Phonology **32** (2015) 318–325. © Cambridge University Press 2015
doi:10.1017/S0952675715000184

D. Robert Ladd (2014). *Simultaneous structure in phonology*. Oxford: Oxford University Press. Pp. xvi + 182.

John A. Goldsmith
University of Chicago

Ladd has written the sort of book that senior scholars ought to write, but rarely do: a monograph in which the central concerns of a field are laid out with the perspective that only a period of time measured in decades can offer.* That alone is reason enough to say that this is a book that any serious phonologist should read. I will try to outline the issues that Ladd reviews, in the hope that this will make clear why I think these reflections deserve the attention of phonologists today.

Ladd's is a short book, composed of six chapters and fewer than 150 pages of text. The chapters hang together well, connected by style and outlook. The first addresses the question of what a feature is, and what an autosegment; the second, the nature of the relationship between phonetics and phonology. The third asks what the word prosody means and what it refers to, while the fourth deals with modulation and gradience in language. The fifth is a careful rereading and explanation of Charles Hockett's notion of duality of patterning, and the book closes with a short discussion of the notion of phonological event, which I will not discuss here.

In the first chapter, Ladd begins by asking us to consider the notions of PHONOLOGICAL SEGMENT and PHONOLOGICAL FEATURE, and of such related notions as AUTOSEGMENT and ASSIMILATION. Phonology began with the observation that the sound stream could be well modelled by a representation that consists of a linear sequence of elements chosen from a small inventory of sounds, the phonemes of the language in question, and it took its first step forward with the realisation that by looking at pairs (or tuples) of related words, we find replacements of one sound by another – a realisation that Baudouin de Courtenay placed at the centre of his view of phonology in 1895 (Baudouin de Courtenay 1972).

From the time of Whitney (1875) up to the 1930s, the traditional organisation given to the phonemes of a language was roughly two-dimensional: one dimension represented sonority (and so segments were ranked from the least sonorous to the most sonorous), and the other represented position, from front to back, in the mouth, though voicing was also recognised as an independent dimension along which obstruents might be distinguished. This is how phonologists displayed a set of sounds in a language on the printed page, and it is clear that that is how they understood sounds to be organised

* E-mail: GOLDSMITH@UCHICAGO.EDU.
I am grateful to Diane Brentari for a number of suggestions that have improved this review.

(and in particular, phonologists before Trubetzkoy did not think of phonemes as an unordered and unstructured set of sounds).

Trubetzkoy (1939) and Jakobson (1971) proposed a much wider range of independent dimensions (which they called FEATURES) along which phonological distinctions could be drawn, and that perspective has become the dominant view today among phonologists. But should a phoneme be thought of as being DECOMPOSABLE into features (as a peach can be decomposed into skin, pulp and seed) or as ANALYSABLE with features (as a student can be analysed as male, undergraduate and red-haired)? The question seemed to be of little more than terminological interest until the development of autosegmental phonology, which made it a central question. In an autosegmental context, a representation is composed of two or more linearly ordered strings of autosegments, each chosen from an inventory of autosegments proper to its tier (tones on one tier, consonants on another tier, etc.). No linear order exists between segments on distinct tiers unless they are associated by association lines, which are essentially formal statements of co-temporality. Types of autosegments on a given tier are distinguished from one another by their features. So segments are conceptually decomposable into autosegments (though really, the traditional phonological segments are not *there*; they are only in the pretheoretic imagination of the linguist), and autosegments are present in the representation, as are association lines. But features are not objects: they are maps from autosegments to the set {+, −}.

Autosegmental phonology, as it was developed in Goldsmith (1976), took it for granted that the kind of phenomena that phonologists should account for was what I would today call (in hindsight) automatic morphophonology and all of the other processes that cannot be reasonably separated from automatic morphophonology.[1] This was not the way generative phonologists in the 1970s described the phenomena that they worked on, but in retrospect I cannot think of a better way to characterise them.

By 1980, then, autosegmental phonology could be characterised by the problems it aimed at solving, and the metaphysical assumptions it made – which is to say, what sorts of things it posited as existing in the phonological world. But there is certainly food for thought here, and for some, fuel for scepticism. If we agree to the notion that autosegments exist in some sense, then we have an account of what phonologists call STABILITY, referring to the fact that when a vowel in a tone language is phonologically deleted, its tone is not deleted. The autosegments exist in the same world in which phonological analyses and derivations exist: and what world is that? It is certainly not the space-and-time world that we live and work in; it is the world of the theoretical linguist, and not the world of the positivist. Is that world real? Opinions differ. If we turn to Philip K. Dick's definition of reality ('reality is that which, when you stop believing in it, doesn't go away'), I would, for one, be willing to say that the world of the phonologist is real.

Ladd's discussion focuses on another aspect of reality and of phonological analysis: the fact that both features and autosegments (across various proposed

[1] The term AUTOMATIC MORPHOPHONOLOGY in post-Bloomfieldian analysis referred to phenomena in which one phoneme is replaced by another in a context that can be entirely described in phonological terms; these processes by definition did not lead the analyst to worry about issues of segmentation, or of blurriness of edges.

analyses) far more often than not correspond to articulatory gestures, and these gestures are certainly events that occur in time and in space (and thus have a strong claim to be real). But just saying that a feature or an autosegment corresponds to a gesture is hardly enough; a gesture is a constantly changing thing whose co-timing with other gestures is typically a highly organised affair. Can phonology (with or without autosegments) be recast as a symphony of articulatory gestures, as Browman & Goldstein have argued in many publications (e.g. 1986)? Such a move would be attractive to someone who hopes to ground phonology in phonetics, and to base phonological naturalness on physical or articulatory simplicity, and such a person is likely to see the central role of autosegmental phonology as a way of describing assimilations without needing recourse to mentioning a feature and its value twice in the statement of a rule (why write C → [−voice] / __ [−voice] when you can write a rule that adds an association line?). But Ladd quite rightly emphasises that assimilation processes are phenomena that are at the margins of autosegmental analysis; the stronger arguments for autosegmental analysis involve more complex phonological phenomena, including unbounded feature (e.g. tone) spreading, contour-valued segments (rising and falling tones) and the morphological status of autosegments (such as floating tones and nasal segments). It is the study of these phenomena that distinguishes the study of autosegmental phonology from, say, a study of the correct notational conventions for writing phonological rules in the *SPE* framework, or a study of the correct formulation of possible constraints in a theory of Optimality Theory (if your version of OT permits different constraints in different languages).

In short, a phonetic interpretation of autosegmental phonology speaks to only a small part of what the approach proposes to handle in phonology, and what it intends by the placement of certain features on distinct tiers. But there is another aspect of autosegmental phonology that Ladd does not mention at all, but which is as important (or so it seems to me): this is the notion that phonological representations have a geometry that can be understood by the linguist, and for which certain patterns of association are more and less well-formed – and that operations exist in languages to maximise the well-formedness of a phonological representation. Conceptually, the metaphor that inspired this was the way in which atoms share electrons in order to achieve an overall lower energy state – and so morphemes placed on separate tiers could become associated if that improved the overall well-formedness of the representation. This notion has remained an important locus of theoretical discussion, as anyone who has read the phonological literature over the last 25 years is well aware.

Chapter 2 is a brilliant discussion of the evolution of the role of phonetics in phonology over the past century, giving proper due to the needs and goals of the developers of orthographies in the 19th century and the changes in the understanding that underlay the IPA. I would cavil with some of the statements: Ladd understates the sophistication in 19th-century phonology, it seems to me, when he writes (p. 37):

> The idea of a universal scheme of classification for phones gives rise to what is perhaps the central theoretical construct of mid-twentieth-century phonology, namely the feature. In an informal way, of course, the dimensions of

> the IPA symbol chart are a kind of feature analysis, but we are concerned here with the place of such classification *in phonology*. Linguists had long been aware that certain kinds of sound changes are common and somehow natural, and that common phoneme inventories across languages are often quite symmetrical if described in terms of phonetic dimensions. But this awareness played no formal role in most Anglo-American phonemic theorizing, which was almost exclusively concerned with the procedures for grouping phones into phonemes.

That last phrase certainly suggests that Ladd has Harris and Hockett in mind, and it is certain that Harris was very resistant to the notion that articulatory phonetics should be allowed in the phonological analysis. But as a more general statement, Ladd's remark is not at all accurate. Whitney's *Life and growth of language*, published in 1875 and highly influential in Leipzig over the next decade or so (when Leipzig was the centre of the linguistic universe), is quite explicit on this point: '*Ofer*, however, has become *over* with us, by the conversion of a surd into its corresponding sonant sound, a phenomenon of very wide range and great frequency in language' (1875: 57) – consonants become voiced, as we say today. But Whitney does much more than just point this out; he goes on to examine examples from English, German and Sanskrit that motivated Grimm's Law (as it had just recently been dubbed):

> This is, indeed, the famous 'Grimm's Law', of the permutation or rotation of mutes in Germanic speech. It is only an example – to be sure, an unusually curious and striking example – of what is universally true between related languages: their sounds, in corresponding words, are by no means always the same; they are diverse, rather, but diverse by a constant difference; there exists between them a fixed relation, though it is not one of identity (1875: 57–58).

If there is any doubt that Whitney is entirely certain that this is a central and not a marginal observation regarding phonology, he eliminates that doubt entirely in what follows: 'and, heterogeneous as the facts may at first sight appear, the student soon finds that they are very far from being a mere confusion of lawless changes; they have their own methods and rules' (1875: 58). Which is what we call phonology, of course. 'One sound passes into another that is physically akin with it: that is to say, that is produced by the same organs, or otherwise in a somewhat similar manner; and the movement of transition follows a general direction, or else is governed by specific causes.' He then goes on to explain that this momentous discovery is sufficient motivation to study linguistic phonetics seriously, and proceeds to introduce his reader to that young science.

But Ladd's discussion of 20th-century phonology, including the specific contributions of Trubetzkoy and Jakobson (and how these contributions related to the work of the English and French phonologists), is solid and important for today's reader. One of the fundamental questions is this: is there a closed universal inventory of possible phonetic segments (i.e. segment types)? The practice of transcribing speech in something like IPA notation assumes that the answer is yes, more or less; the practice of plotting the 'same vowel' on a scatter plot, looking at multiple realisations by the same speaker of the

same utterance, with dimensions defined by formant frequency or any other measurable parameter, typically leads us to the other answer: no, we do not have a separate symbol for each utterance, but the distribution of the pronunciations in some vector space (whose dimensions might be vowel formants) is the right way to characterise how the vowel in *cat* is pronounced by this particular speaker. What is even more interesting is that as we try to characterise the distribution of a sound in this way, we find that there are language-specific correlations between phonetic dimensions: Ladd mentions Pearce's (2007) analysis of Kera (Chadic, Chad), in which covariance of VOT and pitch is part of the analysis of the distribution of the phonetic production of vowels. At this point, we can hope that a good characterisation of speech sounds can be provided for the phones of each language, with the understanding that such a characterisation must employ multidimensional statistics. In any event, the answer to the first question is no: there is no closed, universal inventory of possible phonetic segments.

Another question is whether a speech event can usefully be analysed as a sequence of phone-sized events (by 'phone-sized' I mean something of length 50 to 200 ms). Much of the work in automatic speech recognition since the late 1970s has employed hidden Markov models, which do not make such an assumption. What they do instead is consider a host of linguistically plausible hypotheses as to what is being said, and then they compare the probabilities that each hypothesis assigns to the sound stream that is being analysed: the hidden Markov model selects the hypothesis which assigns the highest probability to the data (also taking into consideration the grammatical probability of the hypothesis based on what was probably said before in the utterance). Now, hidden Markov models have properties that make them implausible candidates for ultimate scientific truth, from most phonologists' point of view, I think; they do not handle issues of timing and suprasegmentals in a natural way, though in fairness it should be added that new models that integrate such concerns are being developed daily. The point to bear in mind is that phonologists share the space of scientific truth with workers with competing hypotheses, and phonologists have to continually rethink the assumptions that they have made, especially if they have been made for convenience in a time gone by. Ladd lays out many of these options clearly, and does not try to oversimplify, or hide from the reader the need for constant refining and retuning of conceptual basics.[2]

The next two chapters are short essays on the growth and use of the terms 'prosody' and 'gradience'. You might think that a discussion of how the term 'prosody' has been used in recent decades would not tell a phonologist something she did not know, but Ladd looks at a good deal of literature to establish that the field continues to slide between (on the one hand) a usage which emphasises the aspects of sound that are anything other than what the

[2] There is an important question lurking close to the surface in Ladd's discussion: to what degree should practical aims and technological capabilities help in our definition of linguistic theory? Both of these considerations – involving fieldwork, development of practical orthographies, development of speech recognition – have played roles in determining, in part, what the central questions are in phonology. Is that good or bad? And if it is not bad (I do not think it is), how important is it to be aware of this?

consonants-and-vowels are in some strict sense and (on the other) a usage which emphasises the relation between earlier and later sounds in the utterance, as well as any hierarchical structure that may be discovered involving constituents larger than individual segments or phones. Maybe not every phonologist needs to read this chapter, but I do suspect that every graduate student working in phonology does.

Ladd's discussion of MODULATION and, especially, GRADIENCE has some surprises in it, most notably the clear sense with which Dwight Bolinger introduced the term gradience into linguistics in 1961 – I, for one, was not aware that Bolinger's use had been muddied in much of the literature since. Bolinger wanted to draw his reader's attention to those phenomena where a continuous physical dimension could be employed by language in order to match up with a meaningful dimension, as when someone says that he 'caught a bi-i-i-i-i-g fish'. He did not have in mind phenomena where judgements of grammaticality were uncertain and variable (though he was certainly the greatest master of that genre in living memory, as his other publications established). Ladd presents an insightful discussion of the advantages and the pitfalls in separating the linguistic content of a statement from the way the statement was said. He helps distinguish the PARALINGUISTIC properties of an utterance (which, if we follow strictly Ladd's definition, include facial and body positioning, but also to some degree pitch, rhythm and loudness) from the INDEXICAL properties of an utterance, which encode information about the speaker and the context of the utterance. Or better: he helps to distinguish the linguistic/paralinguistic difference from the proposition/indexical difference, and he clarifies that neither distinction can simply be tied to a difference between what is language-specific and what is universal.

The last chapter which I will discuss in this review takes as its theme the exegesis of the notion of DUALITY OF PATTERNING in human language, a notion that was developed simultaneously by Hockett (1958) and by Martinet (1960), both drawing on ideas of Hjelmslev (1936). As Ladd stresses, duality of patterning is a characteristic of language that is widely emphasised in introductory courses on language, and rarely addressed in professional scholarship. For those who have forgotten, we may recall that what is at issue here is something like this: within a word of a human language, there may be a structure that is determined solely by phonological considerations, and a structure that is based on lexical and morphological considerations, and these two structures may have little or nothing to do with one another. A word may be monosyllabic (*book, books*) or disyllabic (*stomach, houses*); it may be monomorphemic (*book, stomach*) or bimorphemic (*books, houses*). Ladd explores two assumptions that underlie duality of patterning, as we see it in Hockett's writing: first, the meaninglessness of phonological elements (such as phonemes), and second, the notion that phonological structure is exhaustive. While *most* morphemes have a meaning that cannot be usefully understood as built up out of meanings of their phonological parts, enough exceptions have been documented over the years to establish this area as one in which a watertight wall cannot be erected, and to the extent that there is such leakage, some phonemes may well be said to be at least somewhat meaningful, in at least some corners of the language. Ladd points out that phonological structure is not exhaustive in a way that he takes to be necessary for

language to manifest duality of patterning; as examples, he points to sounds that may be just on the margin of the phonemic inventory of a language, like the velar fricative of *Bach* in English or the retroflex *r* in Montreal French, or the uncertainty in a given dialect as to whether two sounds are allophones or in contrast, or the analytical uncertainty that may arise when a language permits two distinct phonemicisations, something typically found when two neighbouring segments on the surface always agree in some phonological feature.

Ladd extends the discussion of the duality of patterning in language to areas of current research, such as notably the grammatical structure of signed languages of the Deaf. Here there are several issues that come into the picture: (i) the distinction between linguistic and paralinguistic aspects of an ASL utterance is fraught with difficulties. It seems entirely reasonable to exclude the orientation of a speaker's torso or the direction of his gaze from the area of linguistic analysis, and to place it in paralinguistics, but that would almost certainly be the wrong choice in the analysis of ASL. This in turn leads to another point: (ii) it is not at all clear that a complete specification of an ASL sentence could be adequately described by a sequence of words (even if a syntactic tree of constituent structure were added): there are arguably aspects of the linguistic representation which are not the lexical items of the utterance (that is, are in addition to the lexical items of the sentence, and they are not a matter of syntactic constituency). (iii) A very large part of a sign-language utterance can be analysed into meaningful words, and in a number of languages that have been considered, like ASL, these words can be analysed into formal subcomponents (handshape, orientation, location and movement), with each of these subcomponents being drawn from a (largely) fixed inventory of language-specific choices. (iv) It is not clear whether the four subcomponents just mentioned should be taken to be the analogue of the phonological feature or the phonological (auto)segment. (v) Many (but far from most) of the subcomponents have specific meaningful functions. For example, in ASL a number of signs are formed with a path movement of an inverted L; they are all names of American cities (Rochester, Philadelphia and Indianapolis, for example). The movement does not mean 'city' outside of the context of these words, just as the suffix *-ham* or *-by* does not mean 'city', though they are structural parts of the names of many English cities and we recognise them as such. Spoken language does not have the luxury of using some feature specification (such as High tone, or labial-velar) in just one part of the lexicon, while ASL does (since there are so many possible movements, among other things). And finally, the ways in which ASL signs can be iconic is both long and complex.

Ladd's discussion of the relevance of sign-language studies for our understanding of the duality of patterning in language is certainly welcome. I think that it glosses over one or two important points, though, that deserve a bit more elaboration in this context. Ladd notes quite properly that the greater iconicity of sign language 'can plausibly be attributed to the difference of medium between sign and speech: it is easier to make a symbol look like what it denotes than to make a sound like what it denotes' (p. 120). But the difference between sound and space is much greater than this suggests; the difference involves the three-dimensionality of space, and the way joint (or

more generally, body) structure can be mapped to conceptual form (read 'logical form', if you prefer) in sign languages. These differences lead to enormous differences in potential for grammatical development in sign languages, and there are many phenomena that have been studied in sign languages which have only poor cousins (so to speak) among those found in spoken languages. So when Ladd writes that 'as long as we accept that iconicity plays some role in all forms of language, then the fact that it is exploited more heavily in signed than in spoken languages is simply another consequence of the different medium' (pp. 120–121), I think that he greatly understates what separates signed from spoken languages: sign languages have *developed* grammatical systems that exploit both the fact that space is three-dimensional and that signers have an articulated skeleton whose movement at joints can be observed by the hearer – or rather, the viewer. The subject and the object of a verb in a sentence can be visually presented simultaneously and in a particular structural relationship, something that cannot be done in a spoken language, and this is the result of the grammars of sign language developing an opportunity that did not exist before they developed it. Yes, something new has arisen, and it is a consequence of the presence of a new medium (three-dimensional space), as well as the visible joints of the signer, but it is not 'simply a consequence' of those things.

It would not be difficult for a reviewer to write a response to Ladd's book that was twice as long as the book itself. It is an immensely stimulating set of reflections which will, I feel certain, encourage deeper thinking about central questions in phonology.

REFERENCES

Baudouin de Courtenay, Jan (1972). An attempt at a theory of phonetic alternations: a chapter from psychophonetics. In Edward Stankiewicz (ed.) *A Baudouin de Courtenay anthology: the beginnings of structural linguistics*. Bloomington & London: Indiana University Press. 144–212.

Bolinger, Dwight L. (1961). *Generality, gradience and the all-or-none*. The Hague: Mouton.

Browman, Catherine P. & Louis Goldstein (1986). Towards an articulatory phonology. *Phonology Yearbook* **3**. 219–252.

Goldsmith, John A. (1976). *Autosegmental phonology*. PhD dissertation, MIT.

Hjelmslev, Louis (1936). On the principles of phonematics. In Daniel Jones & D. B. Fry (eds.) *Proceedings of the 2nd International Congress of Phonetic Sciences*. Cambridge: Cambridge University Press. 49–54.

Hockett, Charles F. (1958). *A course in modern linguistics*. New York: Macmillan.

Jakobson, Roman (1971). *Selected writings*. Vol. 1*: Phonological studies*. The Hague: Mouton.

Martinet, André (1960). *Éléments de linguistique générale*. Paris: Librairie Armand Colin.

Pearce, Mary (2007). *The interaction of tone with voicing and foot structure: evidence from Kera phonetics and phonology*. PhD dissertation, University College London.

Trubetzkoy, Nikolai S. (1939). *Grundzüge der Phonologie*. Göttingen: Vandenhoeck & Ruprecht.

Whitney, William Dwight (1875). *Life and growth of language*. London: King.