Phonology **24** (2007) 433–459.
doi:10.1017/S0952675707001285 Printed in the United Kingdom

*Constraint weighting and constraint domination: a formal comparison**

Jie Zhang
University of Kansas

The advent of Optimality Theory has revived the interest in articulatorily and perceptually driven markedness in phonological research. To some researchers, the cross-linguistic prevalence of such markedness relations is indication that synchronic phonological grammar should include phonetic details. However, there are at least two distinct ways in which phonetics can be incorporated in an optimality-theoretic grammar: traditional constraint domination and Flemming (2001)'s proposal that the costs of constraint violations should be weighted and summed. I argue that constraint weighting is unnecessary as an innovation in Optimality Theory. The arguments are twofold. First, using constraint families with intrinsic rankings, constraint domination formally predicts the same range of phonological realisations as constraint weighting. Second, with proper constraint definitions and rankings, both the additive effect and the locus effect predicted by constraint weighting can be replicated in constraint domination.

With the advent of Optimality Theory (Prince & Smolensky 1993), there has been a revived interest in articulatorily and perceptually driven markedness in phonological research. The cross-linguistic prevalence of such markedness relations is indication to some researchers that synchronic phonological grammar should include phonetic details (e.g. Boersma 1998, Steriade 1999, Kirchner 2000, Flemming 2001, Zhang 2002). However, there are at least two distinct ways in which phonetics can be incorporated in an optimality-theoretic grammar: while Boersma, Steriade, Kirchner and Zhang adhere to the traditional practice of constraint domination by using universal rankings within constraint families, Flemming (2001) proposes that the costs of constraint violations should be weighted and summed.

I formally compare the predictions of constraint weighting and constraint domination in this paper and argue that constraint weighting is unnecessary as an innovation in Optimality Theory. The arguments are twofold. First, using constraint families with intrinsic rankings, constraint domination formally predicts the same range of phonological realisations

* I am grateful to three anonymous reviewers and an associate editor of *Phonology* for their valuable comments on earlier drafts on the paper. All remaining errors are mine.

as constraint weighting. Second, with proper constraint definitions and rankings, both the additive effect and the locus effect predicted by constraint weighting can be replicated in constraint domination. I also point out two minor calculation errors in Flemming (2001).

1 The Flemming model

The traditional assumption for phonological representations is that they should only include features that can be used contrastively in *some* language, and the details of phonetic implementation are the consequence of universal principles (Chomsky & Halle 1968). The discovery of language-specificness in phonetic patterns such as coarticulation (e.g. Keating 1985, 1990, Manuel 1990, Flemming 1997) put phonologists in a dilemma: either the phonological component in the grammar needs to be augmented to include phonetic details or an additional phonetic component needs to be added to the grammar. Although many researchers, including Keating (1985, 1990), Pierrehumbert & Beckman (1988), Pierrehumbert (1990, 1991), Cohn (1993) and Zsiga (1997), have opted for the latter, other researchers have argued that phonetic knowledge is part of linguistic knowledge proper and should be encoded in the phonological grammar, either directly (Boersma 1998, Steriade 1999, Kirchner 2000, Flemming 2001, Zhang 2002) or mediated via the lexicon (Pierrehumbert 2000, 2001, 2002).[1]

Flemming (2001) makes a significant contribution to this debate by not only presenting arguments for the phonetics-in-phonology approach, for example, the parallel behaviour between phonological and phonetic phenomena and the desire to avoid duplicating sound representations at two different levels, but also proposing an explicit model that provides a unified analysis for parallel phonological and phonetic phenomena such as neutralising assimilation and gradient coarticulation. I review his model below.

1.1 Non-neutralising assimilation

Flemming uses consonant–vowel assimilation in F2 to illustrate the model. He considers a CV sequence in which the F2 targets for the consonant and the vowel are L and T respectively, and the actual realisation of F2 is $F2(C)$ at consonant release and $F2(V)$ at the vowel steady state. The formant transition is schematised in Fig. 1 (from Flemming 2001: 17). The actual realisations $F2(C)$ and $F2(V)$ reflect the compromise between faithful renditions of the consonant locus and vowel target and the least effortful transition between the consonant and the vowel. Flemming formulates three constraints: IDENT(C), IDENT(V) and MINIMISEEFFORT

[1] See Hayes *et al.* (2004) for a review of this debate.

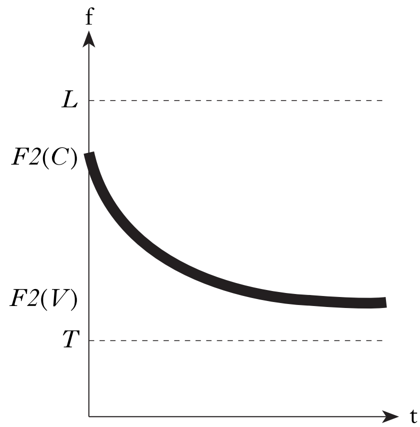


Figure 1

Schematic representation of a CV F2 transition. L is the consonant locus, $F2(C)$ is the actual value of F2 at consonant release, T is the target for the vowel and $F2(V)$ is the actual value of F2 at the vowel steady state.

(MinEffort), whose requirements and costs of violation are defined as in (1) (Flemming 2001 : 19).

(1) *Constraints for consonant–vowel F2 transitions*

	Definition	*Cost of violation*
Ident(C)	$F2(C) = L$	$w_c(F2(C) - L)^2$
Ident(V)	$F2(V) = T$	$w_v(F2(V) - T)^2$
MinEffort	$F2(C) = F2(V)$	$w_e(F2(C) - F2(V))^2$

Intuitively, Ident(C) requires the actual realisation of F2 at consonant release to be identical to the consonant locus, and Ident(V) requires the target F2 of the vowel to be reached at the vowel steady state. MinEffort, on the other hand, requires zero transition between the consonant and vowel F2 realisations. These constraints are not 'ranked' to predict the output of the grammar. Instead, each constraint is associated with a 'cost of violation'; e.g. the cost of violation of Ident(C) is the square of the difference between $F2(C)$ and L, weighted by a coefficient w_c. The grammar, consequently, is constructed to achieve the least overall cost of violation, which is the sum of the cost of violation for each constraint, as shown in (2) (Flemming 2001 : 20).

(2) $Cost = w_c(F2(C) - L)^2 + w_v(F2(V) - T)^2 + w_e(F2(C) - F2(V))^2$

To visualise the effect of $F2(C)$ and $F2(V)$ on the overall cost, Flemming provides a three-dimensional graph in which the x, y and z axes represent $F2(C)$, $F2(V)$ and *Cost* respectively, as in Fig. 2 (Flemming 2001 : 21).

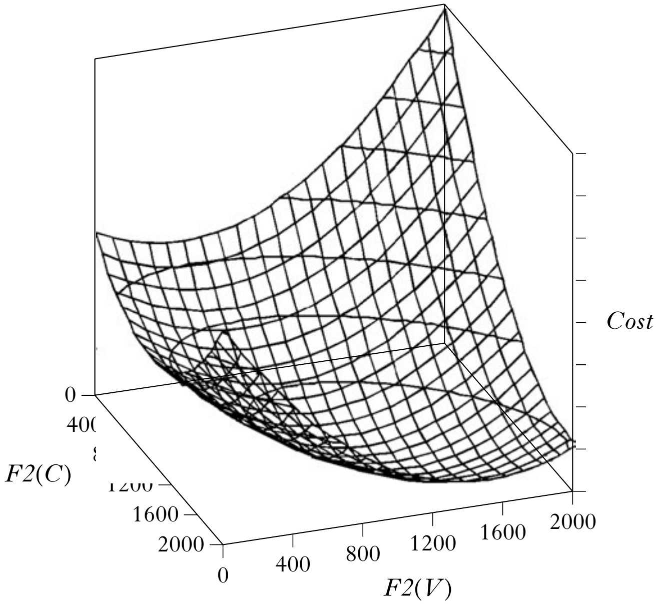


Figure 2
Cost plotted against *F2*(*C*) and *F2*(*V*), with $L = 1700$ Hz, $T = 1000$ Hz, and all weights set to 1. The minimum cost is achieved when *F2*(*C*) = 1467 Hz and *F2*(*V*) = 1233 Hz.

To derive the *F2*(*C*) and *F2*(*V*) values that render the minimum cost, we take the partial derivatives of *Cost* along each dimension, which gives rise to two formulas with respect to *F2*(*C*) and *F2*(*V*), as shown in (3).

(3) a. $$\frac{\partial(Cost)}{\partial(F2(C))} = 0: \quad F2(C) = \frac{w_e}{w_c + w_e}(F2(V) - L) + L$$

b. $$\frac{\partial(Cost)}{\partial(F2(V))} = 0: \quad F2(V) = \frac{w_e}{w_v + w_e}(F2(C) - T) + T$$

The values of *F2*(*C*) and *F2*(*V*) obtained from the two formulas in (3) are shown in (4).[2]

(4) a. $$F2(C) = u_c(L - T) + L, \text{ where } u_c = \frac{-w_e w_v}{w_e w_c + w_v w_c + w_e w_v}$$

b. $$F2(V) = u_v(L - T) + T, \text{ where } u_v = \frac{w_e w_c}{w_e w_c + w_v w_c + w_e w_v}$$

[2] In Flemming (2001), the u_c value is $w_e w_v / w_e w_c + w_v w_c + w_e w_v$; but the result of my calculation requires an extra minus sign, as shown in (4). This minus sign can be justified as follows: according to the first equation in (4), $u_c = (F2(C) - L/L - T)$; since $F2(C) < L$ and $L > T$, u_c must be negative.

This model can predict the cross-linguistic variability of consonant to vowel coarticulation by adjusting the weight coefficients. For example, the greater the w_v value, the more vowel undershoot will contribute to the overall cost; thus it is more important to keep the vowel undershoot small. Such cross-linguistic variability is empirically attested in Flemming's (1997, 2001) phonetic studies of /u/ undershoot between coronals in four languages.

1.2 Neutralisation

To capture neutralising processes that have phonetic bases and their non-neutralising parallels, the model above is supplemented with constraints on contrasts in the forms of MINDIST = Δ and MAXIMISECONTRASTS (MAXCONTRASTS), with the former requiring contrastive entities to be of a minimum perceptual distance, and the latter awarding extra contrasts maintained with negative cost values. The categorical nature of neutralising processes is then the result of a cost calculation in which the advantage afforded by keeping a contrast (negative MAXCONTRASTS cost) is outweighed by the cost induced by it (positive MINDIST = Δ and MINEFFORT costs).

Flemming again illustrates the addition to the model with consonant–vowel assimilation, specifically the F2 realisation of /u/ in the context of a coronal and the possibility of neutralising /u/ with /y/. He considers the following constraints.

(5) *Constraints for consonant–vowel F2 transitions, with possibility of neutralisation*

	Definition	*Cost of violation*
IDENT(C)	$F2(\text{t}) = L_t$	$w_c(F2(\text{t}) - L_t)^2$
MINEFFORT	$F2(\text{t}) = F2(\text{u})$	$w_e(F2(\text{t}) - F2(\text{u}))^2$
MINDIST = Δ	$\lvert F2(\text{y}) - F2(\text{u})\rvert \geq \Delta$	$w_v(\lvert F2(\text{y}) - F2(\text{u})\rvert - \Delta)^2$ for $\lvert F2(\text{y}) - F2(\text{u})\rvert < \Delta$
MAXCONTRASTS		$-w_n$

In this model, MINDIST and MAXCONTRASTS are used in lieu of IDENT(V), so that the vowel is directly constrained by its difference from contrasting vowels rather than by a specific target. Along the same line, the constraint IDENT(C) should only be considered a stand-in for contrast constraints that will yield the $F2$(t) value when interacting with MINEFFORT.[3] MINDIST = Δ is only violated when the absolute difference between $F2$(y) and $F2$(u) is less than Δ, and the cost of violation is proportional to the square of the necessary difference to reach Δ.

[3] In fact, Flemming argues in other work (e.g. Flemming 2002) that with contrast constraints, the model cannot have traditional faithfulness constraints, as otherwise there would be two different mechanisms in the model that derive contrasts – one with contrast constraints, the other with faithfulness outranking relevant markedness – and this has undesirable theoretical consequences.

MAXCONTRASTS deducts w_n from the cost when the /u/~/y/ contrast is maintained. The contrast will only be worth maintaining if the cost that results from keeping the /tu/ syllable is less than w_n – in other words, the *Cost* value as shown in (6) is greater than zero. Otherwise, the grammar will choose to neutralise the /tu/~/ty/ contrast, as the non-existent /tu/ by definition contributes zero to the overall cost.

(6) $Cost = w_c(F2(\mathrm{t}) - L_t)^2 + w_e(F2(\mathrm{t}) - F2(\mathrm{u}))^2 + w_v(|F2(\mathrm{y}) - F2(\mathrm{u})| - \Delta)^2 - w_n$
(for $|F2(\mathrm{y}) - F2(\mathrm{u})| < \Delta$)

To derive the *F2*(t) and *F2*(u) values that render the smallest overall cost, we take the partial derivatives of *Cost* with respect to *F2*(t) and *F2*(u) respectively, as in (7).

(7) a. $\frac{\partial(Cost)}{\partial(F2(\mathrm{t}))} = 0: \quad F2(\mathrm{t}) = \frac{w_e}{w_c + w_e}(F2(\mathrm{u}) - L_t) + L_t$

b. $\frac{\partial(Cost)}{\partial(F2(\mathrm{u}))} = 0: \quad F2(\mathrm{u}) = \frac{w_e}{w_v + w_e}(F2(\mathrm{t}) - (F2(\mathrm{y}) - \Delta)) + (F2(\mathrm{y}) - \Delta)$

The values of *F2*(t) and *F2*(u) obtained from the two formulas in (7) are shown in (8).

(8) a. $F2(\mathrm{t}) = u_c(L_t - (F2(\mathrm{y}) - \Delta)) + L_t$

where $u_c = \frac{-w_e w_v}{w_e w_c + w_v w_c + w_e w_v}$

b. $F2(\mathrm{u}) = u_v(L_t - (F2(\mathrm{y}) - \Delta)) + (F2(\mathrm{y}) - \Delta)$

where $u_v = \frac{w_e w_c}{w_e w_c + w_v w_c + w_e w_v}$

Notice the parallel between (4) and (8): if we replace L with L_t and T with $F2(\mathrm{y}) - \Delta$ in (4), we get exactly (8). Intuitively, the F2 target for /u/ can be construed as $F2(\mathrm{y}) - \Delta$, as this is the optimal outcome perceptually – the MINDIST violation it incurs is zero. Any F2 value further away from *F2*(y) will be no better perceptually in terms of how MINDIST violation is calculated, but will be undesirable articulatorily.

To illustrate how the model predicts neutralisation with a concrete example, Flemming considers the /u/~/y/ contrast in the /t__t/ context, and makes the following simplifying assumptions:

(9) $L_t = 2100$ Hz, $F2(\mathrm{y}) = 2000$ Hz, $\Delta = 1000$ Hz
$w_c = 0{\cdot}25$, $w_e = 0{\cdot}25$, $w_v = 0{\cdot}5$, $w_n = 200{,}000$

Given that the vowel is flanked between two coronals, the cost of violation for MINEFFORT should be doubled to $2w_e(F2(\mathrm{t}) - F2(\mathrm{u}))^2$, as there are two CV transitions, and the IDENT(C) violation should also be doubled to $2w_c(F2(\mathrm{t})-L_t)^2$, as there are two consonants. If we still use (8) to calculate $F2$(t) and $F2$(u), we need to double the w_c and w_e values. The calculation yields an $F2$(t) value of 1733 Hz and an $F2$(u) value of 1367 Hz, which consequently give an overall positive cost of 201,667 (the first three items in (6)). Since the benefit of having the contrast is only 200,000, the contrast is not worth maintaining, and /tut/ must be neutralised to /tyt/.

If we adjust the weight values to $w_c=0{\cdot}2$, $w_e=0{\cdot}25$ and $w_v=0{\cdot}55$, then the optimal $F2$(t) and $F2$(u) become 1665 Hz and 1317 Hz respectively, and the cost of maintaining the contrast goes down to 191,511. It is then worth keeping the contrast, as the cost is smaller than the benefit. If the vowel is not flanked between two coronals, but only adjacent to one (/tu/~/ty/), the same $w_c=0{\cdot}25$, $w_e=0{\cdot}25$, $w_v=0{\cdot}5$ assumption returns optimal $F2$(t) and $F2$(u) values as 1660 Hz and 1220 Hz, which give an overall cost of 121,000 – considerably lower than the threshold of 200,000. This means that the /tu/~/ty/ contrast will be preserved. Flemming's calculation for this case (Flemming 2001 : 28), which shows an overall cost of 117,097, is incorrect, I believe. But it does not affect the conclusion that the /u/~/y/ contrast needs to be preserved in the /t__/context.[4]

1.3 Summary

Flemming's model differs from a traditional optimality-theoretic grammar in two important aspects. First, its constraints refer to phonetically detailed representations rather than only representations that are contrastive in some language; consequently, the model also needs constraints that govern the well-formedness of contrasts so that phonetically minute differences do not emerge as contrastive even though they are represented in the grammar. Second, the constraints in the model are associated with weighted costs and the output of the grammar is the phonetic values that incur the least overall cost of all the constraints. In other words, the harmonic value of an output candidate is not defined by its satisfaction of a ranked constraint hierarchy, but by the weighted sum of the costs associated with each constraint. It is this second innovation that I claim to be unnecessary.

Flemming presents two arguments in favour of constraint weighting as a computationally more feasible alternative to constraint domination. First, in order to capture the trade-off relationship between effort and distinctiveness in constraint domination, the constraints need to be split into a great many ranked sub-constraints, and a very particular ranking of these constraints is needed to derive the observed linear relationship between effort minimisation and contrast distinctiveness. Second, the trade-off between effort and distinctiveness constraints is additive; for example, better vowel distinctiveness together with better consonant

[4] Thanks to an anonymous reviewer for clarifying Flemming's calculation and correcting an error in an earlier draft of the paper.

distinctiveness can motivate more effort expenditure, even though neither distinctiveness is sufficient by itself. To capture all the acceptable trade-offs in constraint domination would require a great many locally conjoined constraints, which would complicate the system too much.

In what follows, I formally compare the predictions of constraint domination with those of constraint weighting by using the same examples as Flemming (2001) – non-neutralising and neutralising consonant–vowel F2 assimilations, and I show that any single output predicted by constraint weighting can also be predicted by constraint domination, and *vice versa*. I also show that with proper constraint definitions and rankings, both the additive effect and the locus effect predicted by constraint weighting can also be replicated in constraint domination. Therefore, I argue that constraint weighting is not a necessary innovation in the architecture of Optimality Theory.

Two caveats should be noted upfront. First, the main thesis of Flemming (2001) is the unification of the treatment of scalar and categorical sound patterns – a point that I do not dispute. My disagreement with Flemming (2001) is on the technical point of how this unification should proceed – via constraint weighting or constraint domination. Second, Flemming in fact acknowledges that the difference between constraint weighting and constraint domination is considerably narrowed by devices such as constraint decomposition and conjunction, and that there remains the possibility that the constraint system is a mixture of weighting and domination relations, with the latter being reserved for patterns that do not revolve around principles of articulatory ease and perceptual distinctness, such as stress. This further narrows the disagreement between Flemming (2001) and the current paper.

2 A formal comparison between constraint weighting and constraint domination

2.1 Non-neutralising assimilation

2.1.1 *Predictions of the constraint-weighting model.* Let us first consider the realm of possible solutions to $F2(C)$ and $F2(V)$ in Flemming's model for non-neutralising assimilation. The $F2(C)$ and $F2(V)$ values that render the minimum cost are given in (4) above. Since $L > T$, and the phenomenon being captured is F2 coarticulation, $F2(C)$ and $F2(V)$ should meet the condition in (10). This condition can be directly observed from the values of $F2(C)$ and $F2(V)$ in (4): since u_c is a negative value and L is greater than T, we deduce that $F2(C) \leq L$; since u_v is a positive value and L is greater than T, we deduce that $F2(V) \geq T$; and since $F2(C) - F2(V)$ is equal to $(1 + u_c - u_v)(L - T)$, which is in turn equal to $(w_v w_c / w_e w_c + w_v w_{c+} w_e w_v) / L - T)$, we deduce that $F2(C) \geq F2(V)$.

(10) $T \leq F2(V) \leq F2(C) \leq L$

Moreover, for *any* $F2(C)\sim F2(V)$ pair that satisfies the condition in (10), I show below that it can be a predicted winner in the grammar under certain w_e, w_c and w_v. Since w_e, w_c and w_v are relative weights, we may set one of them to 1 and solve the other two values from the two equations in (4). If we assume that $w_e = 1$, we can then derive the values w_c and w_v as in (11).

(11) a. $w_c = \dfrac{F2(C) - F2(V)}{L - F2(C)}$ b. $w_v = \dfrac{F2(C) - F2(V)}{F2(V) - T}$

Provided that the condition in (10) is met, we deduce that $w_c \geq 0$, $w_v \leq 0$, indicating that they are both likely weight units for consonant and vowel undershoots. In other words, for any $F2(C)\sim F2(V)$ pair that satisfies the condition in (10), there exists a grammar that predicts it as the winner.

2.1.2 *Predictions of the constraint-domination model.* Let us now consider the predictions of constraint domination with ranked sub-constraints in a constraint family.

I also employ IDENT(C), IDENT(V) and MINEFFORT, but not as constraints that return costs of violation, but as constraint families with universally ranked sub-constraints.

In (12) I define two DIFFERENTIAL LIMEN SCALES with respect to a frequency target F:

(12) *Differential Limen Scales*

a. The Differential Limen Scale above F, denoted as $a_0, a_1, a_2, a_3, \ldots, a_n$, satisfies the following conditions:
   i. $0 = a_0 < a_1 < a_2 < a_3 < \ldots < a_n$.
   ii. $a_n + F$ is the maximum frequency in the perceptual range relevant for speech.
   iii. For $0 \leq i < n$, the frequency difference between $a_i + F$ and $a_{i+1} + F$ is the smallest difference perceivable by listeners.

b. The Differential Limen Scale below F, denoted as $b_0, b_1, b_2, b_3, \ldots, b_m$, satisfies the following conditions:
   i. $0 = b_0 > b_1 > b_2 > b_3 > \ldots > b_m = -F$.
   ii. For $0 \leq i < m$, the frequency difference between $b_i + F$ and $b_{i+1} + F$ is the smallest difference perceivable by listeners.

The PERCEPTUAL STEPS of an arbitrary frequency from F are defined in (13):

(13) *Perceptual steps*

For an arbitrary frequency F' in the perceptual range, if $a_i + F \leq F' < a_{i+1} + F$ or $b_{i+1} + F < F' \leq b_i + F$ $(0 \leq \mathrm{i} \leq m, n)$, then F' is i perceptual steps away from F.

Let us again assume that the locus of consonant F2 is *L*, the target of vowel F2 is *T*, and $L > T$. Then the IDENT(C) and IDENT(V) constraint families can be defined as follows:

(14) *Faithfulness constraints*

a. IDENT(C)-*i*
The locus of a consonant F2 (=*L*) must have an output correspondent *F2*(*C*), which must be fewer than *i* perceptual steps away from *L*.

b. IDENT(V)-*i*
The target of a vowel F2 (=*T*) must have an output correspondent *F2*(*V*), which must be fewer than *i* perceptual steps away from *T*.

The two families of faithfulness constraints have the following universal rankings:

(15) *Universal rankings*

a. … ≫ IDENT(C)-*i* ≫ IDENT(C)-(*i*−1) … ≫ IDENT(C)-2 ≫ IDENT(C)-1

b. … ≫ IDENT(V)-*i* ≫ IDENT(V)-(*i*−1) … ≫ IDENT(V)-2 ≫ IDENT(V)-1

These universal rankings stem from the intuition that a perceptually less accurate rendition of the input is a worse violation of faithfulness. They are parallel to assigning a greater cost to a less accurate realisation of the target in Flemming's model. For a formalisation of this intuition as universally ranked constraints in Optimality Theory, see Steriade (to appear).[5]

Let us note that the basic architecture of constraint domination with families of intrinsically ranked constraints in fact allows the continua of phonetics to be modelled to any accuracy we desire. The above discretisation, based on differential limens, is a possible way to keep the numbers of constraints finite, but is not crucial to the architecture of constraint domination itself.[6]

[5] Based on works by Prince (2000, 2001) and de Lacy (2002, 2004), a reviewer questioned the necessity of fixed constraint rankings. However, we should note that the Prince/de Lacy proposal, with freely rankable constraints in a stringency hierarchy, is more powerful than fixed constraint rankings, in that it makes a wider range of predictions, such as the conflation of adjacent categories in phonological behaviour. Therefore, I deem it more appropriate to show that the same predictions as Flemming's constraint weighting can be made in the more restricted theory, namely one with fixed rankings, especially given that the conflation of adjacent categories is a non-issue in the problem addressed here.

[6] For readers who are interested in concrete numbers, here is an approximation. Although psychoacoustic studies such as Harris (1952), Flanagan & Saslow (1958) and Klatt (1973) have shown that listeners can perceive fundamental frequency differences of less than 1 Hz, 't Hart (1981) and 't Hart *et al.* (1990) have rightly pointed out that the just noticeable differences in psychoacoustic studies are usually elicited under optimal conditions in which the subject's only task is to listen to one particular difference in controlled environments, and the just noticeable differences in real speech should be considerably higher than those elicited in psychoacoustic experiments. 't Hart (1981) studied the differential threshold for pitch changes on a target syllable in real speech utterances in Dutch and reported that only differences

To define the MinEffort constraints, I first define an articulatorily based Frequency Difference Scale, as in (16).

(16) *Articulatorily based Frequency Difference Scale*

The Frequency Difference Scale $\Delta f_0, \Delta f_1, \Delta f_2, \Delta f_3, \ldots, \Delta f_n$, satisfies the following conditions:

a. $0 = \Delta f_0 < \Delta f_1 < \Delta f_2 < \Delta f_3 < \ldots < \Delta f_n$.

b. Δf_n is the maximum F2 difference between a consonant and an adjacent vowel.

c. For $0 \leq i < n$, the difference between Δf_i and Δf_{i+1} is the smallest articulatory difference that can be controllably produced.

Assuming that across the F2 range relevant for speech, the effort expenditure associated with F2 is solely determined by the frequency difference, i.e. the effort for F2 to go from 1600 Hz to 1300 Hz is the same as that from 1400 Hz to 1100 Hz, the MinEffort constraint family can then be defined as in (17a), with its universal ranking shown in (17b).

(17) *Markedness constraints*

a. MinEffort-Δf_i

For $0 \leq i \leq n$, the transition between $F2(C)$ and $F2(V)$ must be less than Δf_i.

b. If $\Delta f_i > \Delta f_j$, then MinEffort-$\Delta f_i \gg$ MinEffort-Δf_j.

Given that research on incomplete neutralisation and the acquisition of second language phonetics has shown that many articulatorily implementable differences cannot be reliably perceived, we can reasonably assume that the Frequency Difference Scale is more finely grained than the Differential Limen Scales. Again, the categorical definition of MinEffort based on the Frequency Difference Scale is only a possible move to keep the number of MinEffort constraints finite. Neither the structure of the constraint family itself nor the results obtained later in the paper depend on the particular numbers of the scale.

It can be proven that under constraint domination, any $F2(C)$~$F2(V)$ pair that can be predicted by the grammar must also satisfy the condition $T \leq F2(V) \leq F2(C) \leq L$. The proof can proceed as follows: consider all situations that do not satisfy the above conditions, e.g. $T \leq F2(V) \leq L \leq F2(C)$, $T \leq L \leq F2(V) \leq F2(C)$, etc., and show that there is an alternative that satisfies the condition and is more harmonic with respect to the constraint hierarchy, in other words, any candidate for which (a) either $F2(C)$ or $F2(V)$ falls outside the range (T, L) or (b) $F2(V) > F2(C)$ is

---

of more than 3 semitones (around 20–30 Hz in the speech range) play a role in communicative functions, while Harris & Umeda (1987) showed that the differential limens for F0 in naturally spoken sentences are between 10 and 50 times greater than those found with sustained synthetic vowels, and that the differential limens varied significantly depending on the complexity of the stimulus and the speaker. Therefore, the number of Ident constraints in a single family based on differential limes is most likely in the magnitude of tens.

harmonically bounded by a candidate that satisfies $T \leq F2(V) \leq F2(C) \leq L$. For reasons of space, I consider this condition as a given.

Let us assume that L and T are n perceptual steps away from each other. In what follows, under the assumption that the Frequency Difference Scale is more finely grained than the Differential Limen Scales, I show that for any i and j between 0 and n and $i+j \leq n$, there is a $F2(C){\sim}F2(V)$ pair such that $F2(C)$ is i perceptual steps away from L and $F2(V)$ is j perceptual steps away from T which can be predicted as the winner by the grammar under some ranking. In so doing, I show that constraint domination and constraint weighting predict the same range of single outputs for non-neutralising assimilation, namely $T \leq F2(V) \leq F2(C) \leq L$, and should thus be considered formally equivalent in this regard.

Consider any i and j such that $0 \leq i,\ j \leq n$ and $i+j \leq n$. For all the $F2(C){\sim}F2(V)$ pairs such that $F2(C)$ is i perceptual steps away from L and $F2(V)$ is j perceptual steps away from T, we can calculate all the $F2(C)-F2(V)$ values, which may fall in different ranges along the Frequency Differences Scale. We can then pick any $F2(C){\sim}F2(V)$ pair that falls in the lowest range, which we assume to be $(\Delta f_k,\ \Delta f_{k+1})$ (therefore $\Delta f_k \leq F2(C)-F2(V) < \Delta f_{k+1}$), and the constraint ranking in (18) will predict this $F2(C){\sim}F2(V)$ pair as the winner. We assume that $\Delta f_m \leq L-T < \Delta f_{m+1}$; thus the highest relevant MinEffort constraint in (18) is MinEffort $-\Delta f_m$.

(18)
$$\begin{array}{llll}
\textsc{Ident}(\mathrm{C})\text{-}n & \gg \textsc{Ident}(\mathrm{C})\text{-}(n{-}1) & \ldots \gg \textsc{Ident}(\mathrm{C})\text{-}i & \gg \textsc{Ident}(\mathrm{C})\text{-}1 \\
\textsc{Ident}(\mathrm{V})\text{-}n & \gg \textsc{Ident}(\mathrm{V})\text{-}(n{-}1) & \ldots \gg \textsc{Ident}(\mathrm{V})\text{-}j & \gg \textsc{Ident}(\mathrm{V})\text{-}1 \\
\textsc{MinEff}\text{-}\Delta f_m & \gg \textsc{MinEff}\text{-}\Delta f_{m-1} & \ldots \gg \textsc{MinEff}\text{-}\Delta f_k & \gg \textsc{MinEff}\text{-}\Delta f_1
\end{array}$$

hierarchies
aligned here

The equal ranking of the three constraints Ident(C)-i, Ident(V)-j and MinEffort-Δf_k on the hierarchy, as shown in (18), is crucial. $F2(C){\sim}F2(V)$ as it stands violates all three constraints: $F2(C)$ and $F2(V)$ are i and j perceptual steps from their targets respectively, but Ident(C)-i and Ident(V)-j require the F2 realisations of consonant and the vowel to be fewer than i and j perceptual steps from their targets respectively (see (14)); $F2(C)-F2(V) \geq \Delta f_k$, but MinEffort-$\Delta f_k$ requires their difference to be less than Δf_k (see (17)). But a smaller $F2(C)$, which is perceptually further away from L, will violate a higher-ranked Ident(C) constraint; a greater $F2(V)$, which is perceptually further away from T, will violate a higher-ranked Ident(V) constraint; an $F2(C){\sim}F2(V)$ pair that fares better in either Ident(C) or Ident(V) will violate a higher-ranked MinEffort constraint, given the assumption that the perceptual scale is less finely grained than the articulatory scale; and an $F2(C){\sim}F2(V)$ pair that does equally well with Ident(C) and Ident(V) will either tie with the current winner on MinEffort as well, or lose due to the violation of a higher-ranked MinEffort constraint according to our assumption that $(\Delta f_k,\ \Delta f_{k+1})$ is the lowest range on the Frequency Difference Scale in which the current $F2(C)-F2(V)$ falls.

To visualise the interaction among the three constraint families, I first define a three-dimensional space (*x*, *y*, *z*), in which the *x* and *y* axes represent *F2*(*C*) and *F2*(*V*) respectively, and the *z* axis represents the RANKING VALUE of constraints that candidates on the *x*-*y* plane violate. The Ranking Value of a constraint can be construed as follows: we assume that every constraint is associated with a Ranking Value, and the more highly ranked the constraint is, the greater the Ranking Value of the constraint. Therefore, $C_1 \gg C_2$ is equivalent to $RV(C_1) > RV(C_2)$, where RV represents Ranking Value.

We can represent IDENT(C), IDENT(V) and MINEFFORT as three surfaces in this three-dimensional space, as in Fig. 3. The surfaces represent the Ranking Value of the highest-ranked constraint in the family that the candidates on the *x*-*y* plane violate, and we only consider the range of *F2*(*C*) and *F2*(*V*) from *T* to *L*, the range within which the winner must fall.[7] The winner predicted by the constraint hierarchies is the (*F2*(*C*), *F2*(*V*)) pair that renders the minimum of the function in (19).

(19) $z = f(F2(C), F2(V)) = maxima(\text{IDENT(C), IDENT(V), MINEFFORT})$

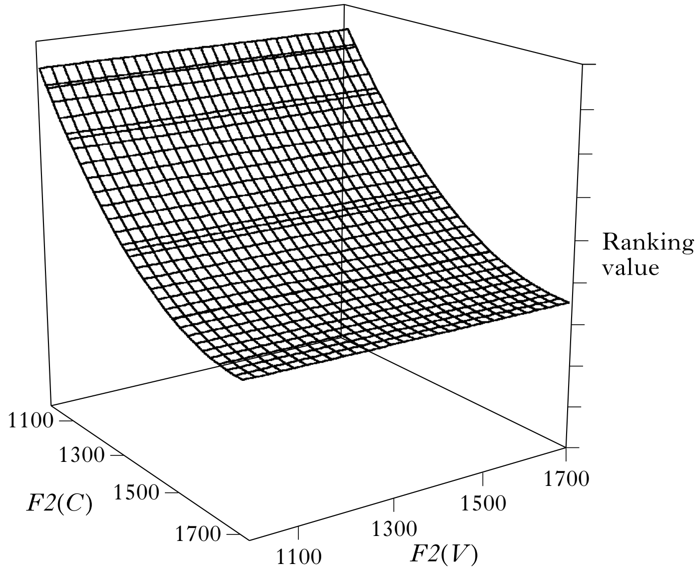
This function is plotted from two different angles in Fig. 4. The second surface is derived by rotating the first surface clockwise approximately 60°, looking down on the *z* axis. The optimal candidate (*F2*(*C*)′, *F2*(*V*)′), which renders the minimum of (19), is indicated in both graphs.

2.1.3 *Comparison between constraint weighting and constraint domination.* Due to the differences in the nature of constraints, there is in theory one difference in the predictions made by constraint weighting and constraint domination.

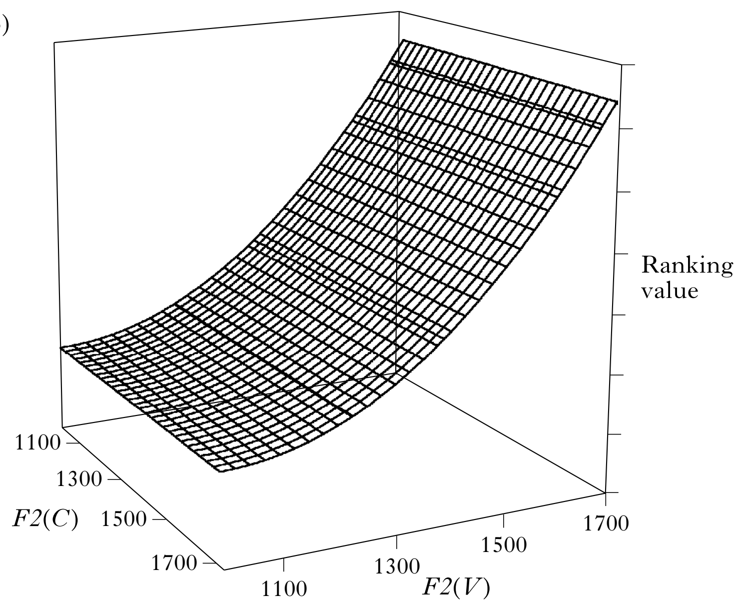
Besides the evaluation mode, the faithfulness constraints here also differ from Flemming's model in that they depict a discrete state of affairs: deviations from the input are only penalised when they cross a particular perceptual threshold; otherwise they do not incur faithfulness violations.

A consequence of his difference is that, for constraint weighting, any *F2*(*C*)~*F2*(*V*) that satisfies $T \leq F2(V) \leq F2(C) \leq L$ can be a possible winner, but for constraint domination, certain *F2*(*C*) and *F2*(*V*) values will tie as winners and certain other *F2*(*C*) and *F2*(*V*) values will never win a competition. Let me illustrate this with a hypothetical example. Let us assume that between 1000 Hz and 1700 Hz, only a frequency difference above 20 Hz is perceivable (i.e. 1000 Hz to 1020 Hz is 0 perceptual steps away from 1000 Hz, 1021 Hz to 1040 Hz is 1 perceptual step away from 1000 Hz, etc.), and a frequency difference of 10 Hz is the smallest detectable difference by the articulatory apparatus. If IDENT(C)-9, IDENT(V)-9 and MINEFFORT-300 Hz are equally ranked, then the winner can be any *F2*(*C*)~*F2*(*V*) pair that satisfies the conditions (a) $1500\text{ Hz} \leq F2(C) < 1520\text{ Hz}$, (b) $1180\text{ Hz} < F2(V) \leq 1200\text{ Hz}$ and (c) $F2(C) - F2(V) < 310\text{ Hz}$, as all such pairs satisfy IDENT(C)-10 and IDENT(V)-10, but violate IDENT(C)-9 and IDENT(V)-9, and also satisfy MINEFFORT-310 Hz,

(a)



(b)



[7] The curve of the plotted surfaces in Fig. 3 is only for aesthetic purposes. The only claim made here is the monotonicity of the surfaces in Fig. 3a, b with respect to *F2*(*C*) and *F2*(*V*) respectively, and the monotonicity of the surface in Fig. 3c with respect to both *F2*(*C*) and *F2*(*V*).

(c)

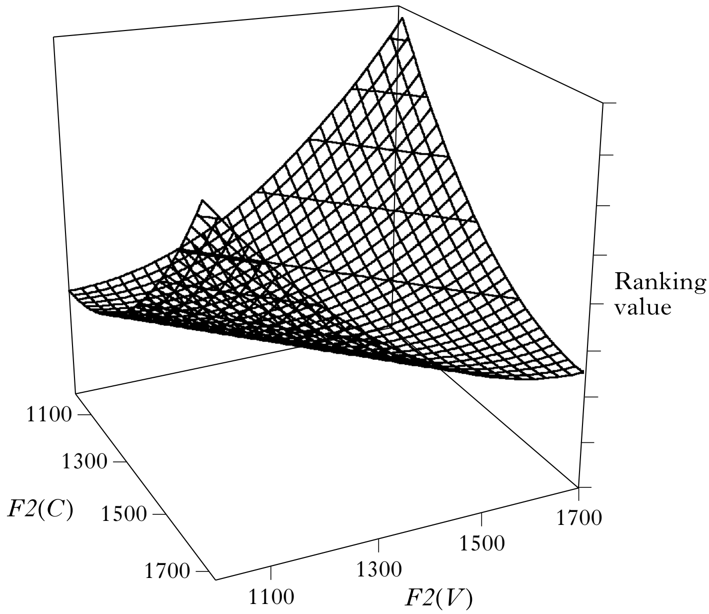


Figure 3

(a) The IDENT(C) surface, (b) the IDENT(V) surface and (c) the MINEFFORT surface plotted against *F2*(*C*) and *F2*(*V*) in the range of (*T*, *L*) for both *F2*(*C*) and *F2*(*V*). *L* = 1700 Hz; *T* = 1000 Hz. In each case, the surface represents the ranking value of the highest-ranked constraint in the relevant family violated by the candidates on the *x*-*y* plane.

but violate MINEFFORT-300 Hz. For example, *F2*(*C*) = 1500 Hz and *F2*(*V*) = 1200 Hz, *F2*(*C*) = 1501 Hz and *F2*(*V*) = 1194 Hz and *F2*(*C*) = 1508 Hz and *F2*(*V*) = 1199 Hz are all possible winners. However, a pair like *F2*(*C*) = 1505 Hz and *F2*(*V*) = 1190 Hz will tie with the winners on faithfulness violations, but do worse with MINEFFORT constraints by violating MINEFFORT-310 Hz. Therefore, it will never be the winner. This hypothetical example is illustrated in (20).

(20)

	IDENT (C)-10	IDENT (V)-10	MINEFF-310	IDENT (C)-9	IDENT (V)-9	MINEFF-300
☞ a. *F2*(*C*)=1500 *F2*(*V*)=1200				*	*	*
☞ b. *F2*(*C*)=1501 *F2*(*V*)=1194				*	*	*
☞ c. *F2*(*C*)=1508 *F2*(*V*)=1199				*	*	*
d. *F2*(*C*)=1505 *F2*(*V*)=1190			*!	*	*	*

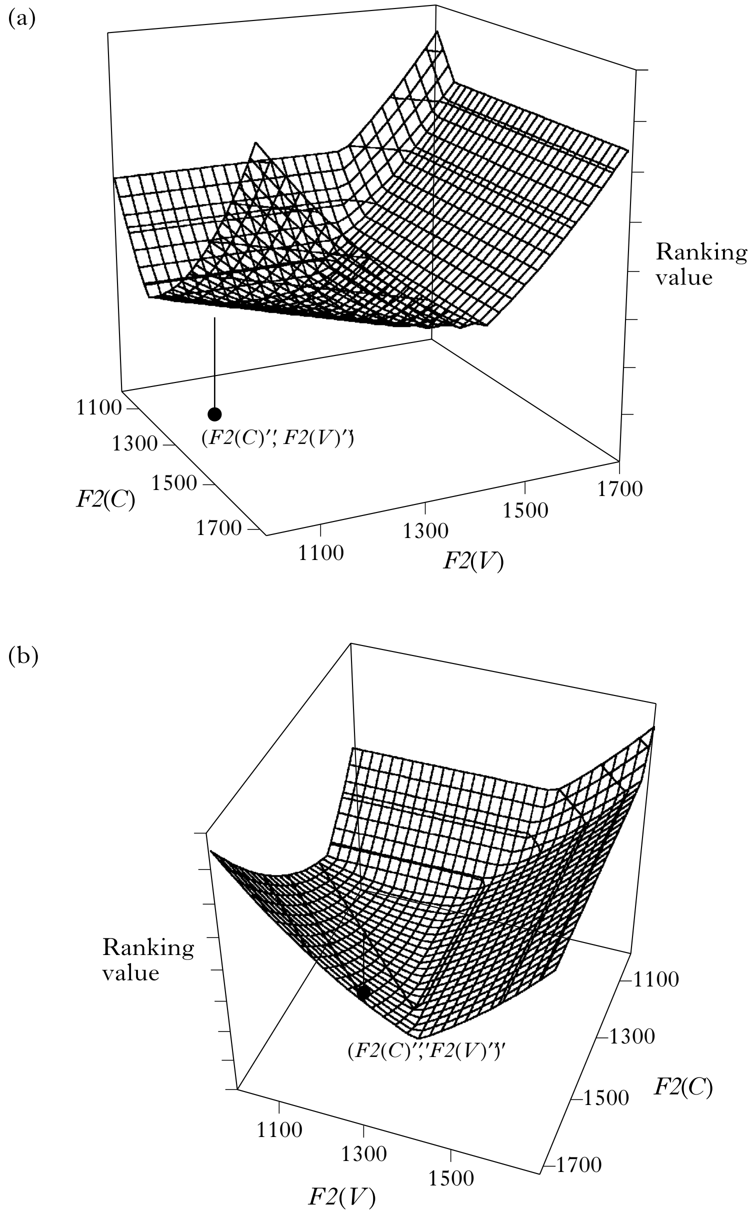


Figure 4

Two different angles of the surface representing the function $z = maxima$(IDENT(C), IDENT(V), MINEFFORT). (*F2*(*C*)′, *F2*(*V*)′), which renders the minimum of the function, is the optimal candidate predicted by the three consonant hierarchies.

I stated that this difference in prediction between constraint domination and constraint weighting is only theoretical because, as a reviewer rightly pointed out, even if the constraint system does not take into account the issue of noticeable differences and generates candidates that the perceptual system cannot differentiate, the differential limens will still play a role in perception, causing the indiscriminable candidates to be learned as one. There may therefore not be an empirical difference between the predictions made by the two approaches.

2.2 Neutralising assimilation

2.2.1 *Predictions of the constraint-weighting model.* Let us now consider the predictions on /u/~/y/ neutralisation in coronal contexts in Flemming's model, taking into account MinDist and MaxContrasts. Assuming that the MinDist requirement is Δ, and the benefit of keeping the /u/~/y/ constrast is $-w_n$, the overall cost for keeping the contrast is then as in (6) above, and the *F2*(u) value that renders the smallest overall cost is as in (21).

(21)
$$F2(\mathrm{u}) = \frac{w_e}{w_e + w_v} F2(\mathrm{t}) + \frac{w_v}{w_e + w_v} (F2(\mathrm{y}) - \Delta)$$

Let us remember that this is a simplified model: the Ident(C) constraint, whose cost of violation is $w_c(F2(\mathrm{t}) - L_t)^2$, is only a stand-in for contrast constraints that will yield the *F2*(t) value when interacting with MinEffort. If we further assume that *F2*(t) and *F2*(y) are known, and that Δ is constant, then we can easily adjust the values of w_e and w_v so that any *F2*(u) value in the range of (0, *F2*(y)) can render the smallest overall cost; and we can adjust the w_n value contributed by MaxContrasts to predict either neutralisation or non-neutralisation, given the *F2*(u) value. This means that the model is able to predict neutralisation of the /u/~/y/ contrast, but also any /u/~/y/ contrast in which the *F2*(u) value is in the range of (0, *F2*(y)).

2.2.2 *Predictions of the constraint-domination model.* These predictions can be easily made by constraint domination as well. The constraint inventory includes: (a) Ident(C) and MinEffort, which are the same as in (14) and (17), (b) MaxContrasts, which I take as a single constraint that penalises /u/~/y/ neutralisation and (c) MinDist, which I define as a constraint family as in (22a), with the universal ranking in (22b).

(22) MinDist *constraints*

a. MinDist-*i*
F2(y) and F2(u) must be at least *i* perceptual steps away from each other.

b. MinDist-1 ≫ MinDist-2 … ≫ MinDist-(*i*–1) ≫ MinDist-*i* ≫ …

For any $F2$(u) in the range of (0, $F2$(y)), assuming that $F2$(y) and $F2$(u) differ by i perceptual steps, $F2$(t) and $F2$(y) are known, and $\Delta f_j \leq F2(\text{t}) - F2(\text{u}) < \Delta f_{j+1}$, the constraint ranking in (23) is sufficient to ensure the preservation of the /u/~/y/ contrast and to give $F2$(u) as the winner.

(23) MaxContrasts, MinDist-$i \gg$ MinEffort-$\Delta f_{j+1} \gg$ MinDist-(i+1) $\gg$ MinEffort-Δf_j

The tableau in (24) illustrates the derivation of a variable $F2$(u) as the winner. $F2$(u) violates MinDist-(i+1) as it is only i perceptual steps away from $F2$(y), and it violates MinEffort-Δf_j as the constraint requires the F2 difference between the /t/ and the /u/ to be less than Δf_j. But a smaller F2 value that can be controllably implemented (the second candidate) will violate MinEffort-Δf_{j+1}; a greater F2 value that is perceptually closer to $F2$(y) (the third candidate) will violate MinDist-i; and neutralisation of /u/~/y/ (the last candidate) will violate MaxContrasts, all of which are more highly ranked than the constraints $F2$(u) violates.

(24)

	Max Contrasts	Min Dist-i	MinEff-Δf_{j+1}	MinDist-(i+1)	MinEff-Δf_j
☞ a. $F2$(u)				*	*
b. $F2(\text{u})-f_1$			*!		*
c. $F2(\text{u})+f_2$		*!		*	
d. u=y	*!				

To predict neutralisation, we only need to ensure that MaxContrasts is ranked lower than either MinDist-(i+1) or MinEffort-Δf_j. The intuition is that any $F2$(u) in the range of (0, $F2$(y)) may not be a worthwhile contrast with $F2$(y), because it does not do well enough in either contrast dispersion or effort minimisation.

2.3 Discussion

2.3.1 *Factorial typological predictions regarding single outputs.* The comparison above between constraint weighting and constraint domination has shown that, in both non-neutralising and neutralising scenarios, they make formally equivalent predictions about the range of possible values as single outputs; their only difference is that the predicted values of constraint weighting are continuous, while those of constraint domination are discrete. This difference stems from the distinct natures of faithfulness constraints in the two approaches: faithfulness violation is calculated in an continuous fashion in constraint weighting, but discretely assessed in the current implementation of constraint domination.

As already mentioned, this difference in prediction may be only theoretical, not empirical, as the differential limens, even if they are not encoded in the grammar, will be relevant to perception and conflate whatever candidates generated by the grammar that cannot be discriminated from each other in the learning process.

2.3.2 *The additive effect.* Related to the predictions on single outputs is the issue of how the additive trade-off between effort and distinctness constraints can be captured. This trade-off relation is captured intuitively by the summation of violation costs of constraints in the weighting model, and Flemming takes this as an argument for the model. There are two scenarios in which additive effects seem relevant. One is that 'better vowel distinctiveness together with better consonant distinctiveness can make up for expending more effort' (Flemming 2001: 33). The other is that when a vowel is flanked between two identical consonants, the consonant effect on the vowel is stronger than when the vowel is only adjacent to one consonant.

But as we have seen in the accounts of both non-neutralising and neutralising assimilation in constraint domination, no reference to constraint conjunction is in fact necessary to capture the first scenario: it is derived from the collective high ranking of IDENT(C) and IDENT(V) constraints, together with the low ranking of MINEFFORT. Put differently, given that the effort depends on the realisations of both the consonant and the vowel, the high ranking of IDENT(C) or IDENT(V) alone plainly does not suffice to justify egregious violations of the MINEFFORT constraint family. This can be further illustrated by the following examples.

A constraint ranking in which the entire IDENT(C) family outranks the MINEFFORT family, which in turn outranks the entire IDENT(V) family, will predict the completely faithful rendition of consonant locus *L* as *F2*(*C*), and an *F2*(*V*) identical to *F2*(*C*), with no MINEFFORT violations whatsoever. This is illustrated in the tableau in (25), and it shows that the high ranking of the IDENT(C) alone does not justify MINEFFORT violations ('(*)' in the tableau indicates that the constraint may or may not be violated by the candidate).

(25)

	IDENT(C) family	MINEFF family	IDENT(V) family
☞ a. *F2*(*C*)=*L* *F2*(*V*)=*L*			*
b. *F2*(*C*)=*L* any *F2*(*V*)<*L*		*!	(*)
c. any *F2*(*C*)<*L* any *F2*(*V*)	*!		(*)

But if the IDENT(V) constraint family also outranks the MINEFFORT family *en masse*, then the constraint ranking predicts entirely faithful renditions of *F2*(*C*) and *F2*(*V*) as *L* and *T*, as shown in (26). This illustrates that the collective high rankings of IDENT(C) and IDENT(V) can motivate MINEFFORT violations. No constraint conjunction is necessary to achieve this effect.

(26)

	IDENT(C) family	IDENT(V) family	MINEFF family
☞ a. *F2(C)=L* *F2(V)=T*			*
b. *F2(C)=L* any *F2(V)>T*		*!	(*)
c. any *F2(C)<L* *F2(V)=T*	*!		(*)

Finally, if the IDENT(C), IDENT(V) and MINEFFORT constraint families interleave with each other as in (18), the winner will be an *F2(C)* which is *i* perceptual steps away from *L* and an *F2(V)* which is *j* perceptual steps away from *T*, where *F2(C)* > *F2(V)*. If we assume *L* = 1700 Hz, *T* = 1000 Hz, *F2(C)* = 1500 Hz and *F2(V)* = 1200 Hz are 9 perceptual steps away from *L* and *T* respectively, the tableau in (27) illustrates how *F2(C)* = 1500 Hz, *F2(V)* = 120 Hz can be selected as the winner. In this case, the MINEFFORT violations incurred by the winner are motivated by the higher ends of both the IDENT(C) and IDENT(V) families.

(27)

	IDENT (C)-10	IDENT (V)-10	MINEFF-310	IDENT (C)-9	IDENT (V)-9	MINEFF-300
☞ a. *F2(C)*=1500 *F2(V)*=1200				*	*	*
b. *F2(C)*=1400 *F2(V)*=1200	*!			*	*	
c. *F2(C)*=1500 *F2(V)*=1300		*!		*	*	
d. *F2(C)*=1600 *F2(V)*=1100			*!			*

What I have shown above is that the ‘ganging up’ effect of consonant distinctiveness and vowel distinctiveness on effort minimisation does not need constraint conjunction to be modelled in constraint domination. As for the additive effect of having two flanking consonants on either side of the vowel, it can be derived in constraint domination if MINEFFORT is evaluated on the syllable level rather than only at the juncture of C and V: the effort expenditure for a CVC syllable is greater than that for CV with the same C and V realisations; therefore, a CVC syllable is more stringently evaluated by MINEFFORT than CV, which would cause a higher likelihood of neutralisation in the former. The minimisation of effort is a general functional principle, and there is thus no *a priori* reason for its evaluation to be restricted to two adjacent segments and not applicable to other linguistic units, such as the syllable. This essentially replicates this second additive effect in constraint domination, without resorting to constraint conjunction.

It remains an independent question whether additive effects that require constraint conjunction are cross-linguistically attested. Works such as Kirchner (1996), Crowhurst & Hewitt (1997), Bonilha (2002), Łubowicz (2002), Moreton & Smolensky (2002) and Tranel & del Gobbo (2002), among others, argue that they are; but McCarthy (2002) and Padgett (2002) believe that such effects can all be dealt with alternatively. I leave this issue open.

2.3.3 *The locus effect.*[8] The final significant point of comparison between constraint weighting and constraint domination concerns their factorial typological predictions of the relation among different outputs in the same grammar. For non-neutralising assimilation, constraint weighting predicts a locus effect between *F2(C)* and *F2(V)* – the linear relation between the F2 realisation of a consonant and the F2 realisation of different vowels that follow the consonant – as a direct consequence of the quadratic evaluation of constraint penalties; but constraint domination needs very particular rankings of the constraints to predict the linearity between *F2(C)* and *F2(V)*. Given that the locus effect has been widely reported in the phonetics literature (e.g. Lindblom 1963, Broad & Clermont 1987, Klatt 1987, Sussman 1989, Sussman *et al.* 1991), Flemming takes this an another argument for the constraint-weighting model.

I show in this section that the locus effect can indeed be derived through particular rankings in constraint domination. Furthermore, I show that any function $F2(C)=f(F2(V))$ predicted by the factorial typology of constraint domination increases monotonically when *F2(V)* falls between the vowel target *T* and the consonant target *L*. This is admittedly a weaker prediction than constraint weighting, but I argue that this weaker prediction might be advantageous when phonetic dimensions other than F2 are considered.

Let us assume that, for a constant consonant target *L* and various vowel targets T_1, T_2, …, T_n, the winning $F2(C){\sim}F2(V)$ realisations are $F2(C)_1{\sim}F2(V)_1$, $F2(C)_2{\sim}F2(V)_2$, …, $F2(C)_n{\sim}F2(V)_n$. To show that the locus effect can be captured, we need to demonstrate that there is a constraint ranking under which the winning *F2(C)* and *F2(V)* values are linearly related; in other words, for any *i*, *j* between 1 and *n*, $F2(C)_j - F2(C)_i = k(F2(V)_j - F2(V)_i)$, where *k* is a positive constant.

Let us first take a winning pair $F2(C)_i{\sim}F2(V)_i$. Clearly, $T_i \leq F2(V)_i \leq F2(C)_i \leq L$. Let us assume that $F2(C)_i$ is C_i perceptual steps from *L*, $F2(V)_i$ is V_i perceptual steps from T_i, and $\Delta f_i \leq F2(C)_i - F2(V)_i < \Delta f_{i+1}$. The constraint ranking that derives the $F2(C)_i{\sim}F2(V)_i$ pair is given in (28), which derives from (18).

[8] Thanks to the associate editor for bringing up this point of discussion.

(28)
$$\begin{array}{lll}
\ldots \gg \text{Ident(C)-(C}_i\text{+1)} \gg \text{Ident} & \text{(C)-C}_i \ldots \gg \text{Ident(C)-1} \\
\ldots \gg \text{Ident(V)-(V}_i\text{+1)} \gg \text{Ident} & \text{(V)-V}_i \ldots \gg \text{Ident(V)-1} \\
\ldots \gg \text{MinEff-}\Delta f_{i+1} \quad \gg \text{MinE} & \text{ff-}\Delta f_i \ \ldots \gg \text{MinEff-}\Delta f_1
\end{array}$$

hierarchies
aligned here

We now need to show that for a different vowel target T_j, there is a constraint ranking consistent with (27) that can derive the winning pair $F2(C)_j \sim F2(V)_j$, such that $F2(C)_j - F2(C)_i = k(F2(V)_j - F2(V)_i)$.

Let us first assume that $T_j > T_i$. I also make the following assumptions about $F2(V)_j$: (a) $F2(V)_j > F2(V)_i$, and (b) if $F2(V)_j$ is V_j perceptual steps away from T_j, then $V_j < V_i$. These are reasonable assumptions, as a vowel target that is closer to the consonant target (since $T_j > T_i$) gives the surface vowel a better chance to be more faithful. From this, we know that $\text{Ident(V)-V}_i \gg \text{Ident(V)-V}_j$.

Since $F2(C)_j = F2(C)_i + k(F2(V)_j - F2(V)_i)$, and $F2(V)_j > F2(V)_i$, $k > 0$, we know that $F2(C)_j > F2(C)_i$. Therefore, $F2(C)_j$ is perceptually closer to L than $F2(C)_j$. If C_j represents the perceptual steps $F2(C)_j$ is away from L, then $\text{Ident(C)-C}_i \gg \text{Ident(C)-C}_j$.

Phonetic research such as Sussman *et al.* (1991) has shown that in the locus equation $F2(C) = k*F2(V) + c$, the value of k is less than 1. Since $F2(C)_j - F2(C)_i = k(F2(V)_j - F2(V)_i)$, we know that $F2(C)_j - F2(C)_i < F2(V)_j - F2(V)_i$, which means $F2(C)_j - F2(V)_j < F2(C)_i - F2(V)_i$. Therefore, if $\Delta f_j \leq F2(C)_j - F2(V)_j < \Delta f_{j+1}$, then $\text{MinEffort-}\Delta f_i \gg \text{MinEffort-}\Delta f_j$.

We have thus far shown that (a) $\text{Ident(V)-V}_i \gg \text{Ident(V)-V}_j$, (b) $\text{Ident(C)-C}_i \gg \text{Ident(C)-C}_j$ and (c) $\text{MinEffort-}\Delta f_i \gg \text{MinEffort-}\Delta f_j$. These rankings determine that it is possible to align the three constraint hierarchies in such a way that Ident(C)-C_j, Ident(V)-V_j and $\text{MinEffort-}\Delta f_j$ are equally ranked, in the meantime leaving the ranking in (28) intact. This is shown in (29), in which the vertical lines indicate the locations of alignment of the three hierarchies. This ranking will not only predict $F2(C)_i \sim F2(V)_i$ as the winning pair when the vowel target is T_i, but also predict $F2(C)_j \sim F2(V)_j$ as the winning pair when the vowel target is T_j.

(29)
$$\begin{array}{llll}
\ldots \gg \text{Id(C)-(C}_i\text{+1)} \gg \text{Id(C)} & \text{-C}_i \quad \ldots \gg \text{Id(C)} & \text{-C}_j \quad \ldots \gg \text{Id(C)-1} \\
\ldots \gg \text{Id(V)-(V}_i\text{+1)} \gg \text{Id(V)} & \text{-V}_i \quad \ldots \gg \text{Id(V)} & \text{-V}_j \quad \ldots \gg \text{Id(V)-1} \\
\ldots \gg \text{MinEff-}\Delta f_{i+1} \gg \text{MinE} & \text{ff-}\Delta f_i \ldots \gg \text{MinE} & \text{ff-}\Delta f_j \ldots \gg \text{MinEff-}\Delta f_1
\end{array}$$

hierarchies aligned at these two locations

For $T_j < T_i$, with the assumptions that (a) $F2(V)_j < F2(V)_i$ and (b) $F2(V)_j$ is further away from the vowel target than $F2(V)_i$, we can similarly show that (a) $\text{Ident(V)-V}_j \gg \text{Ident(V)-V}_i$, (b) $\text{Ident(C)-C}_j \gg \text{Ident(C)-C}_i$ and (c) $\text{MinEffort-}\Delta f_j \gg \text{MinEffort-}\Delta f_i$. We can therefore align the rankings of Ident(V)-V_j, Ident(V)-C_j and $\text{MinEffort-}\Delta f_j$ at a location *higher* than the point of alignment for Ident(V)-V_i, Ident(V)-C_i and

MINEFFORT-Δf_i and this ranking will again predict $F2(C)_i{\sim}F2(V)_i$ and $F2(C)_j{\sim}F2(V)_j$ as the winning pairs for vowel targets T_i and T_j respectively.

To further illustrate this with a concrete example, let us consider a consonant target $L = 1700$ Hz and vowels targets $T = 1000$ Hz, 1100 Hz, 1200 Hz, 1300 Hz and 1400 Hz. If between 1000 Hz and 1700 Hz, only a frequency difference above 20 Hz is perceivable, then the constraint ranking in (30) will derive the outputs shown in Table I.

(30)									
ID(C)-9 ID(V)-9 MINEFF-300	≫	ID(C)-8 ID(V)-8 MINEFF-240	≫	ID(C)-7 ID(V)-7 MINEFF-180	≫	ID(C)-6 ID(V)-6 MINEFF-120	≫	ID(C)-5 ID(V)-5 MINEFF-60	

UR (Hz) L=1700 T=1000	L=1700 T=1100	L=1700 T=1200	L=1700 T=1300	L=1700 T=1400
SR (Hz) $F2(C)$=1500 $F2(V)$=1200	$F2(C)$=1520 $F2(V)$=1280	$F2(C)$=1540 $F2(V)$=1360	$F2(C)$=1560 $F2(V)$=1440	$F2(C)$=1580 $F2(V)$=1520
Highest constraints violated by SR ID(C)-9 ID(V)-9 MINEFF-300	ID(C)-8 ID(V)-8 MINEFF-240	ID(C)-7 ID(V)-7 MINEFF-180	ID(C)-6 ID(V)-6 MINEFF-120	ID(C)-5 ID(V)-5 MINEFF-60

Table I

UR ~ SR pairs predicted by the constraint ranking in (30).

The $F2(C){\sim}F2(V)$ values in Table I indicate that the constraint ranking has succeeded in modelling the following linear relationship: $F2(C) = 0{\cdot}25 * F2(V) + 1200$ Hz. It is also easy to show that any function $F2(C) = f(F2(V))$ predicted by the factorial typology of constraint domination increases monotonically when $T \leq F2(V) < L$. Suppose again that $F2(C)_i{\sim}F2(V)_i$ is the winning output for the input $L{\sim}T_i$, which means that IDENT(C)-C_i, IDENT(V)-V_i and MINEFFORT-Δf_i are equally ranked. For a different vowel target $T_j > T_i$, let us again assume that in the winning pair $F2(C)_j{\sim}F2(V)_j$, with T_j as the vowel target, (a) $F2(V)_j > F2(V)_i$ and (b) if $F2(V)_j$ is V_j perceptual steps from T_j, then $V_j < V_i$. I show here that $F2(C)_j > F2(C)_i$.

All candidates with an identical F2 realisation for the vowel – $F2(V)_j$ – will tie on the entire IDENT(V) hierarchy. Thus the decision will be made by IDENT(C) and MINEFFORT. If $F2(C)_j < F2(C)_i$, then the candidate will violate a higher-ranked IDENT(C) constraint than the cluster of equally ranked constraints: IDENT(C)-C_i, IDENT(V)-V_i and MINEFFORT-Δf_i. But if

$F2(C)_j > F2(C)_i$, not only will the candidate do better on the IDENT(C) hierarchy, but it is also possible to pick an $F2(C)_j$ such that $F2(C)_j - F2(V)_j < F2(C)_i - F2(V)_i$. Thus the candidate will *only* violate constraints that are ranked below the cluster of constraints above.

This reasoning is illustrated in the tableau in (31). $F2(C)_1$, $F2(C)_2$ and $F2(C)_3$ represent $F2(C)$ values that are less than, equal to and greater than $F2(C)_i$ respectively. If they are C_1, C_2 and C_3 perceptual steps from L respectively, then clearly $C_1 > C_2 > C_3$, which means that $F2(C)_1$ and $F2(C)_2$ both violate higher-ranked IDENT(C) constraints than $F2(C)_3$, as IDENT(C)-$C_1 \gg$ IDENT(C)-C_2 (same as IDENT(C)-C_i) $\gg$ IDENT(C)-C_3. And finally, we can pick an $F2(C)_3$ such that it satisfies the following condition: $F2(C)_3 - F2(V)_j < F2(C)_i - F2(V)_i$. Therefore, $F2(C)_3$ only violates a MINEFFORT constraint lower than MINEFFORT-Δf_i. For clarity, I use real frequency differences such as $F2(C)$-$F2(V)_j$ in lieu of the Δf notation for MINEFFORT constraints in the tableau.

(31)

	ID (C)-C_1	ID (C)-C_i	ID (V)-V_i	MINEFF-($F2(C)_i - F2(V)_i$)	ID (C)-C_3	MINEFF-($F2(C)_3 - F2(V)_j$)	MINEFF-($F2(C)_2 - F2(V)_j$)	MINEFF-($F2(C)_1 - F2(V)_j$)
a. $F2(C)_1$ ($<F2(C)_i$) $F2(V)_j$	*!							*
b. $F2(C)_2$ ($=F2(C)_i$) $F2(V)_j$		*!					*	
☞ c. $F2(C)_3$ ($>F2(C)_i$) $F2(V)_j$					*	*		

We have thus far shown that constraint domination can also model the linear locus effect. One potential drawback of the domination model is that although the monotonicity of the $F2(C) \sim F2(V)$ relationship is necessarily predicted, the linearity is not. Constraint weighting, in contrast, necessarily predicts the locus effect due to the quadratic evaluation of constraint costs. But on the other hand, F2 is only one of the many dimensions along which the faithfulness and markedness constraints are evaluated. Although the locus effect is widely attested for F2 in the phonetic literature, it is less clear for other formants such as F3. For example, in a study of American English /CVt/ tokens with initial /b d g/ and ten different vowels, Sussman (1989) showed a linear relationship between the initial consonant and the vowel for F2, but not for F3. Therefore, the necessary prediction of the locus effect for all formants may not be an advantage after all.

3 Conclusion

I have shown in this paper that constraint weighting *à la* Flemming (2001) is not a necessary innovation in Optimality Theory. Using constraint families with intrinsic rankings, constraint domination formally predicts the same range of sound realisations as constraint weighting. Moreover, constraint domination can successfully model both the additive effect and the locus effect predicted by constraint weighting.

It must again be acknowledged that the present paper only takes issue with a minor point raised by Flemming (2001), namely, the mode of constraint evaluation. I do not dispute Flemming's main thesis that both scalar and categorical sound patterns should be accounted for in a unified framework, and the current model accommodates this theoretical standpoint as well. The goal of the paper, therefore, is not to argue for or against phonetics-in-phonology, but to present an alternative model to Flemming's in which this position can be implemented, and to encourage further thinking about the formalisation of phonetic and phonological processes.

REFERENCES

Boersma, Paul (1998). *Functional phonology: formalizing the interactions between articulatory and perceptual drives*. The Hague: Holland Academic Graphics.

Bonilha, Giovana Ferreira Gonçalves (2002). Conjoined constraints and phonological acquisition. Available as ROA-533 from the Rutgers Optimality Archive.

Broad, David J. & Frantz Clermont (1987). A methodology for modeling vowel formant contours in CVC context. *JASA* **81**. 155–165.

Chomsky, Noam & Morris Halle (1968). *The sound pattern of English*. New York: Harper & Row.

Cohn, Abigail C. (1993). Nasalisation in English: phonology or phonetics? *Phonology* **10**. 43–81.

Crowhurst, Megan & Mark Hewitt (1997). *Boolean operations and constraint interactions in Optimality Theory*. Available as ROA-229 from the Rutgers Optimality Archive.

de Lacy, Paul (2002). *The formal expression of markedness*. PhD dissertation, University of Massachusetts, Amherst.

de Lacy, Paul (2004). Markedness conflation in Optimality Theory. *Phonology* **21**. 145–199.

Flanagan, James L. & Michael G. Saslow (1958). Pitch discrimination for synthetic vowels. *JASA* **30**. 435–442.

Flemming, Edward (1997). Phonetic optimization: compromise in speech production. In Viola Miglio & Bruce Morén (eds.) *University of Maryland Working Papers in Linguistics* **5**. *Selected Phonology Papers from the Hopkins Optimality Theory Workshop 1997/University of Maryland Mayfest 1997*. 72–91.

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

Flemming, Edward (2002). *Auditory representations in phonology*. New York: Routledge.

Harris, J. Donald (1952). Pitch discrimination. *JASA* **24**. 750–755.

Harris, M. S. & N. Umeda (1987). Difference limens for fundamental frequency contours in sentences. *JASA* **81**. 1139–1145.

Hart, Johan 't (1981). Differential sensitivity to pitch distance, particularly in speech. *JASA* **69**. 811–821.

Hart, Johan 't, René Collier & Antonie Cohen (1990). *A perceptual study of intonation: an experimental-phonetic approach to speech melody*. Cambridge: Cambridge University Press.

Hayes, Bruce, Robert Kirchner & Donca Steriade (eds.) (2004). *Phonetically based phonology*. Cambridge: Cambridge University Press.

Keating, Patricia A. (1985). Universal phonetics and the organization of grammars. In Victoria A. Fromkin (ed.) *Phonetic linguistics: essays in honor of Peter Ladefoged*. Orlando: Academic Press. 115–132.

Keating, Patricia A. (1990). Phonetic representations in a generative grammar. *JPh* **18**. 321–334.

Kirchner, Robert (1996). Synchronic chain shifts in Optimality Theory. *LI* **27**. 341–350.

Kirchner, Robert (2000). Geminate inalterability and lenition. *Lg* **76**. 509–545.

Klatt, Dennis H. (1973). Discrimination of fundamental frequency contours in synthetic speech: implications for models of pitch perception. *JASA* **53**. 8–16.

Klatt, Dennis H. (1987). Review of text-to-speech conversion for English. *JASA* **82**. 737–793.

Lindblom, B. (1963). Spectrographic study of vowel reduction. *JASA* **35**. 1773–1781.

Łubowicz, Anna (2002). Derived environment effects in Optimality Theory. *Lingua* **112**. 243–280.

McCarthy, John (2002). Comparative markedness. In Angela C. Carpenter, Andries W. Coetzee & Paul de Lacy (eds.) *Papers in Optimality Theory II*. Amherst: GLSA. 171–246.

Manuel, Sharon Y. (1990). The role of contrast in limiting vowel-to-vowel coarticulation in different languages. *JASA* **88**. 1286–1298.

Moreton, Elliott & Paul Smolensky (2002). Typological consequences of local constraint conjunction. *WCCFL* **21**. 306–319.

Padgett, Jaye (2002). Constraint conjunction versus grounded constraint subhierarchies in Optimality Theory. Available as ROA-530 from the Rutgers Optimality Archive.

Pierrehumbert, Janet (1990). Phonological and phonetic representation. *JPh* **18**. 375–394.

Pierrehumbert, Janet (1991). The whole theory of sound structure. *Phonetica* **48**. 223–232.

Pierrehumbert, Janet (2000). What people know about sounds of language. *Studies in the Linguisitc Sciences* **29**. 111–120.

Pierrehumbert, Janet (2001). Exemplar dynamics: word frequency, lenition and contrast. In Joan Bybee & Paul Hopper (eds.) *Frequency and the emergence of linguistic structure*. Amsterdam & Philadelphia: Benjamins. 137–157.

Pierrehumbert, Janet (2002). Word-specific phonetics. In Carlos Gussenhoven & Natasha Warner (eds.) *Laboratory Phonology* 7. Berlin & New York: Mouton de Gruyter. 101–139.

Pierrehumbert, Janet & Mary E. Beckman (1988). *Japanese tone structure*. Cambridge, Mass.: MIT Press.

Prince, Alan (2000). The special and the general. Paper presented at the 24th Penn Linguistics Colloquium. Handout available (September 2007) at http://ling.rutgers.edu/people/faculty/prince.html.

Prince, Alan (2001). Invariance under re-ranking. Paper presented at the 20th West Coast Conference on Formal Linguistics, University of Southern California. Handout available (September 2007) at http://ling.rutgers.edu/people/faculty/prince.html.

Prince, Alan & Paul Smolensky (1993). *Optimality Theory: constraint interaction in generative grammar*. Ms, Rutgers University & University of Colorado, Boulder. Published 2004, Malden, Mass. & Oxford: Blackwell.

Steriade, Donca (1999). Phonetics in phonology: the case of laryngeal neutralization. *UCLA Working Papers in Linguistics: Papers in Phonology* **3**. 25–145.

Steriade, Donca (to appear). The phonology of perceptibility effects: the P-Map and its consequences for constraint organization. In Sharon Inkelas & Kristin Hanson (eds.) *On the nature of the word*. Cambridge, Mass.: MIT Press.

Sussman, Harvey M. (1989). Neural coding of relational invariance in speech: human language analogs to the barn owl. *Psychological Review* **96**. 631–642.

Sussman, Harvey M., Helen A. McCaffrey & Sandra A. Matthews (1991). An investigation of locus equations as a source of relational invariance for stop place categorization. *JASA* **90**. 1309–1325.

Tranel, Bernard & Francesca del Gobbo (2002). Local conjunction in Italian and French phonology. In Caroline R. Wiltshire & Joaquim Camps (eds.) *Romance phonology and variation*. Amsterdam & Philadelphia: Benjamins. 191–218.

Zhang, Jie (2002). *The effects of duration and sonority on contour tone : typological survey and formal analysis*. New York & London: Routledge.

Zsiga, Elizabeth C. (1997). Features, gestures, and Igbo vowels: an approach to the phonology-phonetics interface. *Lg* **73**. 227–274.