Chapter 5

Implications for Phonology

This chapter investigates how the results of the previous experiments might be incorporated into phonological theory. I adopt Optimality Theory (Prince and Smolensky 2004[1993]) in order to facilitate comparison with the typology: from its inception, OT has had as one of its explicit goals the ability to generate all and only attested types of natural languages. Ideally, the 'factorial typology' of a given set of constraints (that is, the set of all language types that can be generated by some ranking of those constraints) should correspond exactly to the set of attested language types.

5.1 Perception: Results of Experiments 2 – 4

5.1.1 The P-Map

As discussed in chapter 4, the P-map (Steriade 2001a,b) is a theory of the relationship between the perceptibility of a given contrast and the ability of the members of that contrast to alternate in a phonological pattern. Differences in perceptibility are translated into universal rankings of OT faithfulness constraints via the P-map, a database that encodes the relative perceptibility of various pairs of contrasts in various environments. Faithfulness constraints are ranked according to the perceptual distance between the underlying forms that they apply to and the output forms that would result if they were violated, with constraints against more perceptible changes to the underlying form ranked above constraints against less perceptible changes.

In a strict interpretation, the P-map predicts that every difference in perceptibility found in Experiments 2 – 4 should map to a ranking of faithfulness constraints, and therefore to a gap in the typology (that is, the repairs corresponding to the higher-ranked constraints should be unattested). The following sections explore whether this prediction is confirmed.

5.1.2 Effects of Voicing and Manner

The main result of Experiment 2 was that the contrast between voiced and voiceless stops is easier to perceive than the contrast between voiced stops and voiced spirants. This fact translates into the following constraint ranking:

(1) IDENT[+voi]/V[___, –cont]V ≫
IDENT[–cont]/V[___, +voi]V

IDENT[+voi]/V[___, –cont]V requires faithfulness to an underlying [+voi] feature for intervocalic noncontinuants. Similarly, IDENT[–cont]/V[___, +voi]V requires faithfulness to an underlying [–cont] feature for voiced segments. I use these highly context-specific constraints in order to limit the domain of discussion to those contexts for which Experiments 2 – 4 provide evidence; see §5.1.4 for discussion of the need to distinguish faithfulness constraints for a given feature according to the underlying value of that feature.

This main result correctly predicts that it is possible to have an alternation by which voiced stops are realized as voiced spirants intervocalically, but not one by which they are realized as voiceless stops intervocalically.

(2) *VDV ≫ IDENT[–cont]: spirantization

/aba/	IDENT[+voi]	*VDV	IDENT[–cont]
[aba]		*!	
☞ [aβa]			*
[apa]	*!		

(3) IDENT[–cont] ≫ *VDV: no change

/aba/	IDENT[+voi]	IDENT[–cont]	*VDV
☞ [aba]			*
[aβa]		*!	
[apa]	*!		

(The relevant environments are assumed but omitted from the IDENT constraints in these tableaus. *VDV stands in for the constraint(s) driving some repair to intervocalic voiced stops; see §5.2 for a proposal based on the results of Experiment 1.) If *VDV outranks IDENT[–cont], as in (2), the result is a language with intervocalic spirantization of voiced stops. If IDENT[–cont] outranks *VDV, as in (3), the result is a language in which intervocalic voiced stops are realized faithfully. Since the ranking IDENT[+voi] ≫ IDENT[–cont] is fixed by the perceptual facts, no other patterns are possible: in particular, it is not possible to have intervocalic devoicing of voiced stops.

In contrast to Experiment 2, Experiment 4 found little or no evidence for a difference in perceptibility between the contrast between voiced and voiceless stops and between voiceless stops and spirants. If the two contrasts really are equally perceptible, then the P-map should not project a universal ranking between Ident[–voi]/V[___, –cont]V and Ident[–cont]/V[___, –voi]V.

(4) *VTV, Ident[–voi] ≫ Ident[–cont]: spirantization

/apa/	*VTV	Ident[–voi]	Ident[–cont]
[apa]	*!		
☞ [afa]			*
[aba]		*!	

(5) *VTV, Ident[–cont] ≫ Ident[–voi]: voicing

/apa/	*VTV	Ident[–cont]	Ident[–voi]
[apa]	*!		
[afa]		*!	
☞ [aba]			*

(6) Ident[–voi], Ident[–cont] ≫ *VTV: no change

/apa/	*Ident[–voi]	Ident[–cont]	*VTV
☞ [apa]			*
[afa]		*!	
[aba]	*!		

The surface pattern depends on which constraint is ranked lowest. If the lowest-ranked constraint is Ident[–cont], the result is spirantization; if it is Ident[–voi], the result is voicing; if it is *VTV, there is no change. Since the ranking between the two Ident constraints is not fixed, this analysis correctly predicts that both spirantization and voicing are possible.

However, recall that Experiment 4 did provide some weak evidence that spirantization might be a slightly more perceptible change for intervocalic voiceless stops than voicing, especially for labials. If there really is a difference between the two contrasts, then in order for the P-map to remain viable it must be able to avoid projecting a universal ranking such as Ident[–cont] ≫ Ident[–voi], which would incorrectly predict that intervocalic voiceless stops may voice but not spirantize. One solution might be to appeal to the fact that the difference between the Z $\sim$ D and T $\sim$ D contrasts found in Experiment 2 was much larger than the difference between S $\sim$ T and D $\sim$ T found in Experiment 4. Perhaps faithfulness constraints are only projected when the difference in perceptibility between two contrasts crosses some threshold; the difference between Z $\sim$ D and T $\sim$ D is sufficiently large, but the difference between S $\sim$ T and D $\sim$ T is not. Although plausible, this approach remains an ad-hoc solution until we find independent evidence for such a threshold. The question of how to prevent the P-map from overgenerating universal rankings among faithfulness constraints is taken up again in §5.1.4.

For the sake of discussion, I will assume that there is no difference in perceptibility between voicing and spirantization for underlying voiceless stops, but I acknowledge that the data is somewhat equivocal on this point. Interestingly, although both voicing and spirantization are attested alternations for intervocalic voiceless stops, voicing seems to be the more common option: as shown in table 2.1, 26 languages in Gurevich's (2004) database of lenition have voicing of voiceless stops, while only 17 have spirantization of voiceless stops. (The difference, however, is not significant; $p = .18$.) Perhaps differences in perceptibility that are too small to be projected as universal rankings among faithfulness constraints are still able to influence typological frequency. Indeed, it is possible that 'hard' typological patterns such as the absence of intervocalic devoicing are simply extreme cases of the type of 'soft' tendency seen here. I leave this very interesting line of investigation to future research.

5.1.3 Effect of Place of Articulation

Experiment 3 showed that the perceptibility of contrasts involving voiced and voiceless stops or voiced stops and spirants interacts with place of articulation. For the voicing contrast, the reaction time data presented in §4.2.2 suggests that the difference between [b] and [p] intervocalically is the most difficult to perceive, while the difference between [d] and [t] is the easiest. This fact translates into the following ranking:

(7) IDENT[–voi]/V[___, –cont, cor]V ≫

IDENT[–voi]/V[___, –cont, dor]V ≫

IDENT[–voi]/V[___, –cont, lab]V

If universal, this ranking predicts that intervocalic voicing can only apply to certain combinations of voiceless stops. If *VTV is undominated, then intervocalic voicing applies across the board:

(8) *VTV ≫
IDENT[–voi]/V[___, cor]V,
IDENT[–voi]/V[___, dor]V,
IDENT[–voi]/V[___, lab]V:
intervocalic voicing of labials, coronals, and dorsals

/apa/ /ata/ /aka/	*VTV	IDENT/[cor]	IDENT/[dor]	IDENT/[lab]
☞ [aba] [ada] [aga]		*	*	*
[aba] [ada] [aka]	*!	*		*
[aba] [ata] [aga]	*!		*	*
[aba] [ata] [aka]	*!*			*
[apa] [ada] [aga]	*!	*	*	
[apa] [ada] [aka]	*!*	*		
[apa] [ata] [aga]	*!*		*	
[apa] [ata] [aka]	*!**			

If *VTV outranks only the IDENT constraints for dorsals and labials, then only /p/ and /k/ are subject to intervocalic voicing:

(9) IDENT[–voi]/V[___, cor]V ≫
*VTV ≫
IDENT[–voi]/V[___, dor]V,
IDENT[–voi]/V[___, lab]V:
intervocalic voicing of labials and dorsals

/apa/ /ata/ /aka/	IDENT/[cor]	*VTV	IDENT/[dor]	IDENT/[lab]
[aba] [ada] [aga]	*!		*	*
[aba] [ada] [aka]	*!	*		*
☞ [aba] [ata] [aga]		*	*	*
[aba] [ata] [aka]		**!		*
[apa] [ada] [aga]	*!	*	*	
[apa] [ada] [aka]	*!	**		
[apa] [ata] [aga]		**!	*	
[apa] [ata] [aka]		**!*		

If *VTV outranks only the IDENT constraint for labials, then only /p/ is subject to intervocalic voicing:

(10) IDENT[–voi]/V[___, cor]V,
IDENT[–voi]/V[___, dor]V ≫
*VTV ≫
IDENT[–voi]/V[___, lab]V:
intervocalic voicing of labials

/apa/ /ata/ /aka/	IDENT/[cor]	IDENT/[dor]	*VTV	IDENT/[lab]
[aba] [ada] [aga]	*!	*		*
[aba] [ada] [aka]	*!		*	*
[aba] [ata] [aga]		*!	*	*
☞ [aba] [ata] [aka]			**	*
[apa] [ada] [aga]	*!	*	*	
[apa] [ada] [aka]	*!		**	
[apa] [ata] [aga]		*!	**	
[apa] [ata] [aka]			***!	

Finally, if *VTV ranks below all three IDENT constraints, then no intervocalic voicing occurs:

(11) IDENT[–voi]/V[___, cor]V,
IDENT[–voi]/V[___, dor]V ≫
IDENT[–voi]/V[___, lab]V ≫
*VTV:
no intervocalic voicing

/apa/ /ata/ /aka/	IDENT/[cor]	IDENT/[dor]	IDENT/[lab]	*VTV
[aba] [ada] [aga]	*!	*	*	
[aba] [ada] [aka]	*!		*	*
[aba] [ata] [aga]		*!	*	*
[aba] [ata] [aka]			*!	**
[apa] [ada] [aga]	*!	*		*
[apa] [ada] [aka]	*!			**
[apa] [ata] [aga]		*!		**
☞ [apa] [ata] [aka]				***

No other patterns are possible: intervocalic voicing of coronals implies voicing of dorsals, and voicing of dorsals implies voicing of labials. However, Gurevich's database contains three counterexamples to these generalizations: Périgourdin French voices [t] only, Apalai voices [k] only, and Lotha voices [t] and [k] only. Thus, for differences by place of articulation, the predictions of the P-map are not borne out.

Experiment 3 also showed that for the continuancy contrast, the difference between [g] and [ɣ] intervocalically is the most difficult to perceive. This fact translates into the following ranking:

(12) IDENT[–cont]/V[___, lab]V,
IDENT[–cont]/V[___, cor]V ≫
IDENT[–cont]/V[___, dor]V

As with intervocalic voicing, these differences by place make predictions about possible patterns of intervocalic spirantization. If *VDV is undominated, then intervocalic spirantization applies across the board:

(13) *VDV ≫
IDENT[–cont]/V[___, lab]V,
IDENT[–cont]/V[___, cor]V,
IDENT[–cont]/V[___, dor]V:
intervocalic spirantization of labials, coronals, and dorsals

/aba/ /ada/ /aga/	*VDV	IDENT/[lab]	IDENT/[cor]	IDENT/[dor]
☞ [aβa] [aða] [aɣa]		*	*	*
[aβa] [aða] [aga]	*!	*	*	
[aβa] [ada] [aɣa]	*!	*		*
[aβa] [ada] [aga]	*!*	*		
[aba] [aða] [aɣa]	*!		*	*
[aba] [aða] [aga]	*!*		*	
[aba] [ada] [aɣa]	*!*			*
[aba] [ada] [aga]	*!**			

If *VDV is dominated only by the IDENT constraint for labials, then spirantization applies only to /d/ and /g/:

(14) IDENT[–cont]/V[___, lab]V ≫
*VDV ≫
IDENT[–cont]/V[___, cor]V,
IDENT[–cont]/V[___, dor]V:
intervocalic spirantization of coronals and dorsals

/aba/ /ada/ /aga/	IDENT/[lab]	*VDV	IDENT/[cor]	IDENT/[dor]
[aβa] [aða] [aɣa]	*!		*	*
[aβa] [aða] [aga]	*!	*	*	
[aβa] [ada] [aɣa]	*!	*		*
[aβa] [ada] [aga]	*!	**		
☞ [aba] [aða] [aɣa]		*	*	*
[aba] [aða] [aga]		**!	*	
[aba] [ada] [aɣa]		**!		*
[aba] [ada] [aga]		**!*		

Analogously, if *VDV is dominated only by the IDENT constraint for coronals, then spirantization applies only to /b/ and /g/ (tableau not shown). If *VDV dominates only the IDENT constraint for dorsals, then spirantization applies only to /g/:

(15) IDENT[–cont]/V[___, lab]V,
IDENT[–cont]/V[___, cor]V ≫
*VDV ≫
IDENT[–cont]/V[___, dor]V:
intervocalic spirantization of dorsals

/aba/ /ada/ /aga/	IDENT/[lab]	IDENT/[cor]	*VDV	IDENT/[dor]
[aβa] [aða] [aɣa]	*!	*!		*
[aβa] [aða] [aga]	*!	*!	*	
[aβa] [ada] [aɣa]	*!		*	*
[aβa] [ada] [aga]	*!		**	
[aba] [aða] [aɣa]		*!	*	*
[aba] [aða] [aga]		*!	**	
☞ [aba] [ada] [aɣa]			**	*
[aba] [ada] [aga]			***!	

Finally, if *VDV is dominated by all three Ident constraints, no intervocalic spirantization occurs at all:

(16) Ident[–cont]/V[___, lab]V,
Ident[–cont]/V[___, cor]V,
Ident[–cont]/V[___, dor]V ≫
*VDV:
no intervocalic spirantization

	/aba/ /ada/ /aga/	Ident/[lab]	Ident/[cor]	Ident/[dor]	*VDV
	[aβa] [aða] [aɣa]	*!	*!	*	
	[aβa] [aða] [aga]	*!	*!		*
	[aβa] [ada] [aɣa]	*!		*	*
	[aβa] [ada] [aga]	*!			**
	[aba] [aða] [aɣa]		*!	*	*
	[aba] [aða] [aga]		*!		**
	[aba] [ada] [aɣa]			*!	**
☞	[aba] [ada] [aga]				***

No other patterns are possible. Because of the rankings fixed by the P-map, spirantization of either labials or coronals implies spirantization of dorsals. Again, though, this prediction does not match the typology. Gurevich's database contains 14 languages with spirantization of labials but not dorsals (Apatani, Assamese, Bashkir, Dahalo, Cardiff English, Kagate, Nepali, Nkore-Kiga, Ayt Seghrouchen Tamazight Berber, and Chitwan Tharu) or coronals but not dorsals (Purki Balti, Dahalo, Périgourdin French, and Purki). For intervocalic spirantization as well, the predictions of the P-map are not confirmed.

5.1.4 Which Faithfulness Constraints Are Projected?

As originally formulated, the P-map appears to be intended to project universal rankings between faithfulness constraints automatically for every difference in perceptibility between two contrasts:

> For *any* two P-map cells, x – y/___K_i and w – z/___K_j, if x – y/___K_i ➢ w – z/___K_j then *any* correspondence constraint referring to x – y/___K_i outranks *any* parallel constraint referring to w – z/___K_j.
> Steriade (2001a, 28, emphasis added)

('➢' denotes 'is more perceptible than'. Parallel correspondence constraints are faithfulness constraints regulating the relationship between the same two representations – IO, OO, BR, etc.) As shown in the previous section, if we project universal faithfulness constraints from *all* of the perceptual results of Experiments 2 – 4, the result is a mixed bag: we make some correct predictions (no intervocalic devoicing) and some incorrect predictions (differences by place of articulation). There are at least two ways to respond to this constellation of results:

1. Conclude that the theory of the P-map is ultimately incorrect. Those cases where the P-map appears to make the correct predictions (such as final devoicing) are due to chance.

2. Allow the database of perceptibility differences to project only rankings of faithfulness constraints that match typological facts.

The P-map *does* appear to make correct predictions in a range of cases – in addition to final devoicing, it has been applied to consonant cluster simplification and epenthesis of [ʔ] and [ə] (Steriade 2001a), the directionality of consonant assimilation (Steriade 2001b), the behavior of voicing in singleton and geminate stops (Kawahara 2006), and laryngeal co-occurrence restrictions (Gallagher 2009). Thus, it seems unwise to reject the P-map entirely until it is clear that a better solution cannot be found.

In addition, restricting the constraint rankings that are projected by the database of perceptibility differences seems desirable independent of these facts. It is unlikely that *every* pair of possible changes to a given input should be associated with a universal constraint ranking. For example, if we were to compare the perceptibility of final devoicing to that of the alternations involved in vowel harmony, it is entirely possible that we would find a difference – and yet the two processes may occur independently. But restricting how constraints are projected is a viable option only if we can find a principled way to do so: simply stipulating that rankings that do not match the typology are not projected only restates the problem. Determining what, if any, restrictions should be placed on how constraint rankings are projected is a project beyond the scope of this dissertation; it ultimately requires us to compare the perceptibility of each possible change for each possible input with every other possible change, and to match those results to the typological facts. However, I will offer some preliminary discussion here, based on the data at hand.

One way in which the universal rankings in (7) and (12) (those that yield problematic predictions for by-place differences in lenition) are different from the one in (1) (whose predictions are correct) is that the constraints within each of the former sets do not apply to the same underlying segments. That is, one constraint in each set applies to labial segments, another to coronal segments, and another to dorsal segments. The original purpose of the P-map, however, was to explain how languages choose among various possible repairs to the *same* marked configuration. The universal ranking in (1) accomplishes this goal: for a given intervocalic voiced stop, the ranking favors spirantization over voicing. The rankings in (7) and (12), however, do not: because the segments to which the constraints in each Ident family apply are disjoint, only one constraint in each family may apply to a given underlying intervocalic voiced stop. Thus, the rankings in (7) and (12) entail implicational relationships among repairs to *different* marked configurations, rather than determining which repair is chosen for a *single* marked configuration.

Therefore, to avoid the undesired rankings of (7) and (12), we might impose a restriction on the projection of rankings from the P-map: a universal ranking between two faithfulness constraints may only be projected if the two constraints regulate competing repairs for the same markedness constraint for some string of segments. Ident[+voi]/V[___]V and Ident[–cont]/V[___]V both regulate potential repairs to sequences that violate *VDV (devoicing and spirantization); therefore, the P-map can licitly project a universal ranking between the two. However, the situation is different for the place-specific Ident constraints. There is no single string of segments such that Ident[–cont]/V[___, lab]V and Ident[–cont]/V[___, cor]V represent two alternative pairs for the same markedness violation: the former applies only to intervocalic labials, while the latter applies only to intervocalic coronals.[1] Thus, the P-map cannot licitly project a universal ranking between the two (or between any of the place-specific constraints considered above), and the problematic rankings in (7) and (12) are ruled out.

[1] Leaving aside the treatment of segments with multiple places of articulation, e.g., labiovelars.

Clearly, more research is required to determine whether this approach is borne out when we examine more types of repairs. In the meantime, it has the benefit of making a principled distinction in this particular case between those predictions of the P-map that appear to be correct and those that we should avoid.

Note that this restriction on the projection of rankings by the P-map also requires that we distinguish among faithfulness constraints that regulate a given feature on the basis of the underlying value of that feature; it is for this reason that I distinguish between constraints such as Ident[+voi] and Ident[–voi]. Ident[+voi] regulates changes in voicing to underlyingly voiced stops (a possible repair for violations of *VDV), while Ident[–voi] regulates changes in voicing to underlyingly voiceless stops (a possible repair for violations of *VTV).

5.2 Production: Results of Experiment 1

5.2.1 Previous Models of Articulatory Effort

As noted in chapter 3, attempts to link articulatory effort and phonological patterns are ubiquitous. This section reviews a sample of research that incorporates articulatory effort into formal phonological models. Note that the authors whose work is reviewed here would probably not claim that their particular proposals are the only or even the best way to account for articulatory effort; this overview is meant only to be representative.

Pater, Hayes, and Kirchner

Pater (2004[1999]) documents a range of processes in several languages that have the effect of eliminating a sequence of a nasal followed by a voiceless stop. Working within Optimality Theory, he proposes a markedness constraint *NC̥ penalizing such sequences and notes that there is evidence that these sequences should be articulatorily disfavored; in particular, Ohala and Ohala (1991, 213[2]) argue that nasal leakage in the first part of the segment is compatible with voiced stops but not voiceless ones.

Hayes (2004[1999]) investigates the status of stop voicing in a number of environments. He derives scores of articulatory effort from Westbury and Keating's (1986) aerodynamic model of the vocal tract and proposes an algorithm for generating markedness constraints that attempts to balance the need for constraints to penalize difficult configurations with the need for constraints to be general and maintain formal symmetry. The result is a set of constraints, including constraints against post-nasal and -sonorant voiceless stops, that correspond well to cross-linguistic patterns of favored and disfavored structures.

Kirchner (2001b) develops a mass-spring model of the vocal tract that assigns a difficulty to a given configuration based on the force exerted throughout the relevant gestures in order to move the required articulators. He posits a Lazy constraint penalizing structures that require more force relative to those that require less force. Among other predictions, the model identifies geminates as more effortful than singletons; thus, Lazy favors degemination. (Kirchner also identifies intervocalic voiced stops as more effortful than their voiceless counterparts, using the same aerodynamic model as Hayes.)

[2] The page number cited by Pater, 273, appears to be a typographical error.

In all three of these accounts, the basic approach involves comparing two members of a phonological contrast (in a particular segmental context) and identifying one member of the contrast as the one that requires more articulatory effort. Con is assumed to contain at least one markedness constraint penalizing the more effortful configuration relative to the less effortful one. However, the results of Experiment 1 do not provide a good match to this type of account. Recall that the dominant pattern of that experiment was not one in which subjects favored one member of a constrast over another (voiced vs. voiceless stops or stops vs. spirants); rather, the common denominator across subjects and segment types was a contraction of the articulatory space such that *both* members of a given opposition moved toward each other. Thus, at least for the segments and environments examined in Experiment 1, articulatory effort minimization does *not* appear to favor one member of a contrast over another. This pattern is unlike that of the analyses of Pater, Hayes, and Kirchner.

Lindblom

Lindblom (1983) makes the case for a general principle of gestural economy in speech, arguing for a wide range of phonetic and phonological patterns that they can be viewed as involving articulatory effort reduction. He places a special emphasis on articulator movement and proposes the mass-spring model developed further by Kirchner (2001b) as a way of quantifying a positive correlation between distance moved by a given articulator and effort expended.

In examining vowels, Lindblom reviews experimental evidence that formant values are related to vowel duration: the shorter the vowel, the less likely it is that its formants will reach their targets. Lindblom interprets this result as reduction of the relevant gesture; when the vowel is short, there is less time for the articulator to reach its target, and the result is gestural 'undershoot'.

The results of Experiment 1 can be interpreted as consistent with Lindblom's proposal that effort reduction leads to gestural undershoot. It is a small step to say that the compression of the articulatory space that was observed across subjects and segment types is a case of undershoot: when subjects are intoxicated, they are less able to execute the full articulatory movement required for a given sound; as a result, extreme gestures of all types are reduced, and the differences between contrasting sounds become smaller. However, the crucial factor determining the degree of undershoot in Experiment 1 is *not* duration of the relevant segment;[3] rather, most important is whether the subject was intoxicated. Indeed, segments were simultaneously *longer* in the intoxicated condition *and* exhibited more undershoot. Thus, while Lindblom's articulatory undershoot seems to provide a good match to the pattern observed in intoxicated speech in Experiment 1, the particular factor he examines (segment duration) is clearly not the relevant one here.

Flemming

Flemming (2002), in his development of Dispersion Theory, proposes three basic and opposing forces that jointly determine the distribution of sounds in a language: the desires to maximize the number of contrasts in a phonological system, to maximize how auditorily distinct those contrasts are from one another, and to minimize articulatory effort. Because his focus is on the role of perception,

[3]Lindblom himself acknowledges (1983; 1990) that segment duration is not the sole or even the most important determinant of whether undershoot occurs, at least in vowels.

he does not formalize any systematic hypotheses as to which segments require more articulatory effort than others; however, he generally assumes (e.g., pg. 16) that the closer a segment is to the periphery of the auditory space, the more effort it requires. Like Lindblom's (1983) proposal, Flemming's is consistent with the results of Experiment 1: intoxicated subjects exhibited contraction of the articulatory space, avoiding extremes for all members of a given contrast.

Under Flemming's account, a language will have segments near the edge of the auditory space only if it must do so in order to maintain a contrast. For example, if a language has a backness contrast for vowels, then its inventory will contain both front vowels (like [i]) and back vowels (like [u]). However, if a language does not have a backness contrast, then its vowels will be neither front nor back, but central (like [ɨ]), possibly subject to variation depending on the segmental environment. This account correctly predicts that 'vertical' vowel inventories, as in Kabardian (Gordon and Applebaum 2010), will involve central vowels rather than front or back ones. Flemming accomplishes this by positing a family of MinDist constraints that require distinct segments to be separated by a certain amount in the auditory space. Since MinDist constraints only apply to phonologically contrastive forms, there is no reason for forms that do *not* contrast to occupy anything but the middle of the auditory space. Thus, the dispersion (or lack thereof) of a given articulatory dimension is related to whether that dimension is used to signal some phonological contrast.

To apply Dispersion Theory to the present results, we must say that intoxication promotes constraints banning 'extreme' articulatory gestures over some MinDist constraints. This account requires that any auditory dimension that exhibits contraction in the intoxicated condition must be used to signal some contrast. Otherwise, there would be no reason for segments to be dispersed along that dimension in the first place; at most, we would see random dispersion due to noisy implementation of the relevant gestures independent of any particular contrast. All of the measurements tested in Experiment 1 showed contraction; therefore, we predict that all of these dimensions are used to signal at least one phonological constrast.

Table 5.1 shows the relationship in Experiment 1 between compression of various dimensions in the intoxicated condition and the use of those dimensions to signal phonological contrast. A gray cell for a given subject and measurement means that that subject exhibited compression along that dimension. A short line segment in the cell denotes a significant difference for some contrast along the dimension in question. A horizontal line means that there is a significant difference between voiced and voiceless stops for that dimension. A vertical line indicates some significant difference by place for that dimension: a line on the left shows a contrast between labials and coronals, a line in the center a contrast between labials and dorsals, and a line on the right a contrast between coronals and dorsals.[4] For significant differences by voicing or place, Holm's correction for multiple comparisons was applied within, but not across, rows (separately for voicing and for place).

A naïve application of Dispersion Theory to Experiment 1 would expect to see two types of cells in table 5.1: cells with *both* compression *and* a significant cue to contrast, and cells with *neither* compression *nor* a significnat cue to contrast. Cells with compression but no contrast, or vice versa, are unexpected. However, as

[4]The position of these lines for place of articulation is meant to be iconic. In a mid-saggital section of the oral tract viewed from the left, the front of the mouth is on the left and the back is on the right.

Table 5.1: Articulatory compression and cues to phonological contrast. A gray cell denotes a regression line with a slope significantly different from 0 and 1. A horizontal line denotes a significant difference between voiced and voiceless stops. A vertical line denotes a significant difference by place: labial-coronal on the left, labial-dorsal in the center, and coronal-dorsal on the right.

Measure	**Subject**							
	00	**01**	**02**	**03**	**04**	**05**	**06**	**07**
Nasal-Stop Stimuli								
Dur. consonant (from V)	├─	│·│	││·	├┼─	│││	│·│	│││	│││
Dur. consonant (from N)	—	─┤	├─	├┼─	├─┤	├─┤	├─┤	├─┤
Dur. voicing (from V)	—	—	—	—	—	—	├┼─	—
Dur. voicing (from N)	│··		—	├┼─	—		—	
Prop. closure voiced (from V)	├─┤	—	—	├─	—	├─	—	├─
Prop. closure voiced (from N)	—	─┤	—	—	─┤	—	├─┤	—
Lenition Stimuli								
Dur. consonant	├─┤	├┼┤	├┼┤	├┼┤	├┼┤	├┼┤	├┼┤	├┼┤
Dur. voicing	—	—	—	├─	—	—	├─┤	—
Prop. closure voiced	├─┤	├─┤	├─┤	—	—	─┤	├─┤	├─┤
Dur. burst	│││	─┼┤	├┼┤	├┼─	├┼┤	─┼┤	─┼┤	├┼┤
Int. ratio, C/V1	·││			│··		─┤	··│	··│
Int. ratio, C/V2	·││		··│	—		··│		│·│
Min. slope, int. contour	├┼┤	├─┤	├─┤	—		├─┤	│·│	├─┤
Max. slope, int. contour	├┼┤	├─┤	├─┤	├┼─		├┼┤	│·│	├─┤

the table shows, both of these unexpected cases are in fact attested. (In addition, there are no cells with neither compression nor a significant contrast!) Let us consider the two types of unexpected cases in turn.

As discussed above, a dimension that shows compression but is not used to signal phonological contrast is unexpected because, if it is irrelevant to contrast, there is no reason for values along the dimension to be dispersed in the first place. There are seven cases of this type in table 5.1, for three auditory dimensions: the duration of voicing in nasal-stop stimuli as measured from the end of the nasal, the intensity of the consonant in lenition stimuli relative to the preceding vowel, and the intensity of the consonant in lenition stimuli relative to the following vowel. These examples may run counter to the most straightforward implications of Dispersion Theory, but we should not press them too far: after all, it is entirely possible that the differences in these cases were too small to be detected by Experiment 1, or that the dimensions in question are in fact used to signal some contrast other than voice or place. For the intensity of the consonant in lenition stimuli, it seems quite likely that we would find significant differences between sonorants and obstruents – a difference not seen here since all of the target consonants in Experiment 1 were obstruents. As for voicing duration in the nasal-stop stimuli, it is surprising that this dimension does not at least cue the voicing contrast! However, statistical power is a real concern for this measure; recall that a separate nasal could not be identified for all tokens.

A dimension that does not show compression but does signal contrast is unexpected because it would show that articulatory effort reduction does not apply to every dimension, or does not apply to every dimension to the same degree. Table 5.1 shows thirteen such cases, scattered across a range of dimensions and subjects.

Certainly, Dispersion Theory is compatible with the idea that the articulatory effort required to disperse sounds is different along different dimensions; or that intoxication affects various dimensions differently, interfering with dispersion for some dimensions but not for others. These cases, then, are like the previous group in being at least theoretically *consistent* with a Dispersion-Theoretic model, but not *explained* by it.

5.2.2 What Makes a Sound 'Effortful'?

As discussed in §3.1, and demonstrated by this brief survey, the literature contains many different views on what it is that makes a given segment (or sequence of segments) articulatorily 'effortful'. Many proposals appeal to biomechanical properties of the gestures involved: a sound is more effortful if the relevant articulator moves farther (Lindblom 1983) or faster (Uchanski 2005, 226), if it must be sustained or is 'tense' (Padgett 2009, 440), if it requires precise execution (Lavoie 2001, 166), or if is relatively 'unstable' when combined with the other gestures required for the sequence (Pouplier 2003). Kirchner's (2001a) mass-spring model attempts to combine several of these ideas; here, effort is defined as the total force exerted throughout a gesture; gestures that move faster or farther, that are longer, or that require several applications of force in order to maintain a precise position all require more total force than gestures that do not. In addition to these proposals, which focus primarily on the oral articulators, Westbury and Keating (1986), Ohala and Ohala (1991), and Hayes (2004[1999]) note the importance of aerodynamic considerations, especially for voicing.

I emphasize again that the results of Experiment 1 do not shed light on the question of *why* particular productions are effortful. Rather, the goal of Experiment 1 was to determine, if possible, *which* productions are more effortful than others; indeed, this is a necessary first step, since we cannot explain the relative difficulty of different productions if we do not know for certain what their relative difficulty is! Since alcohol affects the body in many ways (see §3.2), several of the above-mentioned hypotheses about articulatory difficulty could explain reduction induced by intoxication. For example, subjects might not be able to move their articulators as far or as fast, due to the overall depressive effect of alcohol; cognitive impairment might interfere with subjects' ability to maintain precise gestures or coordinate several gestures simultaneously.

5.3 Putting It All Together

This section offers a unified OT analysis of intervocalic spirantization and voicing that is faithful to the experimental results in the various ways discussed above. This proposal is not meant to be definitive, but rather illustrative. The most important result of this dissertation (especially Experiment 1) is that, when analyzing lenition, we can no longer do 'phonology as usual' – that is, Experiment 1 provides no support for constraints that simply favor the lenited form over the unlenited form. The following discussion illustrates what an analysis might have to look like in order to account for patterns like these *without* relying on constraints that simply say, "Lenite!"

5.3.1 Analysis of Intervocalic Lenition

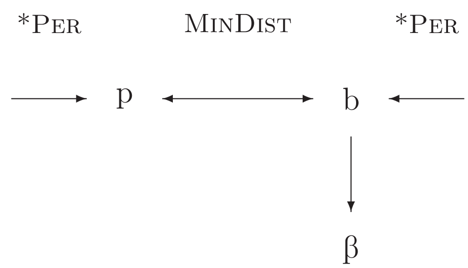
I argued in §5.2.1 that of extant proposals in the literature for modeling articulatory effort, Dispersion Theory (Flemming 2002; Padgett 2003; Ní Chiosáin and Padgett 2009; Padgett 2009)[5] is the closest match to the results of Experiment 1. The following analysis is therefore set within that framework. Dispersion Theory makes use of constraints that evaluate entire systems rather than individual forms; this is necessary because the theory explicitly controls systemic properties such as the phonetic distance between contrasting categories and the number of categories that contrast. Thus, the inputs and the candidates in a Dispersion-Theoretic analysis are idealized sets of forms rather than individuals.

Under my analysis, lenition occurs when effort reduction (represented by a high-ranking markedness constraint) encourages compression of some phonetic dimension (such as voicing), but compression is prevented by another markedness constraint that requires contrasting categories to be sufficiently dispersed along that dimension. To satisfy both constraints, one of two things must happen: either the contrast along that dimension collapses, or one of the two categories lenites by moving along a perpendicular dimension (such as continuancy). In other words, if articulatory effort reduction squeezes the voicing dimension for stops enough, either /p/ and /b/ will merge or one of the two will pop out of the dimension, becoming a spirant. Figure 5.1 sketches how the basic idea applies to spirantization; the constraints named in the figure are discussed in more detail below.

Note that this account requires the grammar to be able to refer to, and regu-

[5]See also Boersma and Hamann (2008) for a similar analysis which assumes that peripheral productions are articulatorily difficult.

Figure 5.1: Sketch of analysis of spirantization



late, subphonemic differences within categories. For example, if [b] lenites to [β] because effort reduction encourages its voicing value to decrease just slightly – bringing it a little too close to [p] but not making it a different category – then constraints must be able to refer to more fine-grained differences in voicing duration than are provided by categories that are known to contrast, such as [p] and [b]. My solution (which, again, I view as preliminary and not definitive) is to have inputs consist of discrete categories – as a first approximation, the categories defined by IPA symbols. Outputs, on the other hand, specify those categories for more fine-grained phonetic detail. Thus, while an input might contain the segment /b/, an output might contain a segment [b] with 50 ms of voicing.

I also make the simplifying assumption that the boundaries of the phonetic space, both its extreme edges and its internal category divisions, are determined by Gen. For example, I assume below that a stop can have between 0 and 60 ms of voicing, and that a voiceless stop is one with less than 20 ms of voicing and a voiced stop has more than 20 ms of voicing. (These values correspond roughly to the behavior of the majority of subjects in Experiment 1.) This is obviously a simplification: not only is there variation *between* subjects with respect to the amount of voicing they produce for different stops, but also *within* subjects the distributions of voicing values for voiced and voiceless stops overlap. One way to account for this inter- and intra-subject variation would be to put these boundaries under the control of constraints.

My analysis uses the following constraints:

- *Merge requires forms that contrast in the input to contrast in the output. Individual input and output forms are linked by a correspondence relation separate from the one that links individual segments. In the tableaux below, corresponding forms are indicated with capital letters when the correspondence is not obvious.

- MinDist constraints require that forms that contrast for a certain feature be separated on the relevant phonetic dimension by a certain distance. Usual practice in Dispersion Theory is to divide the phonetic space into a relatively small number of discrete units, usually corresponding to attested phonemic categories (Flemming 2002; Ní Chiosáin and Padgett 2009), and posit a corresponding family of universally ranked MinDist constraints. For example, suppose F1 is divided into five abstract parts corresponding to the heights of [i], [ɪ], [e], [ɛ], and [æ]. Then MinDist[F1][4] requires that any two vowels contrasting in height be separated by at least 4 units in the F1 space; this constraint would be satisfied by a contrast between [i] and [æ], but by no other pair of front vowels. MinDist[F1][3] is less strict, requiring a distance of only 3 units, and is satisfied by pairs such as [i] and [ɛ]. MinDist constraints are usually assumed to be universally ranked from least to most strict; however, since these constraints are in a strigency relationship, this is not strictly necessary (de Lacy 2002).

Rather than dividing the phonetic space into a small number of abstract categories, the MinDist constraints below refer directly to phonetic dimensions. Thus, MinDist[voi][60ms] requires that segments that contrast for [voice] have voicing durations that are at least 60 ms apart.

- On the basis of the results of Experiment 1, I posit a family of *Periphery constraints that penalize forms that are too close to the periphery of the phonetic space. Like the MinDist constraints, *Periphery constraints refer directly to dimensions; thus, *Periphery[voi][5ms] is violated by any stop with less than 5 ms or more than 55 ms of voicing.

- Clements (2003) argues that languages tend to prefer inventories in which a relatively large number of segments is described by a relatively small number of features; he terms this tendency "feature economy". Although his proposal is focused on phonological features, he presents some evidence (325-326) that a similar principle of "gestural economy", which encourages reuse of a small number of articulatory gestures, operates independently. In a related vein, Ussishkin and Wedel (2003) argue that speakers of a given language have a repertoire of gestural "molecules" which shape the degree to which loanwords are modified to fit the phonotactics of the borrowing language.

  In the spirit of gestural economy, I posit a family of Match constraints that require two instances of the same segment in different contexts to have the same value along a given phonetic dimension (within some margin of error). For example, Match[voi][3ms] assigns one violation to every candidate in which there are two instances of the same segment (e.g., [p]) with voicing durations that are different by more than 3 ms.

- The analysis below also uses the Ident constraints discussed in §5.1, which are evaluated in more or less the usual way. As in §5.1, I restrict these Ident constraints to intervocalic segments. I also employ Ident[–cont]/#___, a positional faithfulness constraint that applies only word-initially (Beckman 2004[1998]).

The tableau in (17) illustrates how this constraint set can derive intervocalic spirantization by imposing restrictions on the voicing dimension. The idealized inventories show the behavior of the full cross-classification of consonants for voicing (voiced vs. voiceless), continuancy (stop vs. spirant), and position (initial vs. intervocalic). I assume that by richness of the base, the input contains all eight possibilities.

The winning candidate, (d), has spirantization of voiced stops intervocalically. The fully faithful candidate (a) loses because it has four violations of *Periphery[voi][5ms]: while (a) has four stops at the edges of the voicing space, (d) has only three, since [b] does not appear intervocalically on the surface. *Periphery must outrank both Ident[–cont] and *Merge, or else (d) would lose, either because it changes the continuancy value of underlying intervocalic /b/ or because it eliminates the contrast between /aba/ and /aβa/.

(17)	/pa/$_A$ /fa/$_C$ /apa/$_E$ /afa/$_G$	/ba/$_B$ /βa/$_D$ /aba/$_F$ /aβa/$_H$	MinDist[voi][60ms]	Ident[−cont]/#___	Match[voi][3ms]	*Periphery[voi][5ms]	Ident[−voi]	Ident[+voi]	Ident[−cont]	*Merge
a.	[p_0a] [fa] [ap_0a] [afa]	[b_{60}a] [βa] [ab_{60}a] [aβa]				****!				
b.	[p_5a] [fa] [ap_5a] [afa]	[b_{55}a] [βa] [ab_{55}a] [aβa]	**!							
c.	[p_0a] [fa] [ap_0a] [afa]	[βa]$_{B,D}$ [aβa]$_{F,H}$		*!		**			*	**
☞d.	[p_0a] [fa] [ap_0a] [afa]	[b_{60}a] [βa] [aβa]$_{F,H}$				***			*	*
e.	[p_0a] [fa] [ap_5a] [afa]	[b_{60}a] [βa] [aβa]$_{F,H}$			*!	**			*	*
f.	[p_0a] [fa] [ap_0a]$_{E,F}$ [afa]	[b_{60}a] [βa] [aβa]				***		*!		*
g.	[p_0a] [fa] [afa]	[b_{60}a] [βa] [ab_0a]$_{E,F}$ [aβa]				***	*!			*

Candidate (b) avoids violating *Periphery at all by moving all four stops slightly towards the center of the voicing space. This candidate thus incurs two violations of MinDist[voi][60ms]: both word-initially and intervocalically, the stops distinguished by voicing are too close together in the phonetic space. MinDist must therefore outrank *Periphery (which the winner (d) does violate) in addition to Ident[–cont] and *Merge. The failure of candidates (a) and (b) demonstrates the squeezing of the voicing dimension in this analysis: *Periphery requires more compression of the voicing dimension than MinDist will allow.

Candidate (c) avoids two of the four violations of *Periphery incurred by the fully faithful candidate by spirantizing all of its voiced stops; thus, only [p] remains to violate *Periphery. However, this candidate is eliminated by a fatal violation of Ident[–cont]/#___: positional faithfulness at the beginning of the word is more important than avoiding a word-initial [b] at the periphery of the voicing space. Ident[–cont]/#___ must outrank *Periphery. (Candidate (c) also does worse than (d) on *Merge; however, we know from candidates (a) and (b) that this constraint ranks *below* *Periphery. Thus, the fatal violation of (c) must come from Ident[–cont]/#___.)

An alternative way to eliminate candidate (c) would be to apply *Periphery only to intervocalic stops; in that case, candidates (c) and (d) would each violate *Periphery once, and the winner (d) would harmonically bound (c). The rationale for this restriction would be that Experiment 1, like Experiments 2 – 4, investigated only intervocalic segments; thus, the intervocalic environment is the only one for which we have evidence of the 'X-pattern'. However, there are hints that articulatory compression is not limited to the intervocalic environment; although final voicing was not systematically manipulated in Experiment 1, §3.3.3 notes that the 'X-pattern' seems to be present for the voicing of the [d] of *said* in the frame sentence as well. Thus, I apply *Periphery to all segments, regardless of environment; but restricting *Periphery would not affect the main point of the present analysis.

Candidate (e) attempts to satisfy both *Periphery and MinDist by spirantizing intervocalic /b/ and adding a small amount of voicing to intervocalic /p/: since intervocalic [b] no longer surfaces, [p] can move towards the middle of the voicing space without violating MinDist. However, with voicing added to intervocalic [p], the two [p]s in candidate (e) no longer have the same amount of voicing, thus violating Match[voi][3ms]. With Match ranked above *Periphery, candidate (e) loses.

Note that without Match, candidate (e) would harmonically bound candidate (d): the two candidates perform identically on all other constraints except *Periphery, where (e) does better. Indeed, if we assume that there is a whole family of *Periphery constraints requiring various distances from the edge of the voicing space, then both candidates would be harmonically bounded by one in which intervocalic [p] has 20 ms of voicing (the closest a voiceless stop can come to the middle of the voicing space without becoming a voiced stop). It is entirely possible that some languages have a 'pull-chain' pattern of this type, whereby intervocalic spirantization of voiced stops enables greater voicing in intervocalic voiceless stops. (Indeed, if a voiceless stop with extra voicing is transcribed in written descriptions as a voiced stop, then perhaps some patterns with both voicing and spirantization intervocalically should be analyzed in exactly this way.) However, in the absence of sufficient data, I do not want to make the strong prediction that *every* language with intervocalic spirantization of voiced stops also has some intervocalic voicing. The Match constraint, highly ranked, allows candidate (d) to surface.

Like (d), candidate (f) attempts to resolve the conflict between MinDist and *Periphery by neutralizing intervocalic /b/ with something else. Where (d) employs spirantization, (f) employs devoicing. The two candidates have the same violation profiles, except that (d) violates Ident[–cont] where (f) violates Ident[+voi]. The universal ranking Ident[+voi] ≫ Ident[–cont] established by the results of Experiment 2 renders candidate (f) unable to win under any ranking.

Finally, candidate (g) neutralizes /apa/ and /aba/ by voicing /p/ rather than by devoicing /b/. Here, the fatal violation is assigned by Ident[–voi]. If we assume that Ident[–voi] and Ident[–cont] are freely rankable (unlike Ident[+voi] and Ident[–cont]), then we can reverse this ranking to allow (g), an attested pattern, to surface. The abbreviated tableau in (18) illustrates.

(18)

			MinDist[voi][60ms]	*Periphery[voi][5ms]	Ident[+voi]	Ident[−cont]	Ident[−voi]	*Merge
	$/\text{pa}/_A$ $/\text{fa}/_C$ $/\text{apa}/_E$ $/\text{afa}/_G$	$/\text{ba}/_B$ $/\beta\text{a}/_D$ $/\text{aba}/_F$ $/\text{a}\beta\text{a}/_H$						
a.	[p_0a] [fa] [ap_0a] [afa]	[b_{60}a] [βa] [ab_{60}a] [aβa]		****!				
b.	[p_5a] [fa] [ap_5a] [afa]	[b_{55}a] [βa] [ab_{55}a] [aβa]	**!					
c.	[p_0a] [fa] [ap_0a] [afa]	[b_{60}a] [βa] $[\text{a}\beta\text{a}]_{F,H}$		***		*!		*
d.	[p_0a] [fa] $[\text{ap}_0\text{a}]_{E,F}$ [afa]	[b_{60}a] [βa] [aβa]		***	*!			*
☞ e.	[p_0a] [fa] [afa]	[b_{60}a] [βa] $[\text{ab}_0\text{a}]_{E,F}$ [aβa]		***			*	*

Thus, squeezing the voicing dimension with MinDist and *Periphery constraints can have at least two different results, depending on the ranking of other constraints: one member of the voicing opposition can move along a perpendicular dimension by spirantizing (illustrated in (17)), or the voicing distinction itself can be eliminated (illustrated in (18)). Squeezing the continuancy dimension has analogous effects: either one member of the contrast changes on a perpendicular dimension (voicing), or the continuancy distinction is collapsed through spirantization. These possibilities are illustrated in the tableaux in (19) and (20), respectively. In these tableaux, the MinDist and *Periphery constraints evaluate only voiceless segments. Since the best measure or combination of measures for continuancy is not obvious, I leave the relevant phonetic dimension for these constraints unspecified. For the sake of discussion, I assume a phonetic space similar to that used for voicing above: a range from 0 (stops) to 60 (spirants).

(19)

	/pa/$_A$ /fa/$_C$ /apa/$_E$ /afa/$_G$	/ba/$_B$ /βa/$_D$ /aba/$_F$ /aβa/$_H$	MinDist[cont][60]	*Periphery[cont][5]	Ident[−cont]	Ident[−voi]	*Merge
a.	[p$_0$a] [f$_{60}$a] [ap$_0$a] [af$_{60}$a]	[ba] [βa] [aba] [aβa]		****!			
b.	[p$_5$a] [f$_{55}$a] [ap$_5$a] [af$_{55}$a]	[ba] [βa] [aba] [aβa]	**!				
☞ c.	[p$_0$a] [f$_{60}$a] [af$_{60}$a]	[ba] [βa] [aba]$_{E,F}$ [aβa]		***		*	*
d.	[p$_0$a] [f$_{60}$a] [af$_{60}$a]$_{E,G}$	[ba] [βa] [aba] [aβa]		***	*!		*

(20)

	/pa/$_A$ /fa/$_C$ /apa/$_E$ /afa/$_G$	/ba/$_B$ /βa/$_D$ /aba/$_F$ /aβa/$_H$	MinDist[cont][60]	*Periphery[cont][5]	Ident[–voi]	Ident[–cont]	*Merge
a.	[p$_0$a] [f$_{60}$a] [ap$_0$a] [af$_{60}$a]	[ba] [βa] [aba] [aβa]		****!			
b.	[p$_5$a] [f$_{55}$a] [ap$_5$a] [af$_{55}$a]	[ba] [βa] [aba] [aβa]	**!				
c.	[p$_0$a] [f$_{60}$a] [af$_{60}$a]	[ba] [βa] [aba]$_{E,F}$ [aβa]		***	*!		*
☞d.	[p$_0$a] [f$_{60}$a] [af$_{60}$a]$_{E,G}$	[ba] [βa] [aba] [aβa]		***		*	*

With MinDist and *Periphery constraints that apply to voiced segments, we can obtain spirantization of intervocalic /b/ in exactly the same way as in (20). However, we cannot cause /b/ to devoice by squeezing the continuancy dimension in the same way as in (19) since Ident[+voi] always outranks Ident[–cont].

Thus, given this set of constraints motivated by the results of Experiments 1 – 4, it is possible to generate both intervocalic voicing and intervocalic spirantization. This analysis demonstrates that an account of these lenition processes does not depend on the existence of a constraint that simply favors the lenited forms over the unlenited forms – exactly the kind of constraint for which Experiment 1 failed to find evidence. In addition, the fixed ranking of IDENT[+voi] over IDENT[–cont] motived by Experiment 2 correctly rules out the unattested pattern of intervocalic devoicing.

5.3.2 Intervocalic Despirantization?

Although the constraint set described in the previous section correctly generates several attested lenition patterns while ruling out intervocalic devoicing, there is another unattested pattern that *is* predicted to occur: intervocalic despirantization. If MINDIST and *PERIPHERY constraints restrict the continuancy dimension and the lowest-ranked IDENT constraint is IDENT[+cont] (a constraint not considered in the tableaux above), the pattern that emerges as the winner is one in which intervocalic spirants become stops.

(21)

	Input		MinDist[cont][60]	*Periphery[cont][5]	Ident[−voi]	Ident[−cont]	Ident[+cont]	*Merge
	/pa/$_A$ /fa/$_C$ /apa/$_E$ /afa/$_G$	/ba/$_B$ /βa/$_D$ /aba/$_F$ /aβa/$_H$						
a.	[p$_0$a] [f$_{60}$a] [ap$_0$a] [af$_{60}$a]	[ba] [βa] [aba] [aβa]		****!				
b.	[p$_5$a] [f$_{55}$a] [ap$_5$a] [af$_{55}$a]	[ba] [βa] [aba] [aβa]	**!					
c.	[p$_0$a] [f$_{60}$a] [af$_{60}$a]	[ba] [βa] [aba]$_{E,F}$ [aβa]		***	*!			*
d.	[p$_0$a] [f$_{60}$a] [af$_{60}$a]$_{E,G}$	[ba] [βa] [aba] [aβa]		***		*!		*
☞ e.	[p$_0$a] [f$_{60}$a] [ap$_0$a]$_{E,G}$	[ba] [βa] [aba] [aβa]		***			*	*

The problem is that the constraints that force lenition, MinDist and *Periphery, are non-directional. MinDist does not care where contrasting elements are located on the relevant scale, as long as they are far enough apart. If there is only one element on the scale, there is no contrast and MinDist has nothing to say at all. *Periphery forbids segments from being too close to the edge of the scale, but it does not favor one end of the scale over another. A stop with no voicing at all is just as bad as a stop with the maximum amount of voicing. Lenition, by contrast, is directional. Intervocalic voiceless stops may become voiced, but not the reverse; stops may become spirants, but not the reverse.

I have argued that we can solve the problem of directionality in the case of voicing by considering a range of possible alternations for underlying voiced stops. Spirantization is such a perceptually 'good' (that is, non-salient) alternation for voiced stops that devoicing never occurs: if a voiced stop alternates at all, there is always a better option than devoicing. The directionality of voicing alternations emerges from restrictions on how one member of the voicing contrast can alternate, not from an asymmetry along the voicing dimension itself.

Is a similar solution possible for the case of (de)spirantization? In principle, yes. We could hypothesize, for example, that spirants never become stops intervocalically because they could undergo an even less perceptible change by becoming approximants. If experimental data supported this proposal, we could posit a universal ranking IDENT[–son] $\gg$ IDENT[+cont] and rule out intervocalic despirantization.

Unfortunately, this explanation reveals an inherent weakness of the approach of the P-map for a series of unidirectional alternations such as lenition processes, as illustrated in figure 5.2. The P-map by itself does not provide the needed directionality: for each segment along the chain, it rules out the possibility of moving to the left by appealing to the superior perceptual consequences of moving to the right instead. Once we reach the rightmost element in the chain, this line of argumentation is no longer possible; and if we simply add another element ot the chain, the problem has only been pushed back a step. If it is the case (and I know of no counterexamples) that spirants may lenite to approximants intervocalically

Figure 5.2: Series of unidirectional leniting intervocalic alternations

$$p \longrightarrow b \longrightarrow \beta \longrightarrow w$$

but approximants do not become spirants, then we have solved the problem for spirants only to create another problem for approximants. It's turtles all the way down.

It is at this point that articulatory considerations are usually brought in to play. Rather than assuming a complex set of perceptual differences that results in a chain of unidirectional alternations, it seems much simpler to suppose that articulatory effort reduction favors some sounds over others, thus encouraging change in only one direction. However, the results of Experiment 1 do not support such a proposal: there is no evidence that subjects in the intoxicated condition were more likely to lenite than subjects in the sober condition. It is certainly possible that further experimental work could produce evidence in support of an articulatory basis for the directionality of lenition, but that is not the data we have at hand.

I do not offer a solution here to the problem of directionality in lenition. The contribution of Experiments 1 – 4 is the observation that phonetic facts may motivate a substantial part of the typology of lenition, but not in the way that is usually assumed. A tendency to reduce articulatory effort may provide the precursors for lenition, but by compressing the articulatory space on both ends, rather than by causing an overall shift in the direction of lenited forms as generally believed. In addition, perceptual facts, long neglected in the study of lenition, can contribute to our understanding of why the alternations involved in lenition only go in one direction, even if perception does not ultimately tell the whole story.

5.4 Conclusion

This chapter has illustrated the implications of Experiments 1 – 4 for any phonological analysis that takes phonetic facts seriously. First, the results of Experiments 2 and 4, when combined with the P-map, provide a very good match to the broad typological facts: although voiceless stops may voice or spirantize intervocalically, voiced stops may spirantize but may not devoice. However, applying the same procedure to the results of Experiment 3 yields incorrect predictions regarding which voiced stops should be more likely to spirantize. This result demonstrates that not every difference in perceptibility should be projected by the P-map into a universal ranking of faithfulness constraints; I have suggested some ways in which we might constrain the P-map in a principled manner.

Second, the results of Experiment 1 do not support the traditional account of lenition as effort reduction; that is, Experiment 1 provides no evidence that reducing articulatory effort makes subjects more likely to produce lenited forms. Rather, the compressed articulatory space exhibited by subjects in the intoxicated condition is reminiscent of approaches that penalize productions in the periphery of the phonetic space, notably Dispersion Theory.

I have also shown that a Dispersion-Theoretic analysis is capable of modeling intervocalic lenition even without a constraint that favors lenited forms over unlenited ones outright. This analysis illustrates the kind of approach phonology must take in order to account for lenition in a way consistent with the results of Experiment 1 – 4. Although some questions remain (such as the best way to rule out intervocalic despirantization), I submit that this approach is a step in the right direction because it adheres more closely to the phonetic facts of lenition than do previous approaches.