1 *Characterizing surface orientedness in phonology*

1.1 Surface orientedness

1.1.1 Surface orientedness in phonology

A long-standing issue of interest in phonology is the extent to which phonological systems can be characterized in terms of conditions on the surface forms (Chomsky 1964, Kiparsky 1971, Kiparsky 1973, Kisseberth 1970).[1] One way to express the issue is in terms of disparities between the (underlying) input representation and the (surface) output representation. If one supposes that the input will be preserved in the output "by default," to what extent are the disparities that are introduced motivated by restrictions on the output form? I will use the term **surface orientedness** to refer to the intuitive property of all disparities being motivated by output restrictions. This is expressed slightly differently in (1.1), which asserts that disparities are introduced only to the extent necessary to satisfy output restrictions.

(1.1) Surface orientedness: disparities between input and output are only introduced to the extent necessary to satisfy output restrictions.

Many if not most phonological phenomena seem to be characterizable in terms of output restrictions. Assimilation acts to ensure that, in the output, adjacent segments of certain kinds (e.g., obstruents) are alike in certain ways (e.g., agree in voicing). Positional neutralization acts to ensure that, in the output, segments of certain kinds (e.g., non-glottal obstruents) do not appear in the neutralized positions (e.g., syllable codas). Minimal-size phenomena act to ensure that, in the output, constituents of certain kinds (e.g., prosodic words) are never smaller than a certain size (e.g., a properly formed metrical foot).

The output restrictions themselves are only part of the story. There is also the matter of which disparities serve to enforce the output restrictions. Nearly any potential violation of an output restriction can be resolved in more than one way, with different ways involving the introduction of different disparities. If

1 For a broader overview of the issue, see McCarthy 2007a.

an input form has adjacent obstruents that differ in voicing, an output restriction requiring adjacent obstruents to agree in voicing can be satisfied by voicing the underlyingly voiceless obstruent, or by devoicing the underlyingly voiced obstruent. If an input form has a labial consonant that is at risk of being syllabified as a coda, an output restriction forbidding labials in coda position can be satisfied by deletion of the consonant (so that it does not appear at all in the output), by changing the place of the consonant to something that is permitted in coda position (e.g., glottal place), or by using an alternative syllabification so that the labial is not in a coda position in the output (e.g., by epenthesizing a vowel immediately after the labial consonant).

Further, the choice of disparities for enforcing output restrictions is not just a matter of cross-linguistic variation. One of the more powerful arguments made for the value of output restrictions in understanding phonology is the existence of "conspiracy" patterns, where different disparities are introduced in different contexts, all to the effect of enforcing the same output restriction (Kisseberth 1970). The use of different strategies in different contexts attests to the presence of multiple output restrictions within a language, so that the preferred disparity for enforcing an output restriction might be abandoned in favor of another one if, in that context, the preferred disparity would run afoul of another output restriction.

An example can be found in Prince and Smolensky's (1993/2004) analysis of Lardil. The presentation here is greatly simplified for expositional purposes. In Lardil, there is a restriction on segments in codas, which Prince and Smolensky label the "Coda Condition": "A coda consonant can have only Coronal place or place shared with another consonant." This is an output restriction, and it is in some cases enforced by the introduction of a deletion disparity. This can be seen in nominative case words. The nominative case itself has no overt phonological realization, so the underlying stem receives no further inflection. In longer stems that end in a consonant, that final consonant is deleted in the output if it would violate the Coda Condition when syllabified as a coda unaltered, as shown in (1.2).[2]

(1.2) Truncation of nominative case forms in Lardil
/ŋɑluk/ → [ŋɑlu]
/wuŋkunuŋ/ → [wuŋkunu]
/waŋalk/ → [waŋal]

Deletion is not the only imaginable way of enforcing the Coda Condition. Epenthesizing a vowel at the end of the word and creating a new syllable would

2 The examples here are taken straight from Prince and Smolensky 1993/2004, who in turn take their data from Hale 1973, Klokeid 1976, Wilkinson 1988, and Kirchner 1992.

also avoid violation of the Coda Condition, as would altering the final consonant to one with coronal place. But in these contexts, Lardil clearly prefers to delete the final consonant.

The situation is rather different in some other contexts. Another distributional observation in Lardil is the Minimal Word Condition: lexical words (such as nouns) always appear as at least bimoraic. This output restriction can be related to the output of short stems ending in a potentially offending consonant: instead of deleting the final consonant, a vowel is epenthesized following the consonant, and the two segments form a separate, final syllable in the output, as shown in (1.3).

(1.3) Augmentation of nominative case forms in Lardil
/yɑk/ → [yɑk**a**]
/relk/ → [relk**a**]

Clearly, deletion and insertion are distinct sorts of disparities. Yet, in Lardil, both serve to enforce the Coda Condition and the Minimal Word Condition. This is a conspiracy of the Kisseberthian sort: deletion and insertion "conspire" to enforce the output restrictions.

It is imaginable that both kinds of disparities would be introduced for sub-minimal stems such as in (1.3), with the underlying final consonant being deleted, and then both a consonant and a vowel being inserted to form a second syllable. But that is not what happens. Abandoning the use of deletion in this instance allows the introduction of a minimal amount of insertion (only one segment, the output-final vowel) while simultaneously satisfying both output restrictions: the word is bimoraic in the output and there are no non-coronal codas in the output. Given that some insertion is unavoidable, deletion would simply necessitate yet more insertion. This illustrates the sense in which disparities are introduced "only to the extent necessary," with respect to surface orientedness. Output restrictions may motivate the introduction of disparities between input and output, but they do not open the floodgates for arbitrary numbers of disparities: only those that contribute to the better satisfaction of the output restrictions will possibly be introduced.

This example, and many others like it, motivate the investigation of surface orientedness in phonological theory. That in turn raises the question of how to properly formalize the intuitive notion, so that the motivating patterns in fact follow as predicted consequences of the theory.

1.1.2 Formalizing surface orientedness

While the statement in (1) may seem intuitively clear, formalizing it in a satisfactory way is not so straightforward. It isn't immediately clear what kind

of thing an output restriction is, or what exactly it means for disparities to be "to the extent necessary" to satisfy output restrictions. Furthermore, it is not immediately obvious how to give such a formalization without delving deeply into the mechanics of a particular theory.

We can abstract away from the mechanics of particular theories by considering the map defined by a phonology. A **phonological map** is a collection of annotated pairings of inputs with outputs (this definition will be revised in Chapter 2, but will work for now). For each linguistic input, a map contains exactly one input–output pair (like a mathematical function), and for each pair an indication of the disparities between them. A map does not make reference to how the grammar determines the output for an input, it only lists the disparities that are applied to the input along with the resulting output. An individual annotated pair is referred to as a **mapping**; a map is a collection of mappings.

A more narrowly focused intuition along the lines of (1.1) is that inputs that are identical to well-formed outputs will map to themselves (with no disparities). If disparities are introduced only to the extent necessary, then if you can do without any disparities, you will do without any disparities. This has sometimes been called the identity map property (Prince and Tesar 2004): a well-formed output, when used as an input, maps to itself. Formally, this is equivalent to the mathematical property of idempotency for a unary operation: applying the operation twice yields the same value as applying it only once.[3]

The canonical example of a phenomenon that violates the identity map property is a chain shift. This can be illustrated with an example based on an analysis by Kirchner (1995) of Etxarri Basque vowel raising in hiatus (this example is examined in greater detail in Section 4.2.1). In the hiatus environment, an input mid vowel e changes to the high vowel i: /e/ → [i]. However, in the same environment, an input high vowel i changes to the raised vowel i^y: /i/ → [i^y]. The high vowel [i] is a grammatical output (setting aside the conditioning environment for purposes of presentation), but when used as an input it does not emerge unaltered; instead, the raised vowel [i^y] is the output. This violates idempotency, which would require that if [i] is the output for some input, like /e/, then it must map to itself.

Intuitively, a chain shift is not output oriented. In the Etxarri Basque vowel raising case, if the single disparity in /e/ → [i] is sufficient to satisfy the output restrictions (whatever they are), then no disparities should be needed for input /i/: it should be the case that /i/ → [i]. If, on the other hand, the single disparity in /i/ → [i^y] is necessary to satisfy the output restrictions, then the disparity of

3 Idempotency has been defined this way in the context of lattice theory, for example (Szász 1963).

/e/ → [i] ought to be insufficient: it should be the case that /e/ → [i^{y}]. The failure to be output oriented makes chain shifts interesting and raises questions about how to best analyze and understand them.

Within the theory of SPE (Chomsky and Halle 1968), input–output disparities result from the application of phonological processes, implemented as rewrite rules. The issue of surface orientedness can then be approached by examining the extent to which the behavior of the processes can be characterized in terms of the final output of the derivation. In this view, surface orientedness is a property of processes with respect to derivations. The conditions determining where a process can apply are solely conditions on the structure before the rule has applied, with no explicit reference to the endpoint of the derivation. Not much can be concluded solely on the basis of the definition of the process itself, in isolation. But one can examine the relationship of a process to an output in the context of a derivation for that output.

Interestingly, in rule-based phonology the focus has been not on directly characterizing what it means to be surface oriented, but on directly characterizing what it means to **not** be surface oriented.[4] The concept of phonological **opacity** has conventionally been used to characterize situations in which the behavior of a phonological process in a derivation fails to be clearly reflected in the output of that derivation (Kiparsky 1971, Kiparsky 1973). Expanding, a process would be opaque in a phonological map if it was opaque in the derivation associated with at least one input–output mapping of the map. Roughly put, a process is said to be opaque if it contributes meaningfully to the analysis of a phonological map (either by applying or by not applying), but the conditions for its (non-)application are not reflected in the output for some derivation. Surface orientedness is then characterized by complementarity: surface orientedness is identified with a lack of opacity.[5]

This can involve either of two somewhat different types of circumstances. In one circumstance, the output contains a component satisfying the conditions for the application of a process. The fact that the component is in the output means that the process has not applied to that component, despite meeting the conditions of the process.[6] In the terminology of McCarthy 1999, such a process is not **surface-true** in that derivation (for that component). In the

4 This suggests a view in which surface orientedness is the "default" situation, holding of most of phonology, and in which the phenomena that seem not to be surface oriented are the more unusual.

5 Processes which are not opaque are often labeled "transparent."

6 Or at least, if the process applied to an earlier form of that component during the derivation, then the component was subsequently returned to the state satisfying the conditions of the process, and the process did not apply again to that component prior to the completion of the derivation.

other circumstance, a process has applied, and the product of that process is present in the output. However, some relevant aspect of the conditions necessary for the application of the process no longer remains. In the terminology of McCarthy, such a process is not **surface-apparent** (in that derivation, for that component). One could then take the view that a phonological map is surface oriented if none of the processes used in defining it are opaque in any of the derivations giving rise to the map. On this view, evaluating the surface orientedness of a phonological map is dependent on the particular processes used to characterize it.

Using phonological processes to assess surface orientedness can be particularly awkward in a theory like Optimality Theory (Prince and Smolensky 1993/2004), in which processes are not primitives of the theory, but are at best descriptive commentaries on the theory, subject to equivocation. Observing that there are no processes,[7] opaque or otherwise, and claiming therefore that OT is vacuously surface oriented, seems merely to duck the issue entirely, rather than address it. OT may not have processes, but it does introduce disparities between inputs and outputs and does so by categorizing (in various ways) both disparities and potential occasions for their introduction. Pursuit of intuitions regarding surface orientedness in such theories seems to require a characterization of surface orientedness that is not dependent on processes. Two possible approaches come to mind: (a) a different characterization in terms of the theoretical constructs of another particular theory (like Optimality Theory); (b) a more abstract characterization that makes reference only to the inherent properties of maps themselves. This book is the result of pursuing option (b).

1.2 Surface orientedness in Optimality Theory

Before outlining the proposed abstract characterization of surface orientedness, I will in this section briefly examine a couple of attempts to characterize surface orientedness in terms of the theoretical constructs of Optimality Theory. I argue that both attempts are ultimately unsatisfactory in that they fail to capture some key aspects of surface orientedness. This provides further motivation for pursuing the more general characterization described in the next section.

7 By "process" I explicitly mean a part of a theory which bundles together specification of a conditioning environment and a change to be made. The term is also sometimes used in a descriptive sense, referring to phonological patterns relating changes to conditioning environments. The latter sense is applicable to Optimality Theory, as it is to any theory which involves mapping inputs to outputs. The descriptive notion of process refers to patterns to be explained by phonological theory, not components of phonological theory which explain patterns.

The two attempts to characterize surface orientedness in terms of Optimality Theory are an interesting pair, because they lead to opposite conclusions about the capacity of Optimality Theory to realize maps that are surface oriented. Both focus on markedness constraints as the particularly important theoretical constructs, because markedness constraints by definition only evaluate outputs, and are quite easily understood as restrictions on outputs.

1.2.1 *Markedness violations as opacity*

One attempt builds on a general notion of opacity as a matter of generalizations that are not surface-true. This is the view put forth by Idsardi (2000): "An opaque generalization is a generalization that does crucial work in the analysis, but which does not hold of the output form." Note that opacity is a property of generalizations, not maps. But one could by extension suggest that a particular analysis of a map is not surface oriented if it involves one or more opaque generalizations. This view of surface orientedness can be taken to suggest that OT is inherently *not* surface oriented: generalizations in OT are expressed by constraints, constraints are inherently violable, and in practice nearly any worthwhile OT analysis involves at least some constraint violation. This includes violation in optimal candidates of constraints that are active in the grammar (constraints that eliminate competitors), active constraints being ones that are clearly doing crucial work. Because markedness constraints only evaluate outputs, violations of markedness constraints in particular constitute instances where generalizations do not hold of the output form. The view that OT markedness constraints which are violated in grammatical outputs constitute generalizations which do not hold of the output form has been stated explicitly by McCarthy (1999: 332) and Idsardi (2000: 342).

The analogy between processes and markedness constraints is less than perfect. Processes are generalizations about derivations, not outputs. One cannot determine if a structure in an output is the result of a particular process based on the output alone; it is necessary to recover the relevant aspects of the derivation. Markedness constraints are generalizations about outputs. A structure in an output does or does not constitute a violation of a markedness constraint regardless of whether it is faithfully preserved from the input or it is in part a consequence of disparities between the input and the output.

It should be emphasized that (to the best of my knowledge) neither McCarthy nor Idsardi has claimed that markedness constraints constitute the *only* linguistically relevant generalizations in an Optimality Theoretic context. But the nature of other generalizations is left unclear, particularly with respect to opacity. The faithfulness constraints might seem like natural instances of such other

generalizations in Optimality Theory, but it is not immediately obvious how a faithfulness constraint can be said to hold or not hold of an output. Simply asserting that any violation of a faithfulness constraint constitutes a failure of the generalization to be surface-true/apparent leads toward the thoroughly unenlightening classification of maps by whether or not they permit any disparities at all. The class of all and only maps that permit no disparities anywhere, while simple enough to describe, completely fails to correspond to the class of surface-oriented maps in any interesting way. Surface orientedness requires that disparities be introduced only to satisfy output restrictions, not that disparities never be introduced.

The statement that markedness constraints which are violated in grammatical forms constitute generalizations that don't hold of surface forms is straightforward. However, the surface-trueness of markedness constraints *alone* doesn't seem to do justice to the intuition of "surface oriented" as a property of maps. To take an example from McCarthy, he offers, as an example of a generalization in an OT analysis that isn't surface-true, an analysis of syllable structure involving the constraint ONSET, which is violated by syllables that lack an onset, and expresses the generalization "syllables have onsets." The constraint plays an active role in the analysis, yet there are grammatical outputs with onsetless syllables: the language permits word-initial syllables to lack onsets.[8] Suppose, for the sake of discussion, that consonants are inserted into onset position to ensure that non-initial syllables have onsets, so the map includes mappings like the ones in (1.4).

(1.4) Mappings in which non-initial syllables must have onsets
/kika/ → [ki.ka]
/ika/ → [i.ka]
/kia/ → [ki.ka]
/ia/ → [i.ka]

While the onsetless word-initial syllables certainly constitute violations of the constraint ONSET, the map itself nevertheless seems surface oriented: consonants are inserted (introducing disparities) only to the extent necessary to ensure that non-initial syllables have onsets (an output restriction). The generalization expressed by ONSET is violated on the surface, but relates to a generalization that is not violated on the surface: all non-initial syllables have onsets. The disparities are characterizable with an output restriction, even if that restriction isn't identical in content to any single markedness constraint.

8 This is a piece of a more complex illustration given by McCarthy 1999, fn. 1, illustrating different kinds of non-surface-true generalizations.

A useful comparison is with a chain shift. Modifying the previous example, suppose we now have a map in which consonants are inserted to ensure that non-initial syllables have onsets (like the previous example), and additionally in which a consonant is inserted into onsetless initial syllables, but only when all other syllables have onsets from the input (not inserted). Such a map would include mappings like the ones in (1.5).

(1.5) Mappings that include a chain shift
/kika/ → [ki.ka]
/ika/ → [ki.ka]
/kia/ → [ki.ka]
/ia/ → [i.ka]

The mappings /ia/ → [i.ka] and /ika/ → [ki.ka] jointly constitute a chain shift pattern. If the introduction of disparities were motivated purely by output restrictions, and /ia/ → [i.ka], then [i.ka] ought to be "good enough" with respect to the output restrictions (after all, it is a valid output). It seems that /ika/ should map to [i.ka], a "good enough" output not requiring any disparities with the input (no disparities are necessary). If [i.ka] is not "good enough" as an output, so that the disparity of inserting an initial consonant is warranted in /ika/ → [ki.ka], then that disparity should be warranted for input /ia/ also: /ia/ should map to [ki.ka]. As McCarthy puts it, the issue is the failure of /ia/ to make a "fell swoop" all the way to [ki.ka] (McCarthy 1999: 364). The disparity of inserting an initial consonant is unavoidably conditioned by something other than the "goodness" of the output itself. If in one instance [i.ka] is good enough, and in the other instance it isn't, then something other than restrictions on the output alone must be conditioning the introduction of the disparity (the inserted consonant).

The distinction between the two examples is not simply a matter of distinct phonotactic regularities. Both examples have the same output inventories: the valid outputs are [ki.ka] and [i.ka]. Both examples have outputs in which the initial syllable lacks an onset. Both examples have no outputs in which a non-initial syllable lacks an onset. The difference concerns the relationships between output restrictions and the disparities that are introduced in mappings.

Note that the intuition concerning chain shifts, as stated here, is about the map itself, not about any generalization/output restriction that one might use in describing the map (such as "non-initial syllables must have onsets"). The key mappings (both grammatical and not) and their disparities are given in (1.6). Since /ia/ → [i.ka] has the inserted [k] in the second syllable as its only disparity, and the first syllable does not have an onset, the input /ika/, which wouldn't require any disparities to reach the same output [i.ka], would

be expected under surface orientedness to map to [i.ka]: any introduction of disparities for input /ika/ is not required by output restrictions, because [i.ka] is a valid output. The fact that /ika/ maps to [ki.ka] instead violates the intuitions of surface orientedness. The violation of surface orientedness can be described in terms of the fact that [i.ka] is a valid output, without reference to the content of any particular output restriction. Indeed, it shouldn't be possible for a pure output restriction to distinguish two cases with identical outputs.

(1.6) Key input–output pairs with their disparities

/ia/ → [i.ka]	Disparities: inserted [k] in second syllable
* /ika/ → [i.ka]	Disparities: none
/ika/ → [ki.ka]	Disparities: inserted [k] in first syllable

The discussions of opacity in the work by Kiparsky, McCarthy, and Idsardi all focus on properties of the generalizations expressed by immediate constructs of phonological theories, e.g., the rules of SPE or the markedness constraints of Optimality Theory. Since phonology is in the business of characterizing generalizations about phonological systems, such a focus is (quite reasonably) to be expected. And yet, it ties the characterization of surface orientedness tightly to the constructs of particular theories and in the process misses the intuition that the maps themselves are or are not surface oriented. The generalization that non-initial syllables must have onsets feels quite surface oriented, in a way that is insensitive to whether the relevant theoretical construct is a statement about syllables (which would not be surface-true) or a statement about non-initial syllables (which would be surface-true).

1.2.2 Markedness constraints cause disparities

Another attempt to characterize surface orientedness in terms of Optimality Theory argues that the only constraints that can cause disparities are markedness constraints, and therefore disparities are always a consequence of the enforcement of output restrictions (the markedness constraints). This leads to the (opposite) conclusion that OT is inherently surface oriented.

This view appears to focus to a somewhat greater extent on the map itself. It makes no reference to the content of markedness constraints, save for the fact that they only evaluate output forms. In particular, it says nothing about whether markedness constraints are actually "surface-true" (unviolated) in a language. The argument implicitly rests on the intuitive assumption that candidates with zero disparities satisfy all faithfulness constraints, and thus can only lose to another candidate on the basis of a markedness constraint (whether or not that markedness constraint is surface-true in the language overall).