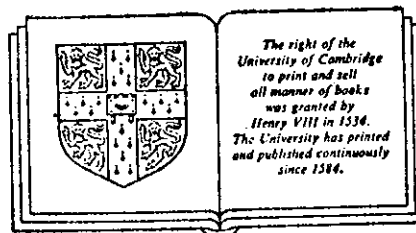


PHONOLOGICA 1984

Proceedings of the Fifth International Phonology Meeting
Eisenstadt, 25 — 28 June 1984

Edited by

Wolfgang U. Dressler, Hans C. Luschützky,
Oskar E. Pfeiffer and John R. Rennison



Cambridge University Press
London New York New Rochelle
Melbourne Sydney

1987

On Phonological Representations

David Stampe

In memory of C.E. Bazell

1. PHONETIC EXPLANATION IN PHONOLOGY

In his 1954 paper "On the choice of criteria in structural linguistics", Bazell pointed out a fundamental flaw in structuralist method, a method he described as accounting for the predictable, irrelevant, and redundant, by deriving it from the unpredictable, functional, and distinctive. Since most phonological analysis is still based on this method, Bazell's remarks are as relevant as ever.

In Japanese, [ɸ] occurs only before [u], while [h] occurs only elsewhere. We explain this by deriving [ɸ] from /h/ before /u/. Similarly, in English, [ŋ] occurs only non-initially, while [h] occurs only elsewhere. Yet we do not derive [ŋ] from /h/. The reason, Bazell says, is not that [h] and [ŋ] are not "phonetically similar" enough. Rather it is that while there is a motivation — a phonetic cause — for [h] to be pronounced as [ɸ] before [u], there is no motivation for [h] to be pronounced as [ŋ] finally. A phonological analysis, Bazell argues, must not only make the features and distributions of sounds mutually *predictable*: it must make them mutually *understandable*.

If phonetically motivated substitution explains why a phoneme is pronounced distinctly in different contexts, we still may ask why these distinct pronunciations are perceived as identical by speakers of the language. Sapir proposed an explanation of such perceptions:

In watching my Nootka interpreter write his language, I often had the curious feeling that he was transcribing an ideal flow of phonetic elements which he heard, inadequately from a purely objective standpoint, as the intention of the actual rumble of speech. (1921:56)

For example, this interpreter wrote Nootka [hɛ] and [hɔ] as *hi* and *hu* (Sapir 1933). Consonants like [h] "favor an a-timbre and cause the [intended] following vowels *i* and *u* to drop to [actual] *ɛ* and *ɔ* respectively" (54). If a Nootka speaker mechanically pronounces /hu/ as [hɔ], he would naturally hear [hɔ] in his speech and the speech of others as a pronunciation of /hu/. Or if the Japanese speaker, as in Bazell's example, pronounces /hu/ as [ɸu], he would hear [ɸu] in the speech of others as /hu/. Thus phonological perception can be understood in terms of phonological production, if phonological production is, in turn, understood in terms of substitutions with synchronic phonetic motivation.

To invoke phonetic motivation is to invoke universal limitations of the innate speech capacity. These processes were part of the traditional subject matter of phonetic science. The handbooks of Sievers (1901), or Grammont (1933), for example, were devoted not merely to the sounds of speech, but to their combinations and substitutions. The substance of such handbooks was the explanatory foundation not only of early historical phonology but also of the earliest synchronic phonology, e.g. Kruszewski (1881), or Baudouin de Courtenay (1895), who described the divergent realizations of phonemes as due to "our phonetic habits and the universal conditions determining the production of phonetic sequences" (159), "living anthropophonic causes operating in the synchronic state of a language" (161), which "arise and are maintained independently of tradition" (162).

2. EXPLANATION AND EVIDENCE

Phonetically motivated principles are not of the language itself, but of the *speaker*. Therefore they govern not only the speaker's production and perception of utterances in his language, but of *any* utterances — novel, foreign, invented, even garbled. Therefore the domain that such principles govern must be wider than the sound patterns a language learner (or a linguist) actually observes. And the domain of evidence is likewise wider.

For example, if we explain Japanese [ϕ] as an assimilation of intended /h/ to /u/, we are invoking a universal process which, in an individual who has not overcome it, should change /h/ before /u/ to [ϕ] not only in Japanese words but in foreign words as well. And indeed the Japanese speaker pronounces English [hʊk] *hook* as [ϕukku]. The same individual should also perceive [ϕ] before /u/ as /h/ not only in Japanese but also in foreign words. And he does, as his spelling /huurusukyappu/ for [fulzkæp] *foolscap* suggests.

An English speaker asked to say [t^hat] *tot* backwards says [t^hat], because [t^h] and [t] are initial and final allophones of English /t/. So if English [h] and [ŋ] were initial and final allophones of a single phoneme, a word like *hang* [hæŋ], said backwards, ought to stay [hæŋ]. It does not, for there is no sound whose phonetically motivated pronunciations are [h] initially and [ŋ] finally.

In fact English speakers find *hang* impossible to say backwards without substituting completely different phonemes, e.g. /n/ for initial /ŋ/ and null for final /h/. These strong constraints could not be guessed from their non-occurrence in the language. For [ay] never occurs before palatals in the language, and yet no English speaker finds it hard to say e.g. *shy* backwards as [ayʃ]. To understand the difference we have to look beyond English. Though some languages with /ŋ/ have it initially (e.g. Welsh) most do not, and though some with /h/ have it finally (e.g. Ojibwa), most do not. This reflects universal phonetic constraints. I have been told that speakers of French, with no /ŋ/ or /h/ at all, find final [ŋ] easier to pronounce than initial, and initial

[h] easier than final. It is also true that languages rarely have /y/ before palatals, and indeed [y] was deleted before palatals in the antecedent of English /ay/ (as in OE [di:ç] 'ditch', ME *[dryč], NE [dič] for expected *[dayč], cf. Stampe, 1972). But [ay] seems to be pronounceable before palatals if /ɪy/ or /ɛy/ are pronounceable, as in *leash*, *age*, whose glides arose after the ME deletion of [y] before palatals. As Sapir said of a similar example, "its articulation and perception have been mastered by implication" (1925 : 45).

The differences between saying *tot*, *hang*, *shy* backwards could not be discovered — either by the linguist, or by the learner — merely by observation of what forms occur in the language. They could not be explained by history or by habit. They show clearly that the phonology of a language resides in a system of universal phonetic principles.

3. STRUCTURAL AND GENERATIVE PHONOLOGY

But during the twentieth century, the rich fabric of explanation and evidence traditional phonology had woven of causality, intention, and consciousness was dismissed as a tissue of unscientific reasoning. The structuralists set out to reconstruct phonology in "objective" terms, based on the view that the significance of phonemes is solely in their difference from other phonemes.

Ce que importe dans le mot, ce n'est pas le son lui-même, mais les différences phoniques qui permettent de distinguer ce mot de tous les autres, car ce sont elles qui portent la signification. (Saussure, 1916:13)

The phoneme was re-defined as the sum of distinctive features of a speech sound, or as a class of sounds in variation or complementary distribution. Variants and complementary features, on this view, were irrelevant. Thus which variants occur where, and why, were equally irrelevant. And so phonology was separated from phonetics, and from any explanation of why there should be phonology at all.

If intention and perception were inadmissible in explanations, they were equally inadmissible in evidence. Evidence such as Sapir cited in his 1925 and 1933 papers was dismissed as superfluous (Twaddell, 1935). Explicitness replaced evidence as a criterion of adequacy. In becoming empiricist, phonology had become virtually non-empirical.

This is most evident in the matter of the archiphoneme. Twaddell (1935) was the first to note that, under the new definition of the phoneme, the [p] of *pin*, which contrasts with the [b] of *bin*, is not the same phoneme as the [p] of *spin*, since it does not contrast with a **sbin*. This result, which, as Swadesh (1935) pointed out, outlandishly proliferates the inventory of phonemes, might normally have been considered a *reductio ad absurdum* of the structuralist definition of the phoneme. Yet the archiphonemic conception persists

(cf. Firth, 1935; Trubetzkoy, 1939; Halle, 1954, 1959; Chomsky & Halle, 1968), though no empirical evidence was ever cited for it. Those who rejected it cited none either, but simply ignored Twaddell's argument. They continued to accept the definition of phonemes in terms of their differences, invoking ad hoc criteria of complementary distribution and phonetic similarity to achieve the desired analyses.

In 1964, Chomsky showed, however, that these criteria do not result in the desired analyses, and that the definition of the phoneme in terms of contrast, opposition, etc. is unsatisfactory. He did not attempt, as Bazell had, to determine *why* it is unsatisfactory, because he concluded, on largely formal grounds (sec. 5), that phonemic representation, as opposed to morphophonemic representation, is not "linguistically significant"

Chomsky's views have been widely accepted, because they pointed out real limitations of empiricist linguistics. But although he defended traditional mentalistic conceptions of language, Chomsky ignored their empirical content. He cited Sapir's psychological reality paper, not for its perceptual conception of phonemes, but simply to note that Sapir had cited both examples of phonetic and phonemic alternations in his paper, without distinguishing between them. He seems unaware of the traditional argument, which Sapir himself accepted, that these differ in perceptibility. For example, contrasting the variation of [t] with [d] before words beginning in a vowel in Upper Chinook and in Sanskrit, Sapir (1925:38) had said:

In Upper Chinook... the two forms of the final consonant are undoubtedly felt to be the "same" sound [because] the Upper Chinook *d* exists only as a mechanical variant of *t*; hence, this alternation is not the same psychologically as the Sanskrit sandhi variation *-t* : *-d*.

In Sanskrit, unlike Upper Chinook, [t] is a separate phoneme from [d].

Nor did Chomsky attend to Sapir's separation of such "mechanical" (phonological) processes from morphological processes, as in *Language* (1921: ch. 8), where Sapir traced the evolution of palatal umlaut in Germanic from a phonological to a morphological process. For Chomsky, virtually any alternation of sounds could be treated as phonological. Thus whereas traditionally such alternations as [k] and [s] in *opaque/opacity* or [d] and [s] in *decide/decisive* were treated as morphological, Chomsky proposed to treat them like the alternations of [t] with [r] in *delight/delighted* or [d] with [r] in *decide/decided*.

4. PHONOLOGICAL AND NON-PHONOLOGICAL ALTERNATIONS

English speakers intend /k/ when they say [c] before /i/, as in *key*. Do they, as Chomsky claimed, also intend /k/ when they say [s] in *opacity*, or /d/ when they say [s] in *decisive*? Let us look at some evidence from speech production, and then speech perception.

Chomsky and Halle (1968:342f) argued that the children's "secret language" Pig Latin is based on a simple cipher rule which, interacting with the ordinary phonology of the language, fully accounts for the bewildering details of utterances in this language, without ad hoc adjustments. There is another secret language, Ob, based on a cipher rule which its speakers describe as inserting /ab/ before every vowel, so that e.g. *Chomsky and Halle* is pronounced [čabamskabi abænd habælabi]. In Ob, *key* is [kabi], not [cabi], and this is explained by the theory that English speakers intend /ki/ when they say [ci]. But in Ob, *opacity* is not [abopabækabitabi] but [abopabæsabitabi] and *decisive* is not [dabisabaydabiv] but [dabisabaysabiv].

It is clear from their discussion of Pig Latin that Chomsky and Halle would say that the Ob rule must be ordered before the rule making /k/ into [c], but after the rules making /k/ and /d/ into [s]. Ob speakers, however, have a better explanation. If asked why they pronounce *opacity* as [abopabæsabitabi], not [abopabækabitabi], they say, because it has an /s/, of course, not a /k/. And similarly for *decisive*. Yet they pronounce *delighted* and *decided*, both ordinarily [—ayrid] in N. America, as [dabilabaytabid] versus [dabisabaydabid] because, they say, *delighted* has a /t/ but *decided* has a /d/. Such intuitions are not explained by ordering.

The traditional explanation (cf. Baudouin, 1895:159) would be that the alternations in *opaque/opacity* and *decide/decisive* are morphological, whereas the alternations in *delight/delighted* and *decide/decided* are phonological. Speakers find *opa[k]ity* and *deci[d]ive* easy to pronounce, even though neither occurs, but *deligh[t]ed* and *deci[d]ed* relatively difficult to pronounce, even though both actually occur as hyperarticulate variants of the usual flapped pronunciations. I see no reason to doubt the traditional explanation.

This is not to deny that alternations in *opacity* and *decisive* are fully productive, but only that they are based on phonological principles. Anything that blocks a phonological process exposes its input. But nothing — secret languages, tongue slips, not even talking backwards — exposes a /-k-/ in *opacity*, or a /-d-/ in *decisive*.

And the evidence of perception seconds this conclusion. Thus, *delighted* and *decided* rhyme, but rhymes with such a morphophonemic mismatch are normally not used. The same is true of *hands* and *plans*, which rhyme as /-ænz/, or *tests* and *mess*, which usually rhyme as /-es/, and so forth. But rhymes like *opacity* and *veracity*, *decisive* and *incisive*, etc., are never balked at, even though there is no /k/ in *veracity* or its alternants, and no /-d-/ in *incisive* (cf. *incise* with /z/). The simplest explanation, of course, is that there is no /-k-/ in *opacity* either, and no /-d-/ in *decisive*.

Or consider the deletion of /t/ after /s/ or /f/ before the verbal suffix *-en* in [krisn] *christen* (cf. [krayst] *Christ*), [čeysn] *chasten* ([čeyst] *chaste*), [sofn] *soften* ([soft] *soft*), etc. This quite regular rule would justify representations

like [krist-n], [čɛyst-n], [sɔft-n] in generative phonology. The Ob rule should block the effect of the suffix on the /t/, so we might expect [krabistabn] for *christen*. But what we get is [krabisabn] for *christen*, [čəbeystabn] for *chasten*, [səbɔfabn] for *soften*, with no trace of the /t/. Again, though the rule is quite productive in derived verbs, e.g. *swiften*, it lacks synchronic phonetic motivation, as is revealed by non-verbs like [pɪstŋ] *piston*, [ɔftŋ] *often* (in one dialect), [bɔstŋ] *Boston*, [aftŋ] *Afton*, etc. And, of course, [krɪsn] and [pɪstn], or [sɔftŋ] and [ɔftn], are not rhymes.

To treat this morphological alternation as if it were phonological only requires us to complicate our explanation of why it does not act phonological. If it were phonological, we might expect /t/ to delete where it secondarily comes before /n/, as in *liftin'* [lɪftŋ] or in *kissed an uncle* [kɪstŋ ʌ̃k]. But it does not. And we would expect /t/ to be deleted in *soften* only if the /n/ abuts it, and not if, as often happens, they are separated by [i]. Yet it deletes regardless. If we treated the deletion as morphological, then these facts would follow from the traditional view that all morphological rules precede all phonological processes (Stampe, 1973).

Contrast a genuine phonological process, the flapping of /t/ as in *delighted*. Flapping does not apply after a homorganic sound, and so it applies in *melted* only if the /l/ is vocalized, e.g. [mɛɹɪd], and in *painted* only if the /n/ is elided, e.g. [pɛ̃jɪd]. Flapping does not apply unless the /t/ is released, and it is never released before homorganic consonants, as in *witness*. Thus in *whiten* or *bitin'*, in northern U.S. dialects where the final syllable is a syllabic /n/, it is never released and therefore never flapped, e.g. [waɪtŋ], [baɪtŋ]. But in southern dialects where the final syllable is /ɪn/, it is released and flapped, e.g. [hwaɪvɪn], [baɪvɪn]. And in words like *mollen* or *paintin'*, flapping will occur only if the union of all these conditions is met, as in southern [moɹɪn], [pɛ̃jɪɪn]. It is completely consistent.

Traditionally, morphological alternations involve phonemes, not features, and, as Baudouin (1895) said, "The degree of phonetic similarity of the alternating phonemes is in such cases completely immaterial." They may involve "leaps which from an anthropophonic point of view are totally incomprehensible" (181). For example, although /d/ in *decide* alternates with /s/ in *decisive*, /t/ does not alternate with /s/ in *combat/combative*, *construct/constructive*, etc., though the [t] is closer to [s] than [d] is. Or, again, although /k/ in *opaque* alternates with /s/ in *opacity*, /g/ in *regal* alternates with /j/ in *regicide*. If this "velar softening" were phonological, one might expect /k/ to alternate with /ç/, or /g/ with /z/, in parallel, since /ç/ and /z/ are both admissible in English. To explain why /k/ in *lucky* or /g/ in *doggy* do not soften, Chomsky and Halle created an exceptional lexical class. They treated softening as conditioned by a following /i/, but the pronunciation of the vowel is immaterial: softening occurs before ^{fore} [i] (in *opacity*) and before [ay] (in *fungus/fungi*). To explain

this, they resorted to a synchronic rule mimicking the Great Vowel Shift. For each problem solved "phonologically", several worse ones spring forth.

This has obscured Chomsky's and Halle's real contributions to the theory of phonology proper: the revival of the process model, its application to phonetic as well as phonemic alternations, and its formulation in terms of phonetic features. Baudouin shied away from processes, for fear of obscuring the fact that he was speaking of synchrony (cf. Baudouin, 1895:154 - 161). Even Sapir rarely spoke of allophonic alternations in terms of processes. And Jakobson's features were intended to serve only phonemic taxonomies, not phonetic variation. Once the non-phonological is eliminated, phonological processes are an indispensable construct not only in the description of particular languages but in an understanding of the dynamic phonetic principles that underlie all phonology. But without such an understanding, the process model is just an empty formalism.

5. LAUT, PHONEMA, MORPHONEMA

Chomsky's process model, unlike Sapir's, assumed that processes are applied in linear order, as in Bloomfield's Menominee morphophonemics (1939). In Chomsky's conception, a "level of representation" must correspond to some "point in the operation of the phonological component" (1964 : 74f). The phonemic level does not conform to this model, as Halle (1959) discovered in Russian, where a process voicing obstruents before voiced obstruents governs phonemic alternations, e.g. /k/ with /g/, and also phonetic ones, e.g. /č/ with [j], which is not a phoneme in Russian. Incorporating a phonemic level in Chomsky's model would require the voicing process to be broken into two parts. Halle took this not as refuting Chomsky, but as refuting the phoneme. And not surprisingly, so did Chomsky.

But as Manaster-Ramer pointed out at this conference, Ulaszyn, in his 1931 paper "Laut, phonema, morphonema", cited a nearly identical example as evidence for the phoneme. In Polish, a single process of final consonant devoicing governs a phonemic alternation in the case of obstruents and an allophonic one in the case of sonorants. But Ulaszyn cited a substantive argument: the alternations (not the process) are psychologically distinct, because they are perceptible to native speakers in obstruents but not in sonorants.

In fact, Baudouin made the same argument in 1895, distinguishing between alternations of a "psychologically homogeneous phoneme", e.g. the voiced and voiceless alternants of /r/ in Polish *Piotra/Piotr*, and alternations in which there is a "split (bifurcation) of the psychological unity", e.g. the alternation of /b/ and /p/ in Polish *ba/leb*, or /t/ and /d/ in Russian *svata/švad'ba*. The former normally "cannot be perceived", the latter can (174f).

And Twaddell (1938), discussing Old High German palatal umlaut, observed

that scribes wrote the umlauted sounds distinctly only if these diverged phonemically from the basic sounds. They wrote umlauted short /a/ as *e* because it coincided with /e/. But they did not write umlauted /u/ as *ü*, because there was no /*ü*/ phoneme until later, when the phonetic conditions for umlaut were obliterated — at which time the scribes did write umlauted /u/ as *ü*. The scribes' orthographies at each period were phonemic, not morphophonemic.

These interpretations rest on the traditional view of the phoneme's role in perception. In the words of Sapir's student Swadesh:

The phonemes of a language are, in a sense, percepts to the native speaker of the given language, who ordinarily hear speech entirely in terms of the percepts. (1934 : 117)

A similar view could be inferred from structuralists' statements that a phonemic representation distinguishes only utterances which to the native speaker sound distinct (e.g. Hockett, 1955 : 144ff).

This view, and the evidence it was based on, was not addressed in Chomsky's critique of the phoneme. His conclusion that phonemic representation is not a significant level of representation is based entirely on his view that such a level must correspond to the input of one process in a linear sequence of processes. Obviously this view was not shared by Baudouin, Ułaszyn, or Twaddell, whose evidence shows that phonemic and phonetic alternations are distinguished perceptually although they are due to one and the same process.

In fact, in any model of phonology, if the phoneme inventory is given as Chomsky and Halle assume, then the phonemic representation of an utterance is uniquely determinable by a simple procedure:

The phonemic representation of any utterance is that which results from applying to its phonetic representation, *in reverse*, the minimal number of substitutions necessary to obtain a representation entirely in terms of phonemes.

We can demonstrate this procedure on examples more challenging than those cited above, namely two forms, *waiting* and *fainting*, whose phonemic representations not only do not arise at the same "point" in their phonological derivations, but in fact do not arise at *any* point. This will be the case whenever an allophonic substitution feeds a phonemic substitution (Stampe, 1973; Donegan & Stampe, 1979).

The flapping process changes the alveolar stops [t], [d], and [n] in *waiting* to [ɾ], [ɾ], and [ɾ], respectively. The voiceless flap (heard in *hothouse*) is usually voiced before voiced segments, merging *waiting* with *wading*. Then, in *fainting*, the vocalism is affected by regressive nasalization of tautosyllabic sonorants, while the nasal consonant is elided before a tautosyllabic homorganic consonant. This allows flapping, and then voicing, to

apply to the |t|. The resultant flap, a posttonic sonorant following a nasal (the nasalized [ẽỹ]), is progressively nasalized, merging *fainting* with *feigning*. The derivations, in linear ordering, are:

Basic Forms:	[weyt-ɪŋ]	[weyd-ɪŋ]	[feynt-ɪŋ]	[feyn-ɪŋ]
Nasalization (Regr.)	ĩ	ĩ	ẽỹ ĩ	ĩ
Nasal Elision			ø	
Flapping	f	r	f	ř
Voicing	r		r	
Nasalization (Progr.)			ř	
Phonetic Forms:	[weyřĩŋ]	[weyřĩŋ]	[fẽyřĩŋ]	[fẽyřĩŋ]

The representation entirely in phonemes which requires reversal of fewest substitutions is, for [weyřĩŋ] (whether *waiting* or *wading*), /weydĩŋ/. For [fẽyřĩŋ] (whether *fainting* or *feigning*), it is /feynĩŋ/. There is no difficulty determining these representations, even though there is no [d] or [n] in the pronunciations or even in the derivations of *waiting* or *fainting*.

And they *are* the phonemic representations. As my nine-year-old daughter put it, "*waiting* is pronounced like *wading*, with a *d*", and "*fainting* is pronounced like *feigning*, with an *n*". They are actually pronounced with [r] and [ř], of course. But they are *perceived as being pronounced* with [d] and [n], because [r] and [ř], in these words, correspond phonemically to /d/ and /n/.

We intend to pronounce *waiting* with |t| and *wading* with |d|, as is apparent in Ob: [wabeytabĩŋ] vs. [wabeydabĩŋ]. Likewise we intend |nt| in *fainting* and |n| in *feigning*: [fabẽỹntabĩŋ] vs. [fabẽỹnabĩŋ]. And we are aware of this lexical intention. To quote my daughter again, "*waiting* only rhymes with *fading* because you pronounce it like *wading*". Poets rarely use a rhyme like *waiting* and *fading*. It is a perfectly good rhyme, but a rhyme like *fading* and *wading*, which rhymes morphophonemically as well as phonemically, is preferable.

This brings us to a revealing asymmetry. We say that *waiting* is pronounced like *wading*, but we would not say that *wading* is pronounced like *waiting*. Yet both are pronounced alike with [r], which matches neither |t| nor |d|. This asymmetry can be understood by recognizing that the [r]'s are not perceived as such, but phonemically, as /d/'s. Since /d/ is a match for the |d| of *wading*, not the |t| of *waiting*, only the latter is perceived as having its intended pronunciation altered.

To sum up the evidence, perception is not in phonetic terms but, as Swadesh put it, entirely in terms of phonemes. This is why *wait* and *skate* rhyme perfectly, regardless of whether one is pronounced in [t] and the other in [t^h] or [g] or [ʒ]: all these are perceived as /t/. And apparently the perception of pronunciation, per se, and thus of homophony, punning, and rhyme, is phon-

emic. This is why *waiting* and *fading*, *pressed* and *pest*, *hands* and *plans*, etc., rhyme, if pronounced as phonemically identical, despite their morphophonemic differences. And it is why *waiting* and *skating* do not rhyme, despite their morphophonemic similarities, if one is pronounced with [c], phonemic /d/, and the other with [g] or [t], phonemic /t/.

But the perception of lexical identity is morphophonemic, because lexical representations are morphophonemic. This is not in conflict with Swadesh's principle that perception is entirely in terms of phonemes, because the phonemic representation of an utterance is not necessarily the only representation in terms of phonemes from which the pronunciation of the utterance can be derived. It is simply that representation which is separated from the pronunciation by the fewest substitutions. It is sometimes forgotten, because of the tendency of phonemic descriptions to be normalized to careful pronunciations (cf. Swadesh, 1934; Hockett, 1955, etc.), that a given word may have as many phonemic representations as it has phonemically distinct pronunciations. Unless all the variants and "automatic" alternants of a form have the same phonemic representation, the form in its lexical representation must be represented morphophonemically, or its variants and alternants will not all be derivable from that representation. Thus, although *water* is pronounced [wɑtɹ] i.e. /wadɹ/, its occasional variants [wɑtɹ] and [wɑtɹ], i.e. /watɹ/, could not be derived from [wadɹ], and so it is derived from [watɹ] instead, as is confirmed by its pronunciation in Ob, [wabatɹ]. Since /t/ as well as /d/ is a phoneme of English, this is still a representation entirely in terms of phonemes.

The main significance of the phoneme in lexicon and grammar is that lexical entries and morphological rules alike are "spelled" in terms of phonemes. What one does not perceive, one does not represent in memory, or invent a rule for.

One does not perceive in archiphonemic terms. Such representations, where the /p/ of *spin* is not identified phonologically with the /p/ of *pin*, are easy to refute. Thus speakers say that the second sound in *spin* is a /p/, they write it with a *p* (even children who do not yet read — Read, 1975), they take it to alliterate better with [p] as in *s'pose* than [b] as in *s'battical* (even though these are all pronounced with [sp]), etc. Asked to say it backwards, they say [nɪps] — not [nɪbz], and not sometimes one, sometimes the other.

1938
^ Although Twaddell endorsed such representations as phonemic, in his paper on Old High German orthography he avoided the term phoneme entirely. Recall that the scribes wrote *e* for unlauded /a/, because it had merged with /e/, but *u* for both [u] and its umlaut [ü], because these remained a single phoneme. But by Twaddell's 1935 definition, [u] and [ü] would not have been the same phoneme: whereas [u] contrasted with [a] and [e], [ü] only contrasted with [e] in unlauded contexts, because here *[a] had merged with [e].

Such conceptions of phonology rest on criteria of distinctiveness, contrast, minimal pairs, etc., which, as Chomsky (1964) demonstrated, presuppose an analysis and therefore, as analytic criteria, are circular. For example, [kæ̃t] *can't* is phonetically a minimal pair with [kæt] *cat*, but, as Chomsky noted, it is not analyzed as /kæ̃t/, but rather as /kænt/, and this is "justified" by saying that it is a minimal pair with /kænd/ *canned* instead!

But note that this is a fact not just about phonologists but about speakers. We perceive minimal pairs in phonemic, not in phonetic or morphophonemic terms. Thus [pæst] *past* vs. [pæs-d] *passed* is perceived not as a minimal pair, but as homophonous, because these are phonemically /pæst/. And two pronunciations of [pæst] *past*, one with and the other without its [t], are not perceived as homophonous, despite their lexical identity, because they are phonemically /pæst/ and /pæst̃/. It is true that judgements of same or different, or rhyme and nonrhyme, presuppose an analysis. But, for the native speaker, the analysis they presuppose is a phonemic one.

6. THE INNATENESS OF PHONOLOGICAL PERCEPTION

Let us pursue the paradox just noted to its conclusion. Unless children differ utterly from adults, they must also perceive utterances in phonemic rather than phonetic terms. And indeed, as Jakobson argued in his *Kindersprache* (1941), this is precisely what the evidence of child language shows us. There is no evidence at all that children perceive utterances and represent them in memory in their phonetic form. But since it is only on the basis of phonetic forms that children could learn the phonological system of a language, then, as Donegan (1984) has argued, it must not be learned.

But this is precisely what is implied by the traditional view that the phonological system of a language is based on phonetically motivated substitutions, and by the view that these substitutions, if not suppressed in learning the pronunciation of a language, persist. As we saw earlier, there are phonetically motivated constraints, like those against initial [ɟ] and final [h], which are not deducible from anything in the utterances of a language, and yet which, if the speaker attempts to pronounce the foreign sounds, manifest themselves in almost irresistible substitutions. Clearly, the phonological system of a language is only superficially determined by the language.

For a suggestion of how and why phonemic perception might occur in the language-innocent speaker, consider the example [kæ̃t] *can't* again. If the child attempted to say this as /kæ̃t/, it would become *[kæt] by context-free vowel denasalization. But if he attempted to say /kænt/, it would become [kæ̃t] by nasalization and nasal elision (sec. 4). Denasalization will not apply here, due to an immutable law that "strengthenings" apply only before "weakenings" (Donegan and Stampe, 1979). Since nothing in English ever

requires us to overcome any of these processes, /kænt/ remains the most natural analysis of [kæ̃t]. In particular, denasalization remains a constraint on vowels — not only on how we pronounce vowels, but how we perceive them. (For additional examples of this sort, and examples showing how phonemic and morphophonemic distinctions might be acquired without reference to phonetic forms except in an utterance just heard and thus still in ephemeral phonetic memory, see Donegan, 1984).

The suggestion that phonological — not morphological — perception might be so breathtakingly simple and deterministic is supported by the observation that, while children obviously confront considerable difficulty in overcoming phonological processes that they impose on the language they are learning, the processes that the language would impose on them seem to manifest themselves spontaneously. And however complex the alternations that these processes cause in the language, children see through these to the correct phonological representations unerringly. This suggests that the innate and universal limitations of the speech capacity are rich and specific, and that they guide the child's perceptions even as they restrict his productions.

REFERENCES

- BAUDOIN DE COURTENAY, JAN. 1895. Versuch einer Theorie phonetischer Alternationen. Strassburg-Cracow. [Cited from Edward Stankiewicz, tr., A Baudouin de Courtenay anthology, pp. 144–212. Bloomington: Indiana University Press.]
- BAZELL, CHARLES E. 1954. On the choice of criteria in structural linguistics. *Word* 10. 6–15.
- BLOOMFIELD, LEONARD. 1933. *Language*. New York: Holt.
- _____. 1938. *Menomini Morphophonemics*. *TCLP* 8. 110–115.
- CHOMSKY, NOAM. 1964. Current issues in linguistic theory. The Hague: Mouton.
- CHOMSKY, NOAM & MORRIS HALLE. 1968. *The sound pattern of English*. New York: Harper & Row.
- DONEGAN, PATRICIA JANE. 1984. The unlearnability of phonology. Wolfgang U. Dressler & Livia Tonelli, eds., *Papers on natural phonology from Eisenstadt*. Padova: CLESP. pp. 19–31.
- DONEGAN, PATRICIA JANE & DAVID STAMPE. 1979. The study of natural phonology. In Daniel A. Dinnsen, ed., *Current approaches to phonological theory*, 126–173. Bloomington.
- FIRTH, J. R. 1935. The use and distribution of certain English sounds: phonetics from a functional point of view. *English studies* 17. 8–18.
- GRAMMONT, M. 1933. *Traité de Phonétique*. Paris: Delagrave.
- HALLE, MORRIS. 1954. The strategy of phonemics. *Word* 10. 197–209.
- _____. 1959. *Sound pattern of Russian*. The Hague: Mouton.
- HOCKETT, C. F. 1955. *Manual of phonology*. (Indiana Univ. Publications in Anthropology and Linguistics, Memoir 11 of IJAL). Baltimore: Waverly Press.
- JAKOBSON, ROMAN. 1941. *Kindersprache, Aphasie, und allgemeine Lautgesetze*. Uppsala. Repr. in *Selected writings I*, pp. 328–401. The Hague: Mouton.

- JOOS, MARTIN, ed. 1957. *Readings in linguistics*. Washington, D.C.: ACLS.
- KRUSZEWSKI, MIKOLAJ. 1881. *Über die Lautabwechslung*. Kazan'.
- MANDELBAUM, DAVID G. 1949. *Selected writings of Edward Sapir*. Berkeley: Univ. of California.
- READ, CHARLES. 1975. *Children's categorization of speech sounds in English*. Urbana: National Council of Teachers of English.
- SAPIR, EDWARD. 1921. *Language*. New York: Harcourt Brace.
- . 1925. Sound patterns in language. *Language* 1. 37-51.
- . 1933. La réalité psychologique des phonèmes. *Journal de psychologie normale et pathologique* 30. 247-265. (English repr. in Mandelbaum 1949:46-60.)
- SAUSSURE, FERDINAND. 1916. *Cours de linguistique générale*. Paris: Payot.
- SIEVERS, EDUARD. 1901. *Grundzüge der Phonetik*. Leipzig: Breitkopf und Härtel.
- STAMPE, DAVID. 1972. On the natural history of diphthongs. *PCLS* 8. 578-590.
- . 1973. A dissertation on natural phonology. PhD dissertation, University of Chicago. (Repr. New York: Garland, 1979.)
- SWADESH, MORRIS. 1934. The phonemic principle. *Language* 10. 117-129. (Repr. in Joos 1957: 32-37.)
- . 1935. Review of Twaddell (1935). *Language* 11. 244-250.
- TRUBETZKOY, N. S. 1939. *Grundzüge der Phonologie*. (=TCLP 7.)
- TWADDELL, W. F. 1935. On defining the phoneme. (*Language Monograph* 16.) Baltimore: LSA. (Repr. in Joos 1957: 55-80.)
- . 1938. A note on Old High German orthography. *Monatshefte für deutschen Unterricht* 30. 177-181. (Repr. in Joos 1957: 85-87.)
- ULASZYN, HENRYK. 1931. Laut, phonema, morphonema. *TCLP* 4. 53-61.