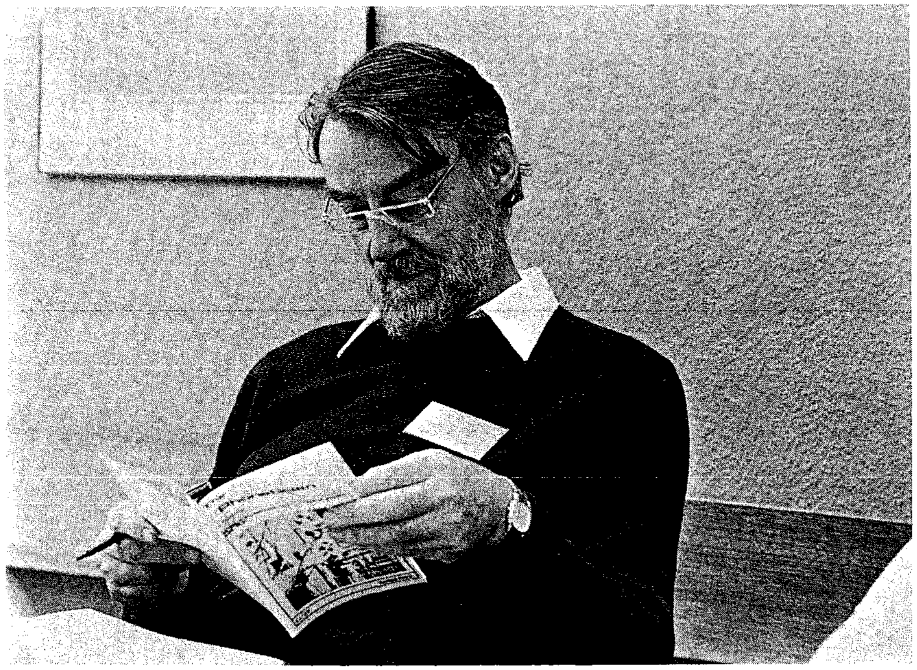

Peter Ladefoged

PHONETIC LINGUISTICS

Essays in Honor of Peter Ladefoged

Edited by

Victoria A. Fromkin

Department of Linguistics
University of California, Los Angeles
Los Angeles, California

1985

ACADEMIC PRESS, INC.
(Harcourt Brace Jovanovich, Publishers)
Orlando San Diego New York London
Toronto Montreal Sydney Tokyo

Leben, W. R. (1978). The representation of tone. In V. Fromkin (Ed.), *Tone: a linguistic survey* (pp. 179–219). New York: Academic Press.

Levin, J. (1983). *Reduplication and prosodic structure*. Unpublished manuscript, MIT.

Liberman, M. Y. (1975). *The intonational system of English*. Unpublished doctoral dissertation, MIT.

McCarthy, J. (1979). *Formal problems in Semitic phonology and morphology*. Unpublished doctoral dissertation, MIT.

McCarthy, J. (1984). A prosodic account of Arabic broken plurals. In I. R. Dihoff (Ed.), *Current approaches to African linguistics*. Dordrecht: Foris Publications.

Pierrehumbert, J. B. (1982). *The phonology and phonetics of English intonation*. Unpublished doctoral dissertation, MIT.

Poser, W. J. (1982). Phonological representations and action-at-a-distance. In H. van der Hulst & N. Smith (Eds.), *The structure of phonological representation* II (pp. 121–158). Dordrecht: Foris Publications.

Pulleyblank, D. G. (1983). *Tone in lexical phonology*. Unpublished doctoral dissertation, MIT.

Rotenberg, J. (1978). *The syntax of phonology*. Unpublished doctoral dissertation, MIT.

Steriade, D. (1982). *Greek prosodies and the nature of syllabification*. Unpublished doctoral dissertation, MIT.

Williams, E. (1976). Underlying tone in Margi and Igbo. *Linguistic Inquiry, 7,* 463–484.

Yip, M. (1982). Reduplication and C-V skeleta in Chinese secret languages. *Linguistic Inquiry, 13,* 637–661.

8

Universal Phonetics and the Organization of Grammars*

Patricia A. Keating

1. INTRODUCTION

Phoneticians have long been interested in the relation between phonetics and phonology, especially since the rather explicit proposals of Chomsky and Halle (1968) in *Sound Pattern of English* (SPE). Much of the attention has focused on the nature and substance of the phonetic feature system; Ladefoged (e.g., Ladefoged 1971, 1980) has been a notable participant in this discussion. However, it is of some theoretical interest that, in the SPE model, phenomena that might be called "phonetics" are found in two separate places. On the one hand, the phonetic rules that convert binary into scalar feature values are part of the phonological component of the grammar. On the other hand, the part of phonetics actually called phonetics is not technically in the grammar. It is a largely universal and predictable component that translates a segmental phonetic transcription into continuous physical parameters. Broadly speaking, this extragrammatical physical phonetics is the locus of many of the traditional (as well as current) concerns of phoneticians—articulation, timing, and coarticulation, for example. In this chapter I consider the division of labor between phonology and phonetics in more detail and suggest a direction for revision in the model.

Figure 8.1 gives a schematic view of the relevant parts of the SPE model. First, the phonological component of the grammar contains both phonological rules that operate on binary-valued features and language-

---

* Earlier versions of this work were presented as talks to helpful groups at the University of British Columbia, the University of California at Irvine, the 1983 International Congress of Phonetic Sciences, and the Phonetics Lab at UCLA. I would especially like to thank Peter Ladefoged, Vicki Fromkin, Marie Huffman, Bruce Hayes, and John Ohala for comments on the manuscript, and John Westbury for our collaborations.

ISBN 0-12-268990-9

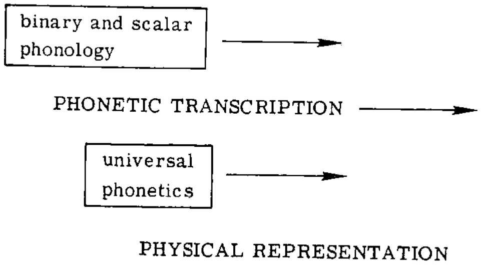

Figure 8.1 The SPE Model.

specific phonetic-detail rules. The phonetic-detail rules convert binary phonological feature specifications into quantitative phonetic values, called the "phonetic transcription," or systematic phonetic representation. These rules in part depend on universal phonetic constraints concerning possible combinations and contrasts. "Given the surface structure of a sentence, the phonological rules of the language interact with certain universal phonetic constraints to derive all grammatically determined facts about the production and perception of this sentence. These facts are embodied in the 'phonetic transcription' " (Chomsky & Halle 1968:293). According to Chomsky (1964), phonological rules apply until a representation in a universal phonetic alphabet results. The phonetic transcription represents "what the speaker of a language takes to be the phonetic properties of an utterance, given his hypotheses as to its surface structure and his knowledge of the rules of the phonological component"; it is not a direct record of the speech signal and is only one "parameter determining the actual acoustic shape of the tokens of the sentence" (Chomsky & Halle 1968:293, 294). The physical utterance itself is not generated by the grammar; the phonetic transcription is the terminal output of the grammar.

As such, the phonetic transcription must be further interpreted (translated, spelled out, realized) as a physical phonetic representation by a phonetic component that is not technically part of the grammar. The assumption here is that, with the right phonetic representation, any utterance in any language can be interpreted by a set of phonetic conventions. The translation from discrete segments to articulations that exist in time is treated as being automatic; the phonetic component includes, for example, "the different articulatory gestures and various coarticulation effects—the transition between a vowel and an adjacent consonant, the adjustments in the vocal tract shape made in anticipation of subsequent motions, etc." (Chomsky & Halle 1968:295).[1] Chomsky and Halle further suppose that these phonetic conventions are universal rules, that the same phonetic rules can interpret a phonetic transcription in any language. Although

not strictly necessary within the general model, this view certainly is an appealing one. The phonology contains the language-specific statements required to produce a detailed enough transcription to allow phonetic interpretation, and the phonetics converts that transcription into a physical utterance in a quite automatic way. This distinction between language-specific rules and automatic low-level phonetic rules is in some ways similar to the distinction between extrinsic and intrinsic allophones (MacNeilage 1970) or between soft and hard coarticulation (Fujimura & Lovins 1978). The phonetic rules are often thought to be directly motivated by, or be identical to, physical constraints on articulation or perception, although Chomsky and Halle apparently do not hold this view.

The SPE model of grammar thus specifies a very constrained relation between phonological and phonetic representations. The use of phonetic features in phonological representations ensures that lexical representations and phonological rules can be evaluated for their phonetic naturalness (the Naturalness Condition, Postal 1968). However, the phonetic representation may, under this theory, be much broader or much narrower than a traditional transcription. Here the phonetic transcription is defined by its position in the model between the phonological and phonetic components. It follows, then, that the fewer universals of phonetic realization posited, the narrower the phonetic transcription will have to be, while the more such phonetic universals there are, the broader the transcription will be. Suppose, for example, that speakers interpreted all phonetic properties as grammatical ones. In that case, the phonetic transcription as defined by the theory and generated by the grammar would be much narrower than the traditional segmental phonetic transcription. Such a view is found in Pierrehumbert's work (1980) on intonation. The language-specific quantitative rules in her system directly output something very close to a physical utterance.

How could a more traditional phonetic transcription be maintained if speakers interpreted phonetic properties as grammatical? Clearly we would have to say that the interpretive rules of the phonetic component are part of the grammar. This move would recast the phonetic component so that it is no longer mainly the domain of universal conventions. Would such a revision be completely arbitrary, just to preserve the phonetic transcription? We might propose that, rather than provide automatic aspects of interpretation, the phonetic component derives any aspect of a representation in which continuous time is involved (Anderson 1974). Then the phonetic component could still do much of the work of deriving the physical phonetic form of an utterance, and the transcription could be a fairly broad one.

Some revision in the SPE model is required in one of these directions simply because in fact there does not appear to be a well-defined body of phonetic universals that operate automatically across languages. Phoneticians are aware that many supposed universal rules of phonetic interpretation have exceptions. What is to be made of these exceptions, of the phonetic rules, and of the phonetic component of the SPE model? Is there any role for a universal phonetic component, in or out of the grammar? And what is the relation of near-universals to physical phonetic constraints?

Questions about the nature of phonetic rules do not disappear if we reject the SPE model of grammar. In fact, the importance of the issue is only increased when we look at alternative theories proposed in the 1970s. Some rejections of the SPE model have been based on phonetic naturalness as a defining property of phonology. In these views, naturalness, in turn, is generally linked to the mechanisms of speech production and to phonetic universals (e.g., natural generative phonology, Hooper 1976). Phonological rules are constrained to be those natural rules that are exceptionless because they are directly physiologically motivated.

In this paper I examine three known phonetic patterns in light of the SPE model and the discussion above. With the first pattern, I discuss the fact that phonetic patterns are not necessarily automatic results of speech physiology. With the second pattern, I illustrate that they need not be universal and that they can operate as abstract phonological rules. With the third, I consider the limitations of physiology in determining phonetic patterns. While none of these types of observations is original, taken together they lead to some tentative proposals about the place of phonetic patterns in grammars.

2. INTRINSIC VOWEL DURATION

In most if not all languages, low vowels such as [a] and [æ] are longer than high vowels such as [i] and [u], all things being equal (Lehiste 1970). Not only can this phonetic pattern be observed across languages, but a physical explanation has been suggested. As Lehiste notes, lower vowels require a greater articulatory movement; if movement velocity is nearly the same across vowels, lower vowels are longer. Furthermore, the observed differences in vowel duration could be accounted for by automatic biomechanical effects rather than by deliberate temporal control. Lindblom (1967) provided an explicit account of such an automatic effect with a mechanical model of jaw activity. Of course, since Lindblom mentioned that compensations in vowel durations could be made, he did not intend the model to account automatically for all aspects of intrinsic vowel

duration. However, his work is important because in principle it could provide such an account. Lindblom's model showed that if the force input to jaw lowering muscles had the same duration but different amplitudes for different vowels, biomechanical sluggishness would automatically result in the correct vowel-duration pattern. That is, if the jaw gets a harder, but not longer, send-off for lower vowels, which translates automatically into longer movements, then no explicit timing representation is needed. By hypothesis, all vowel heights have the same representation for duration at every point in their production. Thus no information about intrinsic vowel duration need be included in a grammar, since intrinsic vowel duration patterns can be accounted for automatically in a universal component. Such a view is compatible with what Fowler (1980) called "extrinsic timing" models, in which time is never specifically included in the plan for an utterance but is introduced only in production. That is, here is a case where, apart from any cross-linguistic data, modeling of the speech-production mechanism supports an automatic phonetic universal.

An experiment by Westbury and Keating (1980) investigated this claim about speech production in a physiological study of spoken vowels. Electromyographic (emg) techniques were used to study the force input to a jaw-lowering muscle for vowels, the anterior belly of the digastric (ABD). Three American speakers read items of the form /sVts/ 15 times each, where V = each of ten English vowels. We recorded simultaneously on channels of FM tape the speech signal from a microphone, the mandible displacement from a strain-gauge device attached to a tooth splint, and the emg signal from the ABD as measured with hooked wire electrodes. We then measured the acoustic vowel duration from the speech signal, the extent and timing of jaw displacement from the movement signal, the emg duration from the emg waveform, and the emg maximum amplitude from the rms time envelope.

Our results replicated the earlier finding by others (see Lehiste 1970 for literature review) of intrinsic vowel duration: Lower vowels are longer in acoustic duration than higher vowels, especially (but not crucially) if the English phonological distinction between tense–lax or long–short vowels is taken into account. The measurements also showed (as expected) that lower vowels have a lower jaw position than higher vowels. In addition, the two measures were statistically correlated: The vowels with the lower jaw position had longer acoustic durations. The longer durations were due to longer travel times, not longer steady states. Thus we obtained the data enabling us to address the question of force input. We found that the emg duration and maximum amplitude both showed the same pattern across vowels, with low vowels having longer durations and higher amplitudes of emg activity, correlating with jaw displacement. That is, more extensive and longer movements are made with a force input that is both longer in duration and higher in amplitude: The abd muscle fires longer and more actively; loosely speaking, it pushes both longer and harder to go farther.

We conclude, then, that at the level of neural control tapped by measuring emg activity, vowels are represented as having different durations, since the muscle firing varies in duration, like the acoustic vowel signal. Thus vowel duration differences are not due directly to sluggishness of the jaw; rather, they are controlled as such. This result in itself does not show that vowel-duration differences must be language specific or represented in the grammar. Durations could be provided at a very late stage in production before motor commands are issued, by the phonetic component. However, if vowel duration is a controllable parameter, it is in principle available for language-specific manipulation. Thus we should expect to find languages with different vowel-duration patterns. Such patterns could be seen, for example, when low vowels are considered phonologically short and are produced with short travel times and high velocities.

We may still wonder why so many languages have similar patterns of intrinsic vowel durations. Though the pattern is not a necessary one, it must be convenient in some sense. It may be that some physical patterns and movements are preferred over others because of general principles of economy of effort and motor control (e.g., Nelson 1980). The point here, though, is that such principles must be more subtle than absolute mechanical constraints of the sort that might have been proposed. Physical factors clearly influence vowel duration, but they do not control it.

3. EXTRINSIC VOWEL DURATION

Consider next the general finding that vowels are shorter before voiceless obstruents than before voiced obstruents or sonorants.[2] Chen (1970) surveyed a number of languages, including some described in the literature, and found such vowel duration differences in all of them. Of seven languages studied, all showed at least a 10% difference in vowel duration. This was so whether the vowel and consonant were word final and tautosyllabic or word medial and heterosyllabic. Chen suggested that some contextual durational difference is universal and physiologically determined, although languages may individually exaggerate this difference by rule, for example, English. Although there were problems with Chen's cross-language comparisons,[3] it has generally been accepted that vowel-length differences depending on consonant voicing constitute a phonetic universal, albeit one whose mechanism is not understood. Fromkin (1977) uses this result to argue that the vowel duration effect is given by phonological rule in English (that is, is represented in the phonetic transcription) but is automatically supplied by universal phonetic conventions in other languages without the exaggeration. However, the pattern is not a universal one, and it must be given by rule even in some languages that do not exaggerate the effect.

As part of a study on Polish voicing contrasts (Keating 1979), Polish vowel durations before voiced and voiceless consonants were measured for the pair *rata–rada*. Polish, like other Slavic languages, has a rule of word-final devoicing, so there are no voicing contrasts at the end of isolated words. Thus the durational phenomenon can only be studied in medial position, although based on Chen's survey and on the English studies cited below a robust difference is still to be expected. Twenty-four speakers in Wrocław, Poland, were recorded reading this pair, and durations of the stressed syllabic nuclei were measured from a computer-implemented oscillographic display at the Brown University Phonetics Lab. The mean duration of [a] before [t] was 167.4 msec and of [a] before [d], 169.5 msec. The ratio of these two means is .99. In addition, the ratio of the two vowel durations was computed for each individual speaker. The mean of these 24 ratios is 1.0. These data indicate that Polish vowel duration does not vary systematically according to the voicing of the following consonant.

Comparable data for English vowels before medial stops have been collected by Sharf (1962) and Klatt (1973). The pre-voiceless to pre-voiced ratios they obtained were .75 and .79, respectively. A higher ratio, .89, was obtained by Port (1977), using sentence contexts rather than word lists. (Since English flaps its medial alveolar stops before stressless vowels, these data are for vowels before labials and velars.)

The finding that Polish, unlike English, does not shorten vowels before voiceless consonants was extended by recording speakers of Czech. As both Czech and Polish are West Slavic languages, they are similar in many ways, but Czech has phonemic vowel-length contrasts. Thus it seemed possible that Czech would also fail to differentiate vowel durations according to consonant voicing, so that vowel duration could be reserved for the phonemic length contrast. Three native speakers of Czech read several words of the following form:

$$\begin{Bmatrix}\text{p}\\ \text{ml}\end{Bmatrix}\begin{Bmatrix}\text{a}\\ \text{a:}\end{Bmatrix}\begin{Bmatrix}\text{t}\\ \text{d}\end{Bmatrix}\ \text{V (C)}$$

The number of phonemic short and long vowels was balanced. The mean duration of vowels before [t] was 193.7 msec, before [d], 204.2 msec. The ratio of these two means is .95, and the mean of the individual ratios is .98. Thus there is a slight tendency for vowels to be shortened before voiceless consonants, but the difference in durations did not reach statistical significance ($t_{30} = -.37$, $p > .20$). In sum, neither Czech nor Polish disyllables show the supposed universal vowel shortening before voiceless consonants.

One line of explanation that has been offered for vowel length differences involves the fact that closure-interval durations also vary with voicing and are inversely related to the vowel durations. That is, voiceless stops have longer closure intervals than do voiced stops. In English and presumably other languages with vowel lengthening, the two ratios, vowel and closure, essentially balance each other, so that the syllable duration is relatively constant. What happens to closure, and syllable, durations in Polish? Closure durations for the same 24 pairs were also measured. The mean duration for [t] was 130.1 msec and for [d], 91.5 msec. The ratio of these means is 1.42, and the difference is statistically significant ($t_{23} = 8.81$, $p < .001$). Comparable data for English labial stops are Lisker's (1957) ratio of 1.60, and Port's (1977) ratio of 1.35 (again, from sentence contexts). Thus Polish, like English, has longer closure durations for voiceless stops. Because Polish shows the closure but not the vowel effect, its syllable durations are not balanced.

This finding indicates that the vowel-shortening effect, in those languages where it occurs, is not physiologically determined by the closure-duration effect. Of course, there could still be some nonphysiological relation between closure and vowel duration that some languages could choose to implement. For example, language-specific prosodic factors like stress or rhythm could make it desirable to balance intrinsic syllable durations. This factor may operate more powerfully in a language like English, with variable stress and vowel reduction, than in a language like Polish, with fixed stress.

Thus the possibility that vowel shortening before voiceless consonants is an (automatic) phonetic universal is not supported by an investigation of Polish and Czech. Further counterevidence from Saudi Arabic is found in Flege (1979). He found that long /aː/ was not significantly longer before word-final /d/ than /t/. Therefore, we know that this rule cannot be placed in a universal phonetic component because it does not occur universally across languages. Rules of phonetic vowel duration as a function of voicing of a following consonant must be language specific.

Furthermore, Chen's study shows that exceptions to the phonetic pattern take still another form. For example, Chen found a vowel-duration difference in Russian completely comparable to that in other languages, although a footnote indicates that all the final consonants determining the vowel durations were voiceless, Russian having a rule of final devoicing.[4] The duration pattern was apparently determined by underlying values of the voicing feature. A similar difference was found for German and was shown to be marginally perceptible by Port *et al.* (1981). In the same way, vowel duration for speakers of some English dialects varies before voiced flaps according to underlying stop-voicing values (Fox & Terbeek 1977). Clearly, if vowel durations can be determined by underlying phonological values for voicing, then the relation between vowel duration and voicing cannot be automatic and physiological. It is important to realize that these cases are actually counterexamples to the pattern at the systematic phonetic level, since in Russian longer vowels occur before voiceless consonants, and in the English dialects shorter vowels occur before voiced consonants. The pattern is clear only at some point in the derivation before the phonetic transcription.

At the same time, there is obviously a trend across languages and across phonological rules that must be accounted for. We can summarize the possibilities as follows: Languages can show no vowel durational differences, or they can show some kinds of differences that relate shorter vowels to following voiceless obstruents. If they do show such a pattern, they can do so at either the phonetic or phonological level. No language shows durational effects in which vowels are shortened before all voiced consonants and lengthened before all voiceless consonants. It is as if there were a possible patterning available to languages: Vowels may be shorter before voiceless consonants. The reverse pattern is not available in this way. Thus we find languages like Polish and Czech with no difference, languages like French and some English dialects with shorter vowels before phonetically voiceless consonants, and languages like Russian and German with a phonologically conditioned pattern.

This example of extrinsic vowel-duration patterning shows that a supposed phonetic universal is not in fact universally attested. Because of this fact, and because the extent and level of duration differences varies across those languages with the pattern, the pattern cannot be automatic or predictable. Each language must specify its own phonetic facts by rule. Possibly, following Fromkin (1977), we could say that languages with an exaggerated pattern and languages with no pattern must include a rule in their grammars. In addition, languages whose patterns are not exaggerated but operate on phonological representations must also include a rule in their grammars. This leaves languages with an unexaggerated duration difference that is entirely phonetically conditioned. Following Fromkin, this pattern could be provided by the phonetic component.

Obviously, however, such a treatment entails a change in the conception of the phonetic component. Rather than a phonetic universal that is predictable and automatic, that phonetic statement would represent one special case, simply a kind of "elsewhere" condition on phonetic detail. Alternatively, the phonetically conditioned cases could be treated exactly like the phonologically conditioned cases, by a grammatical rule.

As in the intrinsic duration case, it appears that the role of the phonetics is to provide a pattern that might be preferred. Within any one language, however, vowel duration is controlled by the grammar, even though it is a low-level phonetic phenomenon. While it is a good idea to continue looking for phonetic universals that would support a model of automatic phonetic interpretation, it seems more likely that our eventual model will incorporate phonetic rules of timing into the phonology.

4. VOICING TIMING

In the case of the occurrence and timing of stop-consonant voicing, each of the investigative methods considered above has been employed by a number of people. Cross-language surveys have revealed patterns that must be explained, and modeling studies have tried to provide some explanations. What is interesting is that none of the patterns found is universal, yet each is a good example of phonetic naturalness. Thus these patterns are a key to the relation among physical motivations, phonetic rules, and the grammar.

The sort of patterns I have in mind are exemplified as follows. Surveys of phoneme inventories (e.g., Maddieson 1984) produce two major observations. First, voiceless stops are generally preferred to voiced stops, especially for geminates. Second, the extent of this stop-consonant preference with regard to voicing varies according to place of articulation, with further front stops more likely to be voiced. Thus some languages have /b/ but no /p/ (labials favor voicing), or /k/ but no /g/ (velars favor voicelessness). Surveys of allophone occurrence and detail lead to similar conclusions. In most environments, voiceless unaspirated stops are favored, even in intervocalic position, contrary to popular belief (Houlihan 1982; Keating, Linker, & Huffman 1983). Place-of-articulation effects on the duration of voicing and of aspiration can be observed across languages (Lisker & Abramson 1964), although various exceptions have been noted. In this section I confine discussion to the more categorial effects on voicing discussed in Keating *et al.* (1983), namely, the position-in-utterance preferences seen in unrelated languages.

The best-known work on physiological motivations for voicing patterns in general is probably that of Ohala (much of it summarized in Ohala 1983). Ohala has used a simple model of breath-stream dynamics to illustrate the common observation that voicing requires glottal airflow, while stop occlusion impedes such airflow; in this sense stop occlusion and voicing are at odds with each other. Thus it is understandable that voiceless stops should be more common than voiced. He also used the model to reason about the further patterns found. Drawing on other modeling work by Rothenberg (1968) and Muller and Brown (1980), Ohala stressed the role of passive and active expansion of the vocal tract walls in allowing airflow, and hence voicing, to continue during stop occlusion. Wall expansion is related to findings about place of articulation in that the further front the occlusion, the more expandable vocal tract wall area there is between glottis and occlusion. Thus we should expect further-front places to allow voicing continuation more easily than further-back places.

Westbury and I, together and separately, have looked in more detail at effects of place of articulation, position in utterance, and stress on Voice Onset Time (VOT) and closure voicing duration. A model of voicing based on Rothenberg's was devised and is described in more detail in Westbury (1983). It allows us to vary over time the subglottal pressure, the position of the vocal cords, the oral constriction in three dimensions, and the stiffness of the vocal tract walls. We use results from an x-ray study (Westbury 1979) and a tracheal puncture study (conducted by Westbury and others at the University of Texas Phonetics Lab) for constriction and pressure data, and other published data on factors such as glottal opening as inputs (see Westbury 1983 for references). The computer program takes these inputs and calculates the resulting airflows and air pressures in the vocal tract. From the airflow through the larynx we can see exactly when voicing should occur. In the case of position-in-utterance effects, our results (Westbury & Keating 1984) were clear. We compared initial, intersonorant, and final positions, assuming that the only difference across them was the subglottal pressure being generated. We assumed that the vocal cords were equally ready to vibrate in all three positions and that the closures and velocities of the oral gestures were the same, except (noncrucially) that initial closures were longer. Such modeling showed that the pressure differences result in three different acoustic patterns. In initial position, voicing does not occur until after consonant release with these inputs; in medial position, voicing continues from the preceding sonorant through most but not all of the stop occlusion; in final position, voicing continues into the beginning of the occlusion but ceases earlier than in medial position.

A preference of languages for voiceless unaspirated initial and final allophones is thus seen to arise from the physical operation of the speaking device. What does this preference explain? It may be useful to compare our account of final stop voicelessness with Dinnsen's (1980) discussion of supposed aerodynamic explanations of phonological rules of final devoicing. He distinguishes explaining the structural description of a rule (here, that it affects final stops) from explaining its structural change (here, that it devoices them) and from explaining why there should be any rule in the first place (here, some difficulty posed by final voiced stops); he says that only the structural description is explained by, for example, the work of Ohala. Our explanation is different from Ohala's,[5] however, and goes further in illuminating the structural change of the final devoicing rule and arguably its motivation. This improvement comes from carefully quantifying the articulatory conditions that hold before the stop consonant, the acoustic characteristics of a stop in which those conditions are changed only minimally, and the acoustic characteristics of stops in which those conditions are changed more drastically. A devoicing rule specifies a structural change most in accord with the result of a minimal change in articulatory conditions.

However, as Dinnsen emphasizes, a phonological rule exists independently of such a phonetic motivation. That this must be so in the case of final devoicing is shown by the fact that our motivation applies only to position-in-utterance effects. Position in utterance is not the same as position in word, and many linguistic rules and constraints operate in the word domain. For example, instances of word-final devoicing in utterance-medial position are phonetic counterexamples to the patterns generated by the model. They may serve to demarcate word boundaries in running speech, for example, but are no longer directly physically motivated. At best, then, physiology motivates one basic case that can be incorporated arbitrarily into phonological rules.

Does that mean that those cases where a linguistic voicing pattern does correspond directly to outputs of the model are in fact automatic? The answer must be, only if controlling articulation in the way we have assumed is automatic. It is important that specific sets of articulatory inputs are required to produce the outputs discussed here. The speech production system must be controlled by a real speaker in a way that ensures those inputs and no others. Possibly, as was suggested for the extrinsic vowel duration case, such control of the phonetic pattern is provided outside the phonology, with only the exceptions given in the phonology. But already that means that some very low-level phenomena of timing are to be included in the phonology. Consider, for example, the pattern for place of articulation to correlate with VOT, presumably due in part to differences in the movement velocity of the various articulators. Suppose that in some language this pattern were counterexemplified by having apical stops with lower VOT values than labials, and that the reason was that the upper lip did not participate in the labial gesture (giving a lower net labial-movement velocity). This would mean that in this language the place of articulation counterpattern would be specified in the grammar, though it is concerned with mere milliseconds of timing difference. Thus, if every time we find an exception to a phonetic generalization we state that exception in the grammar, our notion of grammar will be much expanded. In fact, the grammar will include all the kinds of statements that remain in the phonetic component, for no kind of generalization appears to be exceptionless.

On analogy with the use of an articulatory model, we can think of preferred articulatory values as being "default" values of the articulatory system and the outputs that result from these inputs as default outputs of the system. Speakers are not physically constrained to use these default inputs, and it is clear that across languages a wide variety of articulatory values are used.[6] In those cases the language has chosen to override the default settings and substitute more marked settings. Possibly the more substitutions a given output requires, the more marked it will be. Nonetheless, the default settings, where found, must still be specified at some point in the production of an utterance.

5. DISCUSSION

Three candidates for inclusion in the set of phonetic universals have been considered: intrinsic vowel duration, extrinsic vowel duration, and voicing timing. None of them is an automatic consequence of articulatory biomechanics, the strongest view of what a set of universals might be. None of them is necessarily universal. Thus it cannot be the case that a segmental phonetic transcription is automatically interpreted by phonetic conventions, at least with respect to such timing variables. Rather, language-specific rules extend further into phonetics than was assumed in the constrained SPE model. There are two ways that the model can be revised. If the phonetic component still consists of universals, or even just default cases, then almost everything is in the phonology, and the phonetic transcription will be quite narrow. If the phonetic component can include language-specific rules, then the phonetic transcription need not be so narrow, but some independent way of deciding what is in the phonetic component is needed, for example, all timing rules. Phonetic experiments will not determine which of these possibilities is preferable. Only actually trying to devise grammars to include new phonetic data is relevant to that question.

What phonetic experiments can do is identify those parameters that must be controlled by the speaker and default values for those parameters, by studying recurrent phonetic patterns. These patterns exist as options available to languages as physical conveniences but not necessities. Languages must choose whether to incorporate the default and at what level of the grammar. It is not the phonetic patterns themselves that constitute universals; rather, what is universal are the general principles that dictate the default articulatory settings.

Lindblom (1983), in discussing the concept of economy of effort as a factor in the development of sound systems, arrives at a similar overall conclusion. He stresses that speech typically underexploits the capabilities of the speech production system. In his view, more economical speech gestures are favored but are not inevitable. Thus the occurrence and extent of consonant–vowel coarticulation, for example, may differ across speakers or be specified phonologically. Patterns found across languages are due to minimizing the expenditure of energy per unit time; lack of a pattern in a given language indicates a greater level of performance effort of the speech system. Lindblom also concludes that the physiological mechanisms underlying economy of effort are not yet understood.

Previous approaches to language-specific exceptions to phonetic patterns have given a special grammatical role to the exceptions. Stampe (1973) proposed that a child begins acquisition with a set of phonetic processes and replaces some of them with rules on the basis of learning. Hyman (1975) developed the idea of phonologization, that some universal phonetic processes get incorporated into the phonologies of certain languages by being made arbitrary in some way and then play a role in the grammar. But it seems more plausible that every aspect of phonetic control must be learned, for example, the patterns of rise and fall of subglottal pressure that give rise to consonant voicing patterns. I am suggesting here that we consider all phonetic processes, even the most low level, to be phonologized (or grammaticized) in the sense that they are cognitively represented, under explicit control by the speaker, and once-removed from (that is, not automatic consequences of) the physical speaking machine.

Where this account seems unmotivated, as discussed before, are those cases where the default pattern actually occurs without exception phonetically. In these cases it would be possible to say that the default pattern is not controlled by a phonologized rule but that a value is filled in by a phonetic component after all rules have applied. Consider, however, such a phonologization account of extrinsic vowel duration in various languages. That account will distinguish languages like Russian and German (with final devoicing and opaque vowel-length differences) from languages like French (with phonetically transparent vowel-length differences). Russian and German will have phonologized vowel length, while French will not; it will have durations supplied by the phonetics. Suppose now that French acquires a rule of final-consonant devoicing like that of German or Russian, and that, as in German and Russian, vowel length is sensitive to phonological voicing. The phonologization account would have to say that at the moment the devoicing rule is added to the French grammar, vowel length also becomes a grammatical rule, as opposed to a default option or pattern. Since the only change in the vowel-length pattern is that it has changed from phonetically transparent to opaque, then rule transparency must be criterial in assigning phonetic patterns to the phonetics or the phonology. On the other hand, if all phonetic patterns, including transparent vowel length, are represented in the grammar, then the only change in the French grammar is the addition of the devoicing rule. In the absence of arguments for the transparency criterion, then, the phonologization account seems more complex than required.

The view that all phonetic phenomena are controlled by rule has a further interesting implication. Anderson (1981) argued that phonological rules by definition are not natural; they are what is left when everything else is factored out. As a response to various theories of natural phonology, this argument is valid. But it leaves the frequent phonetic naturalness of rules—even rules with exceptions on the surface—unexplained. It sounds ad hoc that some rules (most low-level ones) should actually be natural, while other rules (the opaque ones) only look natural. But once we recognize that all phonetic patterns are rule governed and once-removed from the physical machine, then naturalness can be seen as a more abstract and general property of rules, wherever they are in the phonology. Various rules will have in common the fact that they embody default patterns. Some of these rules will apply transparently; others will apply opaquely. Naturalness is not directly a fact about the speaking machine. It is a fact about the phonological component: The phonology values highly rules that in form indulge the preferences of the speaking machine.

Patterns of phonetic detail are interesting, then, not because they constitute a special universal component outside of grammars, one whose workings are quite different from those of phonology, but rather because they are an integral part of phonology. It seems likely that there are no true linguistic phonetic universals and that the grammar of a language controls all aspects of phonetic form.

NOTES

1. Also taken into account at this stage are nongrammatical suprasegmental parameters, both for languages (base of articulation) and for individuals at a given moment (voice quality, rate of utterance).

2. It does not matter for this discussion whether the pattern is seen as shortening of vowels before voiceless obstruents or lengthening of vowels in converse environments.

3. Chen's comparisons confounded language and position of the vowel + consonant in a word: Some languages were represented mainly by monosyllables, others mainly by disyllables with a medial vowel + consonant. The degree of vowel-duration difference is known to vary even within a single language according to position (compare Sharf 1962 and Klatt 1973 with Lehiste 1970).

4. The rule of devoicing does not guarantee that the neutralized consonants themselves are identical (Dinnsen 1982).

5. Ohala links final devoicing to an observed lengthening of final consonants, that is, they devoice for the same reason geminates do. Notice that in our modeling we have not lengthened final consonants, showing that such lengthening is not required, though of course it would have the enhancing effect Ohala describes.

6. Although these settings have some absolute limits in the physical world (e.g., how fast the tongue can move), it is interesting that these limits are not typically approached in speaking. For example, the changing volume of the oral cavity is relevant in any consideration of voicing maintenance for stop consonants, as we have seen, and obviously there is some finite limit on how large an individual's oral cavity can become. But this limit is probably never approached in speaking. Westbury (1983) shows that the set of possible maneuvers to expand the oral cavity makes so much expansion possible that from the point of view of speaking the oral cavity seems to have unlimited potential volume. When a speaker exploits these maneuvers is a separate question, of course.

REFERENCES

Anderson, S. (1974). *The organization of phonology*. New York: Academic Press.

Anderson, S. (1981). Why phonology isn't 'natural'. *Linguistic Inquiry, 12,* 493–589.

Chen, M. (1970). Vowel length variation as a function of the voicing of consonant environment. *Phonetica, 22,* 129–159.

Chomsky, N. (1964). Current issues in linguistic theory. In J. A. Fodor & J. J. Katz (Eds.), *The structure of language* (pp. 50–118). Englewood Cliffs, NJ: Prentice Hall.

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

Dinnsen, D. A. (1980). Phonological rules and phonetic explanation. *Journal of Linguistics, 16,* 171–191.

Dinnsen, D. A. (1982). *Abstract phonetic implementation rules and word-final devoicing in Catalan*. Talk presented at the annual meeting of the Linguistic Society of America, San Diego.

Flege, J. E. (1979). *Phonetic interference in second language acquisition*. Unpublished doctoral dissertation, Indiana University.

Fowler, C. A. (1980). Coarticulation and theories of extrinsic timing. *Journal of Phonetics, 8,* 113–133.

Fox, R. A., & Terbeek, D. (1977). Dental flaps, vowel duration and rule ordering in American English. *Journal of Phonetics, 5,* 27–34.

Fromkin, V. A. (1977). Some questions regarding universal phonetics and phonetic representations. In A. Juilland (Ed.), *Linguistic studies offered to Joseph Greenberg on the occasion of his sixtieth birthday* (pp. 365–380). Saratoga, CA: Anma Libri.

Fujimura, O., & Lovins, J. (1978). Syllables as concatenative phonetic units. In A. Bell & J. Hooper (Eds.), *Syllables and segments* (pp. 107–120). Amsterdam: North-Holland.

Hooper, J. B. (1976). *An introduction to natural generative phonology*. New York: Academic Press.

Houlihan, K. (1982). *Is intervocalic voicing a natural rule?* Talk presented at the annual meeting of the Linguistic Society of America, San Diego.

Hyman, M. (1975). *Phonology: Theory and analysis*. New York: Holt, Rinehart and Winston.

Keating, P. A. (1979). *A phonetic study of a voicing contrast in Polish*. Unpublished doctoral dissertation, Brown University.

Keating, P., Linker, W., & Huffman, M. (1983). Patterns in allophone distribution for voiced and voiceless stops. *Journal of Phonetics, 11,* 277–290.

Klatt, D. (1973). Interaction between two factors that influence vowel duration. *Journal of the Acoustical Society of America, 54,* 1102–1104.

Ladefoged, P. (1971). *Preliminaries to linguistic phonetics*. Chicago: University of Chicago Press.

Ladefoged, P. (1980). What are linguistic sounds made of? *Language, 56,* 485–502.

Lehiste, I. (1970). *Suprasegmentals*. Cambridge, MA: MIT Press.

Lindblom, B. (1967). Vowel duration and a model of lip mandible coordination. *STL-QPSR, 4,* 1–29.

Lindblom, B. (1983). Economy of speech gestures. In P. F. MacNeilage (Ed.), *The production of speech* (pp. 217–246). New York: Springer-Verlag.

Lisker, L. (1957). Closure duration and the intervocalic voiced-voiceless distinction in English. *Language, 33,* 42–49.

Lisker, L., & Abramson, A. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word, 20,* 384–422.

MacNeilage, P. (1970). Motor control of serial ordering of speech. *Psychological Review, 77,* 182–196.

Maddieson, I. (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.

Muller, E. M., & Brown, W. S., Jr. (1980). Variations in the Supraglottal Air Pressure Waveform and their Articulatory Interpretation. In N. Lass (Ed.), *Speech and language: Advances in basic research and practice:* Vol. 4 (pp. 317–389). New York: Academic Press.

Nelson, W. L. (1980). Performance bounds in speech motor control of jaw movements. *Journal of the Acoustical Society of America, 68,* Suppl. 1, S32 (A).

Ohala, J. J. (1983). The origin of sound patterns in vocal tract constraints. In P. F. MacNeilage (Ed.), *The production of speech* (pp. 189–216). New York: Springer-Verlag.

Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation*. Unpublished doctoral dissertation, Massachusetts Institute of Technology.

Port, R. (1977). *The influence of speaking tempo on the duration of stressed vowel and medial stop in English trochee words*. Doctoral dissertation, University of Connecticut, published by the Indiana University Linguistics Club.

Port, R., Mitleb, F., & O'Dell, M. (1981). *Neutralization of obstruent voicing in German is incomplete*. *Journal of the Acoustical Society of America, 70,* Suppl. 1, S13(A).

Postal, P. (1968). *Aspects of phonological theory*. New York: Harper and Row.

Rothenberg, M. (1968). *The breath-stream dynamics of simple-released-plosive production* (Bibliotheca Phonetica 6). Basel: Karger.

Sharf, D. (1962). Duration of post-stress inter-vocalic stops and preceding vowels. *Language and Speech, 5,* 26–30.

Stampe, D. (1973). *A dissertation on natural phonology*. Doctoral dissertation distributed by Indiana University Linguistics Club as Stampe (1979).

Westbury, J. R. (1979). *Aspects of the temporal control of voicing in consonant clusters in English. Texas Linguistic Forum, 14,* 1–304.

Westbury, J. R. (1983). Enlargement of the supraglottal cavity and its relation to stop consonant voicing. *Journal of the Acoustical Society of America, 74,* 1322–1336.

Westbury, J., & Keating, P. (1980). Central representation of vowel duration. *Journal of the Acoustical Society of America, 67,* Suppl. 1, S37 (A).

Westbury, J., & Keating, P. (1984). On the naturalness of stop consonant voicing. Unpublished manuscript.

9

Computation of Mapping from Muscular Contraction Patterns to Formant Patterns in Vowel Space*

Yuki Kakita
Osamu Fujimura
Kiyoshi Honda

1. INTRODUCTION

In previous papers (Fujimura & Kakita 1978; Kakita & Fujimura 1977), we demonstrated that a three-dimensional static model of the tongue can be used to explain basic characteristics of vowels. At that time we specifically discussed some characteristics of the vowel [i]. In the present chapter we present some preliminary results of our study of various vowels, in particular the five tense vowels of American English, /i/, /e/, /ɑ/, /o/, /u/. This study is in part based on the emg data reported by Alfonso *et al.* [1982]. The computational method, after Kiritani's original work, is an application of the finite-element method (Kiritani, Miyawaki, & Fujimura 1976). The anatomical data on which the design of our model is based are largely due to Miyawaki (1974).

2. A COMPUTATIONAL MODEL OF THE TONGUE

Figure 9.1 shows our tongue model in three different views. The anatomical structure in the front view (to the right) shows a frontal section in the plane indicated by the vertical line in the top view (upper left)

---

* We are grateful to our colleagues at Haskins Laboratories for their permission to use the emg and speech data that originally appeared in Alfonso, Honda, Baer, and Harris (1982) and in Alfonso, Baer, and Honda (1983), respectively.

PHONETIC LINGUISTICS