Chapter 16: Mid-century American phonology: the post-Bloomfieldians

D. Robert Ladd

Revised April 2018
[Corrected March 2019]

Prepared for inclusion in the *Oxford History of Phonology*, edited by Elan Dresher and Harry van der Hulst.
Abstract

From the early 1940s to the early 1960s, American phonology, led by Bloch, Trager, and Hockett, was remarkably monolithic. Its practitioners nominally followed Bloomfield, but in practice rejected his clearly expressed views by treating the elements of phonetic transcription (‘phones’) as an intermediate level of abstraction between the phoneme and the physical signal. They also accepted a strong version of what Hockett called ‘duality of patterning’, according to which individual utterances are simultaneously arrangements of phonemes and arrangements of meaningful units; this led to their insistence that phonemic analysis had to be carried out without ‘mixing levels’, i.e. without any reference to grammatical and lexical categories. Among other things, this principle made it difficult to deal insightfully with morphophonological regularities. This was perhaps the major source of disagreement between the post-Bloomfieldians and the generative phonologists who followed; by contrast, reliance on phonetic transcription represents a major source of continuity between the two.
16.1 Introduction

For about 25 years in the middle of the 20th century – roughly 1940-1965 – American linguistics was dominated by a remarkably uniform theory of phonology, which I will refer to as ‘post-Bloomfieldian phonemics’. The theory took on its canonical shape in the 1940s among a group of scholars who identified themselves as followers of Leonard Bloomfield, and was given textbook presentation in the two major introductory texts of the period (Gleason 1955a/1961 and Hockett 1958). The link to Bloomfield’s own ideas about phonology was actually quite tenuous, but the sense of common endeavour among the group was very real.

The most influential members of this group during the period under consideration were undoubtedly Bernard Bloch (1907-1965), George Trager (1906-1992), and Charles Hockett (1916-2000). Other contributors to this phonological worldview included Martin Joos (1907-1978), Henry Allan Gleason Jr. (1917-2007), Henry Lee Smith, Jr. (1913-1972), Rulon S. Wells (1918-2008), and William G. Moulton (1914-2000). These men (the cast of characters is virtually all male) were mostly scattered in language departments at universities all over the eastern half of the United States, but they were united by their almost missionary enthusiasm for the new scientific linguistics propounded in Bloomfield’s Language (1933), and beginning in the mid-1930s they also had a regular opportunity to work together informally at the Linguistic Society of America’s annual summer institutes. Many of them also shared the experience of working on strategically relevant language teaching materials during the Second World War (see Hall 1991; Murray 1994: 144-151).

---

1 Joos seems to have played a variety of behind-the-scenes roles in mid-century American linguistics. He was influential in bringing wartime research on the sound spectrograph to the attention of linguists (Joos 1948) and bringing linguistics to the attention of engineers (Joos 1950); he also edited the influential collection Readings in Linguistics (Joos 1957). The influence of Gleason’s workbook of analysis problems (1955b), which initiated a generation of linguistics students into phonemic and morphophonemic analysis, should not be underestimated. Smith collaborated with Trager for several years and their Outline of English Structure (1951/1957) was the standard description of American English for at least a decade. Wells and Moulton are less central but specific contributions of theirs are discussed later in the chapter.

2 It is true that the institutes started in 1928, but after 1931 there was a hiatus until 1936. At the time of the 1928-1931 institutes, many of the scholars we are discussing were still in high school or college.
There were two other scholars working on phonological questions in North America at the time who require separate mention. Kenneth Pike (1912-2000) raised important theoretical objections to the mainstream Bloch-Trager-Hockett orthodoxy throughout the 1940s and early 1950s, but he was treated as an outsider, or at best as a devil's advocate (Hockett 1949); I return to discuss his contributions in section 16.3.4. Zellig Harris (1909-1992) was a regular participant in the theoretical discussions of the 1940s and 1950s and, as Chomsky's PhD supervisor, provides a link from the post-Bloomfieldian period to what followed. Several friendly critics have suggested to me that my account here unjustly slight Harris's work, but there are both practical and substantive reasons for my focus on Bloch, Trager and Hockett. The practical reason is that Harris's contributions are clearly sketched in Dresher and Hall's chapter on “Developments leading toward generative phonology” (this volume). The substantive reason is my view that, during most of the period under discussion, Harris was more respected than genuinely influential. Hockett's and Gleason's textbook presentations of phonology owe much to Bloch and Trager and little to Harris; Harris’s influence was primarily on subsequent developments, which is why he is more appropriately discussed in connection with the origins of generative phonology. However, I concede that my understanding of the intellectual history may be distorted by my own experience.3

The rapid growth of American linguistics immediately following the war involved developments across the whole breadth of the field. American linguists

---

3 Although it is neither necessary nor usual to provide an account of one’s qualifications for writing on a particular topic, I offer a brief autobiographical footnote to anchor my own phonological career in the era I discuss in this chapter. I studied linguistics as an undergraduate at Brown University from 1964 to 1968, where my teachers included Freeman Twaddell (1906-1982) and Nelson Francis (1910-2002) and where my textbooks included both Hockett (1958) and to a lesser extent Gleason (1955a/1961). Among other things, this undoubtedly makes me one of the youngest linguists still working (b. 1947) to have been taught the Trager-Smith analysis of English phonology (Trager & Smith 1951/1957) as current truth rather than historical curiosity. From 1970-1972 and 1975-1978 I was a PhD student at Cornell, and there I spent a lot of time in classes and in discussion with Hockett, though he was never officially involved in supervising my graduate work. Both Twaddell and Hockett were fairly free with anecdotes and opinions about their colleagues and about the field, and my understanding of what went on in the heyday of post-Bloomfieldian phonemics is unquestionably coloured by things I heard from them during my time as their student. Some of the flavour of the interactions among the mid-century post-Bloomfieldians can be gleaned from Joos's obituary of Bloch (Joos 1967), from Hockett's obituary of Trager (Hockett 1993), and from the extensive comments in Joos 1957; see also the summary in Makkai 1972: 3-6.
continued to engage in generalist fieldwork both as a continuation of the Boas/Sapir tradition of anthropological linguistics (on which see Silverstein, this volume) and through Christian missionary work, and a number of basic ideas in morphology and syntax, such as immediate constituents (Wells 1947a) and ‘morpheme alternants’ (Harris 1942a) or ‘allomorphs’ (Nida 1948), were first clearly aired at this time. Nevertheless, much of the theoretical and empirical focus of mid-century American linguistics lay in phonology, and there was a clear sense that genuine progress had recently been made in understanding sound systems. This sense is conveyed by Bloch’s article ‘A Set of Postulates for Phonemic Analysis’ (1948), which attempted to codify the consensus that had emerged in the previous 15 or 20 years. In his introduction Bloch emphasizes that he is not presenting anything new, but says that a statement of postulates is ‘a form of stocktaking: a pausing, at some crucial point in the development of a science, to look more closely into the substructure of its methods and to repair whatever logical flaws may appear in it’. In his conclusion, having finally presented the definitive version of his phonemic transcription of one short utterance, he says:

To ask whether this is a ‘correct’ or a ‘true’ transcription of the given utterance is meaningless. The transcription either is, or is not, an accurate record of the phonemes that we believe to occur in this utterance, and of their order. And the analysis on which the transcription is based either is, or is not, in accord with our set of postulates.

Any objections to such a transcription ... must therefore be stated and answered wholly in terms of these postulates. Whoever prefers a different transcription ... must show either that our analysis violates one or more of the postulates that we have stated, or else that these postulates are untenable. If he takes the latter position, we may reasonably ask him to state his own assumptions in equal detail, and put off all argument until he has done so.

His confidence in the importance of the new synthesis is clearly conveyed by the uncompromising tone of these final paragraphs.
In this chapter I will sketch the main tenets of post-Bloomfieldian phonemics, drawing in particular on the work of Bloch, Trager, and Hockett, on Pike's critiques of their work, and on Joos's commentaries on the articles anthologized in *Readings in Linguistics I* (Joos 1957), as well as on my own direct experience. I use the term 'phonemics' to draw attention to the fact that the theory covered a rather narrow subset of the phenomena that might reasonably be regarded as falling within the scope of 'phonology'. Among other things, I will attempt to explain what was behind these scholars' willingness to promote such a narrow perspective on sound patterns.

16.2 Phonemes, allophones, and the phone

It is rather generally taken for granted that generative phonology superseded post-Bloomfieldian phonemics and that what went before is now of primarily historical interest. However, there is a good deal of continuity despite the stormy theoretical upheavals of the 1960s. Even though their theoretical legitimacy has never really recovered from the generative critique of 'taxonomic phonemics' (especially Chomsky 1964, but also Halle 1959), the basic descriptive constructs of post-Bloomfieldian phonemic theory are still thoroughly familiar. This is because, ignoring for a moment the specific aspects of the mid-century orthodoxy that attracted Chomsky's and Halle's critical fire, some version of the phonemic principle remains at the heart of phonology (see Schane 1971 for an early generative statement of this assessment).

The core phonemic idea may be summarized as follows. The sound system of a language involves an inventory of abstract units of sound that stand in contrast with one another and that can be realized phonetically in different ways depending on the phonetic and/or structural context. In the terminology of the post-Bloomfieldians, the abstract units are *phonemes* and the different realizations are their *allophones*; their version of contextually conditioned variation was conceptualized as *complementary distribution*. Standard examples from English are still used in beginning linguistics courses throughout the Anglophone world to illustrate these fundamental concepts. They include the difference between clear and dark allophones of /l/ found in many varieties of
English (clear [l] in onset position in a syllable, dark [h] in coda position) and the difference between aspirated and unaspirated allophones of the voiceless stops (aspirated stops in absolute initial position, unaspirated stops following syllable-initial /s/). Some such notion of contextually conditioned variation in the phonetic realization of abstract phonological elements remains central to the phenomena that occupy phonologists’ attention.

A less obvious but equally important element of the post-Bloomfieldian synthesis, which also remains central to much phonological thinking, is another abstraction: the phonetic segment or phone. The theory took for granted the scientific validity of a segment-based idealised phonetic representation of speech. I have discussed the emergence of the phone concept and its role in 20th century phonology at greater length elsewhere (Ladd 2011), and I only briefly summarize that discussion here. With the exception of Pike’s treatise on phonetics (1943), no one involved in developing phonemic theory – on either side of the Atlantic – seems to have worried much about the basis of the phone idealization. Pike stated the problem succinctly:

Speech, as phoneticians well agree, consists of continuous streams of sound within breath groups; neither sounds nor words are separated consistently from one another by pauses, but have to be abstracted from the continuum. Phonemicists concur in the belief that some unit of speech, the phoneme, can be discovered as the basic constituent of a linguistic system. ... Is there a significant halfway point between the continuum and the phoneme? Is there a real, nonfictitious segment of sound which is not a phonemic one? (p. 42)

He devoted several pages of his monograph to developing ‘a workable method for the delineation of natural phonetic segmentation’ in which the ‘segmental unit is to be determined entirely apart from phonemic function’, and he understood that having such a method would mean that ‘an impressionistic phonetic record of a new language proves theoretically legitimate as well as practically valuable ... for the phonemicist ...’ (p. 53, emphasis added). However, for the most part the ‘phonemicists’ continued to use a string of segments –
phones, that is – as their representation of the primary speech data without worrying very much about its theoretical legitimacy. Bloch built the theoretical legitimacy of phonetic transcription into his Postulate 11:

The series of perceptible articulations of any given vocal organ during an utterance can be divided without a residue into successive parts .... In [this] postulate we do not [disregard the instrumental evidence that the articulators are in continuous movement]; rather, we imply that a phonetically trained observer can interpret the auditory fractions of an utterance in terms of articulations that seem (to his perception) to be static or unidirectional. (Bloch 1948: §11.1)

Since the emergence of the phone concept, one component of virtually all phonological theories has been a segment-based symbolic transcription of speech – what Chomsky (1964) called 'systematic phonetics'. That is, the centerpiece of many phonological theories is a mapping between a symbolic representation expressed in terms of abstract elements such as phonemes or features and another symbolic representation expressing the phonetic data. This characterisation applies very clearly to Prague School phonology, to the post-Bloomfieldians, to classical generative phonology, and to more recent developments from classical generative phonology such as lexical phonology and various versions of Optimality Theory. These differ in their conception of the mapping (derivational or declarative?), in the nature of the abstract elements (phonemes or features? ‘taxonomic’ or ‘systematic’ phonemes?), and in various other ways, but all assume that the primary phonetic data can be expressed in terms of a segmented symbolic representation.

Given this conception of phonology, the study of how systematic phonetic representations are manifested in quantitative acoustic and articulatory data is relegated to the phonetics laboratory and a different set of theoretical and methodological precepts. (In effect, the phone concept has relieved generations of phonologists of the need to come to grips with fine phonetic detail; see Pierrehumbert & Beckman 1988:1-5 for some discussion of this point). The most prominent current exception to this generalization is Articulatory Phonology.
(Browman & Goldstein 1986, 1989, and numerous colleagues and followers since then), which assumes that the surface phonetic data – the 'output', as it were – can best be expressed in terms of quantitative model parameters. But the central role played by the phone in post-Bloomfieldian theory is otherwise widely reflected more than half a century later.

It is worth pointing out that the idea of anchoring phonology to the phone seems not to have been shared by Bloomfield himself. In fact, Bloomfield drew attention to the manifold inconsistencies of close phonetic transcription and insisted that ‘only two kinds of linguistic records are scientifically relevant. One is a mechanical record of the gross acoustic features, such as is produced in the phonetics laboratory. The other is a record in terms of phonemes, ignoring all features that are not distinctive in the language …’ (Bloomfield 1933: 85; and cf. Hockett 1965, footnote 23). Among his followers, however, the phone occupied a central theoretical role. Speech – the primary data – was assumed to consist of a string of phones, and in some sense the goal of phonological analysis was to assign phones to phonemes. The term ‘complementary distribution’ is especially revealing, because it refers to the distribution of phones. Phones were real; phonemes were analytic constructs4.

The broad acceptance of systematic phonetics by both the post-Bloomfieldians and the early generativists is a major point of continuity. It seems to me that it outweighs an issue that became one of the major points of dispute between them, namely the question of whether phonemes (and allophones) were atoms or were themselves composed of smaller components that are now universally known as distinctive features. There are at least two reasons for seeing the latter issue as secondary. First, the post-Bloomfieldians were certainly

---

4 This is not to suggest that the phonemicists were unaware that the relation between phones and the speech signal might be rather complex. Joos, as noted in footnote 2, wrote a monograph (1948) aimed at informing linguists of the new developments almost as soon as wartime secrecy surrounding the sound spectrograph was lifted. Hockett, in his Manual of Phonology (1955: 180-211), attempted to come grips with the fact that, despite the 'natural a priori expectation ... [that] each allophone will appear on a spectrogram as a recognizable pattern ... the results of actual experimentation are staggering at odds with any such expectations.' Nevertheless, they did not abandon the phone idealization, only sought to refine it or justify it. Bloch in particular (postulate 9) insisted that 'an observer can be trained to make a phonetic description of the utterances of any dialect, or of a sufficient sample thereof, without the aid of laboratory devices, that will be adequate and valid for the purposes of phonological analysis' (1948.§9.1).
aware of work by Trubetzkoy, Jakobson, and others, and considered what place features might have in their developing understanding of the phoneme; Bloch devotes an entire section of his 1948 article to six postulates (§§ 47-52) on ‘features’. Second, despite the ostensible grounding of the distinctive features in the physical details of speech, the actual practices with respect to phonetics differed little between most generative phonologists and most of their predecessors. Instead, the more fundamental rift between post-Bloomfieldian and early generative views involves the degree of abstraction involved in defining the phoneme, and the consequent relation between phonemes (whether post-Bloomfieldian ‘taxonomic’ phonemes or generative ‘systematic’ phonemes) and the phones that make up the surface phonetic representation. This is the thread running through the topics discussed in the next section.

16.3 The heart of the matter

16.3.1 Biuniqueness

Within post-Bloomfieldian phonemics, the theoretical importance of the phone was enhanced or exaggerated by the principle of biuniqueness. (This term, which is also used in mathematics, seems to have been first used in connection with phonology by Harris, e.g. 1942b, 1944.) According to this principle, every phoneme of a language can be manifested by a variety of phones, following regular distributional (‘allophonic’) rules; every phone in an utterance can be identified as the phonetic manifestation of a specific phoneme of the language; and crucially, in a given context no phone can manifest more than one phoneme. Together, these principles meant that there is a ‘biunique’ or one-to-one correspondence between any string of phones (i.e. any ‘phonetic representation’, in present-day terminology) and any string of phonemes. For example, given the predictable allophonic distribution of aspirated and unaspirated manifestations of voiceless stops in English, the phonemic string /kit/ necessarily begins with the phonetic segment [kʰ] and the string /skt/ necessarily includes the phonetic segment [k]. In the other direction, the phonetic segment [k] following the phoneme /s/ in /skt/ can only correspond to the phoneme /k/.
The first clear statement of this view came in Bloch’s short paper ‘Phonemic overlapping’ (1941). He began by discussing several cases in which the realizations of phonemes might appear to intersect, but which can be analysed in other ways, or where the realizations are distinguished by being in complementary distribution. (His example of complementary distribution contrasts the alveolar flap as a manifestation of the phoneme /r/ after /θ/ as in *three* and as a manifestation of /t/ between vowels as in *butter.* He then develops his ‘most seductive example of apparent intersection’, involving ostensibly allophonic variation in vowel length in some varieties of American English. It is necessary to quote him at some length to convey the essence of the argument.

The pairs of words *bit bid, bet bed, bat bad, but bud, bite bide, beat bead*, etc. have respectively the same vowel phoneme, but exhibit a regular and fairly constant difference in the length of the vowel allophones. ... The alternation between longer and shorter allophones runs through the whole phonemic system. The vowel of *pot* is affected by the same automatic alternation: ... there is nothing, so far, to show that a pair like *pot pod* is not in every way comparable to *bit bid.*

In my speech *bomb* is different from *balm, bother* does not rime with *father,* and *sorry* does not rime with *starry:* the vowel quality is the same in all these words, but in the first word of each pair the vowel is short (just as it is in *pot*), and in the second noticeably longer. Since the difference in length cannot be explained as an automatic alternation (like the difference in *bit bid*), we conclude that *bomb* and *balm, bother* and *father, sorry* and *starry* have different vowel phonemes; and we naturally identify the vowel of *bomb, bother, sorry* with the phoneme of *pot.* The vowel of *balm, father, starry* appears also in *alms, palm, pa, star, card.* Again, there is nothing, so far, to show that the phonemic organization is in any way abnormal. But now comes a hitch.

In the sentence *Pa’d go (if he could),* the utterance fraction *pa’d* must be analyzed, according to what we have just said, as containing the phoneme of
balm. In the sentence *The pod grows*, the utterance fraction *pod* must be analyzed, again according to what we have said, as containing the phoneme of *pot*. But *pod ...* is phonetically identical with *pa’dl*. Two occurrences of x under the same conditions have been assigned to different phonemes.

Bloch then goes on to conclude that the only solution to the problem is to conclude that the vowels of *pot* and *pod* are *not* allophonic variants of the same phoneme, but must be phonemically different. By doing so, he acknowledges, ‘we destroy the neat parallelism of the pairs *bit bid, bet, bed, bite bide, pot pod ...* [b]ut by sacrificing this symmetry we are able to account for all the facts of pronunciation, which is surely the more important requirement.’

In practice, at least one clear class of partial exceptions to biuniqueness was acknowledged by the post-Bloomfieldians, namely ‘free variation’. The usual example in American English was the variation between released and unreleased final stops. It was acknowledged that both [bitʰ] and [bit] reflect the same string of phonemes (/biyt/ in the Trager-Smith analysis) and that the variation between the two types of utterance-final stop was of no phonological significance. Bloch discussed free variation in his Postulates (§§27-29).

Other phenomena that might have led to a weakening of the precept of biuniqueness, however, did not do so. The most obvious example is neutralization. If we assume that American English ‘flapping’ neutralizes the /t−d/ distinction in pairs like *betting* and *bedding* or *latter* and *ladder*, we cannot, given the principle of biuniqueness, treat this as the consequence of allophonic rules whereby both /t/ and /d/ are realized intervocally as a flap (which I will represent here as [ɾ]). The theory would require the phone [ɾ] to be assigned unambiguously as the phonetic manifestation of one phoneme only. Such an analysis was rigorously applied in the pronunciation transcriptions given in the Third International edition of the unabridged Merriam-Webster dictionary (1961), where words normally pronounced with a flap are transcribed with /d/; this choice is justified at some length in the dictionary’s ‘Guide to pronunciation’ (p. 41a). This analysis means that morphemes like *bet* have two phonemically
distinct allomorphs, /bet/ and /bed/, and the alternation between the two would have to be treated as a matter of morphophonemics or morphology\(^5\).

16.3.2 Separation of levels

The architects of the theory did not regard these consequences as flaws. On the contrary: to put it in modern terms, they saw the theoretical consequences of biuniqueness not as a bug but as a feature. Counterintuitive conclusions like Bloch’s analysis of *pot* and *pod*, which might reasonably be regarded as a problem or even a *reductio ad absurdum*, were sometimes held up as evidence of the insight afforded by rigorous adherence to the logic of the theory. In the specific case of phonemic overlapping, the insight was the principle that generally went by the name of ‘separation of levels’. As Joos put it in assessing the impact of Bloch’s paper (Joos 1957: 96), Bloch ‘made clear, as it never had been before, that phonemics must be kept unmixed from all that lies on the opposite side of it from phonetics’. According to this principle, the sound system of a language was to be analysed entirely in its own terms, without any reference to the grammar or the lexicon. The two levels were to be kept rigidly separate.

Failing to separate them was committing the sin of ‘mixing levels’, a temptation that proved difficult to stamp out. To quote further from Joos's commentary, ‘the ghost of the slain dragon continued to plague the community of linguists under such names as “grammatical prerequisites to phonemic analysis” [a direct swipe at Pike; see section 3.4 below] and has not been completely exorcized to this

---

\(^5\) My discussion of this point is hedged with conditionals because, remarkably, this issue was seldom discussed in post-Bloomfieldian work even in analyses of the phonology of American English. Neutralization (and the construct of the archiphoneme) loomed large in theoretical discussions in Eastern Europe, where final devoicing in many languages means that neutralization is widespread, but the lone example of flapping in American English was simply ignored or assumed not to involve neutralization at all. Bloch’s discussion of phonemic overlapping (1941) clearly implies that the ‘alveolar flap’ in *betting* or *kitty* contrasts with [d] in *bedding* and *kiddy*. Similarly, Trager & Smith’s analysis (1957) claims on p. 32 that ‘In *butter* we find a voiced fortis [t] in most American speech’, and on p. 34 mention only [d] as a medial allophone of /d/. Dogged empirical investigation by e.g. Haugen (1938) and Oswald (1943), showing clearly that intervocalic /t/ and /d/ were indistinguishable by listeners, seems to have had little impact on post-Bloomfieldian theoretical thinking. By the time of Gleason’s textbook, however, we find a clear statement that ‘in some dialects the contrast between /t/ and /d/ has been lost between vowels’. Gleason lays out the possible analyses, noting that ‘American linguists generally have preferred to match [the flap with either /t/ or /d/ but not both] wherever possible; some Europeans, to maintain units like /T/ in which the voiced : voiceless contrast is said to be neutralized.’ (Gleason 1961: 295). The solution in Webster’s Third is certainly of Gleason’s first type.
Seen from the point of view of an individual hearing a sentence, the idea of separating phonology and grammar makes a certain amount of sense. The phonemicists’ assumption was that hearers must first identify the phonemes they are hearing and then parse them into words. This was clearly stated by Wells:

Phonemics takes the point of view of the hearer. Now the hearer, in order to interpret correctly an utterance that he hears, must rely on two separate sources of information: (a) the heard sounds ...; (b) the extra-linguistic context .... [P]honemics makes a point of recording nothing but what is conveyed by (a). All else belongs to grammar (and lexicography.) (Wells 1947b: 271.)

If this is indeed the hearer’s task, then there is a conundrum unless something like biuniqueness is at work. Consider the input string [bɛɾɪŋ]. Confronted with [ɾ] in a phonology where it could be an allophone of either /t/ or /d/, the hearer has no basis for deciding which phoneme is intended without knowing whether the intended message is *betting* or *bedding*. But – given the assumption that sentence understanding proceeds strictly from the bottom up – the only way to know the intended message is to have identified the phonemes first. The circularity is apparent.

The importance of the hearer’s point of view was a theme of Hockett’s thinking⁷. For example, part of his response to the success of Chomsky’s early

---

⁶ The moralistic tone of Joos’s comments here can be found repeatedly in mid-century discussions of phonemic theory. Later in the same commentary, for example, Joos goes on to preach the virtues of post-Bloomfieldian descriptivism: ‘Trubetzkoy phonology tried to explain everything from articulatory acoustics and a minimum set of phonological laws ... , and offer[ed] too much of a phonological explanation where a sober taxonomy would serve as well. Children want explanations, and there is a child in each of us; descriptivism makes a virtue of not pampering that child.’ Trager’s reaction (1950: 158) to Pike’s monograph *Phonemics: A technique for reducing languages to writing* (Pike 1947a) echoes similar themes: he ‘condemn[es] the book as a theoretical work, and even more as a text-book – since as the latter it will lead astray many who might otherwise be valuable workers in linguistic science.’ Note also Bloch’s own characterization of his *pod pa’d* example as ‘seductive’.

⁷ While I was at Cornell, Hockett suggested more than once that an important difference in outlook between generative and post-Bloomfieldian linguistics was that the former took the point of view of the speaker and the latter the point of view of the hearer. However, except for
work was an article entitled 'Grammar for the Hearer' (Hockett 1961a), in which he argued, in strikingly contemporary-sounding terms, that for the hearer, parsing can indeed be seen as a stochastic process and modelled by a finite state grammar.

‘Chomsky has shown that, if we accept certain very reasonable empirical assumptions about English, then English is not a finite state language. He has also claimed that no finite state approximation to English can match the known facts of the language closely enough to be of any interest. This second point is, I believe, false. It will be shown later in this paper that it is in theory possible to match the facts of English as closely as we wish with a finite Markov chain.’ (1961a: 220)

Yet from the outset of the article he effectively exempted the sound system from this outlook, explicitly adopting the following assumption: ‘The hearer always hears correctly: that is, he hears two words or sequences of words as different if and only if they are phonemically distinct.’ (1961a: 221). He described this assumption as ‘customary’, though he also noted that it is ‘rarely stated explicitly’. This assumption encapsulates the view expressed by Wells as quoted just above: hearers must first identify the phonemes they are hearing and then parse them into words. Given this assumption, the input to the stochastic process is not an acoustic signal but a phonemic representation.

With hindsight based on decades of psycholinguistic research (comprehensively summarized by Cutler 2012), it is clear that this assumption cannot be maintained. Speech processing – all speech processing, not just grammatical parsing – is in many respects a stochastic process. We know now that hearers continuously generate hypotheses about the incoming speech stream, and try to identify words as they come in; they do not – or not necessarily – try to identify phonemes first. Indeed, as anyone who has ever listened to the stimuli in a ‘gating’ experiment knows, it is often difficult or impossible to identify phonemes until after one has identified the words.

the 1961a paper, he did not to my knowledge develop this idea until many years later (Hockett 1987).
Furthermore, in almost any context, it will be clear whether [bɛɾɪŋ] is intended to represent *betting or bedding*; the phonemic identity of the [ɾ] (to the extent that this even matters to the hearer) follows from that. But the assumption that hearers somehow hear phonemes directly left the post-Bloomfieldians unable to acknowledge the interdependence of the phonological and the grammatical levels of linguistic structure.\(^8\)

16.3.3 Morphophonemics

Nowhere is the interdependence of phonology and grammar more obvious than in the case of phonologically conditioned alternations in the forms of words or morphemes. The principle of biuniqueness made it impossible for post-Bloomfieldian phonemics to treat most automatic alternations as part of a unified account of phonology. Neutralization, as we saw earlier, could not be analysed in terms of two different phonemes converging on the same allophonic realization; it was theoretically necessary to treat the neutralized allophone as belonging only to one phoneme or the other. In the same way, phonologically exceptionless alternations like the [-s]/[-z]/[-iz] variation in the form of the English plural and 3rd person singular morphemes can only be treated as part of morphology or a specific morphophonemic component of the grammar, not as part of the phonology. Once the theoretical definition of the phoneme was established, the descriptive consequences for neutralization and other automatic alternations did not lead to any reconsideration.

This is not *a priori* unreasonable. For example, as discussed by Iosad (this volume), the Moscow and Leningrad schools of Soviet linguistics differed precisely on the matter of what to do about neutralization; Leningrad privileged the surface sound of a neutralized segment, and posited a single phonemic representation, while Moscow focused on lexical identity, and accepted that distinct phonemes could be realized in the same way. I would state the underlying question here as follows: is the sound system of a language primarily

\(^8\) A residual sense that phonological and grammatical description should be kept separate may be responsible for the fact that, several decades later, Kaisse's study of connected speech (1985) was criticized for making direct reference to syntactic structure, and that most subsequent work on these problems (e.g. Selkirk 1984, Nespor & Vogel 1986, and much work since then) has posited a similar but separate 'prosodic structure' *in the phonology*. 
a set of phonetic generalizations over the language’s lexicon, or is it a template that plays a role in guiding listeners’ percepts of speech sounds independently of the lexicon? There is plenty of evidence for both positions; the post-Bloomfieldians came down firmly on the side of phonetic identity rather than lexical identity.

However, there was plenty of discussion before general agreement on this issue was reached. Bloomfield himself, notably in his 1939 paper on Menominee, distinguishes clearly between alternations that are phonologically conditioned and those that are lexically idiosyncratic, and seems to suggest that the phonologically conditioned ones are part of phonetics; he explicitly states – only two years before the publication of Bloch’s paper banning phonemic overlapping – that there is ‘some overlapping between phonemes’ (1939: §38). Hockett (1993: 787) says that ‘early on [Trager] believed that morphophonemics belongs, with phonemics, in phonology; partly under my influence, he later changed that opinion.’ Joos (1957: 92) describes the debates about this issue as follows:

When we look back at Bloomfield’s work, we are disturbed at this and that, but more than anything else Bloomfield’s confusion between phonemes and morphophonemes disturbs us. Bloomfield kept himself out of trouble here, usually, by describing just one language at a time, or one area within each at a time, adjusting for the effects of the confusion. ...

The escape from this confusion was, naturally, itself a confused as well as an arduous journey, like that of the Israelites from Egypt. Most of it remains undocumented, consisting of endless hours of discussion and of private pondering. ...

He then gives his own definition of the distinction, which again makes reference to assumptions about what hearers do:

The native listener may be said to perceive – to somehow exploit for message-understanding ends – items in what he hears. Insofar as this process does not depend on understanding, the items are phonemic; insofar as items cannot be perceived without understanding, morphophonemics at least (perhaps more) is involved. I borrow an example from C. F. Hockett:
Once he heard someone say ‘She has poise’ and, momentarily insufficiently attentive, innocently said ‘What’s a poy?’ The phonemic items had been apprehended perfectly, but, through a lapse in understanding, the morphophonemic items had not.

Nevertheless, although the issue was temporarily settled, the place of morphophonemics in post-Bloomfieldian work remained awkward. Gleason’s textbook mentions the term *morphophonemics* (or *morphophonemic rules*) at several scattered places in its coverage of morphology, but it never mentions the issue of why the line between phonology and morphology is drawn where it is, simply taking it for granted that the morpheme alternants under discussion are phonemically different. Hockett’s textbook provides slightly more unified coverage of the actual phenomena in a group of chapters entitled ‘Morphophonemic systems’, and furthermore makes a serious attempt to specify the place of morphophonemics in the overall design of language (1958: 137-144); cf. also the more technical discussion in Hockett (1961b). But for both writers the definition of the phoneme involving biuniqueness was taken as settled, and the treatment of morphophonemics followed as a consequence.

This was the basis of Halle’s famous critique of the post-Bloomfieldian phoneme in *The Sound Pattern of Russian* (1959). Under the phonemicists’ definitions, phonologically conditioned automatic voicing alternations in Russian may be treated in a few cases as involving the complementary distribution of allophones, but must be treated in most cases as morphophonemic effects on the phonemic representation of morphemes. Most Russian obstruents are part of phonemically distinct voiced/voiceless pairs (/p b, t d/ etc.), and in these cases automatic voicing alternation will change the phonemic representation. Three obstruent phonemes, however, namely /ts, tʃ, x/, have no contrastive voiced counterpart, so here the occurrence in complementary distribution of corresponding phonetically voiced and voiceless obstruents can be analysed as involving voiced and voiceless allophones of a single phoneme. Something that is transparently the same fairly low-level process of phonetic assimilation has to be treated in two different ways. The case is almost perfectly analogous to

---

9 For a more detailed examination of Halle’s critique see Dresher and Hall (this volume), sec. 3.2.
Bloch’s *pod Pa’d* example, except that Halle and Bloch drew opposite conclusions. Halle privileged the unity of the phonetic assimilation process, and argued for ignoring the presence or absence of surface contrasts of voicing in phonemic analysis. Bloch privileged surface contrasts and argued for ‘sacrificing the symmetry’ of allophonic duration effects in a phonemic analysis, insisting that the difference between *pot* and *pod* must involve the same phonemic contrast in duration found in *bomb* and *balm*.

What is noteworthy about the difference between the post-Bloomfieldian and the generative approaches is that both approach morphophonemic analysis in very similar ways – based, as it happens, on general principles that were already clearly spelled out by Bloomfield himself (1933, chapter 13). Both Gleason and Hockett give summaries of how to describe morphophonemic alternations in terms of ‘base forms’ (i.e. underlying forms) and sketch the principles involved in selecting an appropriate base form; Hockett (pp. 277-278) discusses the difference between internal and external sandhi and the sense in which one process ‘precedes’ the other; Gleason’s workbook (1955b) includes a long section of data problems that give students practice in identifying base forms and using ordered rules. The main difference between post-Bloomfieldian morphophonemics and early generative phonology concerned the theoretical status of the elements manipulated in these analyses, not actual descriptive practices. For some discussion of the continuity involved here see Kilbury 1976.

16.3.4 ‘Grammatical prerequisites’

One consequence of the radical separation of levels espoused by the post-Bloomfieldians is that there are no words in the phonology. Phonemic contrasts serve to distinguish one *utterance* (or ‘utterance fraction’) from another; words are part of grammar. Bloch built this assumption into his Postulates (especially §2) and this is why, for example, he thought it relevant and important to compare *pod* with *Pa’d* in his essay on phonemic overlapping. So what do we do when we find direct phonetic evidence for grammatical boundaries?

The most obvious cases of this sort involve allophonic variation
apparently conditioned by position in morphological structure. The standard example at the time was the difference in pronunciation between nitrate and night rate, where there are two different allophones of the phoneme /t/. Since it was theoretically out of bounds to define the conditioning factors for this allophonic difference in terms of grammatical facts like word boundaries, the solution was to posit the existence of a ‘juncture’ phoneme between the /t/ and the /r/ in night rate. That is, a phenomenon that might reasonably be analysed as the phonetic effect of a grammatical boundary was recast as ordinary allophonic variation conditioned by the presence or absence of a phonological entity, a ‘phoneme’ usually symbolised by /+/ and referred to as ‘plus-juncture’. This idea was first tentatively proposed by Trager & Bloch (1941 §4) and was eventually widely adopted, though reluctantly by some (e.g. by Wells, who noted that ‘the validity of juncture phonemes is open to grave doubts on phonetic grounds’ 1947a: §64).

For example, plus-juncture played a crucial role in the standard post-Bloomfieldian analysis (Moulton 1947) of the German diminutive suffix –chen ([çən]). In general, [x] (or [χ]) and [ç] are in complementary distribution, the former occurring after back vowels and the latter elsewhere, including (rarely) word-initially. The diminutive suffix posed a theoretical problem by using the palatal allophone [ç] even after back vowels, in a word like Frauchen ‘mistress’ [in the sense of ‘female dog owner’]. Plainly, the existence of the morphological boundary between stem and suffix is relevant in some way, but this could not be acknowledged in direct morphological terms; instead, a plus-juncture phoneme was said to occur between the /au/ diphthong and the dorsal fricative phoneme, conditioning the palatal allophone [ç]. Moulton’s analysis drew immediate critical comment (Leopold 1948), and the two papers are reprinted together in Joos (1957), but Joos’s own editorial comments (1957:216) make clear that he regards the junctural analysis as correct.

Analyses of this sort – and with them, the whole doctrine of the separation of levels – were called into question by Pike in two articles on what he called ‘grammatical prerequisites’ to phonological analysis (Pike 1947b and 1952). Pike’s argument was that ‘when phonological and grammatical facts are
mutually dependent, the treatment of phonology without reference to grammar is a concealment of part of a most important set of structural facts pertinent to phonology’ (1947b: §0).10 He emphasized that in practice the classical minimal-pair test presupposed lexical or grammatical units; he pointed out that by focusing on utterance-initial contrasts, as fieldworkers often do, ‘the analyst can be certain that he is at the beginning of a phoneme, of a syllable, of a stress or rhythm group, of an intonation contour, of a phonological sequence of some type’ and that ‘sounds at the beginning of utterances ... are simultaneously at the beginning of a word, and at the beginning of a construction.’ (§3) Specifically with regard to juncture phonemes, he posed several questions that never received satisfactory answers: ‘If a juncture is a phoneme, can one describe its variant forms or indications as allophones? And how will one treat allophones of a juncture phoneme if they have nothing physically in common with each other...?’ (§4) He cited a number of papers in which analysts posited juncture phonemes, and he drew attention to contradictions between their supposed theoretical basis and the way they were actually deployed in the analysis.

Many of Pike’s specific criticisms of juncture phonemes were raised by others, and the reliance on meaning as a ‘shortcut’ to determine phonemic contrast was generally recognized as something of a problem. But his explicit rejection of the separation of levels made him a heretic rather than just a critic. Bloch, as editor of Language, rejected the first grammatical prerequisites paper (see Ladd 2015: 133). Pike’s work was completely excluded from the selection of papers in Readings in Linguistics (Joos 1957) and, as we saw above, Joos lamented the fact that the notion of grammatical prerequisites continued to ‘plague the community of linguists’11. In fact, while Pike called for more explicit recognition of the role of grammatical factors in phonemic analysis, there were other voices calling for an even more radical limitation of phonemic analysis to what was conveyed by the sounds alone.

10 Essentially this view lay at the foundation of Chomsky, Halle and Lukoff’s paper (1956) on the phonology of English stress. For fuller discussion see Dresher and Hall (this volume).
11 It seems possible that Pike was marginalized in part because of his commitment to Christian missionary work (Murray 1994:174, 189f), but comments I heard as a student suggest that he was also regarded as rather unsophisticated; among other things, his prose was inelegant and prone to occasional malapropisms.
Specifically, some scholars drew a distinction between ‘word phonemics’ and ‘utterance phonemics’, implying that by focusing on contrasts between words phonologists were not fulfilling the promise of Bloch’s postulates. There were never clearly laid-out competing theories of word phonemics and utterance phonemics, and published references to the distinction are rare\textsuperscript{12}; for the most part, everyone theoretically believed in utterance phonemics but in practice worked mostly with contrasts between words. This theoretical disconnect was what led Pike to argue for accepting the theoretical legitimacy of word-level phonology, but it also encouraged a fundamentalist contrary view that saw word phonemics as, at best, a preliminary step on the way to true understanding, and at worst a betrayal of basic principles. Trager, for example, grumbled that ‘many linguists have remained content to do word phonemics’ (1962: 13).

16.3.5 Duality of Patterning

Somewhere behind the post-Bloomfieldians’ insistence on the separation of levels lay the notion of duality of patterning. This was one of the several ‘design features’ that Hockett later (Hockett 1958, ch. 64; Hockett 1960; Hockett & Ascher 1964) identified as characteristic of human language, and as a property that distinguishes language from most or all other communication systems in the natural world. The key idea is that all utterances have both a grammatical structure and a phonological structure, and that these structures are inherently distinct and even incommensurate. In this conception, an utterance can be studied either as a string of words (or morphemes, or grammatical formatives of some sort) or as a string of phonemes. Although Hockett’s work on design features appeared some time after the debates over grammatical prerequisites and the separation of levels, there is plenty of reason to believe that he was thinking about such issues earlier (Radick 2016), and plenty of reason to believe that he helped lead the post-Bloomfieldians to the conviction that phonological and grammatical description must be entirely independent of one another (cf.

\textsuperscript{12} Probably the most thorough discussion of this issue in print appeared many years later (Hockett 1987, esp. section 5.3).
the clear early statement of this principle in Hockett 1942, and Hockett’s own opinion, cited earlier, that he had brought Trager round to this point of view).

Treating the separation of grammar and phonology as a fundamental design feature or theoretical precept was also under discussion in Europe, in the form of Hjelmslev’s distinction between the ‘content plane’ and the ‘expression plane’ (Hjelmslev 1935, 1975) and Martinet’s notion of ‘double articulation’ (Martinet 1949, 1960). Although Hjelmslev’s work was influential primarily in his native Denmark, it definitely influenced Hockett’s thinking about duality of patterning (Ladd 2014: 108f). Martinet’s ‘double articulation’ and Hockett’s ‘duality of patterning’ are both now widely understood to refer to the same property of language, and are still widely taken for granted, though seldom considered in any depth. It is obvious that if duality of patterning is a fundamental design feature of language – that is, if grammatical structure and phonological structure are completely distinct and incommensurate aspects of an utterance – then the requirement of separating levels in a linguistic description follows logically.

However, closer consideration of the concept of duality of patterning (Ladd 2014, ch. 5) suggests that the two types of structure are actually not completely distinct and incommensurate. Rather, I have argued that there is an intrinsically hierarchical relation between them, which is implicit in Martinet’s term ‘double articulation’ but not in Hockett’s or Hjelmslev’s versions of the same general idea: in the ‘primary articulation’, utterances can be segmented into (or built up from) words, and in the ‘secondary articulation’ words are then segmented into (or built up from) phonemes. That is, it may be possible to describe whole utterances as sequences of phonemes, but our description is simpler and more insightful if we treat words as sequences of phonemes and utterances as sequences of words. If we do that, we implicitly accept Pike’s arguments for ‘grammatical prerequisites’: for example, we have no difficulty talking about word-initial or word-final allophones; we have no need for ‘juncture’ phonemes; and we have no problem describing internal sandhi and external sandhi as inherently ‘ordered’.
So long as duality of patterning is understood as involving completely parallel and independent structures, then Trager’s lack of regard for those who ‘have remained content to do word phonemics’ is justified; but if we acknowledge the inherently hierarchical relationship between grammatical structure and phonological structure, then ‘word phonemics’ may be not only legitimate but also appropriate. At the very least, it seems clear that the issues of separation of levels, grammatical prerequisites, and word phonemics are ultimately related to the question of whether ‘phonology’ is primarily about perceptibly different sounds in utterances or about networks of contrast and generalizations over the forms of words. The post-Bloomfieldians took the first view, and most of the details of their theory followed from that.

16.4 Postscript

If post-Bloomfieldian phonemics emerged fairly rapidly around 1940, there was nevertheless no decisive break with the developments that preceded it. We could cite Trager & Bloch 1941 or Bloch 1941 as the starting point, but that would ignore the extent to which linguists in the 1920s and 1930s, on both sides of the Atlantic, were all actively engaged in coming to grips with the concept of the phoneme. Bloch (1948) listed a number of works that his ‘postulates’ attempted to synthesize; these included not only several from within the post-Bloomfieldian fold, but also several notable predecessors (e.g. Sapir 1933, Chao 1934, Swadesh 1934, Twaddell 1935) and two works by Pike (1945 and a preliminary version of 1947a).

The end of the post-Bloomfieldian period was more abrupt. Bloch died (at only 58) in 1965, and Trager effectively withdrew from theoretical linguistics in the mid-1960s (Hockett 1993: 786). As for Hockett, his presidential address to the Linguistic Society of America in December 1964 (published as ‘Sound change’ in Language in 1965) made one last attempt to overcome the bitterness of the years that followed the publication of Chomsky’s Syntactic Structures (1957).\textsuperscript{13}

\textsuperscript{13}In his presidential address, Hockett drew a comparison between the atmosphere of the early 1960s and that of the 1870s, when the neogrammarians were expounding their ideas. He made it
He identified four key breakthroughs that define modern linguistics, the third and fourth being roughly the phonemic principle and the mathematical formalization of linguistic theory; the talk seems to have been intended as an olive branch to the generativists, acknowledging that their work had led to ‘the first nontrivial mathematizing of linguistics’, but also pleading for recognition of the importance of what had been achieved in phonology in the 1940s and 1950s. If this was indeed the intention, it failed, and only a few years later Hockett (1968: 3) burned his bridges, dismissing ‘Chomskyan-Hallean “phonology”’ [scare quotes in the original] as ‘completely bankrupt’. A few years later he moved his Cornell office from the linguistics department to the anthropology department (he had always had a joint appointment), and turned his attention to other projects such as an introductory anthropology text (1973) and other topics such as the evolution of language (e.g. Hockett 1978).

Nevertheless, his presidential address suggests that he was ready to abandon some of the dogmas of the post-Bloomfieldian years. He concedes that we need two representations, one respecting lexical identity and the other respecting perceptual identity, and says that it is of no consequence whether we call the mapping between the two ‘phonology’ or ‘morphophonemics’. He also suggests that the representation respecting perceptual identity needs to be expressed in terms of local frequency maxima in a multidimensional phonetic space, implying a significant shift away from the phone-based strings of the post-Bloomfieldian years and toward a representation that can more readily accommodate the relationship between the surface categories and the continuous signal. In the end, he could not overcome the bitterness that grew out of the transition to the generative ascendancy, or the indifference of generative phonologists to the question of how phonological categories relate to the physics of speech. But his faith that progress was possible may yet be vindicated.

clear that he objected to both the tone and the content of some of the early generative work, but at the same time he acknowledged that ‘the attack on the new dispensation from some quarters is vicious indeed’. He declined to repeat the substance of these attacks on the grounds that they ‘were largely made in informal conversation’. Adopting the same policy here, I note only that the temper of the times was sufficiently polarized that eminent professors felt it appropriate to share scurrilous anecdotes with undergraduates.
References


