Phonology **14** (1997) 133–141. Printed in the United Kingdom
© 1997 Cambridge University Press

Reviews

Steven Bird (1995). *Computational phonology: a constraint-based approach.* (*Studies in Natural Language Processing.*) Cambridge: Cambridge University Press. Pp. xv+203.

John Goldsmith
University of Chicago

What is computational phonology, and what does it bring to phonology?* Does it offer a new set of tools for phonological analysis, or a new conception of data or a new set of goals? These are the questions that come up in reading Steven Bird's recent book, and I would like to examine some of the perspectives Bird offers.

The type of phonological analysis that Bird develops in this book has been explored most notably by several phonologists in Great Britain, including John Coleman and James Scobbie, in addition to Bird himself. In some interesting respects it has resonances with the influential work on syntax developed by Gerald Gazdar, Ewan Klein, Geoff Pullum and Ivan Sag under the rubric of Generalised Phrase-structure Grammar (GPSG; Gazdar *et al.* 1985), and it is surely no coincidence that Bird expressed his indebtedness to Ewan Klein's guidance in the writing of this book, itself a development of a doctoral dissertation at Edinburgh University (Bird 1990).

Computational phonology is organised into six chapters and an appendix. Chapter 1 offers an introduction, a brief discussion of autosegmental formalism and a survey of computational phonology and of constraints in phonology. Chapter 2 presents a formalism appropriate for an autosegmental phonology, while Chapter 3 discusses apparent problems for a non-destructive theory of phonology (that is, a theory of phonology in which specifications can never be undone by rules, roughly speaking). Chapter 4 presents a version of feature geometry linked up to Browman & Goldstein's gestural score model (Browman & Goldstein 1989), Chapter 5 presents some aspects of Bird's computer implementation of his system and Chapter 6 presents some conclusions. The appendix covers some technical areas sensibly left out of the main text.

1 The name computational phonology

The dissertation which provided the roots for this book was entitled *Constraint-based phonology*, and I think that that would have been a better title for the present book.[1] What, after all, is 'computational phonology'? Is this a term we are free to invent and provide with an idiosyncratic meaning, or does it have a sense by virtue of the meaning of its parts? Surely the latter; we already have an idea of what computational linguistics is, so we are within our rights to expect

https://doi.org/10.1017/S095267579700328X Published online by Cambridge University Press

computational phonology to be some version of phonology that does one or more of the following things:[2]

(i) Computational phonology could increase the base of the empirical data, archived in some fashion, used to inform and test our theories. Computational syntax has been doing this for decades now, with increasing sophistication and effect.

(ii) Computational phonology could enter into a new relationship with the data of phonology, by using computers to directly interact with sound, rather than passing from phonological representations to sounds via a stage with a phonetic representation.

(iii) Computational phonology could use the techniques derived from knowledge of programming languages and data structures to increase the sophistication of the toolbox used by the working phonologist. Many of the ideas used to good effect in computational syntax, most strikingly various versions of feature inheritance, have been borrowed from programming languages.

(iv) Computational phonology could be phonology informed by the needs of researchers who are trying to develop computational devices that can speak and understand spoken natural languages, in much the way that computational syntax has been heavily influenced by the work of researchers who are actively trying to develop programs that parse and generate texts in English, French and so forth. In this sense, computational phonology would highlight those aspects of phonology that help in voice recognition and in speech synthesis.

Which of these four images of computational phonology does Bird focus on? Primarily the third, the ways in which certain kinds of programming languages, those which do not encourage serial derivations, can inform our thinking about phonology. As we will see shortly, Bird also mentions that first point, the fact that computers can be used to test our predictions in a tougher way. But while this is an important point (I completely agree with Bird on this), it is not a major part of the story that he has to tell.

A certain conception of how linguistic information is organised, a conception that can be modelled relatively easily in certain programming languages, leads to the model of phonology that Bird explores. This model is a 'non-destructive' one, one in which phonological representations are achieved by various sorts of unification processes, summing and reorganising information without anything like deletion or erasure, whether of segments or of feature specifications.

In §1.3, 'Computational phonology', Bird reviews in the briefest of ways a bit of the work of some of the automatic rule testers that were written to apply *SPE*-style rules, and mentions that speech technology and phonology ought to have something to say to each other, at least in principle. But he doesn't say many of the things here that I think really ought to be said. For example, much of the work of contemporary phonology (especially in the European languages) concerns what traditionally has been called morphophonology, the relationship between words, and much of the area of computerised language technology feels that it can get along fine with little or no help along these lines, thank you very much, since the vocabulary of English is large, but by no means overwhelming, and that vocabulary can be treated as a very large list. Whether *omen* and *ominous* are linguistically related is a matter with a remarkably low priority in the speech technologies. Even syllabification is a matter of relatively low priority in the fields of speech recognition and synthesis, despite the fact that that's an area of real-time processing in English.

https://doi.org/10.1017/S095267579700328X Published online by Cambridge University Press

When we take a look at what real computational linguists (or engineers who work on language in some fashion) are interested in, it brings home a striking fact: the revolution in linguistic theory, the one we often identify as the Chomskian revolution, has put an extremely heavy emphasis on evaluating empirical exploration simply by the measure of how well that exploration answers questions set up by the theoretical interests of the moment. (Graduate students internalise this early on: a good conference paper is one that supports, or better still, attacks an analysis made by a known theoretician a few months ago in an unpublished paper. I know linguists who look at me funny when I make this observation: of course that's true, they say – how could linguistics be otherwise and still be a science?) An immediate consequence of this is that very little effort goes into trying to develop a complete or systematic description of the sound pattern of even as well-known a language as English. Example? Well, what are the most common underlying segments in English? What are the most common sequences of two segments in English? Do you know the answer, or even where to look to find out the answer? (I recently found out the answer to those questions by exploring some corpora, but I don't think the answer would have been easy to find in any published literature.) Better yet, consider this: if we determine the savings that a lexical phonological rule offers us by virtue of its allowing us to leave features unspecified in this lexicon, and if we take that notion seriously, can we find any published research on English (or any other language) which actually measures the savings provided by one analysis *versus* the savings offered by another?[3] (What is occasionally assumed in the literature on underspecification is that saving one feature specification on an 'i', for example, is worth just as much as saving one feature on a 'u'; but of course the number of instances of 'i' in the lexicon is overwhelmingly larger than that of 'u'. Isn't it? Yes, it is.) These questions (which are admittedly rhetorical observations) point out two things: we as phonologists have not, by and large, actively worked on providing general accounts of the sound patterns of even the languages we know well, and second, a greater familiarity with the techniques of computational linguistics would allow us to answer our own theoretical questions more adequately and in a more sophisticated fashion. While none of the alternatives that Bird discusses in the first chapter are discussed in depth – this would hardly have been feasible – the discussion serves to situate Bird's goals and interests against the backdrop of other work going on in the field, much of it done with other sets of assumptions and guidelines.

2 Formalisation

Bird begins his development with a discussion of the notion of explicitness and formalisation in linguistic theory. In a word, he's for it, and he feels that many leading phonologists in the past thirty years have not in practice lived up to their own professed support for explicit formal theories. I think it is not only helpful but important to recognise that one of the most important factors in the tremendous interest in phrase-structure approaches to syntax in the 1980s (which included not only GPSG, but developments and descendants of GPSG as well, and Bresnan's lexical-functional grammar, to mention only the two most prominent examples) was a disappointment in the growing lack of precision in the syntactic theories that developed out of Chomsky's (1973) 'Conditions on transformations' and (1981) *Lectures on government and binding*.

https://doi.org/10.1017/S095267579700328X Published online by Cambridge University Press

All too often, work within that framework seemed to consist essentially of making a clear and sharp statement in terms that were as distant from observational properties as one could manage, with concomitant research amounting to finding a way to interpret the sweeping generalisation in such a way that it would not be contradicted by the observed facts. The Case filter (as in Chomsky 1981, originally proposed in Siegel 1974 in a different context) was as clear an example of this enterprise as one might want to find. While it hardly needs to be added that proponents of this principles and parameters approach to syntax would challenge my characterisation, the fact remains, I believe, that it was this perception that lent support to work on GPSG and other frameworks, work whose explicitness was guaranteed by the commitment to implement the grammars in a computational context. And as Bird knows well, once one is committed to developing a computational grammar, fuzzy principles and handwaving simply can't sneak into the theory in the same way. Bird wants to see no such illusions in phonology; and his effort in this respect is the heart of this book: to make clear what the price of admission is for a theory that wants to be taken seriously as a formal theory of phonology.

I am forced to confess that I am of two minds regarding the lack of formal rigour in phonological analysis and description in the last twenty years. I know perfectly well that he is right in bringing our attention to this lack of rigour.[4] But I'm not sure about the particular examples he chooses to make his point. Is current phonology flawed by a lack of formal precision – and if so, what parts? What can be done to remedy the situation?

Bird's answer to these questions is unmistakable. Yes, the flaws are there: 'given the central status of notation in autosegmental theory, the imprecise definition and widespread abuse of notation in the literature is somewhat disconcerting' (p. 12). Bird explains his own concerns in this way: 'the reader seeking a *precise* understanding of the basic elements of autosegmental representations and rules will often be left floundering'. In particular, he says, 'Is the association relation reflexive, symmetric or transitive? What does the no-crossing constraint really mean? What does the *absence* of an association line mean?'. The reader will recall that the association relation is a relation between elements on distinct autosegmental tiers. That fact alone is sufficient to tell us that the association relation is not reflexive: an autosegment is on no more than one tier, and hence cannot be associated with itself. The relation is symmetric, a conclusion never hidden in the foundational work on the subject, where the formal symmetry between the tonal and the vocalic tiers (for example) is emphasised at length. It follows that if a relation is symmetric and non-reflexive (and non-null!), then it cannot be transitive. There's a formalisation of the notion of tier, association line and the Well-formedness Condition in Goldsmith (1976: 28ff) which uses notions of point set topology which Bird does not cite (nor seem to show any awareness of here). Bird later (p. 66ff) develops an alternative approach, extending some ideas of Sagey (1988).[5]

Bird focuses on the 'no-crossing constraint', and he says that there has been confusion as to whether this constraint blocks a rule from creating a situation in which lines are crossed, or whether the constraint repairs violations. Now, this is a point whose importance I have often emphasised; Bird cites Goldsmith (1990: 30) as his source for the first interpretation, but a search there shows no reference to the constraint in question (we find rather a reference to a rule formalism for unbounded spreading). A long section (Goldsmith

https://doi.org/10.1017/S095267579700328X Published online by Cambridge University Press

1990: 319–331) discusses this very issue in excruciating detail, detailing the connection between the Well-formedness Condition and a phonological theory employing violable constraints.[6]

When all is said and done, we phonologists will have to answer Bird's requirement that we provide a full and explicit theory. Curiously, one thing that Bird does not require of phonology (though it seems to me that we should be required to ante up on this one) is a theory of where underlying forms come from. Given his surface-orientation (see the next section), his approach has an easier time accounting for where underlying forms come from than most approaches do. But many contemporary theories operate as if the problem of language acquisition were primarily a problem of grammar acquisition, with no parallel concern for how underlying forms are abstracted.[7]

3 Declarative, constraint-based phonology

The focus of Bird's attention in this book is declarative, constraint-based phonology. The central idea of this point of view is that all generalisations that we might wish to make about the phonology of a language, involving both statements true across a language and statements specific only to a single lexical item (such as what its pronunciation is), are simultaneously valid and form a consistent set of propositions. Thus to watch a declarative phonology provide an analysis of any particular utterance is to watch a process in which statements of varying degrees of generality are brought together and in which constraints are resolved. Bird in addition is committed to a one-level or monostratal view of phonology, permitting only a single level of representation on which all of the generalisations and constraints will be statable and stated.

My own judgement of such a view is rather negative, but that judgement has nothing to do with the general computational context in which Bird's proposal is made. I'm as little convinced by this view of phonology in this context as I'd be in any other (I say it in this context to acknowledge prejudices on my part, not as a critique of Bird!); equally important is the fact (for it seems to me that it is a fact) that being computational doesn't force one's hand in the direction of a monostratal view of phonology. There may be a danger lurking somewhere that some quite non-obvious link will emerge in phonologists' minds that doing computational phonology is doing monostratal phonology, and I think that would be unfortunate (though Bird does make a reasonable effort to give equal air time to other approaches in Chapter 1, as I noted above).

Chapter 3 is 'A critique of destructive processes', those processes countenanced by most theories of phonology today which go well beyond what is permitted by a constraint-based approach to phonology. Bird tackles the question of how to deal with phonological phenomena that have often been called 'deletions', in view of the fact that with only one phonological level at his theoretical disposal, a segment must either be there or not be there: there's no way to say it's 'deleted'. The consequence of this is that deletions that involve lexeme-specific material (your usual deletion, that is) have to be reinterpreted as complex lexical items. Bird does not consider that phonological material entered in a lexical entry forms part of a representation on a linguistic level, so a given lexeme can have a complex of segments in it, with complex conditions determining which of those segments appears on the (unique, monostratal) representation in a given utterance. For example, if *petit* has a stem-final 't' in

https://doi.org/10.1017/S095267579700328X Published online by Cambridge University Press

its lexical entry, that 't' could have associated with it the condition that it appears if and only if the following segment is a vowel.

At this point, one would be within one's rights to make two observations: first, such a model is what has traditionally been called a two-level model, and unless there's some good reason to change the terminology and call this a one-level model, principles of fair advertising suggest that we recognise that this is what a model with a morphophonemic and a second phonological or phonetic level looks like. If your lexical representation is distinct from your phonological representation, then you've got two levels in your model. You may not like that terminology, but that's how the terms have traditionally been used.

Second, what of the many cases in the literature where derivations seem necessary? Kenstowicz's (1994: 95) textbook, for example, cites a case in Tangale where a voicing assimilation rule fails to apply between a suffix and a preceding stem just in case the stem ends underlyingly with a vowel. As Kenstowicz notes, the voicing assimilation arguably applies at a level (some level) where the stem-final vowel is present – that's why the voicing assimilation rule 'fails' to apply. I expect that if someone is going to redo phonological theory without any fancy stuff like derivations, then a few of the tough cases will be dealt with, and it will be made clear that the methods can be extended to familiar cases. In this case, I don't see that Bird has offered us that ability to view familiar cases in a monostratal way. He does note that some deletions can profitably be viewed as articulatorily hidden, following accounts by a number of people, including Browman & Goldstein, such as when *ten pins* is realised *tem pins*: the coronal gesture may take place but be hidden behind the labial gesture, but this is surely a small fraction of the cases of deletions that phonologists have studied.

If it sounds like I'm defending the case for derivations in phonology, it's certainly not because of any desire to see them there, but just because of the complexity that many generations of phonologists have unearthed in their studies.[8] The reader will share with me a curiosity and an interest in how Bird treats some other well-known cases, such as the *rider/writer* neutralisation of *t/d* as flap in American English, which seems to follow (and counterfeed) the rule lengthening the vowel before a voiced consonant. I expected that he would follow the path suggested by Lakoff (1993) and others, and state both Flapping and Lengthening as rules which directly relate the lexical and the surface form, Flapping conditioned by the surface environment and Lengthening by the lexical representation.[9] As I've discussed in Goldsmith (1991), for example, this case is certainly not the best one that can be made for rule ordering. But Bird chooses instead to attack the particular formulations of Lengthening and Flapping. He notes, following Selkirk, that *italic* and *idyllic* cast doubt on the formulation of Lengthening that he gives, though as far as I can see, what these examples show is that an effect of a consonant on the preceding [ay]'s quality only occurs within a foot, and not across a foot boundary. Bird notes as well that a cross-linguistic survey will support a phonetic basis for Flapping, but he only implies that this weakens the case for rule ordering, without spelling out why he thinks so – yet surely one can agree that both Flapping and Lengthening have a phonetic basis, in some sense, without thereby agreeing that there is or isn't such a thing as rule ordering.

This brings us to the aspect of Bird's argumentation that appeals to me the least, which is this: as I read him, he holds an overly heightened sense of the

https://doi.org/10.1017/S095267579700328X Published online by Cambridge University Press

divide that separates phonology and phonetics, believing that if some linguistic phenomenon (Flapping, word-final devoicing, etc.) shows a property that we would usually associate with the business of phoneticians, then phonologists no longer need to worry about it, and they'd even be better off if the theory of phonology were constrained so as not to permit phonologists to worry about it. Bird does not accept this characterisation, but this reader found it all too easy to read his strategy in chapter 3 this way: for any rule that his non-destructive phonology really can't handle, he points out some aspect of the phenomenon that seems to indicate that a phonetician would be interested in the problem as well. German word-final devoicing? Robert Port (Port & Crawford 1989) has noted that there are statistically significant differences in the phonetics of a devoiced obstruent and an underlying voiceless obstruent. No matter that the difference that remains is functionally inaudible. The same view seems to carry over to variable rules, as I read Bird: any evidence of a rule's variability is evidence that the phonological theory need not concern itself with the process.

4 Conclusion

I have chosen for comments just a few of the topics that Steven Bird addresses in *Computational phonology*. Anyone with an interest in the formalisation of phonological theory, and especially in the ways in which computational implementation of phonological theory must respond to the innovations of the last twenty-five years, would be served well by reading this book. Its views regarding phonology are, from my perspective, idiosyncratic, but it comes to grips with interesting issues that arise out of the attempt to implement phonology computationally, and that will be of interest to more and more phonologists, we may hope.

It would not surprise me if the issue of destructive processes in phonology – Bird against them, derivational phonologists for them – became a very important issue in the next decade or so, as researchers try to integrate phonological theory into speech recognition devices. At the moment, speech recognition systems use a formal model that is almost completely unknown in the phonological community, that of hidden Markov models,[10] devices which permit zero realisations of phonological material, but at a probabilistic cost. Formalisations of the sort that Bird calls for will have to be accomplished if phonological theory is to be integrated into that work.

NOTES

* This review was written while I was a visitor at the Microsoft Corporation and the University of Washington during a sabbatical year, and I am grateful for the support they provided. Some influence of that year on this review may be noticeable. This review has also been greatly improved by Steven Bird's criticisms and suggestions based on a prior draft. I do not believe he is satisfied with my revisions, however.

[1] Bird's decision to change the title makes perfect sense, though, in view of the very specific sense assigned to the term 'constraint' within Optimality Theory, and his desire not to mislead the potential reader into thinking that the book at hand was concerned with OT constraints.

[2] Steven Bird reminds me that he has dealt with this question as well, in §3 of the preface to the special issue of *Computational Linguistics* (**20:3**, 1994).

https://doi.org/10.1017/S095267579700328X Published online by Cambridge University Press

[3] Steven Bird points out to me that Mark Ellison's (1992) PhD dissertation does just this.

[4] One of the more egregious examples, in this writer's view, is the treatment of whatever it is that might be called the OCP. It seems to me that it just won't do to invoke the OCP when something like the OCP helps the analysis, and ignore it when the same interpretation undoes the good work in the next language – for example, on Scott Myers' account of High Tone spreading in Shona, presented in Kenstowicz (1994: 335ff), a High tone spreads rightward towards the next High, but stops spreading one toneless syllable short of that High so that there will be a toneless buffer syllable in between the domains of the spreading H and the next H. On the other hand, in Makua, essentially the same spreading takes place (Odden 1990), but the High spreads right up to the next High, leaving no toneless buffer syllable, OCP or not. It would be helpful to bear in mind that some terms (like the OCP, at times, or 'Meeussen's Rule' or 'feeding order') function primarily as helpful tools for communication, but don't provide the basis for an explanation. Some serious efforts have been made to do better than this with the OCP, to be sure – but if one is looking for examples of theoretical laxity, citations of the OCP are a rich lode to mine.

[5] There is some irony in that this approach was explored in Goldsmith (1974: 8ff), but I later rejected it in favour of what struck me as a more elegant approach, that which I gave in Goldsmith (1976). I had some unsympathetic remarks about this approach in Goldsmith (1990: 352).

[6] 'This type of notion of a "soft" – a violable – well-formedness condition is extremely important to the approach being suggested here, and in outlook is at odds with the classical generative approach. Nonetheless, it has clear antecedents in the literature...' (Goldsmith 1990: 357).

[7] I recently read a (carefully done) paper on feature geometry which was at great pains to show how (given the underlying form that the author had chosen) it would not be difficult for the language learner to figure out how to formulate rules to derive the correct output forms involving a particular morpheme, treating the surface forms as if they were the mystery that the grammar builder had to figure out how to achieve. This is backwards, of course; assuming feature geometry is an adequate representational system for surface phonetics, the learner really has no difficulty at all coming up with the correct surface representations: they are within her epistemological hand's reach; she hears them directly. What's difficult for the language learner is convincing herself that a single underlying form could be responsible for the variety of forms that are actually heard as realisations of the single morpheme in question. In this sense, a formal phonological theory without a formal theory of selection of underlying representations is only half the story.

[8] The situation is not nearly so simple. It seems to me that anyone who spends some time reading the literature on Bantu tone systems (Clements & Goldsmith 1984 is as good a place to start as any) knows that the discrepancy between underlying and surface representations is a good deal more distant and complex. Look at Massamba (1984), for example, for a case of deletion of (in some cases, quite long) strings of adjacent High tones. That case is typical of the situation found in Bantu systems, often under the rubric of 'Meeussen's Rule'.

[9] That is, the choice of vowel length and quality for [ay] would depend on whether the lexical representation contained a voiced stop as one of the options for the following consonant.

[10] See Rabiner (1989), Charniak (1993) or Kornai (1995), for example.

REFERENCES

Bird, S. (1990). *Constraint-based phonology*. PhD dissertation, University of Edinburgh.

https://doi.org/10.1017/S095267579700328X Published online by Cambridge University Press

Browman, C. & L. Goldstein (1989). Articulatory gestures as phonological units. *Phonology* **6**. 201–251.
Charniak, E. (1993). *Statistical language learning*. Cambridge, Mass.: MIT Press.
Chomsky, N. (1973). Conditions on transformations. In S. Anderson & P. Kiparsky (eds.) *A Festschrift for Morris Halle*. New York: Holt, Rinehart & Winston. 232–286.
Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.
Clements, G. N. & J. Goldsmith (eds.) (1984). *Autosegmental studies in Bantu tone*. Dordrecht: Foris.
Ellison, M. (1992). *Machine learning of phonological structure*. PhD dissertation, University of Western Australia.
Gazdar, G., E. Klein, G. Pullum & I. Sag (1985). *Generalized Phrase-structure Grammar*. Oxford: Blackwell.
Goldsmith, J. (1974). An autosegmental tonology of a fragment of Igbo. Ms.
Goldsmith, J. (1976). *Autosegmental phonology*. PhD dissertation. MIT. Published 1979, New York: Garland.
Goldsmith, J. (1990). *Autosegmental and metrical phonology*. Oxford: Blackwell.
Goldsmith, J. (1991). Phonology as an intelligent system. In D. J. Napoli & J. Kegl (eds.) *Bridges between psychology and linguistics: a Swarthmore Festschrift for Lila Gleitman*. Hillsdale, NJ: Erlbaum. 247–267.
Kenstowicz, M. (1994). *Phonology in generative grammar*. Oxford: Blackwell.
Kornai, A. (1995). *Formal phonology*. New York: Garland.
Lakoff, G. (1993). Cognitive phonology. In J. Goldsmith (ed.) *The last phonological rule*. Chicago: University of Chicago Press. 117–145.
Massamba, D. (1984). Tone in Ci-Ruri. In Clements & Goldsmith (1984). 235–254.
Odden, D. (1990). Tone in the Makonde dialects: Chimaraba. *Studies in African Linguistics* **21**. 61–105.
Port, R. & P. Crawford (1989). Incomplete neutralization and pragmatics in German. *JPh* **18**. 257–282.
Rabiner, L. A. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the Institute of Electrical and Electronics Engineers* **77**. 257–285.
Sagey, E. (1988). On the ill-formedness of crossing association lines. *LI* **19**. 109–118.
Siegel, D. (1974). *Topics in English morphology*. PhD dissertation. MIT.

https://doi.org/10.1017/S095267579700328X Published online by Cambridge University Press