.

(5) $\ln F_0(t) = A \min[1-(1+\beta t)e^{-\beta t}),0.9]$

The target for a tone specifies target F_0 through the duration of a syllable, as illustrated by the dashed lines in figure 2. A faithfulness constraint, IDENT-T(ONE), then requires that the actual F_o at a given time t_i, $F_0(t_i)$, must be equal to the target value specified for that time point, $T(t_i)$ (6). The cost of violating this constraint is a weighted sum of the squared deviations between actual and target F_o, summed over the duration of the target (7). For the purposes of this implementation, deviation is calculated at 10 ms intervals. The weighting function, $w_\sigma(t)$, allows implementation of the idea that faithfulness to tone targets is more important in the rhyme than in the onset by varying the cost of deviation from tone targets according to position in the syllable. In other words, $w_\sigma(t)$ is higher during the rhyme than during the onset. A series of tone targets is then realized by a series of F_0 commands to raise or lower F_o, and the timing of these commands is selected so that the resulting F_0 contour best satisfies the IDENT-T constraint.

(6) IDENT-T: $F_0(t_i) = T(t_i)$ for all t_i in target T.

(7) Cost of violation: $\sum_{t_i} w_\sigma(t_i)(T(t_i) - F_0(t_i))^2$

We apply this analysis to the derivation of carryover coarticulation in the realization of a H.L tone pattern on a sequence of two syllables, each 0.2 s in length. The targets for each tone are level throughout the syllable. The realization of this pattern involves setting the timing of the F_o lowering command associated with the L tone so as to minimize violation of IDENT-T. If we were to adopt a uniform weighting function where $w_\sigma(t) = 1$ at all times, then the optimal realization would be as in Fig. 2(b), with the transition from H to L beginning well before the end of the first syllable. If rhyme priority is imposed by setting $w_\sigma(t) = 0.01$ for the first half of each syllable and $w_\sigma(t) = 1$ for the second half ('rhyme') of each syllable, the optimal realization is as in Fig. 2(a), with the transition beginning just before the end of the first syllable.

[1] Fujisaki actually analyzes F_o contours as the sum of a phrase component and a series of tone components. We are only interested in local syllable to syllable contours here, so we only consider the tone component.

2.2 Priority for High tones

We hypothesize that rhyme priority in tone faithfulness is universal, accounting for the existence of a cross-linguistic tendency for carryover tonal coarticulation to dominate. However, this is not the only constraint governing tone realization, so we can account for the existence of contrary patterns like Kinyarwanda where substantial anticipatory coarticulation is observed. We hypothesize that the additional factor that accounts for the Kinyarwanda pattern is the possibility of differential faithfulness to High and Low tones: in Kinyarwanda faithfulness to High tone targets is weighted more heavily than faithfulness to Low tone targets. This is achieved by adding a second weighting term to the cost of violating IDENT-T, w_T, which can be assigned different values for different tone categories (8). In this case the weight for High tones, w_H, is greater than the weight for low tones, w_L.

(8) $$\sum_{t_i} w_T w_\sigma(t_i)(T(t_i) - F_0(t_i))^2$$

This priority of faithfulness to H over L tones is not unique to Kinyarwanda – in Chichewa, L tone targets are only realized to the extent that there is sufficient time to lower F_0 between H tone targets – i.e. realization of L tone targets is sacrificed in favor of realization of H tone targets (Myers, 1998)[2]. There is also evidence for greater faithfulness to High tones in phonological processes. For example, where tones compete to be realized on a single vowel, e.g. as a result of vowel deletion, High vowels are preserved in preference to Low or Mid tones in Ogori (Chumbow, 1982) and Yoruba (Pulleyblank, 2004).

So there is broad evidence for a bias in favor of preserving and realizing High tones over Low or Mid tones. But in Chichewa and Kinyarwanda there is an additional factor that could also contribute to the higher weight given to faithfulness to H tone targets: both languages permit only one H tone per morpheme, so it is possible that the higher weight given to faithfulness to H tone targets actually represents greater faithfulness to more informative tones. That is, the restricted distribution of H tones in these languages means that the identification of H tones contributes more to distinguishing morphemes than identifying L tones, which are in effect default tones. In Mandarin and the other SE Asian languages cited above, there is no such distributional asymmetry between tones and all tones show a consistent pattern of carryover coarticulation. If this hypothesis is correct, it indicates that patterns of phonological distribution affects patterns of coarticulation.

We are proposing here that the preference to realize H tones over L tones outweighs rhyme priority in Kinyarwanda, but both factors are hypothesized to be applicable cross-linguistically, so we should expect to find effects of both in the same language. A possible case is Edo, a language in which H tone is described as spreading onto a following L (H.L → H.HL), as in Yoruba, but L is not described as spreading onto a following H tone (Akinlabi & Liberman, 2001). In this case, the rightward spread of H could be analyzed as the result of the combination of rhyme priority together with H priority, while in the case of an L.H sequence, rhyme priority conflicts with H priority, resulting in less carryover coarticulation, so the result is not perceived as a contour tone.

[2] Myers actually argues that L tones are phonetically underspecified in Chichewa, but this proposal is implemented by positing a default low F_0 target in the absence of tonal specification, which is essentially the same as positing specification of L tones with low-weighted faithfulness to L targets.

2.3 Phonetics and phonology of tone spread

The systematic cross-linguistic variation in patterns of tone coarticulation shows that they are not simply an unintended side effect of speech production. Rather we have argued that tonal coarticulation is shaped by the interaction of articulatory constraints on rate of F_0 change with perceptually-motivated constraints requiring the faithful realization of tone targets. The weighting of faithfulness violations depends on position in the syllable and the phonological status of the tones (H vs. L, or more vs. less informative). So if rightward tone spread is the result of phonologization of carryover coarticulation, it is phonologization of a grammatically specified pattern, not misinterpretation of an extra-grammatical phenomenon.

An alternative account of the parallels between tone spreading and coarticulation is that both are shaped by the same constraints (cf. Flemming, 2001) – that is, the rightward bias in tone spreading arises because it is motivated by the same basic articulatory and faithfulness constraints that shape tonal coarticulation. The F_0 contour that results from tone spread in a Yoruba H.L sequence looks very similar to the F_0 contour that results from tonal coarticulation in a Mandarin H.L sequence, so it seems reasonable to analyze both in similar terms. Differences between the two languages can be analyzed in terms of additional constraints that interact with the tone constraints proposed so far. For example, Laniran & Clements's (2003) observation that rising tone can be preserved even where the conditioning L tone is deleted (4, above) can be analyzed as resulting from a constraint requiring that the realization of a word in phrasal contexts should correspond to its realization as a citation form. That is, the realization of /ófɛ́ ̀igbá/ as [ófɛ́gbǎ] results from correspondence to the citation form of /ìgbá/, which is [ìgbǎ], with a final rising tone. In other words, these data do not show that Yoruba tone spread is phonological as opposed to phonetic, instead they may show that the effects of tonal coarticulation are subject to Output-Output Correspondence constraints (cf. Steriade, 2000).

3. Vowel fronting by coronals

Our second case study involves coarticulatory fronting of back vowels conditioned by anterior coronal consonants. This effect arises because the tongue body usually moves forward in anterior coronals to facilitate formation of the constriction between tongue tip and the teeth or alveolar ridge (Manuel & Stevens, 1995, Öhman, 1966:167). Coarticulatory assimilation of adjacent vowels to this fronted tongue body position results in fronting of back vowels. As a result back vowels have higher F2 adjacent to anterior coronals than in labial and velar contexts (e.g. Stevens & House, 1963). The coarticulatory fronting effect of anterior coronals appears to be universal – it has been observed experimentally in at least Mandarin, English, French and German – but the magnitude of the effect is language-specific (Stevens & House, 1963, Hillenbrand, Clark & Nearey, 2001, Oh, 2002, Strange et al, 2007) . This language specificity is replicated here in an acoustic study of coarticulatory fronting of high back rounded /u/ vowels in four languages: English, French, German and Hindi[3]. We will see that this variation can be analyzed in terms of differences in the precise nature of the anterior coronals in the four languages and differences in the weighting of a small set of universal constraints.

Coarticulatory fronting was measured in each language by comparing the acoustic realization of /u/ in a neutral context that is expected to have little coarticulatory influence on /u/, and in a context between anterior coronal stops (alveolars in English, dentals in the other three languages). The neutral contexts were either isolated /u/ vowels (as in French *ou*) or adjacent to a laryngeal,

[3] Thanks to Hee-Sun Kim for assistance in collecting these data.

as in English *who* /hu/. The coronal contexts were represented by words such as /tut/ (e.g. English *toot*). The minimum value of F2 in /u/ in the neutral context was taken as an estimate of the F2 target for /u/. The coarticulatory fronting effect of the coronal stops was then quantified by measuring the minimum F2 during /u/ in the coronal context, and taking the difference between this value and the target F2 for /u/. We will refer to this difference as 'vowel undershoot' (cf. Lindblom, 1963).

There were four speakers for each language, two male and two female. Each produced six repetitions of each word in a carrier phrase. The carrier phrases were constructed so that the segmental contexts of the target words were similar across languages. In addition they recorded six repetitions of words containing /i/ in neutral and coronal contexts, e.g. /hi/ and /tit/.

The basic result is that /u/ undershoot differs significantly between languages (fig. 3): English > French, Hindi > German. Comparable differences between English, French and German are also reported by Strange et al (2007). To analyze these differences we need to look at a broader picture of the realization of these syllables, taking into account the precise nature of the coronal consonants in each language.

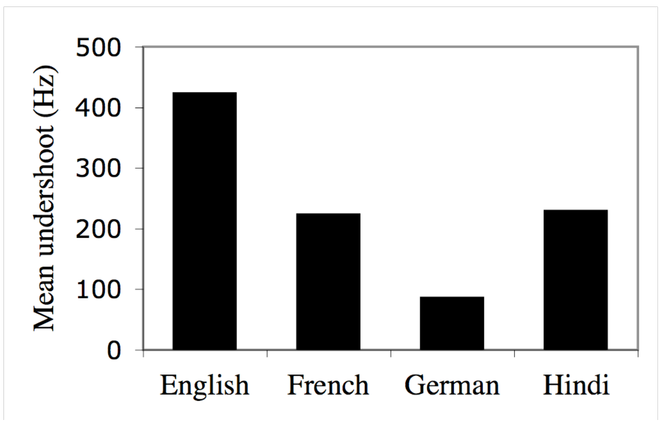


Fig. 3. Mean undershoot of /u/ in four languages.

The universal fronting effect of coronals on back vowels is analyzed as being due to a constraint disfavouring fast articulator movements. That is, in a /tut/ sequence the target tongue body position for the back vowel is far from the target positions for the adjacent consonants so reaching all three targets would require rapid articulator movements. Limiting speed of movement can result in undershooting the vowel target, but rate of articulator movement can also be reduced by undershooting consonant targets, so one source of cross-linguistic variation in coarticulatory fronting lies in differences in the resolution of the conflict between the requirements that consonant and vowel targets be realized and the dispreference for rapid articulator movements. This line of analysis is formalized following Flemming (2001). In addition, we will see that the consonant and vowel targets differ on a language-specific basis, so the severity of the constraint conflict presented by a /tut/ sequence differs somewhat from language to language.

Each segment in a CV or VC sequence has a target F2 value: the vowel target is T, while the consonant target (or locus) is L. The actual F2 frequencies adjacent to the consonant ($F2_C$), and at the vowel mid-point ($F2_V$) are selected so as to best satisfy the three constraints in (9). This situation is illustrated schematically in figure 4.

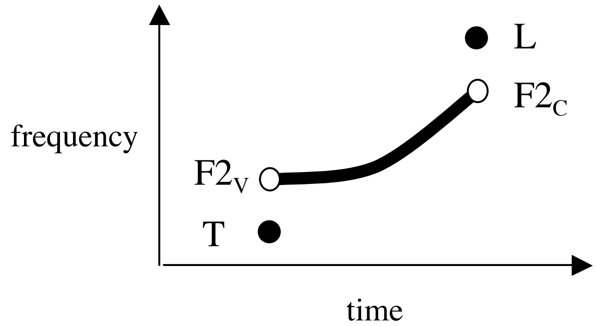


Fig. 4. Schematic representation of a VC transition, showing vowel target (T), consonant locus (L), an the actual F2 at vowel midpoint ($F2_V$) and adjacent to the consonant ($F2_C$).

The constraint *EFFORT expresses the dispreference for the effort involved in producing fast articulator movements (cf. Nelson, 1983, Perkell, 1997), and requires no movement through the vowel, i.e. F2 adjacent to the consonant should be equal to F2 at vowel midpoint. The cost of violating this constraint is equal to the square of the magnitude of the transition, multiplied by a positive constraint weight, w_E. The constraints IDENTC and IDENTV require that the vowel and consonant F2 targets be realized. The cost of violating these constraints is equal to the square of the deviation from the target, multiplied by the relevant constraint weight. The values of $F2_C$ and $F2_V$ are selected so as to minimize the summed violation costs imposed by these three constraints (10).

(9) Constraint		Cost of violation
*EFFORT	$F2_C = F2_V$	$w_E(F2_C\text{-}F2_V)^2$
IDENTV	$F2_V = T$	$w_V(T\text{-}F2_V)^2$
IDENTC	$F2_C = L$	$w_C(L\text{-}F2_C)^2$

(10) $w_V(T\text{-}F2_V)^2 + w_E(F2_{C2}\text{-}F2_V)^2 + w_C(L\text{-}F2_C)^2$

The differences in coarticulatory patterns between the four languages are analyzed in terms of differences in constraint weights, and differences in the targets, L and T. The segment targets require an independent analysis – for example they might be derived from selection of an optimal inventory of contrasting segments (cf. Lindblom, 1986, Flemming, 2004) – but for present purposes we will simply estimate them from the data.

As discussed above, the vowel target T is estimated from the realization of /u/ in a neutral context. The locus of the coda coronal stops in each language can be estimated by comparing $F2_C$ in /tut/ and /tit/ contexts. This only feasible for the final consonants because the initial consonants differed in VOT across languages (aspirated in English, unaspirated in other languages), so it was difficult to make comparable measurements of $F2$ adjacent to the onset consonant. The locus is taken to be the frequency at which $F2_C = F2_V$ (Klatt, 1987), on the assumption that this situation only arises where vowel and consonant targets are equal, because if the targets differed there would be some movement in F2. This frequency is estimated by fitting a straight line to $F2_C$ and $F2_V$ measurements from /tut/ and /tit/. The consonant locus is then taken to be where this regression line crosses the line $F2_C = F2_V$. The use of a straight line fit to identify the locus is justified by the consistent finding the relationship between $F2_C$ and $F2_V$ is linear (e.g. Krull, 1987, Sussman et al, 1993). *L* was calculated separately for male and female speakers in each language, then averaged t give a mean locus frequency for each language. Target F2 for /u/, *T*, and F2 locus for the coronals are listed in (11). It can be seen that both vary substantially across languages.

(11)

	T (Hz)	*L* (Hz)
English	1079	2192
French	786	2086
German	755	1793
Hindi	736	1690

Given *L* and *T*, we can calculate undershoot of both targets. The mean pattern for each language is summarized in figure 5. The top of each bar corresponds to *L* and the bottom of the bar corresponds to *T*, so the total heights of the bars indicate the distance between *L* and *T* in each language. The bars are then sub-divided to indicate the amount of C and V undershoot, and the magnitude of the F2 transition. It can be seen that the languages differ in all of these quantities.

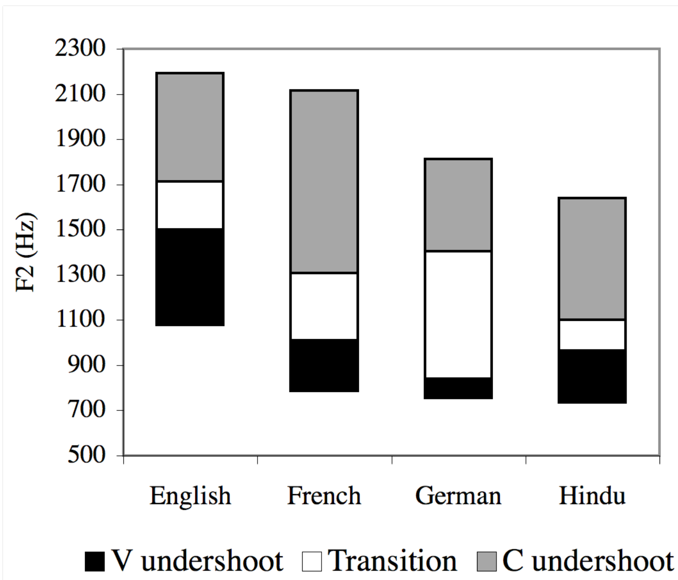


Fig. 5. Mean vowel undershoot, consonant undershoot and F2 transition magnitude in four languages (see text for details).

Since we do not have measurements of *F2* adjacent to C_1, we just model the vowel and C_2, using the three constraints on the realization $F2_V$ and $F2_C$ in (9). In other words, $F2_V$ and $F2_C$ are selected so as to minimize the cost function in (10). Since C_1 is not directly modeled, its coarticulatory effects on the vowel are effectively collapsed together with C_2. This would be valid if the loci of both consonants were the same, and the weights of their respective IDENTC constraints were the same, but it should still be a reasonable approximation as long as any difference in weights of IDENTC for C_1 and C_2 is comparable across the languages (for example IDENTC might be weighted higher for onsets).

The model in (10) implies that the interval between vowel target *T* and consonant locus *L* is divided into vowel undershoot, F2 transition and consonant undershoot, as illustrated in fig. 5, in the proportions w_Ew_C: w_Vw_C: w_Ew_V (Flemming, 2001:22), so the ratios of the constraint weights can be calculated directly from these measurements. To convert the ratios into absolute values for each weight, we impose the condition that the weights sum to 1 in each language. The resulting constraint weights are shown in (12).

(12)

	w_V	w_C	w_E
English	0.26	0.22	0.52
French	0.50	0.14	0.37
German	0.74	0.15	0.10
Hindi	0.32	0.14	0.54

According to this analysis, the amount of coarticulatory fronting of /u/ is given by the expression in (13). In other words, it is determined by two basic factors that differ across languages. The first is the distance between consonant locus and vowel target, *L-T*. The larger this distance, the greater the potential for undershoot, since it is not possible to approximate both targets without a rapid transition. This difference is greatest in French and smallest in Hindi, where the dental stop has a lower locus than in the other languages. So the relatively low vowel undershoot in Hindi can be attributed in part to the relatively small distance between Locus and Target in this language. The second factor concerns the constraint weights. As can be seen from (13), higher w_V results in less vowel undershoot, since if w_V is higher then w_C and w_V are lower, reducing the numerator and increasing the denominator. This is expected since w_V is the weight of constraint IDENTV and thus represents the importance of realizing the vowel target. French and German have higher values of w_V than English and Hindi. So English has the most vowel undershoot because it has a relatively large separation between Target and Locus and a relatively low w_V. Hindi has a comparable value of w_V, but this does not translate into large vowel undershoot because of the low consonant locus of the Hindi dental.

(13) $$\text{vowel undershoot} = \frac{w_E w_C}{w_E w_C + w_E w_V + w_C w_V}(L - T)$$

Although we are proposing to account for cross-linguistic differences in coarticulatory fronting in terms of differences in constraint weights, these weights are probably not free parameters of variation between languages, rather they seem to correlate to some extent with independent properties of the languages. The differences in w_V correlate with differences in the system of vowel contrasts in these languages: French and German contrast front [y] with back [u] while English and Hindi only have unrounded front vowels. That is, the higher values of w_V may reflect the importance of resisting fronting of /u/ where that would bring it too close to front rounded /y/. Between the two languages that have an /u/-/y/ contrast, German /y/ has a lower F2 than French /y/ and so is closer to back /u/ on this dimension. Correspondingly, w_V is higher in German than in French (see Strange et al (2007) for a similar observation).

The hypothesis that coarticulation can be restricted by the need to maintain the distinctiveness of contrasts has been proposed by Manuel (1990), but in the present case the preence of a contrast does not correlate directly with the absolute magnitude of coarticulatory fronting. French and Hindi show comparable fronting of /u/, but only French has the front rounding contrast. The analysis here suggests that this is because Hindi dental stops have a much lower locus than French dentals, which offsets the difference in w_V. So the relationship is between the system of contrasts and the constraint weight w_V.

It is not likely that the existence of an /u/-/y/ contrast necessarily causes higher w_V since contrasts can be neutralized due to coarticulatory effects. For example, Cantonese has contrasts between front and back rounded vowels, but these contrasts are neutralized between anterior coronals (Cheng, 1991), a pattern that can be attributed to the coarticulatory fronting effect of the coronals on back rounded vowels making the contrasts with front rounded vowels insufficiently distinct (Flemming, 2001). It is more plausible that patterns of coarticulation and the distribution of contrasts must be compatible: if a contrast between /u/ and /y/ is to be maintained between coronals, the fronting effect on /u/ must be limited, and conversely, if /u/ is subject to significant coarticulatory fronting between coronals, then there should not be a contrast with /y/ in that context. This type of coordination between phonetic realization and phonological distribution implies bi-directional interaction between the two components of grammar – the distribution of contrasts is sensitive to patterns of coarticulation, and patterns of coarticulation are sensitive to the system of contrasts.

Another striking difference in constraint weights across the four languages is the much lower value of w_E in German (0.1) compared to the other three languages (0.37-0.54). However, this may not be a real difference in constraint weighting but rather a reflection of the fact that the /u/ was longer in German (mean 147 ms compared to 106-116 ms for the other three languages). Greater duration means that it is possible to produce a larger F2 transition without moving the articulators as rapidly. In other words, since we have hypothesized that *EFFORT properly penalizes rapid articulator movement, a full formulation of the constraint should take vowel duration into account, and this might yield more comparable values of w_E for French and German.

4. Conclusions

Attempts to explain phonology in terms of phonetics leads us to the task of explaining the phonetic patterns themselves. The analyses of coarticulatory phenomena developed here motivate a model based on interacting constraints where the constraints resemble ones that are familiar from phonetically-based analyses of phonological patterns: effort constraints and perceptually weighted faithfulness constraints. So not only is coarticulation part of grammar, but the form and content of the grammar of coarticulation is similar to models of phonological grammar.

The analysis of coarticulation serves to clarify the nature of the influence of mechanical properties of speech physiology on linguistic sound patterns. These properties impose constraints on language, such as limits on rates of articulator movement, but these constraints only give rise to sound patterns in interaction with other types of linguistic constraints, such as perceptually-grounded constraints on the realization of targets. So mechanical properties of speech production anatomy cannot in themselves give rise to sound patterns that could serve as the basis for a process of phonologization, sound patterns only result from grammars.

Thanks to the participants at *La Coarticulation* workshop and an audience at the MIT Phonology Circle for useful comments and stimulating questions on this research.

References

Akinlabi, A. & Liberman, M. (2001), Tonal complexes and tonal alignment. M. Kim and U. Strauss (eds.), *Proceedings of the North East Linguistic Society,* 31, pp. 1-20. Amherst: GLSA.

Barnes, J. (2006), *Strength and Weakness at the Interface: Positional Neutralization in Phonetics and Phonology*, Berlin: Mouton de Gruyter.

Beddor, S.P., Harnsberger, J.D. & Lindemann, S. (2002), Language-specific patterns of vowel-to-vowel coarticulation: acoustic structures and their perceptual correlates, *Journal of Phonetics*, 30, pp. 591-627.

Bermúdez-Otero, R. (2006), Phonological change in Optimality Theory, K. Brown (ed.), *Encyclopedia of language and linguistics*, 2nd edn, vol. 9, pp. 497-505. Oxford: Elsevier.

Blevins, J. (2004), *Evolutionary Phonology: The Emergence of Sound Patterns*, Cambridge: Cambridge University Press.

Brunelle, M. (2003), Coarticulation effects in Northern Vietnamese tones. Ms., University of Ottawa.

Cheng, L. (1991), Feature geometry of vowels and co-occurrence restrictions in Cantonese, *WCCFL* 9, pp. 107-124.

Chumbow, B.S. (1982), Contraction and tone polarization in Ogori, *Journal of West African Languages*, XII, 1, pp. 89-103.

Clumeck, H. (1976), Patterns of soft palate movements in six languages. *Journal of Phonetics,* 4, pp. 337-51

Fastl H, & Weinberger, M (1981), Frequency discrimination for pure and complex tones. *Acustica*, 49, pp. 77–78.

Flemming, E. (2001), Scalar and categorical phenomena in a unified model of phonetics and phonology, *Phonology*, 18, pp. 7-44.

Flemming, E. (2004), Contrast and perceptual distinctiveness, B. Hayes, R. Kirchner & D. Steriade (eds.) *Phonetically-Based Phonology*, pp. 232—276. Cambridge: Cambridge University Press.

Fowler, C. A. (1981), Production and perception of coarticulation among stressed and unstressed vowels, *Journal of Speech and Hearing Research*, 46, pp. 127-139.

Freyman, R.L. & Nelson, D.A. (1991), Frequency discrimination as a function of signal frequency and level in normal-hearing and hearing-impaired listeners, *Journal of Speech and Hearing Research,* 34, pp. 1371-1386.

Fujisaki, H. & Hirose, K. (1984), Analysis of voice fundamental frequency contours for declarative sentences of Japanese, *Journal of the Acoustical Society of Japan*, 5, pp. 233-242.

Fujisaki, H., Wang, C., Ohno, S. & Gu, W. (2005), Analysis and synthesis of fundamental frequency contours of Standard Chinese using the command-response model, *Speech Communication* 47, pp. 59-70.

Gandour, J., Potisuk, S. & Dechongkit, S. (1994), Tonal coarticulation in Thai, *Journal of Phonetics*, 22, pp. 477-492.

Hale, M. & Reiss, C. (2000). Substance abuse and disfunctionalism: current trends in phonology, *Linguistic Inquiry*, 31, pp. 157-169.

Hillenbrand, J.M., Clark, M.J. & Nearey, T.M. (2000), Effects of consonant environment on vowel formant patterns, *Journal of the Acoustical Society of America*, 109, pp. 748-763.

Huffman, M. K. (1988), Timing of contextual nasalization in two languages, *UCLA Working Papers in Phonetics* 69, pp. 68-76.

Hyman, L.M. & Schuh, R.G. (1974), Universals of tone rules: evidence from West Africa, *Linguistic Inquiry*, 5, pp. 81-115.

Hyman, L.M. (1976), Phonologization, A. Juilland (ed.), *Linguistic Studies Offered to Joseph Greenberg*, pp. 407-418, Saratoga, California: Anma Libri.

Hyman, L.M. (2007), Universals of tone rules: 30 years later, C. Gussenhoven & T. Riad (eds.), *Tones and Tunes, vol. 1: Studies in Word and Sentence Prosody*. Berlin: Mouton de Gruyter.

Keating, P. & Cohn, A. (1988), Cross-language effects of vowels on consonant onsets," *Journal of the Acoustical Society of America*, 84, p. S84.

Klatt, D.H. (1987), Review of text-to-speech conversion for English, *Journal of the Acoustical Society of America*, 82, pp. 737-793.

Krull, D. (1987), Second formant locus patterns as a measure of consonant-vowel coarticulation. *PERILUS,* 5, pp. 43-61.

Laniran, Y.O. & Clements, G.N. (2003), Downstep and high raising: interacting factors in Yoruba tone production, *Journal of Phonetics*, 31, pp. 203-250.

Li, Y, Lee, T, & Qian, Y. (2004), Analysis and modeling of F_0 contours for Cantonese text-to-speech, *ACM Transactions on Asian Language Information Processing*, 3(3), pp. 169-180.

Lindblom, B. (1986), Phonetic universals in vowel systems. J.J. Ohala and J.J. Jaeger (eds.) *Experimental Phonology*. Orlando: Academic Press.

Magen, H. (1984), Vowel-to-vowel coarticulation in English and Japanese, *Journal of the Acoustical Society of America*, 75, p. S41.

Manuel, S. Y. & Krakow, R. A. (1984), Universal and language particular aspects of vowel-to-vowel coarticulation, *Haskins Status Report on Speech Research*, SR-77/78, pp. 69-78.

Manuel, S. Y. (1990), The role of contrast in limiting vowel-to-vowel coarticulation in different languages, *Journal of the Acoustical Society of America*, 88, pp. 1286-98.

Manuel, S.Y., & Stevens, K.N. (1995), Formant transitions: Teasing apart consonant and vowel contributions, *Proceedings of the International Conference of Phonetic Sciences 1995,* Vol. 4, pp. 436-439.

Myers, S. (1998), Surface underspecification of tone in Chichewa, *Phonology* 15, pp. 367–391.

Myers, S. (2003), F_0 timing in Kinyarwanda, *Phonetica*, 60, pp. 71-97.

Nelson, W.L. (1983), Physical principles for economies of skilled movement, *Biological Cybernetics*, 46, pp. 135-147.

Oh, E. (2002), Fronting of back vowels in coronal contexts: a cross-language study, *Studies in Phonetics, Phonology and Morphology*, 8, pp. 239-254.

Ohala, J.J. (1981), The listener as a source of sound change, C.S. Masek, R.A. Hendrick & M.F. Miller (eds.) *Papers from the Parasession on Language and Behavior*, pp. 178-203, Chicago: Chicago Linguistic Society.

Ohala, J.J. (1992). What's cognitive, what's not, in sound change, G. Kellerman & M.D. Morrissey (eds.), *Diachrony within Synchrony: Language History and Cognition*, pp. 309-355, Frankfurt: Peter Lang Verlag.

Öhman, S.E.G. (1966), Coarticulation in VCV utterances: Spectrographic measurements, *Journal of the Acoustical Society of America*, 39, pp. 151–168.

Parker, S. (2008), Sound level protrusions as physical correlates of sonority, *Journal of Phonetics*, 36, pp. 55-90.

Perkell, J.S. (1997), Articulatory processes, W.J. Hardcastle & J. Laver (eds.) *The Handbook of Phonetic Sciences*. pp. 333-370, Oxford: Blackwell.

Prince, A. & Smolensky, P. (2004). Optimality Theory: Constraint Interaction in Generative Grammar, Malden: Blackwell.
Pulleyblank, D. (2004), A note on tonal markedness in Yoruba, *Phonology*, 21, pp. 409-425.
Roengpitya, R. & Ohala, J.J. (2001), Tonal assimilation in Thai. *Journal of the Acoustical Society of America*, 110:5, p. 2655.
Sefton, S. & Beddor, P. S. (2005), Nasals and nasalization: the interplay between segmental duration and coarticulation, poster presented at the 149th Meeting of the Acoustical Society of America, Vancouver.
Steriade, D. (2000), Paradigm uniformity and the phonetics-phonology boundary, M. Broe & J. Pierrehumbert (eds.), *Papers in Laboratory Phonology 5: Acquisition and the Lexicon*. pp. 313-334, Cambridge: Cambridge University Press.
Stevens, K.N. & House, A.S. (1963), Perturbation of vowel articulation by consonantal context: an acoustical study, *Journal of Speech and Hearing Research*, 6, pp. 111-128.
Strange, W., Weber, A., Levy, E.S., Shafiro, V., Hisagi, M. & Nishi, K. (2007), Acoustic variability within and across German, French, and American English vowels: Phonetic context effects, *Journal of the Acoustical Society of America*, 122, pp. 1111-1129.
Sussman, H.M., Hoemeke, K. & Ahmed, F. (1993), A cross-linguistic investigation of locus equations as a source of relational invariance for stop place categorization, *Journal of the Acoustical Society of America*, 94, pp. 1256-1268.
Xu, Y. & Liu, F. (2006), Tonal alignment, syllable structure and coarticulation: Toward an integrated model, *Italian Journal of Linguistics*, 18, pp. 125-159.
Xu, Y. (1997), Contextual tonal variations in Mandarin, *Journal of Phonetics*, 25, pp. 61-83.
Zhang, J. (2004), The role of contrast-specific and language-specific phonetics in contour tone distribution, B. Hayes, R. Kirchner & D. Steriade (eds.) *Phonetically-Based Phonology*, pp. 157-190, Cambridge: Cambridge University Press.