DE GRUYTER MOUTON The Linguistic Review 2015; 32(3): 573–599

Bert Vaux* and Bridget Samuels

Explaining vowel systems: dispersion theory vs natural selection

DOI 10.1515/tlr-2014-0028

Abstract: We argue that the cross-linguistic distribution of vowel systems is best accounted for by grammar-external forces of learnability operating in tandem with cognitive constraints on phonological computation, as argued for other phonological phenomena by Blevins (2004). On this view, the range of *possible* vowel systems is constrained only by what is computable and learnable; the range of *attested* vowel systems is a subset of this, constrained by *relative* learnability (Hale and Reiss 2000a, Hale and Reiss 2000b; Newmeyer 2005). A system that is easier to learn (e.g., one whose members are more dispersed in perceptual space) is predicted by our model to become more common cross-linguistically over evolutionary time than its less learnable competitors. This analysis efficiently accounts for both the typological patterns found in vowel systems and the existence of a non-trivial number of "unnatural" systems in the world's languages. We compare this model with the leading forms of Dispersion Theory (notably Flemming's (1995) implementation in Optimality Theory), which seek to explain sound patterns in terms of interaction between conflicting functional constraints on maximization of perceptual contrast and minimization of articulatory effort. Dispersion Theory is shown to be unable to generate the attested range of vowel systems or predict their interesting properties, such as the centralization typically found in two-vowel systems and the quality of epenthetic segments.

Keywords: Vowel systems, dispersion theory, evolutionary phonology

1 Introduction

> "That 'unnatural' systems, with attributes that in no way follow from the conflicting forces of 'ease of articulation' and 'maintaining contrasts', exist is, I hope, obvious to anyone who has examined the matter in any reasonable detail. Such 'unnatural' (but computationally tractable) systems reveal that the computational system of human phonology is *more* than the interaction of these two forces." (Hale 1997)

***Corresponding author: Bert Vaux,** Department of Theoretical & Applied Linguistics, Cambridge University, Cambridge, England, E-mail: bv230@cam.ac.uk
Bridget Samuels, Department of Linguistics & Cognitive Science, Pomona College, Claremont, CA, USA; Center for Craniofacial Molecular Biology, University of Southern California, Los Angeles, CA, USA, E-mail: bridget.samuels@gmail.com

Phonologists since at least Jakobson and Trubetzkoy (e.g., Trubetzkoy 1939) have been concerned with why languages display their particular array of sound patterns. The central questions raised in treatments of the distribution of vowel systems in particular are (i) why vowels fall into the particular zones of the perceptual/ articulatory space that they do, (ii) why some sounds are more common in vowel inventories, and (iii) why 'unnatural' vowel systems exist, such as the vertical ones of the sort we find in Abkhaz (Hewitt 1979) and Marshallese (Choi 1992).

The theories in (1) have been proposed to answer these questions concerning vowel inventories:

(1) Theories of vowel inventory typology
   i. **Quantal Theory** prioritizes perceptual invariance: the point vowels {a i u} are typologically preferred because they represent regions of the articulatory space in which variability in production has relatively small acoustic impact (Stevens 1972, Stevens 1989; Lang and Ohala 1996).
   ii. **Traditional markedness theory** derives its hierarchy largely from phonological behavior and cross-linguistic distribution, or by counting features.
   iii. **Dispersion Theory** (DT) maintains that the structure of inventories derives from functional principles such as the minimization of effort and maximization of contrast.

Our focus here will be on Dispersion Theory (1.iii). As has been noted in the literature, Quantal Theory (1.i) makes several incorrect predictions. First of all, the fourth most common vowel, [ɛ], is not stable as predicted by the theory (Carré 1996). Second, Livijn's (2000) survey of 28 differently-sized inventories showed no evidence for acoustically-favored hotspots, contrary to the core prediction of Quantal Theory. Third, vertical systems such as those of Abkhaz, Kabardian, and Marshallese are missed by Quantal Theory, which predicts that such small vowel systems should select their members from the quantal set {a i u}. Finally, Quantal Theory wrongly predicts that the point vowels should be invariant cross-linguistically (Disner 1983).

Traditional markedness theory (1.ii) has faced the objection that it merely formalizes the attested facts, rather than explaining them in terms of constraints on human articulation, perception, or processing. Consequently, Dispersion Theory has gained support as it incorporates these principles. One of the early dispersion-based approaches was Lindblom's purely computational phonetic theory of Adaptive Dispersion (Liljencrants and Lindblom 1972; Lindblom 1990, etc.), followed by Flemming's (1995) implementation of dispersion in Optimality Theory (referred to here as OT-DT; see (2.v) for references). Some advocates of dispersion limit its application to the diachronic axis; Lindblom and Maddieson (1988: 72) for example maintain that "consonant inventories tend to evolve so as to achieve maximal perceptual distinctiveness at minimum articulatory cost." OT-DT, however, incorporates the mechanisms of dispersion directly into synchronic phonological systems, as we shall see.

Here we present a number of empirical and theoretical arguments that Dispersion Theory cannot adequately describe the range of attested vowel systems, and we present an alternative based on Evolutionary Phonology (Blevins 2004). The main focus of our critique will be on OT-DT, since it attempts to account for phonological facts as well as static phonetic distributions. As McMahon (2000) and Jansen (2002) have pointed out, linguistic phenomena that incorporate products of history, including the vocalic patterns we discuss here, highlight the difficulties with universalist theories like synchronically-implemented DT. Our overarching point is therefore that the structure of vowel systems is largely an evolutionary and historical problem, not a synchronic one. Less is more, in this case: a more permissive theory than DT is required to account for the rich variation in vowel systems seen cross-linguistically.

2 Early dispersion theories

We begin by summarizing the range of forms that DT has taken, dating back to the pre-generative era. There are six major incarnations of DT in the literature, summarized in (2).

(2) Incarnations of Dispersion Theory

i. **Structuralist functionalism** (Jakobson 1941; Martinet 1955): vowels are maximally dispersed.

ii. **Theory of Adaptive Dispersion**: *maximal dispersion* (Liljencrants and Lindblom 1972), *sufficient dispersion* (i.e., articulatory economy balanced with perceptual distinctiveness; Lindblom 1975, Lindblom 1986, Lindblom 1990, Lindblom 1992, Lindblom 2000; Kawasaki 1982; Lindblom and Maddieson 1988; Carré et al. 1994; Yang 1996; Traunmüller 1998; Cho et al. 2000, and many others).

iii. **Lexically-based Dispersion**: distance is proportional to the functional (= lexical) load borne by the contrast (ten Bosch 1991, Bosch 1995; cf. van Son et al. 1998).

iv. **Focalization Theory** (a combination of Quantal Theory and Dispersion): *Distinctive Regions Model* (Mrayati et al. 1988; Carré and Mrayati 1992, Carré and Mrayati 1995; Carré and Mody 1997;

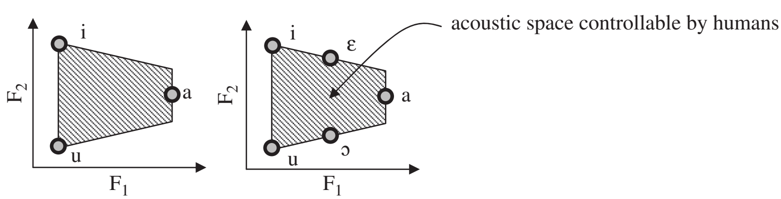
Carré 1996); *Dispersion-Focalization Theory* (prefer vowels showing a convergence between two formants; Vallée 1994; Schwartz et al. 1997a, Schwartz et al. 1997b, Schwartz et al. 1999).

v. **Optimal Dispersion Theory** (Flemming 1995, Flemming 1996, Flemming 2004; Ní Chiosáin and Padgett 1997; Boersma 1997; Riggle 1999; Ahn 2000, Ahn 2002a, Ahn 2002b; Harris and Lindsey 2000; Sanders 2002; Padgett 2004).

vi. **Self-organizing systems theory**: eschews linguistic rules and constraints (de Boer 2001).

All variants of DT consider maximization of distinctiveness to be the main method by which languages select phonemic distinctions. The elements involved in this dispersion process have been claimed to include **maximal contrast in the articulatory space** (Jakobson 1941), **minimization of effort** (Lindblom 1975[1]),[2] **focalization** (Stevens 1972; Schwartz et al. 1997b; Roark 2001), **maximal overall contrast compensating for F_0** (Ryalls and Lieberman 1982; Diehl et al. 1996; Traunmüller 1998), **lexical frequency** (ten Bosch 1995), and **formant normalization** (Roark 2001). The plots in (3) show the basic idea of how vowels in three- and five-member systems can be maximally dispersed in a given acoustic space according to their F_1 and F_2 values.

(3) Maximal dispersion of 3- and 5-vowel systems in a given acoustic space



The first major treatment of dispersion in the generative era was proposed by Liljencrants and Lindblom (1972), who used a computer model designed to maximize the distance between a given number of vowels within a formant space based on a theoretical construct from Lindblom and Sundberg (1971) merging F_2 and F_3 into an "effective second formant" conveying both backness and rounding. Perceptual effects were simulated using mels. This computer model was used to predict F_1 and F_2/F_3 values, which were assigned to corresponding phonetic

1 Cf. also Traunmüller (1998), who asserts that men have less dispersed formants than women because it requires more energy for them to speak, and they do not need the extra contrast to be understood, thanks to their lower fundamental frequency.

2 Cf. Lang and Ohala (1996) on the difficulty of measuring articulatory cost.

categories and compared to known typology. This procedure yielded what the authors deemed "approximately correct" results for systems of three to six vowels.

This early model and its Adaptive Dispersion (AD) relatives fell short of predicting the attested range of vowel systems, for both small and large inventories. Overall, and perhaps most importantly, human languages permit more variation than the model predicts: it only accurately generates a small number of the attested inventories of languages with fewer than six vowels (de Boer 2001), and it is limited to generating systems of 3–9 vowels. Secondly, it does not predict the most common four-vowel configuration (Roark 2001: 2). Roark's (2001) version of AD, which incorporates quantality and normalization, correctly predicts the four-vowel system but does not get the facts right for six- and seven-vowel systems. For systems with six or more vowels, AD wrongly predicts that inventories should commonly include an increasing number of high central vowels, contrary to fact, and in general it models the behavior of central vowels poorly. As Vallée et al. (2001) point out, AD also fails to generate the /i y u/ series within the high vowel set.

The unattested proliferation of vowels between /i/ and /u/ generated by AD in larger systems includes four contrasting high vowels in systems with as few as seven vowels. However, attested languages never contrast more than three degrees of backness (Lindau 1975: 13). There are moreover significant differences between the high vowels of real systems and the theory's predictions. Figure (4) exemplifies for seven-vowel systems how the inventory predicted by Lindblom (1990) is quite unlike the two most commonly attested inventories. The AD-predicted inventory is asymmetric and disorganized from a featural perspective, as Boersma (1997) has noted: the reason is that the distance function in the model favors a height asymmetry between front and back vowels, because an F_1 difference always contributes positively to the perceptual difference between a pair of vowels.

(4) Predicted and actual 7-vowel systems

a. Lindblom's predicted optimal system (1990)

i	ʉ		u
		ɤ	
ɛ			
a		ɑ	

b. The most common 7-vowel systems according to Crothers (1978)

i	ɨ	u
e	ə	o
	a	

i		u
e		o
ɛ		ɔ
	a	

Although in this case (4a) AD predicts a system that is too asymmetric, a certain degree of asymmetry may not be undesirable in general. Disner (1983) points out that symmetrical systems are the exception rather than the rule, and indeed, the systems predicted by DT tend to be less clustered than those found in nature. Swedish, for example, clusters vowels in the high front region. Disner further points out that AD "predicts that the vowels in one nine-vowel system ought to be phonetically similar to the vowels in another nine-vowel system. Yet not even such closely related vowel systems as those of Norwegian and Swedish are correctly predicted by the theory" (1983: 115).

Dispersion-Focalization Theory (DFT) attempts to address some of these shortcomings of Dispersion Theory by incorporating elements of Quantal Theory. The dispersion component of DFT proposes that sounds are dispersed so as to maximize the perceptual distance between them, and the focalization component (drawing on Quantal Theory) states that vowels which show a convergence between two formants are preferred. The general procedure in DFT is to associate an energy function to each system. The optimal system has the lowest energy function, either locally (in which case it is a stable system) or globally (in which case it is the best system).

While DFT is useful in explaining the appearance of the high round vowel /y/, it does not correctly generate mid front round vowels (Carré 1996). Like AD, it generates too many high vowels and not enough central vowels in larger systems (Carré 1996), and it is also limited to making predictions for three- to nine-vowel systems (Vallée et al. 2001).

Despite the surface differences between these variants of DT, they suffer from many of the same problems. First, there are many cases like that of Navajo, described by McDonough et al. (1993), which has a vowel inventory consisting of the unbalanced set {i e æ o}. Note that these vowels are not equidistant, thus contradicting the prediction of AD – /i/ and /e/ are too close together. DT in general fails to predict or account for the fact that /u/ shows more variance than the other vowels. McDonough et al. propose in response that the upper back corner of the vowel triangle may be less "sharp" or perhaps high back vowels are less defined. This implies that, while /i/ and /æ/ may be considered to be quantal, there is no obvious high back quantal vowel.

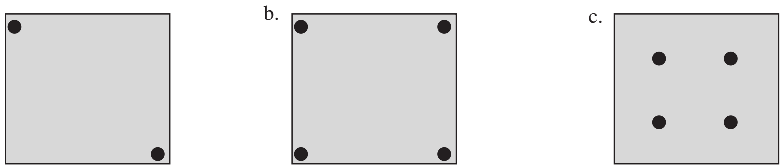
In summary, the forms of Dispersion Theory presented thus far have four significant shortcomings: (i) they cannot or do not generate sub-optimal systems, though such systems are widely attested; (ii) they do not perform well with central vowels; (iii) they work only for three- to nine-vowel systems; and (iv) they make particularly poor predictions regarding systems of six or more vowels.

3 Dispersion theory in OT

Flemming (1995, 1996, 2004) addresses the problems just mentioned and incorporates dispersion effects into OT, thereby using independently-required phonological machinery (i.e., ranked constraints on surface representations) to produce a theory of vowel systems.

The functional principles that drive OT-DT are the following (Flemming 1996, Flemming 2004): maximize the distinctiveness of contrasts (using the MINDIST family of constraints), minimize articulatory effort (*EFFORT), and maximize the number of contrasts (MAXCONTRAST). Markedness is treated as being paradigmatic (i.e., a property of contrasts), not syntagmatic (a property of features), which is claimed to account for vertical systems as well as predict the quality of epenthetic vowels. Also, unlike typical OT analyses of other phonological phenomena, inputs and outputs are sets of words rather than individual words. Figure (5) illustrates the possible interactions of these principles, formalized as MINDIST, *EFFORT, and MAXCONTRAST ("#").

(5) Interaction of MINDIST, *EFFORT, and MAXCONTRAST (Flemming 1996)



a. two categories MINDIST >> *EFFORT, #

b. four categories # >> MINDIST >> *EFFORT

c. four categories *EFFORT >> MINDIST

Before we delve into empirical matters, it should be pointed out that these constraints are controversial at a basic level. McCarthy (2005: 40) argues that there is a fundamental difficulty with constraints like MAXCONTRAST and MINDIST, which represent "more of an intuition than a usable phonological principle. . . . Injunctions like 'minimize the differences' or 'be as metrically consistent as possible' are too ill-defined to serve as constraints." While such principles may have heuristic value, the difficulty in formulating them specifically enough to be incorporated fruitfully into EVAL must be appreciated. For example, Lang and Ohala (1996) note that it is difficult to evaluate the hypothesis that consonant inventories tend to evolve so as to achieve maximal perceptual distinctiveness at minimum articulatory cost (Lindblom and Maddieson 1988), due to uncertainty over how to measure articulatory cost. In the remainder of this section, we set these issues aside to focus on the predictions made by OT-DT.

3.1 Height systems

Let us examine a key case for DT, namely that of vertical vowel systems. Vertical systems do not employ an underlying [back] contrast, but instead oppose height among a series of (underlyingly) central or perhaps underspecified vowels. One example is provided by Kabardian, which has an underlying inventory consisting of the two short vowels {ə, a}; we will discuss the surface manifestations of these phonemes later in the text. In the case of small, vertical vowel systems like this, there is a difference between on the one hand Quantal Theory and traditional markedness theory, which predict that a subset of the cardinal vowels / ɑ i u æ/ should always be preferred in these systems regardless of size, and on the other hand DT, which does not make this prediction. Some languages putatively lacking back-front vowel contrasts in their underlying inventories include (according to Flemming 2004: 251) Marshallese (Bender 1968; Choi 1992), the Northwest Caucasian languages Kabardian (Colarusso 1988, Colarusso 1992), Shapsug (Smeets 1984), Ubykh (Colarusso 1988), and Abkhaz (Hewitt 1979; Vaux and Psiypa 1997), and some Ndu languages of Papua New Guinea, such as Iatmul (Laycock 1965; Staalsen 1966).[3]

Flemming (1995, 2004) and Ní Chiosáin and Padgett (1997) observe that universal markedness rankings of the sort *ɨ >> *u >> *i, which have been postulated to account for the implicational hierarchies of vowel systems containing three or more vowels, incorrectly predict that a language with only one high vowel, such as Abkhaz, should select [i]. However, attested languages of this type choose [ɨ] rather than [i]. On the basis of this fact, Flemming and Ní Chiosáin & Padgett argue for a universal family of Contrast constraints requiring that segments maintain at least a specified perceptual distance. To account for the high vowel case, Ní Chiosáin and Padgett (1997: 22) postulate a constraint Contrast(Colour), which states “maintain a colour contrast with sufficient perceptual distance.” Minimally sufficient distance is defined in this case as the perceptual distance between [i] and [u] on the scale i ↔ ɨ ↔ u. Given this scale, the pairs {i ɨ} and {ɨ u} violate Contrast(Colour), because the perceptual distance between the two members of each of these sets is smaller than the distance between [i] and [u]. By varying the ranking of Contrast(Colour) relative to the universal hierarchy of segmental markedness constraints, systems with one or two high vowels can be obtained.

3 One should not be too quick to form generalizations about two-vowel systems since our data come from only about six languages, four of which are closely related. However, let us assume for the sake of argument that the patterns found in these languages reveal a general trend.

An issue with the OT-DT system of constraints with respect to height systems is that the constraints described above must necessarily hold over surface representations, not underlying representations, as Flemming and Ní Chiosáin & Padgett explicitly note.[4] This is a logical requirement given that perceptual distance is a coherent notion only when applied to concrete surface forms with specific acoustic properties, and in an OT context, Richness of the Base prohibits constraints on underlying representations. However, the languages typically cited as having only two vocalic phonemes, Abkhaz and Kabardian, actually have rich surface vowel inventories: Abkhaz has 13 vocalic allophones (Hewitt 1979), while Kabardian has 22 (Halle 1970; Colarusso 1989). It is only at the most abstract level of phonemic analysis that these languages can be said to possess just two vocalic phonemes. In fact, this is true for all of the reduced-inventory languages that have been cited in the literature: the *underlying* inventory is limited, but the inventory of *surface* allophones is quite large (see Choi 1992; Hale and Reiss 2008: 150ff on Marshallese). Since OT-DT countenances only surface phonetic representations, this abstract level at which only two vowels exist is neither available nor relevant.

With respect to the surface inventories of these languages, the contextual nuances of vowel color that adorn the underlyingly two-way contrast can be explained as linear interpolations between the F_2 values of the preceding and following consonants, as argued by Choi (1992) for Marshallese. Flemming (2004) outlines a simplified version of such a system in OT-DT, using a simplified version of Kabardian (call it K´) based on the generalizations made by Anderson (1978) and Choi (1992). In K´, front vowels appear after palatalized consonants (C^j) and central vowels appear after labials and coronals. On the basis of these generalizations Flemming draws up the simplified phonetic scale in (6), with the gradient F_2 (≈ backness) dimension discretized into six values:

4 Flemming's theory of phonology more generally contains only surface phonetic representations, and no abstract underlying representations. This creates a number of additional problems. Underlying representations are needed to capture the relations between categories in different idiolects (e.g., to identify a man's /a/ as belonging to the same phonemic category as a woman's /a/, even though the two occupy different phonetic spaces) and to correctly generate effects referring to neutralized contrasts, such as final devoicing. Moreover, surface-only models ignore the ample phonological and psycholinguistic evidence for abstract underlying and intermediate phonological representations (see Hyman 1970; Campbell 1986; Baumann 1996; Pallier et al. 1999; Hallé et al. 2000; Itô and Mester 2003 inter alia, as well as the discussion in Section 3.2).

(6)

F_2:	6	5	4	3	2	1
	i	i̠	ɨ̟	ɨ	ɯ	u
	e	ə̟	ə	ɤ	o	
	Cʲ				Cˠ	

The inventory of surface CVC syllables in K´ can then be derived via the tableau in (7), provided that one considers only extra-short vowels, secondary articulations of consonants, and F_2 contrasts among vowels. The basic idea in (7) is that backness contrasts separated by three or fewer columns in the chart in (6) are assigned a violation of the perceptual dispersion constraint MINDIST=F_2:4, but violations of this constraint are trumped by violations of the articulatory effort constraint *HIGHEFFORT, which penalizes sequences of a consonant followed by a vowel not drawn from its column in (6). In other words, any CV sequence other than Cʲi, Cʲe, Cˠɯ, or Cˠo violates this constraint. MAXCONTRAST assigns a single asterisk for each member of a system, so for example the candidate system in (7d) receives two asterisks by virtue of containing two members, and the number of asterisks is potentially infinite. For reasons discussed by Karttunen (1998) and McCarthy (2003), gradient constraints of this sort are computationally intractable in a finite-state system; space considerations prevent us from considering this issue further here.

(7) Selection of a CVC inventory for K´ (Flemming 2004)

		*HIGH EFFORT	MINDIST = F_2:4	MAXCONTRAST
a.	CʲiCʲ CʲuCʲ CʲiCɤ CʲuCɤ CˠiCʲ CˠuCʲ CˠiCˠ CˠuCˠ	*!*****		8
b.	CʲiCʲ Cʲɨ̟Cʲ Cʲi̠Cˠ CʲɨCˠ Cˠi̠Cʲ CˠɨCʲ Cˠɨ̟Cˠ CˠɯCˠ		*!***	8
c.	☞ CʲɨCʲ CʲɨCˠ CˠɨCʲ CˠɨCˠ			4
d.	CiC CuC			2!

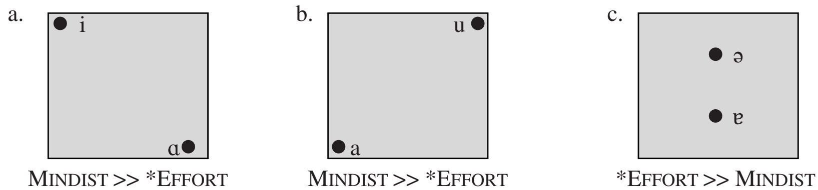
A further problem with Flemming's analysis of K´ is that in real Kabardian, plain laryngeals are followed by back vowels (Colarusso 1992), not central vowels as predicted by (7). Furthermore, both Abkhaz and Kabardian contrast front and back vowels in loanwords, which freely contain {a e i o u}. One could propose a different ranking of constraints for the loanword stratum in which most tokens of these vowels occur, but then the vowel distribution would not be synchronically derivable from just the inventory of surface vowels – and with no underlying representations in OT-DT, surface vowels are all one has to work with. On top of this, Abkhaz contrasts front [ej] and back [aj] in native words (Hewitt 1979).

Elaborating on Choi's interpolation theory, Flemming (2004) states that "backness and rounding are governed by minimization of effort. This means that they are realized as smooth transitions between preceding and following consonants, which frequently results in central or centralized vowel qualities." Deriving centralization via smooth articulatory transitions is not the same as generating central vowels from acoustic dispersion constraints, though. The restriction imposed on very short vowels by *HIGHEFFORT also forces these "smooth transitions" of tongue body and lip positions between the two flanking consonants, but this should mean that a vowel between two plain labials should be round and central (with the tongue in the default rest state, since the consonants specify no tongue position), rather than non-round and high. OT-DT essentially predicts systems like the Kabardian one – modulo the laryngeal/back vowel problem mentioned earlier – where the vowels receive all features except height from neighboring consonants. Such a system should never produce central vowels except perhaps with labials; coronals should yield front vowels, dorsals should yield back vowels, and so on. The Abkhaz system is therefore a problem, as Flemming rightly notes, because there is no coloring by plain consonants. Moreover, if an unspecified vowel occurs between two non-coloring consonants (e.g., [h]), why does it not surface as a copy of the preceding and following vowels, especially if they are identical?

The length of vowels in Abkhaz provides another challenge. Flemming (2004) ties reduction of contrast to shortness, saying "a given limit on effort effectively makes a smaller portion of the vowel space available in a short vowel than in a long vowel," which wrongly predicts that languages should always have more contrasts in long vowels than in short vowels. Because of historical changes, the opposite is often true, as in the Shamaxi dialect of Armenian (Vaux 1998). Harris and Lindsey (2000) also observe that vowel neutralization is not necessarily linked to stress or length; Chumash for example has a 6-vowel contrast in roots but a 3-vowel contrast in suffixes. This is independent of stress, which in Chumash is typically penultimate. The same holds for Turkish and many other familiar cases. The effects of vowel shortening also produce unpredictable results which can go against Flemming's claim: only short front vowels centralize in Nawuri (Casali 1995) and shortening significantly affects only F_2 (i.e., backness, not height) in Chickasaw (Gordon et al. 2000).

Finally, as we saw in the previous section, OT-DT predicts that two-vowel systems without a phonological [back] contrast should not have to surface with central vowels; ranking MINDIST above *EFFORT should produce a diagonal system of {a u} or {i ɑ}, invalidating the OT-DT claim to a correct prediction of centralization in two-vowel systems. Flemming (1996) includes the possibility of an {i ɑ} system (cf. our figure (5a), reproduced here as (8a)), but the problem that this poses with respect to predicting two-vowels systems is not noted. It should also be mentioned that the same constraints could generate {a u}, as in (8b).

(8) Two-vowel systems in OT-DT (Flemming 1996)



A diagonal system in effect involves enhancement (Stevens and Keyser 1989), which Flemming (2004) allows for other contrasts such as voicing. It is not clear in this system why enhancement would be invoked in the preferred 3-vowel system, {a i u}, but not in the preferred vertical 2-vowel system. It is important to notice that OT-DT does not actually predict that 2-vowel systems will contain central vowels; free ranking of the relevant dispersion constraints yields both vertical and diagonal systems, and even horizontal systems. Many linguists assume horizontal systems to be impossible, and have proposed a variety of patches to rule out such systems and answer the related question of whether distinctness of F_2 universally has precedence over distinctness of F_1. Ten Bosch (1991) achieves such an F_1/F_2 asymmetry by stipulation, assigning a value of 0.3 to the importance of F_2 distance vs. F_1 distance; proponents of DFT similarly use a weighting of 0.25. Traunmüller (1998: 4) suggests that speakers are more concerned with contrast in F_1 than in F_2 because F_2 is less sensitive to variations in F_0 than F_1 is. Lindblom attributes the disparity to proprioception: jaw height can be felt more accurately than tongue-body position. Boersma (1997: 14) asserts that F_1 is more important because it is louder: "the second spectral peak has a larger chance of drowning in the background noise." Until one can show that a horizontal system is not learnable by children, though, the hypotheses just reviewed here are purely speculative. The important points are that the factorial typology yielded by OT-DT both overgenerates and misses the target with respect to the attested two-vowel systems, the surface orientation of the system is also problematic given the dramatic disparity between underlying phonemes and surface phones, and on the implementational level, there are considerable problems with formulating a number of the constraints that are crucial to the system.

3.2 Faithfulness problems in OT-DT

In the previous section we were chiefly concerned with height systems, since DT models differ from Quantal Theory and traditional markedness theory in the predictions they make regarding such inventories. We now move on to evaluating OT-DT on matters of faithfulness and markedness more generally. Faithfulness poses a challenge for OT-DT since, as Flemming correctly observes,

> "it is not possible to combine dispersion constraints with the faithfulness-based account of allomorphic similarity[,] because the two are fundamentally incompatible ... The inclusion of faithfulness constraints subverts the intended effect of the MINDIST and MAXIMIZE CONTRASTS constraints, because it makes the selected inventory of vowel height contrasts dependent on the input height under consideration." (Flemming 2004)

As a result, OT-DT uses surface-driven Output-Output correspondence. While this remains a contentious issue, as we discussed in the previous section (see also Hale et al. 1998; Kiparsky 2000; McCarthy 2007), a phonological theory that includes only comparison of output forms without any consideration of underlying or intermediate representations cannot adequately account for a variety of generalizations (e.g., opacity). Boersma argues further that "inventory grammars ... do not explain how a random input is filtered into a well-formed utterance ... Flemming's global inventory evaluation procedure is not a model of grammar; it just shows that inventories can be described with a strict ranking of principles" (1997: 16).

3.3 Markedness problems in OT-DT

The OT-DT model has a conceptual problem with markedness as well as faithfulness, arising from its assumption that markedness is based on paradigmatic contrast rather than intrinsic properties of segments. Flemming (2004) claims that "there is no analogous notion of effort involved in perceiving a sound – perceptual difficulties don't arise because particular speech sounds tax the auditory system, the difficulty arises in categorizing sounds." However, there are many instances where auditory difficulties are obviously involved – very complex sounds, quiet sounds, and vowels that are fully or partially devoiced, for example.

As Rice (1999) notes, and as should already be clear from previous sections, the number of attested systems that are unexpected in DT is vast. Rice summarizes the issue as follows:

> There is a simplicity to this theory which does not always meet with phonetic facts. ... [O]ne would expect maximal differentiation of contrasts in phonetic terms. However, this is often not the case. For instance, some languages have a three-way place contrast [along the F_2

dimension] between i-ü-u rather than i-ɨ-u (e.g. Rukai; Austro-Tai 417); some have a two-way contrast between i-ɯ (Adzera, Austro-Tai 419; Japanese, Ural-Altaic 71) rather than the expected i-u; some have a single high vowel which is front rather than central (e.g. Navajo, Athapaskan; Klamath, Penutian). These facts call into question the claim that articulatory simplicity governs in the absence of contrast and maximal distance in the presence of contrast." (Rice 1999: 414)

One reason for this, as Rice notes, is that markedness can vary even among identical systems. This is incompatible with the paradigmatic notion of markedness employed by OT-DT, tied as it is to the existence of particular contrasts.

As mentioned earlier, one claimed advantage of OT-DT is that it can predict the quality of epenthetic segments in a particular inventory by "minimization of effort or other contextual markedness constraints" (Flemming 2004). As Rice (1999) notes, though, it is not necessarily true that articulatory concerns win out in positions of reduced contrast, such as epenthetic contexts. Vaux and Samuels (2003) and Blevins (2008) discuss a wide variety of epenthetic segments cross-linguistically, and though their focus is on consonants, their overarching point remains valid: epenthetic segment choice is highly idiosyncratic and often owes to reanalysis of historical deletion patterns. For example, the arbitrariness of epenthetic segments is illustrated by the difference between Standard Armenian and the Muslim Hamshen dialect, which have identical phonemic inventories but epenthesize different vowels: Standard Armenian selects [ə], whereas Hamshen selects [ɛ] (Vaux 1998). In short, the choice of epenthetic vowel is not predictable from the vowel inventory.[5]

OT-DT also suffers because its relational notion of markedness rules out traditional markedness constraints like *ə. How then can OT-DT generate vowel systems with gaps, or choose between {a i u} and {a i o} for example? Instead of traditional markedness constraints, Flemming (2004) appeals to "constraints favoring less confusable contrasts over more confusable contrasts." Innate constraints for this purpose seem superfluous given that the principles they embody are well-motivated properties of general cognition independent of any encoding in the grammar (cf. Boersma 1997; Ohala 2005).

5 Dupoux et al. (2011) note that the epenthetic vowel in Japanese is /u/, which is the shortest vowel in the language, while the epenthetic vowel in Brazilian Portuguese is /i/, which is the shortest vowel in that language. However, the identity of the shortest vowel in a language is itself not predictable from the inventory.

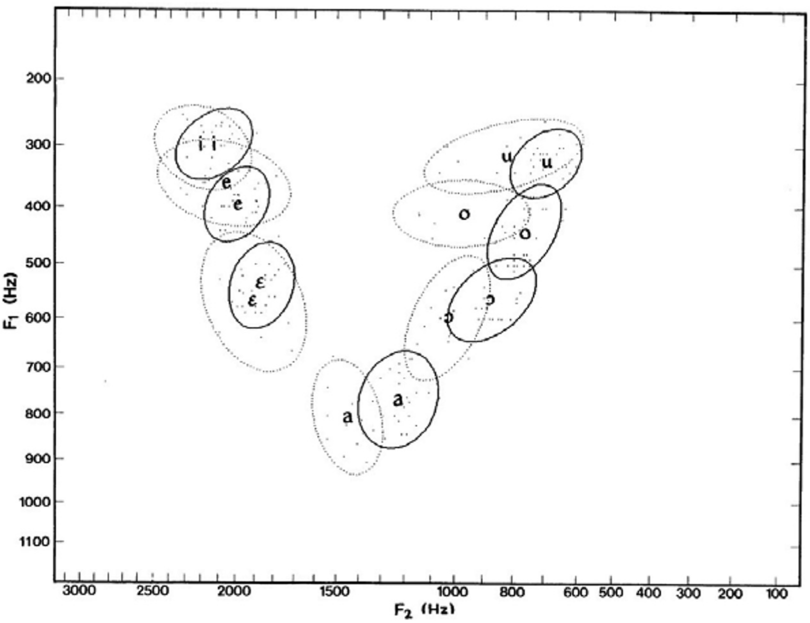
4 Problems with all forms of dispersion theory

We have seen so far that pre-OT and OT-based implementations of DT each encounter a number of problems, some distinct to the particular implementation in question. Before concluding, we will remark on some other, more general concerns that are common to all forms of DT.

One major empirical issue is that DT is based on generalizations about vowel systems that are not phonetically reliable. The familiar figures from the UPSID database (Maddieson 1984) cited by Flemming (2004) and others, such as the generalization that 94% of front vowels are unrounded and 93.5% of back vowels are rounded, are now rendered outdated by newer and more reliable phonetic data. For instance, we now know that contrary to UPSID, English back vowels are [-round] and centralized for many speakers (Joanisse and Seidenberg 1998: 4), that American English /u/ nearly overlaps with Korean /ɨ/ (Yang 1996: 259), and that there is a consistent difference in height and backness between the high front unrounded vowels of German and of Norwegian, although both are transcribed as [i:] in the Stanford archive (Disner 1983: 4).

Recasens and Espinosa (2009) present a literature review as well as new data, which point in general to a conclusion that undercuts DT: there is typically not a correlation between the number of vowels in an inventory and the size of its phonetic categories. The DT principle of minimal distance predicts that an inventory with more vowels should have smaller categories, entailing more precise phonetic realization. This is not borne out by the cross-linguistic data. Nor is it clear that larger inventories cover a larger acoustic space, as would seem to be demanded by the DT principle of maximal contrast (see Becker-Kristal 2009: 14ff for an overview). Since Lindau and Wood (1977), it has been recognized that even closely related languages with the same vowel systems do not distribute their vowels the same way in the acoustic space. To take one example, in (13) it can be appreciated that the same seven-vowel system is implemented very differently in Italian and Yoruba (Disner 1983: 17ff). Each of these seven-vowel systems derived from an earlier nine-vowel inventory with the additional categories of *ɪ and *ʊ. However, they differ in the historical developments by which they lost these two categories. In Yoruba, *ɪ and *ʊ merged with the high vowels and a pull chain raised /e/, /ɛ/ and /o/, while in Italian *ɪ and *ʊ merged with /e/ and /o/. Analysis of variance and Duncan post hoc analysis confirm that the F1 values of /e, o, ɔ, a/ are significantly different in the two languages, and there is a significant interaction between the language and vowel variables, which Disner (1983: 22) calls a “pattern effect.”

(13) Italian (solid; data from Ferrero 1972) and Yoruba (dotted; data from Lindau and Wood 1977) vowel systems (figure reproduced from Disner 1983: 23)



Even within a single language, there are differences that dispersion theories do not predict. Indeed, it has been known since Quillis and Esgueva (1983) that vowel qualities differ across varieties of Spanish; O'Rourke (2010) has shown that the same {a e i o u} system in speakers of Peruvian Spanish from Lima even differs significantly from that of their countrymates in Cuzco. Yang (1996: 259) notes that Korean males and females have the same phonemic systems, but "Korean female speakers produced vowels with a wider range of jaw movement than the male speakers." Traunmüller (1984) discusses how the proportionally longer pharynx in males can produce differences in F_2, while Ryalls and Lieberman (1982) and Diehl et al. (1996) attribute female peripherality to compensation for higher F_0. Traunmüller (1998) in turn attributes this higher F_0 to the relation of speaking energy and energy supply to linear body size (1:5 and 1:2 respectively). It has also been shown that bilingual speakers often implement the same phoneme in their two languages differently, which proves that the difference cannot be attributed to the variation between individual vocal tracts (Disner 1983; Bradlow 1995).

Dispersion theories also wrongly predict the non-existence of languages of the sort in (14), which lack some or all of the point vowels.

(14) Systems lacking point vowels:
   a. Tagalog and Berber have none of the point vowels (Disner 1984)
   b. Squamish, Alabama, and Amuesha have only [a] of the point vowels (Disner 1984)
   c. Many languages have [o] rather than [u]: Chickasaw (Gordon et al. 2000), Mura-Pirahã (Sheldon 1974), Hupa (Golla 1970; Gordon 1996), Navajo (McDonough and Austin-Garrison 1994), Banawá (Ladefoged et al. 1997), Nuu-chah-nulth (Sapir and Swadesh 1939) and Wari' (MacEachern et al. 1997).

These facts are particularly problematic for OT-DT, which requires that all vowel systems include high vowels.

Systems consisting entirely of point vowels also pose problems for DT. The fact that {i e a u} is favored over {i a o u} is not derivable within DT, as Joanisse and Seidenberg (1998: 3) point out: both systems are equally dispersed, and thus they should be equally favored. Joanisse and Seidenberg propose that {i e a u} is preferred because the ability to stiffen the genioglossus muscle while making high front vowels reduces variation of [i], making the [i]~[e] contrast easier than the [u]~[o] contrast. They then argue that the system with the easier-to-distinguish contrast is more likely to be learned successfully, which results in the front-heavy system being favored cross-linguistically. To preview our discussion in the next section, this is precisely the type of approach that we espouse: it is an argument about learnability, not synchronic grammar, and it is able to decide between two systems that are not differentiable by DT.

DT also suffers from a conflation of acoustic distance with confusability. Diehl et al. (1996) have shown that, for a given set of formants, higher F_0 increases confusability. By virtue of this fact, two vowel pairs wherein the acoustic distance between the two members is identical can differ in confusability, if the F_0 for the vowels is not identical. Also, the potential for confusion is not symmetrical; short vowels tend to be perceived with lower first and second formants, which means that misinterpreting [ɑ] as [ɔ] is more likely than vice versa (ten Bosch 1995). The asymmetry of confusability and its uncertain relationship with acoustic distance is problematic for DT because maximum perceptual distinctiveness can no longer be equated with maximum dispersion, which calls the whole DT enterprise into question.

Pulling back from the details, one can see that the basic problem is that DT, as a surface-driven, functionalist typological analysis, misses both underlying generalizations and idiosyncratic subtleties. In the first case it misses deep generalizations of the sort that link the underlyingly identical[6] but superficially different inventories of Italian and Yoruba shown in (13), for example, and in the second it fails to deal with gaps in languages like Navajo, which typically result from accidents of history. The notion that synchronic sound systems are best understood through their diachronic origins dates back to the Neogrammarians, Otto Jespersen, Joseph Greenberg, and perhaps most notably Baudouin de Courtenay. It has long been understood that sound changes can produce a vowel system with no felicitous synchronic explanation. One such example involves an ATR contrast that spreads from consonants to adjacent vowels and then disappears entirely in the consonants, as happened in many dialects of Armenian (Vaux 1998). This change resulted in an F_1 contrast produced not from the structure of the vowel system itself, but from a now-lost property of the consonant system. As Lang and Ohala (1996) state, "[w]e should not ignore a purely historical factor in the shaping of languages' segment inventories: less common sounds such as glottal stops, aspirated stops, etc. may evolve from – split off from – the more common ones."

5 Evolutionary phonology as an alternative to dispersion theory

Taken together, the arguments presented in the previous sections and throughout several decades of literature on various incarnations of DT combine to suggest that DT is inherently too restrictive to account for the observed variety of vowel systems. We may nevertheless retain the basic intuition that a more dispersed system, or an easier to pronounce system, is easier to learn than one that is less so, as ten Bosch (1995) and Joanisse and Seiderberg (1998) have suggested; however, this does not entail that dispersion needs to be built into rankable synchronic constraints, nor does it entail that a less dispersed or harder to pronounce system *cannot* be learned, as Joanisse and Seiderberg claim. In fact, the entire array of attested vowel systems can be explained solely via the acquisition process, as argued by Ohala (1992), Hyman (1998), Hale and Reiss (2000a, 2000b), Harris and Lindsey (2000), and Blevins (2004). If we can explain the attested linguistic patterns with independently needed acquisition processes, there is no need to duplicate them in the grammar (Hale and Reiss 2000a, Hale and Reiss 2000b; Blevins 2004; Samuels 2011). With respect to vowel systems in particular, Harris and Lindsey (2000)

6 A reviewer suggests that Yoruba, which has ATR harmony, may utilize different features in the mid vowels than Italian, which does not (and, e.g., could use [tense/lax] instead). Even if these features are indeed different, and if it can further be argued that Italian uses [tense/lax], this does not explain all the differences observable in (13).

observe that "it is not immediately obvious how positional asymmetries in vowel distribution can simultaneously 'arise from' grammar-internal constraints (Beckman 1997: 7) and be motivated by grammar-external speech functioning. The constraints in question could have at best some intermediate place in a chain of causation." De Boer (2000: 2) adds that synchronic constraints do not make acquisition easier, or help explain it (see also Hale and Reiss 2008):

> "These functional explanations [NB: of the distribution of vowels] are not the full explanation, either. They assume that the systems of speech sounds one finds are the result of an optimization of one or more of the proposed criteria. However, it is not clear who is doing the optimization. Certainly children that learn a language do not do an optimization of the system of speech sounds they learn. Rather, they try to imitate their parents (and peers) as accurately as possible."

How then should the attested range of vowel systems be accounted for? We believe that the answer lies in Evolutionary Phonology (EP; Blevins 2004), which we view an extension of Ohala's listener-based analyses of sound change, Kiparsky's (1972) account of apparent phonological conspiracies, Pierrehumbert's (2001) exemplar-based model, and to some degree Boersma's (1997) Functional Phonology.[7] The core of EP is simple and intuitive: a system that is easier to learn than a competitor is predicted to become more common than that competitor over evolutionary time. This is a theory of "probable languages" in the sense of Newmeyer (2005); it is essentially a form of survival of the fittest. Crucially, just as application of survival of the fittest in the animal kingdom does not require that a less fit creature disappear, so its linguistic analogue does not require that harder-to-learn systems disappear (or fail to appear in the first place); they are merely predicted to become less common over evolutionary time.

At the heart of EP is an exemplar-based model of sound change (based on Pierrehumbert 2001) overlaid on a continuous psycho-acoustic vowel space, thus permitting gradual change (as opposed to theories that permit only discrete changes in terms of phonological features). In this model, when a speaker wants to produce a vowel – /u/, for example – he attempts to produce the "best" exemplar of /u/ that he has previously heard. Crucially, "best" in this context means "most likely to be categorized as /u/." The likelihood of being categorized as [u] comes

7 The model proposed here diverges from Boersma's (1997) theory of Functional Phonology, which purports to be able to explain the symmetries and gaps in inventories of vowels and consonants. In Boersma's model, the attested symmetries are language-specific results of general human limitations on the acquisition of perceptual categorization and motor skills. The gaps are results of local hierarchies of articulatory effort, perceptual contrast, and perceptual confusion. The proposal developed here is largely in agreement with Boersma with regards to how asymmetries may arise, but departs from his theory's explanation of the gaps, for which we seek no functionalist synchronic explanations.

from the exemplar's summed similarity to other exemplars of [u] taken as a fraction of its summed similarity to exemplars of all vowels (Pierrehumbert 2001). The other members of the vowel system thus become crucial to the categorization process. In a five-vowel system of /i, e, a, o, u/, /u/ may begin to front because it is closer in the acoustic space to /o/ than to /i/. This may explain why /u/ exhibits some degree of fronting in a wide variety of languages and dialects (e.g., French, Greek, Swedish, São Miguel Portuguese, and most varieties of American English; see Haudricourt and Juilland 1949; Labov et al. 1997), as it is to be expected when /i/ is not the closest neighbor to /u/.[8] This creates a feedback loop enhancing distinctness in perception and production, with the effect of distributing the vowels evenly throughout the perceptual space, because the system will reach equilibrium only when the best exemplar of each category is also its mean (see Samuels 2007 on the implications of this view for vocalic chain shifting).

EP categorizes phonetically-motivated sound changes into three types: CHANGE, CHANCE, and CHOICE. CHANGE includes cases in which a learner analyzes an utterance as having an underlying representation that is perceptually similar to but distinct from the underlying representation of the speaker, due to mishearing. For example, /anpa/ may be heard as [ampa] and therefore analyzed by the listener as /ampa/ due to the weak cues indicating the place of the preconsonantal nasal. CHANCE changes are those in which the listener analyzes an ambiguous signal in a way that differs from the speaker's underlying form. For example, [ʔa̰ʔ] may be analyzed as /ʔa̰ʔ/, /ʔa/, /aʔ/, or /a̰ʔ/. Finally, CHOICE can occur when there is variation in the pronunciation of a particular underlying form and the listener posits a different underlying form along with a different phonological rule governing the variation. An example would be an underlying form /kakata/ with an optional vowel shortening/deletion rule producing [kăkata] and [kkata] instead being analyzed as /kkata/ with an optional epenthesis rule.

To take an example of how CHOICE can operate in the evolution of vowel systems, consider the case of Yurok /e/ and /a/ (Blevins 2003). The former **e* phoneme in Yurok has split into /e/ and /a/ following the loss of the conditioning environment for a rule lowering /e/ to [a]. Specifically, the loss of /h/ before laryngealized consonants rendered opaque the complementary distribution of these erstwhile allophones in some classes of verbs, resulting in phonemicization of /a/. Blevins further reports on a sound change in progress that lowers /e/ to /a/ before rhotics, /w/, and the fricatives /s, c, ɬ/. This is a gradient phenomenon

8 Note that /u/ fronting to [y] is unexpected from a markedness perspective; see Calabrese (2005) on this diachronic development as an instance of "emergence of the marked" and on the possibility that this fronting process stems in part from the articulatory difficulty of producing a [+ATR] back vowel.

that produces considerable allophonic variation on the continuum from [e] to [a] through [ɛ] and [æ], sometimes even produced as a diphthong in careful speech. In the aftermath of the /e/~/a/ phonemic split, the opportunity for CHOICE emerged: listeners can now posit either /e/ or /a/ as an underlying representation for the variable vowels. This may explain the variation observed across speakers. Blevins hypothesizes that the establishment of a phonemic contrast between /e/ and /a/ opened the door for more instances of **e* to shift into the /a/ category. Note that /a/ did not emerge out of any dispersion-related principles such as maximal distinctiveness of contrasts, but rather due to loss of the conditioning environment for a predictable allophonic distribution of [e] vs. [a] in one small corner of the grammar.

It is also emphasized in EP that common synchronic sound patterns are those that arise from common sound changes, while rare sound patterns arise from unusual sound changes or rare sequences of sound changes. We emphasize in this regard that what is "diachronically possible" (i.e., what phonological patterns can be produced by possible linguistic changes) may be a subset of what is "computationally possible" (i.e., see Hale and Reiss 2000a, Hale and Reiss 2000b). The overarching factor that we have called "ease of learnability" crosscuts this distinction, since cognitive factors are intimately involved in the language acquisition process, from which linguistic change results. As our understanding of perception and cognition increases, the exact interplay of these factors will become better defined.

EP has the advantage over all-or-nothing synchronic models of providing a straightforward account for cross-linguistic tendencies, while simultaneously allowing a place for accidents of history. This problem of how to deal with 'unnatural' patterns is endemic in phonology, as has been recognized and debated for decades. Idiosyncrasies of history inevitably intrude in language; the synchronic phonological component must be able to accommodate these, but need not incorporate them as synchronic functional constraints.

6 Conclusions

We have argued here that Dispersion Theory is insufficient to characterize the variety of vowel systems found cross-linguistically. The issue at the core of many of the arguments we have presented here – above and beyond the issues we have raised concerning the implementation of various types of dispersion theories – is that the properties of vowel systems are not predictable simply on the basis of opposed articulatory and perceptual concerns. That is to say, at least partially due to accidents of history, inventories are not always symmetrical, do not make ideal use of the articulatory/perceptual space, and so forth. Attempts to force synchronic explanations onto phenomena that are historical or evolutionary miss important phonological generalizations, fail to account for the peculiarities of individual languages, produce unnecessary redundancy, and obscure the actual driving force behind both the similarities and the differences that we observe in human language. An Evolutionary Phonology account of the evolution of vowel systems is more permissive than Dispersion Theory, which is desirable, yet still predicts that some vowel systems will be more common than others for reasons of articulation, perception, and other properties that may affect their learnability.

References

Ahn, Sang-Cheol. 2000. Lowering of nasal vowels in French: Enhancement and dispersion. *Proceedings of the 8th National Conference on Speech Sciences*, 113–120. Seoul: The Association of Speech Sciences, Seoul National University.

Ahn, Sang-Cheol, 2002a. A dispersion account on middle Korean vowel shifts. In Norkio M. Akatsuka & Susan Strauss (eds.), *Japanese/Korean Linguistics 10*, 237–250. Stanford: Center for the Study of Language and Information.

Ahn, Sang-Cheol. 2002b. An optimality approach to the Great Vowel Shift. *Korean Journal of Linguistics* 27. 153–170.

Anderson, Stephen R. 1978. Syllables, segments, and the Northwest Caucasian languages. In Alan Bell & Joan Hooper (eds.), *Syllables and Segments*, 57–89. Amsterdam: North Holland.

Baumann, Monika. 1996. Lexical syllables precede postlexical syllables: Psycholinguistic evidence for two levels of syllable structure. Paper presented at the Eighth International Phonology Meeting, Vienna.

Becker-Kristal, Roy. 2009. *Acoustic typology of vowel inventories and Dispersion Theory: Insights from a large cross-linguistic corpus*. Los Angeles, CA: University of California dissertation.

Beckman, Jill. 1997. Positional faithfulness, positional neutralisation, and Shona vowel harmony. *Phonology* 14. 1–46.

Bender, Byron. 1968. Marshallese phonology. *Oceanic Linguistics* 7. 16–35.

Blevins, Juliette. 2003. One case of contrast evolution in the Yurok vowel system. *International Journal of American Linguistics* 69. 135–150.

Blevins, Juliette. 2004. *Evolutionary phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press.

Blevins, Juliette. 2008. Consonant epenthesis: Natural and unnatural histories. In Jeff Good (ed.), *Language Universals and Language Change*, 79–107. Oxford: Oxford University Press.

Boer, Bart de. 2000. Emergence of vowel systems through self-organisation. *AI Communications* 13. 27–39.

Boer, Bart de. 2001. *The origins of vowel systems*. Oxford: Oxford University Press.

Boersma, Paul. 1997. *Inventories in functional phonology*. Manuscript, University of Amsterdam.

Bosch, Louis ten. 1991. *On the structure of vowel systems: Aspects of an extended vowel model using effort and contrast*. Amsterdam: University of Amsterdam dissertation.
Bosch, Louis ten. 1995. Lexically-based vowel dispersion: A case study for Dutch. *Institute of Phonetic Sciences* 19. 39–50.
Bradlow, Ann R. 1995. A comparative acoustic study of English and Spanish vowels. *Journal of the Acoustical Society of America* 97. 1916–1924.
Calabrese, Andrea. 2005. *Markedness and economy in a derivational model of phonology*. Berlin: Mouton de Gruyter.
Campbell, Lyle. 1986. Testing phonology in the field. In John J. Ohala and Jeri J. Jaeger (eds.), *Experimental Phonology*, 163–173. Orlando, FL: Academic Press.
Carré, René. 1996. Prediction of vowel systems using a deductive approach. *International Conference on Spoken Language Processing (ICSLP)* 96. 434–437.
Carré, René, Björn Lindblom & Peter MacNeilage. 1994. Acoustic contrast and the origin of the human vowel space. *Journal of the Acoustical Society of America* 95. S2924.
Carré, René & Maria Mody. 1997. Prediction of vowel and consonant place of articulation. *Special Interest Group on Computational Phonology (SIGPHON)* 3. 26–32.
Carré, René & Mohamad Mrayati. 1992. Distinctive regions in acoustic tubes: Speech production modeling. *Journal d'Acoustique* 5. 141–159.
Carré, René & Mohamad Mrayati. 1995. Vowel transitions, vowel systems, and the Distinctive Region Model. In Carmen Dobrovie-Sorin (ed.), *Levels in Speech Communication*, 73–89. Amsterdam: Elsevier.
Casali, Roderic F. 1995. An overview of the Nawuri verbal system. *Journal of West African Languages* 25. 63–86.
Cho, Taehong, Sun-Ah Jun, Seung-Chul Jung, & Peter Ladefoged. 2000. The vowels of Cheju. *UCLA Working Papers in Phonetics* 98. 81–94.
Choi, John. 1992. *Phonetic underspecification and target interpolation: An acoustic study of Marshallese vowel allophony*. Los Angeles: University of California dissertation.
Colarusso, John. 1988. *The Northwest Caucasian languages: A phonological survey*. New York: Garland.
Colarusso, John. 1989. East Circassian (Kabardian dialect). In Hewitt, B. George (ed.), *The Indigenous Languages of the Caucasus, Vol. 2: The Northwest Caucasian Languages,* 261–355. Delmar, NY: Caravan Books.
Colarusso, John. 1992. *A grammar of the Kabardian language*. Calgary: University of Calgary Press.
Crothers, J. 1978. Typology and universals of vowel systems. In Joseph Greenberg (ed.), *Universals of Human Language vol. 2*, 93–152. Stanford: Stanford University Press.
Diehl, Randy, Björn Lindblom, Kathryn Hoemeke, & Richard Fahey. 1996. On explaining certain male-female differences in the phonetic realization of vowel categories. *Journal of Phonetics* 24. 187–208.
Disner, Sandra. 1983. Vowel quality: The relation between universal and language-specific factors. *UCLA Working Papers in Phonetics* 58. Los Angeles: UCLA Dept. of Linguistics.
Disner, Sandra. 1984. Insights on vowel spacing. In Ian Maddieson (ed.), *Patterns of Sounds*, 136–155. Cambridge: Cambridge University Press.
Dupoux, Emanuel, Erika Parlato, Sonia Frota, Yuki Hirose, & Sharon Peperkamp. 2011. Where do illusory vowels come from? *Journal of Memory and Language* 64. 199–210.
Ferrero, Franco E. 1972. Caratteristiche acustiche dei fonemi vocalici italiani. *Parole e Metodi* 3. 9–31.

Flemming, Edward. 1995. *Auditory representation in phonology*. Los Angeles: University of California dissertation.
Flemming, Edward. 1996. Evidence for constraints on contrast: The dispersion theory of contrast. *UCLA Working Papers in Phonology* 1. 86–106.
Flemming, Edward. 2004. Contrast and perceptual distinctiveness. In Bruce Hayes, Robert Kirchner, & Donca Steriade (eds.), *The Phonetic Bases of Phonological Markedness*, 232–276. Cambridge: Cambridge University Press.
Golla, Victor. 1970. *Hupa grammar*. Berkeley, CA: University of California dissertation.
Gordon, Matthew. 1996. The phonetic structures of Hupa. *UCLA Working Papers in Linguistics* 93. 164–187.
Gordon, Matthew, Pam Munro & Peter Ladefoged. 2000. Some phonetic structures of Chickasaw. *Anthropological Linguistics* 42. 366–400.
Hale, Mark. 1997. Otiosity. *Linguist List* 8. 690.
Hale, Mark, Madelyn Kissock & Charles Reiss. 1998. Evaluating the empirical basis of output-output correspondence. *West Coast Conference on Formal Linguistics (WCCFL)* 23. 137–147.
Hale, Mark & Charles Reiss. 2000a. Substance abuse and dysfunctionalism: Current trends in phonology. *Linguistic Inquiry* 31. 157–169.
Hale, Mark & Charles Reiss. 2000b. Phonology as cognition. In Noel Burton-Roberts, Philip Carr, & Gerard Docherty (eds.), *Phonological Knowledge*, 161–184. Oxford: Oxford University Press.
Hale, Mark & Charles Reiss. 2008. *The phonological enterprise*. Oxford: Oxford University Press.
Halle, Morris. 1970. Is Kabardian a vowel-less language? *Foundations of Language* 6. 95–103.
Hallé, Pierre A., Céline Chéreau, & Juan Segui. 2000. Where is the /b/ in “absurde” [apsyrd]? It is in French listeners’ minds. *Journal of Memory and Language* 43. 618–639.
Harris, John & Geoff Lindsey. 2000. Vowel patterns in mind and sound. In Noel Burton-Roberts, Philip Carr, & Gerard Docherty (eds.), *Phonological Knowledge*, 185–205. Oxford: Oxford University Press.
Haudricourt, André-Georges & Alphonse G. Juilland. 1949. *Essai pour une histoire structural du phonétisme français*. Paris: Klincksieck.
Hewitt, B. George. 1979. *Abkhaz*. Amsterdam: North Holland.
Hyman, Larry. 1970. How concrete is phonology? *Language* 46. 58–76.
Hyman, Larry, 1998. Positional prominence and the ‘positional trough’ in Yaka. *Phonology* 15. 41–75.
Itô, Junko & Armin Mester. 2003. Lexical and postlexical phonology in optimality theory: Evidence from Japanese. In Gisbert Fanselow & Caroline Féry (eds.), *Resolving Conflicts in Grammars*, 183–207. *Linguistische Berichte* Sonderheft 11.
Jakobson, Roman. 1941. *Kindersprache, Aphasie, und allgemeine Lautgesetze*. Mouton: The Hague.
Jansen, Wouter. 2002. Review of de Boer 2001. *Linguist List* 13. 541.
Joanisse, Marc & Mark Seidenberg. 1998. Functional bases of phonological universals: A connectionist approach. *Berkeley Linguistics Society (BLS)* 18. 335–345.
Karttunen, Lauri. 1998. The proper treatment of optimality in computational phonology. In FSMNLP ’98. *International Workshop on Finite-state Methods in Natural Language Processing*. Bilkent University, Ankara, Turkey. cmp-lg/9804002. 1–12.
Kawasaki, Haruko. 1982. *An acoustical basis for universal constraints on sound sequences*. Berkeley, CA: University of California dissertation.

Kiparsky, Paul. 1972. Explanation in phonology. In Stanley Peters (ed.), *Goals of Linguistic Theory*, 189–227. Englewood, NJ: Prentice Hall.
Kiparsky, Paul. 2000. Opacity and cyclicity. *The Linguistic Review* 17. 351–367.
Labov, William, Sharon Ash, & Charles Boberg. 1997. A national map of the regional dialects of American English. http://www.ling.upenn.edu/phono_atlas/NationalMap/NationalMap.html (accessed 22 July, 2014).
Ladefoged, Peter, Jenny Ladefoged, & Daniel Everett. 1997. Phonetic structures of Banawá, an endangered language. *Phonetica* 54. 94–111.
Lang, Carrie & John Ohala. 1996. Temporal cues for vowels and universals of vowel inventories. *International Conference on Spoken Language Processing (ICSLP)* 96. 434–437.
Laycock, Donald. 1965. *The Ndu language family*. Canberra: Australian National University.
Liljencrants, Johan & Björn Lindblom. 1972. Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language* 48. 839–862.
Lindau, Mona. 1975. Features for vowels. *UCLA Working Papers in Phonetics* 30. Los Angeles: UCLA Phonetics Lab.
Lindau, Mona & Patricia Wood. 1977. Acoustic vowel spaces. *UCLA Working Papers in Phonetics* 38. 41–48.
Lindblom, Björn. 1975. Experiments in sound structure. Revue de Phonétique Appliquée, Université de l'État Mons, Belgique, *International Congress of Phonetic Sciences (ICPhS*) 51. 155–189.
Lindblom, Björn. 1986. Phonetic universals in vowel systems. In John J. Ohala & Jeri J. Jaeger (eds.), *Experimental Phonology*, 13–44. Orlando: Academic Press.
Lindblom, Björn. 1990. On the notion "possible speech sound". *Journal of Phonetics* 18. 135–152.
Lindblom, Björn. 1992. Phonological units as adaptive emergents of lexical development. In Charles A. Ferguson, Lise Menn, & Carol Stoel-Gammon (eds.), *Phonological Development: Models, Research, Implications*, 131–163. Timonium, MD: York.
Lindblom, Björn. 2000. Developmental origins of adult phonology: The interplay between phonetic emergents and the evolutionary adaptations of sound patterns. *Phonetica* 57. 297–314.
Lindblom, Björn & Ian Maddieson. 1988. Phonetic universals in consonant systems. In Larry Hyman & C. Li (eds.), *Language, Speech, and Mind*, 62–78. London: Routledge.
Lindblom, Björn & Johan Sundberg. 1971. Acoustical consequences of lip, tongue, jaw, and larynx movement. *Journal of the Acoustical Society of America* 50. 1166–1179.
Livijn, Peder. 2000. Acoustic distribution of vowels in differently sized inventories – hot spots or adaptive dispersion? *Swedish Phonetics Conference* 13. 93–96.
MacEachern, Margaret, Barbara Kern, & Peter Ladefoged. 1997. Wari' phonetic structures. *Journal of Amazonian Languages* 1. 3–28.
Maddieson, Ian. 1984. *Patterns of sounds*. Cambridge: Cambridge University Press.
Martinet, André. 1955. *Economie des changements phonétiques*. Berne: Francke.
McCarthy, John. 2003. OT constraints are categorical. *Phonology* 20. 75–138.
McCarthy, John. 2005. Optimal paradigms. *Linguistics Department Faculty Publication Series* 55. Amherst, MA: University of Massachusetts.
McCarthy, John. 2007. *Hidden generalizations: Phonological opacity in Optimality Theory*. London: Equinox.
McDonough, Joyce & Martha Austin-Garrison. 1994. Vowel enhancement and dispersion in the vowel space of Western Navajo: A study of traditional Navajo speakers. *UCLA Working Papers in Phonetics* 87. 93–104.

McDonough, Joyce, Peter Ladefoged, & Helen George. 1993. Navajo vowels and universal phonetic tendencies. *UCLA Working Papers in Phonetics* 84. 143–150.
McMahon. April 2000. *Change, chance, and optimality*. Oxford: Oxford University Press.
Mrayati, Mohamad, René Carré, and B. Guérin. 1988. Distinctive regions and modes: A new theory of speech production. *Speech Communication* 7. 257–286.
Newmeyer, Frederick. 2005. *Possible and probable languages: A generative perspective on linguistic typology*. Oxford: Oxford University Press.
Ní Choisáin, Máire & Jaye Padgett. 1997. *Markedness, segment realization, and locality in spreading*. Report LRC-97-01, Linguistics Research Center, Department of Linguistics, University of California Santa Cruz.
Ohala, John. 1975. Phonetic explanations for nasal sound patterns. In Charles Ferguson, Larry Hyman & John J. Ohala (eds.), *Nasalfest: Papers from a Symposium on Nasals and Nasalization*, 289–316. Stanford: Language Universals Project.
Ohala, John. 1992. What's cognitive, what's not, in sound change. In Günter Kellermann & Michael Morrissey (eds.), *Diachrony within Synchrony: Language History and Cognition*, 309–355. Frankfurt: Peter Lang Verlag.
Ohala, John. 2005. Phonetic explanations for sound patterns. Implications for grammars of competence. In William Hardcastle & Janet Mackenzie Beck (eds.), *A Figure of Speech. A Festschrift for John Laver*, 23–38. London: Erlbaum.
O'Rourke, Erin. 2010. Dialect differences and the bilingual vowel space in Peruvian Spanish. *Conference on Laboratory Approaches to Spanish Phonology* 4. 20–30.
Padgett, Jaye. 2004. Russian vowel reduction and Dispersion Theory. *Phonological Studies* 7. 81–96.
Pallier, Christophe, Nuria Sebastian-Galles, & Angels Colome. 1999. Phonological representations and repetition priming. *European Conference on Speech Communication and Technology* 4. 1907–1910.
Pierrehumbert, Janet B. 2001. Exemplar dynamics: Word frequency, lenition and contrast. In Joan Bybee & Paul Hopper (eds.), *Frequency and the Emergence of Linguistic Structure*, 137–157. Amsterdam: John Benjamins.
Quillis, Antonio & Manuel Esgueva. 1983. Realización de los fonemas vocálicos españoles en posición fonética normal. In M. Esgueva & M. Cantarero (eds.), *Estudios de Fonética*, 159–251. Madrid: Consejo Superior de Investigaciones Cientificas.
Recasens, Daniel & Aina Espinosa. 2009. Dispersion and variability in Catalan five and six peripheral vowel systems. *Speech Communication* 51. 240–258.
Rice, Keren. 1999. Featural markedness in phonology: Variation. *Glot International* 4(7). 3–6, 4 (8). 3–7.
Riggle, Jason. 1999. *Relational markedness in Bantu vowel height harmony*. Santa Cruz: University of California thesis.
Roark, Brian. 2001. Explaining vowel inventory tendencies via simulation: Finding a role for quantal locations and formant normalization. *Northeast Linguistic Society (NELS)* 31. 419–434.
Ryalls, John H. & Philip Lieberman. 1982. Fundamental frequency and vowel perception. *Journal of the Acoustical Society of America* 72. 1631–1634.
Samuels, Bridget. 2007. Review of Blevins (2004) *Evolutionary phonology*. *Biolinguistics* 1. 130–134.
Samuels, Bridget. 2011. *Phonological architecture: A biolinguistic perspective*. Oxford: Oxford University Press.

Sanders, Nathan. 2002. Dispersion in OT: Color contrast in Middle Polish nasal vowels. *West Coast Conference on Formal Linguistics (WCCFL)* 21. 415–428.
Sapir, Edward & Morris Swadesh. 1939. *Nootka texts: Tales and ethnological narratives, with grammatical notes and lexical materials*. Philadelphia: Linguistic Society of America.
Schwartz, Jean-Luc, Christian Abry, Louis-Jean Boë, Nathalie Valle & M-A. Cathiard. 1999. Economy and Phonology in the perceptual Dispersion-Focalization Theory and the Theory of Perception for Action-Control. *First International Symposium on Linguistics (LICSSOL1)*. http://www.isc.cnrs.fr/schwartz.htm (accessed 22 July, 2014).
Schwartz, Jean-Luc, Louis-Jean Boë, Nathalie Valle, & Christian Abry. 1997a. Major trends in vowel system inventories. *Journal of Phonetics* 25. 233–254.
Schwartz, Jean-Luc, Louis-Jean Boë, Nathalie Valle, & Christian Abry. 1997b. The Dispersion-Focalization Theory of vowel systems. *Journal of Phonetics* 25. 255–286.
Sheldon, Steven N. 1974. Some morphophonemic and tone perturbation rules in Mura-Pirahã. *International Journal of American Linguistics* 40. 279–282.
Smeets, Rieks. 1984. *Studies in West Circassian Phonology and Morphology*. Leiden: Hakuchi Press.
Son, R. van, Florien Koopmans-van Beinum, & Louis Pols. 1998. Efficiency as an organizing principle of natural speech. *International Conference on Spoken Language Processing (ICSLP)* 98.
Staalsen, Philip. 1966. The phonemes of Iatmul. *Papers in New Guinea Linguistics* 5. 69–76.
Stevens, Kenneth. 1972. The quantal nature of speech: Evidence from articulatory-acoustic data. In Edward E. David & Peter B. Denes (eds.), *Human Communication: A Unified View*, 51–66. New York: McGraw-Hill.
Stevens, Kenneth. 1989. On the quantal nature of speech. *Journal of Phonetics* 17. 3–45.
Stevens, Kenneth & S. Jay Keyser. 1989. Primary features and their enhancement in consonants. *Language* 65. 81–106.
Traunmüller, Hartmut. 1984. Articulatory and perceptual factors controlling the age- and sex-conditioned variability in formant frequencies of vowels. *Speech Communication* 3. 49–61.
Traunmüller, Hartmut. 1998. Modulation and demodulation in production, perception, and imitation of speech and bodily gestures. *FONETIK* 98. 40–43.
Trubetzkoy, Nikolai. 1939. *Grundzüge der Phonologie*. Prague: s. n.
Vallée, Nathalie. 1994. *Systèmes vocaliques: de la typologie aux prédictions*. Grenoble: Université Stendhal dissertation.
Vallée, Nathalie, Louis-Jean Boë, Jean-Luc Schwartz & Pierre Badin. 2001. The weight of substance in phonological structures of the world's languages. Manuscript, Université Stendahl, Grenoble.
Vaux, Bert. 1998. *The phonology of Armenian*. Oxford: Oxford University Press.
Vaux, Bert & Zihni Psiypa. 1997. The Cwyzhy Dialect of Abkhaz. In Susumu Kuno, Bert Vaux, & Steve Peter (eds.), *Harvard Working Papers in Linguistics* 6. Cambridge, MA: Harvard University Linguistics Department.
Vaux, Bert & Bridget Samuels. 2003. Consonant epenthesis and hypercorrection. Manuscript, Harvard University.
Yang, Byunggon. 1996. A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics* 24. 245–261.