On the Acquisition of Phonological Representations

B. Elan DRESHER
Department of Linguistics
University of Toronto
Toronto, Ontario
Canada M5S 3H1
dresher@chass.utoronto.ca

Abstract

Language learners must acquire the grammar (rules, constraints, principles) of their language as well as representations at various levels. I will argue that representations are part of the grammar and must be acquired together with other aspects of grammar; thus, grammar acquisition may not presuppose knowledge of representations. Further, I will argue that the goal of a learning model should not be to try to match or approximate target forms directly, because strategies to do so are defeated by the disconnect between principles of grammar and the effects they produce. Rather, learners should use target forms as evidence bearing on the selection of the correct grammar. I will draw on two areas of phonology to illustrate these arguments. The first is the grammar of stress, or metrical phonology, which has received much attention in the learning model literature. The second concerns the acquisition of phonological features and contrasts. This aspect of acquisition turns out, contrary to first appearances, to pose challenging problems for learning models.

1 Introduction

I will discuss the extent to which representations are intertwined with the grammar, and consequences of this fact for acquisition models. I will focus on phonological representations, but the argument extends to other components of the grammar.

One might suppose that phonological representations can be acquired directly from the acoustic signal. If, for example, children are equipped with innate phonetic feature detectors, one might suppose that they can use these to extract phonetic features from the signal. These extracted phonetic features would then constitute phonological representations (*surface*, or *phonetic, representations*). Once these are acquired, they can serve as a basis from which learners can acquire the rest of the grammar, namely, the phonological rules (and/or constraints) and the *lexical*, or *underlying, representations*.

This idea of acquisition by stages, with representations preceding rules, has enduring appeal, though details vary with the prevailing theory of grammar; versions of this theory can be found in (Bloch, 1941) and (Pinker, 1994:264–5). The idea could not be implemented in American Structuralist phonology, however (Chomsky, 1964), and I will argue that it remains untenable today. I will discuss two areas of phonology in which representations must be acquired together with the grammar, rather than prior to it. The first concerns the grammar of stress, or metrical phonology. The second concerns the acquisition of phonological features. These pose different sorts of problems for learning models. The first has been the subject of considerable discussion. The second, to my knowledge, has not been discussed in the context of formal learning models. Though it has often been assumed, as mentioned above, that acquisition of features might be the most straightforward aspect of phonological acquisition, I will argue that it presents challenging problems for learning models.

2 Representations of stress

Phonetic representations are not simply bundles of features. Consider stress, for example. Depending on the language, stress may be indicated phonetically by pitch, duration, loudness, or by some combination of these dimensions. So even language learners gifted with phonetic feature detectors will have to sort out what the specific correlates of stress are in their language. For purposes of the ensuing discussion, I will assume that this much can be acquired prior to further acquisition of the phonology.

But simply deciding which syllables have stress does not yield a surface representation of the stress contour of a word. According to metrical theory (Liberman and Prince 1977, Halle and Idsardi 1995, Hayes 1995), stress results from grouping syllables into feet; the strongest foot is assigned the main stress, the other feet are associated with secondary stress. Moreover, some syllables at the edges of the stress domain may be designated as extrametrical, and not included in feet.

For example, I assume that learners who have sorted out which acoustic cues signal stress can at some point assign the stress contours depicted in (1) to English words. The height of the column over each syllable, *S*, indicates how much relative stress it has. However, these are not the surface representations. They indicate levels of stress, but no metrical organization.

(1) Representations of stress contours before setting metrical parameters
a. *América* b. *Mànitóba*

```
  x            x      Line 2
  x        x   x      Line 1
x x x x    x x x  x   Line 0
S S S S    S S S  S
America    Manito:ba
```

According to conventional accounts of English stress, the metrical structures assigned to these words are as in (2).

(2) Acquired representations
a. *América* b. *Mànitóba*

```
  x              x      Line 2
 (x)       (x    x)     Line 1
x(x x)<x>  (x x)(x)<x>  Line 0
L L L  L    L L  H  L
Ameri ca    Mani to:ba
```

Looking at the word *America*, these representations indicate that the first syllable *A* is unfooted, that the next two syllables *meri* constitute a trochaic foot, and that the final syllable *ca* is extrametrical. *Manitoba* has two feet, hence two stresses, of which the second is stronger than the first. The *L*s and *H*s under the first line of the metrical grid designate light and heavy syllables, respectively. The distinction is important in English: The syllable *to:* in *Manitoba* is heavy, hence capable of making up a foot by itself, and it receives the stress. If it were light, then *Manitoba* would have stress on the antepenultimate syllable, as in *America*.

How does a learner know to assign these surface structures? Not just from the acoustic signal, or from the schematic stress contours in (1). Observe that an unstressed syllable can have several metrical representations: it can be footed, like the first syllable in *America*; it can be the weak position of a foot, like the second syllable of *Manitoba*; or it can be extrametrical, like the final syllables in both words. One cannot tell from the sound which of these representations to assign. The only way to know this is to acquire the grammar of stress, based on evidence drawn from the observed contours in (1).

Similar remarks hold for determining syllable quantity. English divides syllables into light and heavy: a light syllable ends in a short vowel, and a heavy syllable contains either a long vowel or is closed by a consonant. In many other languages, though, a closed syllable containing a short vowel is considered to be light, contrary to the English categorization. Learners must decide how to classify such syllables, and the decision cannot be made on phonetic grounds alone.

3 Acquisition of metrical structure

How, then, are these aspects of phonological structure acquired? Following Chomsky (1981), I will suppose that metrical structures are governed by a finite number of parameters, whose value is to be set on the basis of experience. The possible values of a parameter are limited and given in advance.[1]

Parameter setting models must overcome a basic problem: the relation between a parameter and what it does is indirect, due to the fact that there are many parameters, and they interact in complex ways (Dresher and Kaye, 1990). For example, in English main stress is tied to the right edge of the word. But that does not mean that stress is always on the last syllable: it could be on the penultimate syllable, as in *Manitoba*, or on the antepenultimate, as in *America*. What is consistent in these examples is that main stress devolves onto the strong syllable of the rightmost foot. Where this syllable and foot is in any given word depends on how a variety of parameters are set. Some surprising consequences follow from the nontransparent relationship between a parameter and its effects.

The first one is that a learner who has some incorrectly set parameters might know that something is wrong, but might not know which parameter is the source of the problem. This is known as the *Credit Problem* (cf. Clark 1989, 1992, who calls this the *Selection Problem*): a learner cannot reliably assign credit or blame to individual parameters when something is wrong.

There is a second way in which parameters can pose problems to a learner. Some parameters are stated in terms of abstract entities and theory-internal concepts that the learner may not initially be able to identify. For example, the theory of stress is couched in terms of concepts such as heavy syllables, heads, feet, and so on. In syntax, various parameters have been posited that refer specifically to anaphors, or to functional projections of various types. These entities do not come labelled as such in the input, but must themselves be constructed by the learner. So, to echo the title character in Plato's dialogue *The Meno*, how can learners determine if main stress falls on the first or last foot if they do not know what a foot is, or how to identify one? This can be called the *Epistemological Problem*: in this case we know about something in the abstract, but we do not recognize that thing when it is front of us.

Because of the Credit Problem and the Epistemological Problem, parameter setting is not like learning to hit a target, where one can correct one's aim by observing where previous shots land. The relation between number of parameters correct and apparent closeness to the target is not *smooth* (Turkel, 1996): one parameter wrong may result in forms that appear to be way off the target, whereas many parameters wrong may produce results that appear to be better (Dresher, 1999). This discrepancy between grammar and outputs defeats learning models that blindly try to match output forms (Gibson and Wexler, 1994), or that are based on a notion of goodness-of-fit (Clark and Roberts, 1993). In terms of Fodor (1998), there are no unambiguous triggers: thus, learning models that seek them in individual target forms are unlikely to be successful.

I have argued (Dresher, 1999) that Plato's solution – a series of questions posed in a specified order – is the best approach we have. One version of this approach is the *cue-based* learner of (Dresher and Kaye, 1990). In this model, not only are the principles and parameters of Universal Grammar innate, but learners must be born with some kind of a road map that guides them in setting the parameters. Some ingredients of this road map are the following:

First, Universal Grammar associates every parameter with a *cue*, something in the data

[1] For some other approaches to the acquisition of stress see (Daelemans Gillis and Durieux, 1994), (Gupta and Touretzky, 1994), (Tesar, 1998, 2004), and (Tesar and Smolensky, 1998).

that signals the learner how that parameter is to be set. The cue might be a pattern that the learner must look for, or simply the presence of some element in a particular context.

Second, parameter setting proceeds in a (partial) order set by Universal Grammar: this ordering specifies a *learning path* (Lightfoot 1989). The setting of a parameter later on the learning path depends on the results of earlier ones.

Hence, cues can become increasingly abstract and grammar-internal the further along the learning path they are. As learners acquire more of the system, their representations become more sophisticated, and they are able to build on what they have already learned to set more parameters.[2]

If this approach is correct, there is no parameter-independent learning algorithm. This is because the learning path is dependent on the particular parameters. Also, the cues must be discovered for each parameter. Thus, a learning algorithm for one part of the grammar cannot be applied to another part of the grammar in an automatic way.[3]

4. Segmental representations

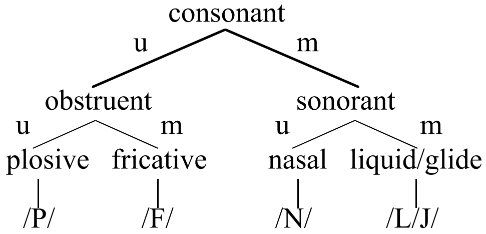
Up to now we have been looking at an aspect of phonological representation above the level of the segment. I have argued that acquisition of this aspect of surface phonological representation cannot simply be based on attending to the acoustic signal, but requires a more elaborate learning model. But what about acquisition of the phonemic inventory of a language? One might suppose that this be achieved prior to the acquisition of the phonology itself.

Since the pioneering work of Trubetzkoy and Jakobson, phonological theory has posited that phonemes are characterized in terms of a limited set of *distinctive features*. Therefore, to identify a phoneme one must be able to assign to it a representation in terms of feature specifications. What are these representations? Since Saussure, it has been a central assumption of much linguistic theory that a unit is defined not only in terms of its substance, but also in negative terms, with respect to the units it contrasts with. On this way of thinking, an /i/ that is part of a three-vowel system /i a u/ is not necessarily the same thing as an /i/ that is part of a seven-vowel system /i ɪ e a o ʊ u/. In a three-vowel system, no more than two features are required to distinguish each vowel from all the others; in a seven-vowel system, at least one more feature is required.

Jakobson and Halle (1956) suggested that distinctive features are necessarily binary because of how they are *acquired*, through a series of ‘binary fissions’. They propose that the order of these contrastive splits, which form what I will call a *contrastive hierarchy* (Dresher 2003a, b) is partially fixed, thereby allowing for certain developmental sequences and ruling out others. This idea has been fruitfully applied in acquisition studies, where it is a natural way of describing developing phonological inventories (Pye Ingram and List, 1987), (Ingram, 1989), (Levelt, 1989), (Dinnsen et al., 1990), (Dinnsen, 1992), and (Rice and Avery, 1995).

Consider, for example, the development of segment types in onset position in Dutch (Fikkert, 1994):

(3) Development of Dutch onset consonants (Fikkert 1994)



At first there are no contrasts. The value of the consonant defaults to the least marked (*u*) onset, namely an obstruent plosive, desig-

[2]For details of parameter ordering, defaults, and cues in the acquisition of stress, see (Dresher and Kaye, 1990) and (Dresher, 1999).

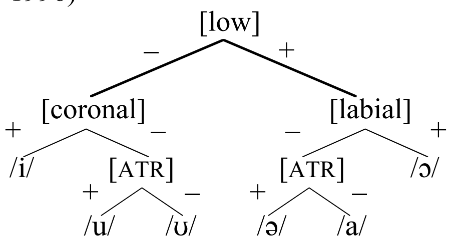
[3] For further discussion and critiques of cue-based models see (Nyberg, 1991), (Gillis Durieux and Daelemans, 1995), (Bertolo et al. 1997), and (Tesar, 2004).

nated here as /P/. The first contrast is between obstruent and sonorant. The former remains the unmarked (*u*), or default, option; the marked (*m*) sonorant defaults to nasal, /N/. At this point children differ. Some expand the obstruent branch first, bringing in marked fricatives, /F/, in contrast with plosives. Others expand the sonorant branch, introducing marked sonorants, which may be either liquids, /L/, or glides, /J/. Continuing in this way we will eventually have a tree that gives all and only the contrasting features in the language.

5. Acquiring segmental representations

Let us consider how such representations might be acquired. To illustrate, we will look at the vowel system of Classical Manchu (Zhang, 1996), which nicely illustrates the types of problems a learning model will have to overcome. Zhang (1996) proposes the contrastive hierarchy in (4) for Classical Manchu, where the order of the features is [low]> [coronal]>[labial]>[ATR].

(4) Classical Manchu vowel system (Zhang 1996)[4]



Part of the evidence for these specifications comes from the following observations:

(5) Evidence for the specifications in (4)
  a. /u/ and /ə/ trigger ATR harmony, but /i/ does not, though /i/ is phonetically [+ATR], suggesting that /i/ lacks a phonological specification for [ATR].
  b. /ɔ/ triggers labial harmony, but /u/ and /ʊ/ do not. Though phonetically [+labial], there is no evidence that /u/ and /ʊ/ are specified for this feature.

Acquiring phonological specifications is not the same as identifying phonetic features. Surface phonetics do not determine the phonological specifications of a segment. Manchu /i/ is phonetically [+ATR], but does not bear the feature phonologically; /u/ and /ʊ/ are phonetically [+labial], but are not specified for that feature. How does a learner deduce phonological (contrastive) specifications from surface phonetics?[5]

It must be the case that phoneme acquisition requires learners to take into account phonological processes, and not just the local phonetics of individual segments (Dresher and van der Hulst, 1995). Thus, the phonological status of Manchu vowels is demonstrated most clearly by attending to the effects of the vowel on neighbouring segments.

This conclusion is strengthened when we consider that the distinction between /u/ and /ʊ/ in Classical Manchu is phonetically evident only after back consonants; elsewhere, they merge to [u]. To determine the underlying identity of a surface [u], therefore, a language learner must observe its patterning with other vowels: if it co-occurs with [+ATR] vowels, it is /u/; otherwise, it is /ʊ/. The nonlocal and diverse character of the evidence bearing on the feature specifications of segments poses a challenge to learning models.

Finally, let us consider the acquisition of the hierarchy of contrastive features in each language. Examples such as the acquisition of Dutch onsets given above appear to accord well with the notion of a learning path, whereby learners proceed to master individual feature contrasts in order. If this order were the same for all languages, then this

[4]Zhang (1996) assumes privative features: [F] vs. the absence of [F], rather than [+F] vs. [–F]. The distinction between privative and binary features is not crucial to the matters under discussion here.

[5]Phonological contrasts that play a role in phonological representations are thus different from their phonetic manifestations, the subject of studies such as (Flemming, 1995).

much would not have to be acquired. However, it appears that the feature hierarchies vary somewhat across languages (Dresher, 2003a, b). The existence of variation raises the question of how learners determine the order for their language. The problem is difficult, because establishing the correct ordering, as shown by the active contrasts in a language, appears to involve different kinds of potentially conflicting evidence. In the case of metrical parameters, the relevant evidence could be reduced to particular cues, or so it appears. Whether the setting of feature hierarchies can be parameterized in a similar way remains to be demonstrated.

6 Conclusion

I will conclude by raising one further problem for learning models that is suggested by the Manchu vowel system. We have observed that in Classical Manchu, /ə/ is the [+ATR] counterpart of /a/. Both vowels are [+low]. Since [low] is ordered first among the vowel features in the Manchu hierarchy, we might suppose that learners determine which vowels are [+low] and which are not at an early stage in the process, before assigning the other features. However, a vowel that is phonetically [ə] is ambiguous as to its featural classification. In many languages, including descendants of Classical Manchu (Zhang, 1996, Dresher & Zhang, 2003) such vowels are classified as [–low]. What helps to place /ə/ as a [+low] vowel in Classical Manchu is the knowledge that it is the [+ATR] counterpart of /a/. That is, in order to assign the feature [+low] to /ə/, it helps to know that it is [+ATR]. But, by hypothesis, [low] is assigned before [ATR]. Similarly, the determination that /i/ is contrastively [+coronal] is tied in with its not being contrastively [–labial]; but [coronal] is assigned prior to [labial].

It appears, then, that whatever order we choose to assign features, it is necessary to have some advance knowledge about classification with respect to features ordered later. Perhaps this paradox is only apparent. However it is resolved, the issue raises an interesting problem for models of acquisition.

7 Acknowledgements

This research was supported in part by grant 410-2003-0913 from the Social Sciences and Humanities Research Council of Canada. I would like to thank the members of the project on Contrast in Phonology at the University of Toronto (http://www.chass.utoronto.ca/~contrast/) for discussion.

References

Stefano Bertolo Kevin Broihir Edward Gibson and Kenneth Wexler. 1997. Characterizing learnability conditions for cue-based learners in parametric language systems. In Tilman Becker and Hans-Ulrich Krieger, editors, *Proceedings of the Fifth Meeting on the Mathematics of Language*. http://www.dfki.de/events/ mol/.

Bernard Bloch. 1941. Phonemic overlapping. *American Speech* 16:278–284. Reprinted in Martin Joos, editor, *Readings in Linguistics I*, Second edition, 93–96. New York: American Council of Learned Societies, 1958.

Noam Chomsky. 1964. Current issues in linguistic theory. In Jerry A. Fodor and Jerrold J. Katz, editors, *The Structure of Language*, 50–118. Englewood Cliffs, NJ: Prentice-Hall.

Noam Chomsky. 1981. Principles and parameters in syntactic theory. In Norbert Hornstein and David Lightfoot, editors, *Explanation In Linguistics: The Logical Problem of Language Acquisition*, 32–75. London: Longman.

Robin Clark. 1989. On the relationship between the input data and parameter setting. In *Proceedings of NELS 19*, 48–62. GLSA, University of Massachusetts, Amherst.

Robin Clark. 1992. The selection of syntactic knowledge. *Language Acquisition* 2:83–149.

Robin Clark and Ian Roberts. 1993. A computational model of language learnability and language change. *Linguistic Inquiry* 24:299–345.

Walter Daelemans Steven Gillis and Gert Durieux. 1994. The acquisition of stress: A data-oriented approach. *Computational Linguistics* 20:421–451.

Daniel A. Dinnsen. 1992. Variation in developing and fully developed phonetic inventories. In Charles Ferguson Lise Menn and Carol Stoel-Gammon, editors, *Phonological Development: Models, Research, Implications*, 191–210. Timonium, MD: York Press,.

Daniel A. Dinnsen Steven B. Chin Mary Elbert and Thomas W. Powell. 1990. Some constraints on functionally disordered phonologies: Phonetic inventories and phonotactics. *Journal of Speech and Hearing Research* 33:28–37.

B. Elan Dresher. 1999. Charting the learning path: Cues to parameter setting. *Linguistic Inquiry* 30:27–67.

B. Elan Dresher. 2003a. Contrast and asymmetries in inventories. In Anna-Maria di Sciullo, editor, *Asymmetry in Grammar, Volume 2: Morphology, Phonology, Acquisition*, 239–257. Amsterdam: John Benjamins.

B. Elan Dresher. 2003b. The contrastive hierarchy in phonology. In Daniel Currie Hall, editor, *Toronto Working Papers in Linguistics (Special Issue on Contrast in Phonology) 20*, 47–62. Toronto: Department of Linguistics, University of Toronto.

B. Elan Dresher and Harry van der Hulst. 1995. Global determinacy and learnability in phonology. In John Archibald, editor, *Phonological Acquisition and Phonological Theory,* 1–21. Hillsdale, NJ: Lawrence Erlbaum.

B. Elan Dresher and Jonathan Kaye. 1990. A computational learning model for metrical phonology. *Cognition* 34:137–195.

B. Elan Dresher and Xi Zhang. 2003. Phonological contrast and phonetics in Manchu vowel systems. Paper presented at the Twenty-Ninth Annual Meeting of the Berkeley Linguistics Society, February 2003. To appear in the Proceedings.

Paula Fikkert. 1994. *On the Acquisition of Prosodic Structure (HIL Dissertations 6).* Dordrecht: ICG Printing.

Edward Flemming. 1995. Auditory representations in phonology. Doctoral dissertation, UCLA.

Janet Dean Fodor. 1998. Unambiguous triggers. *Linguistic Inquiry* 29:1–36.

Edward Gibson and Kenneth Wexler. 1994. Triggers. *Linguistic Inquiry* 25:407–454.

Steven Gillis Gert Durieux and Walter Daelemans. 1996. A computational model of P&P: Dresher & Kaye (1990) revisited. In Frank Wijnen and Maaike Verrips, editors, *Approaches to Parameter Setting*, 135–173. Vakgroep Algemene Taalwetenschap, Universiteit van Amsterdam.

Prahlad Gupta and David Touretzky. 1994. Connectionist models and linguistic theory: Investigations of stress systems in language. *Cognitive Science* 18:1–50.

Morris Halle and William J. Idsardi. 1995. General properties of stress and metrical structure. In John Goldsmith, editor, *The Handbook of Phonological Theory*, 403–443. Cambridge, MA: Blackwell.

Bruce Hayes. 1995. *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago Press.

David Ingram. 1989. *First Language Acquisition: Method, Description and Explanation.* Cambridge: Cambridge University Press.

Roman Jakobson and Morris Halle. 1956. *Fundamentals of Language.* The Hague: Mouton.

Clara C. Levelt. 1989. An essay on child phonology. M.A. thesis, Leiden University.

Mark Liberman and Alan Prince. 1977. On stress and linguistic rhythm. *Linguistic Inquiry* 8:249–336.

David Lightfoot. 1989. The child's trigger experience: Degree-0 learnability (with commentaries). *Behavioral and Brain Sciences* 12:321–375.

Eric H. Nyberg 3rd. 1991. A non-deterministic, success-driven model of parametric setting in language acquisition. Doctoral dissertation, Carnegie Mellon University, Pittsburgh, PA.

Steven Pinker. 1994. *The Language Instinct*. New York: William Morrow.

Plato. *Meno*. Various editions.

Clifton Pye David Ingram and Helen List. 1987. A comparison of initial consonant acquisition in English and Quiché. In Keith E. Nelson and Ann Van Kleeck, editors, *Children's Language (Vol. 6)*, 175–190. Hillsdale, NJ: Erlbaum.

Keren Rice and Peter Avery. 1995. Variability in a deterministic model of language acquisition: A theory of segmental elaboration. In John Archibald editor, *Phonological Acquisition and Phonological Theory*, 23–42. Hillsdale, NJ: Lawrence Erlbaum.

Bruce Tesar. 1998. An iterative strategy for language learning. *Lingua* 104:131–145.

Bruce Tesar. 2004. Using inconsistency detection to overcome structural ambiguity. *Linguistic Inquiry* 35:219–253.

Bruce Tesar and Paul Smolensky. 1998. Learnability in Optimality Theory. *Linguistic Inquiry* 29:229–268.

William J. Turkel. 1996. Smoothness in a parametric subspace. Ms., University of British Columbia, Vancouver.

Xi Zhang. 1996. Vowel systems of the Manchu-Tungus languages of China. Doctoral dissertation, University of Toronto.