Blackwell
Reference Online

22. Phonological Acquisition

MARLYS A. MACKEN

Subject	Theoretical Linguistics » Pholonogy
DOI:	10.1111/b.9780631201267.1996.00024.x

0 Introduction

Over the years, Chomsky has encouraged us to ask, how is it that human beings, whose contacts with the world are brief, personal, and limited are nevertheless able to know as much as they do? How does the child come to master a complex, abstract system like language when the evidence available to the child is so sparse? This question is one of the great scientific puzzles of our time, and phonology is only one of the many disciplines that have attacked the problem, but found the answer elusive. The field is large, inter-disciplinary, and relatively young. Each discipline generates its own theories, ideologies, and heated debates, but this theoretical diversity is warranted by the complexity of the puzzle, for language is indeed complex, and many descriptions are possible for each phenomenon. Language acquisition data are at times like Rorschach ink blots. Phoneticians see in a CVCV transcription (already multiply removed from the original object) evidence for articulatory primitives, gestures, or mandibular open-close frames, while phonologists look at the same ink configuration, see abstract units, and debate about syllables, moras, features, nodes, and feet. Inherent descriptive equivalence or indeterminancy is compounded because of the dual nature of a representation that is input to the two different systems of phonology and phonetics.

To the descriptive problems, one must add the complication that the child is also complex and changing at a remarkable rate. Studying the developing child presents many challenges and requires dealing with a large number of correlations that can be misinterpreted. Yet a theory of language acquisition is incomplete without a theory of cognition and learning that is compatible with what we know about the mind and human development.

Finally, one must factor in the child's world. Is a particular structure observed because language possesses that form, because the mind does, or because the world is structured that way? Language, mind, environment: these three aspects of our puzzle make “explanations” easy to come by and shortlived. The problem is compounded by the goal of separating the structure of language from the way it is used and acquired. As the mind is simultaneously both structure and process, our data conflate the two, but we must attempt the difficult, perhaps impossible, task of separating the grammar from the processor, declarative from procedural knowledge, propositional content from images. Different fields have different versions of this problem: in phonology, there is no grammaticality judgement technique to separate the underlying grammar from other levels of description or other systemic effects.

Much of the acquisition field is predictably, then, a vigorous debate over what the child knows and in what format the knowledge is represented. Does the infant who distinguishes two stimuli “know” the syllables [ba] and [da] or the segments [b], [d], or [a]? Does the child who says “wanna go” have an optional subject rule or know a particular kind of verb complement structure? How, we must ask, do these questions differ from the question of what a thermostat “knows” about temperature and raising / lowering? And does language structure exist separately from the time, acquisitional distance traveled, and situations of its use?

Unfortunately, some divisive issues, like innateness, continue to impede progress. The field is still polarized between empiricists, who tend to be phoneticians and / or general learning theorists (usually today connectionists), and rationalists, who tend to be formal phonologists or cognitive psychologists. To show that phonetics and learning are sufficient to account for phonology, the former theorists concentrate on developmental continuity from infancy through the transition to speech. In contrast, the latter theorists focus on ages two to four (and up), on the properties of phonologically more complex systems, on developmental discontinuities from earlier phonetic stages, and on nonlinearities and other evidence of reorganization in terms of abstract units or rules.[1] This polarization, the legacy of the Chomsky wars, still shackles the field, pointlessly dividing some groups of people and obscuring two indisputable points. First, phonology subsumes phonetics. Articulatory and perceptual systems play key roles in acquisition, while part of phonology is an abstract and semiformal system with objects, constraints, and principles not fully determined by phonetic content nor fully explained by phonetic theory. Second, some aspects of language are learned and some are innate. The question of how learning is accomplished in the presence of incomplete and contradictory input is still the central question to be asked, and the answer lies in part on the a priori structures that determine the speed of acquisition, constraints on variation, and the independence from limiting factors like intelligence. Empirical evidence showing the need for innate, domain-specific structure has also come from a variety of other sources, like the failure of animal language-learning and the structure and acquisition of sign languages, etc. Empirical evidence showing the significant role of learning includes individual variation and cross-language differences; research using computational modeling of acquisition phenomena provides intriguing support for some empiricist claims.

The challenge for us is to partition the domains properly between phonology and phonetics, determine the interplay between learning and innate constraints, and separate general learning from domain-specific linguistic process and structure. The multiple aspects of language acquisition are complex and indeed very apparent today, yet that variation falls within strict limits. It is the task of the theory to explain both the freedom that the variation documents and the constraints imposed by the innate structure. Though there is controversy over each point, the evidence to date shows that acquisition constraints are in the mind, not in the world, and that they are specifically linguistic, explicitly represented structures.

Granted that acquisition is a very hard problem for the scientist to study and resolve, why is it an important problem? First, studying real-time acquisition will provide the clearest answers to the central questions on language variation constraints and the interaction between innate structure and the environment. Second, language acquisition data can directly influence the theory of phonological structure. For many cognitive systems, the end state necessarily is constrained by the way in which it is acquired: earliest phonological capacities structure what we learn, thus setting boundaries for what is learned and perhaps leaving an imprint on the phonological knowledge acquired during the final stages as well; in addition, these earliest capacities remain basic to the adult phonological system and provide one of the clearest windows on the core structure of phonology.

This paper examines (1) the structure that is acquired and (2) the relationship between acquisition and theory. I shall argue that the capacity of children and adults is the same - the strong identity thesis - and that phonological principles explain variation among children and particular differences between children and adults. Thus, acquisition data can provide direct answers to certain core phonological issues, and any phonological theory that fails in principle to account for acquisition data fails as a theory of phonology. Given the focus (and length) of this chapter, much of the literature - which is overwhelmingly descriptive - is not covered here. For example, there is much sophisticated research on infant perception and the transition from babbling to speech that shows, respectively, the innate status of phonetic features (e.g., Kuhl 1987) and the infant's sensitivity within the first year to specific properties of the environmental language (e.g., Boysson-Bardies 1993). For an overview of the descriptive literature, acquisition data, stages and acquisition theories, see Ingram (1989), while Smith (1973) remains the best theoretical study of phonological acquisition. For representative research, one may see the papers in Yeni-Komshian, Kavanaugh, and Ferguson (1980) and Ferguson, Menn, and Stoel-Gammon (1992). The model of phonological acquisition presented here follows on the work of Kiparsky and Menn (1977b), Macken and Ferguson (1983), and others, notably Smith (1973). For alternative views, the reader may look to evolutionary, self-organizing theory (e.g., Lindblom 1992; MacNeilage and Davis 1993), performance theories (e.g., Stemberger 1992a), natural phonology theory (Stampe 1969; Edwards and Shriberg 1983), and Firthian phonology (Waterson 1987).

1 Nature of the Relationship

Children begin saying first words around 12 to 18 months of age. Early on, they may use long, prosodically sentential jargon or invented, idiosyncratic protowords not clearly based on words of their language, but generally, for at least this first year, the form of words in their native language is reduced, highly restricted, and somewhat variable. For the next several years, their speech continues to differ substantially from the speech of adults. Assuming that the speech of adults reflects a uniform underlying grammar, we may ask how the speech of children is related to the language of their parents. In nature, we find two contrasting developmental relations. In a relation manifesting essential continuity, the young are unskilled and simpler, yet they are fundamentally like the adults of the species in key respects; in a qualitatively different kind of developmental relation (which we might call “nonlinear”), there is a radical difference between the beginning and end states and a major discontinuity in development.

Children acquiring phonology do change over time, going through a recognizable set of stages. This progression is an indication, along with independence from the limiting effects of ability and environment and the presence of a critical period, that shows that language acquisition is a biologically controlled behavior. We are interested here in the nature of those changes: are they qualitative, in the sense that the basic structures or capacities change, or quantitative, in the sense that the information or knowledge of a specific domain changes. If the principles and objects of phonology are present at the outset of language learning, and thus instantiated at each stage and in each interim grammar constructed by the learner as in Chomsky's theory, then the developmental model is one of basic continuity. We would then look to nonqualitative factors to explain the developmental stages. If, on the other hand, some phonological principles or objects are not present at the outset, then there is no necessary relationship between a developmental stage of the child and the properties of phonological systems: the developmental model will then be one of discontinuity, and we will explain the qualitative characteristics of each stage in terms of the maturation of new linguistic skills or changes in other cognitive capacities, as presented in Piaget's theories.

Roman jakobson's hallmark monograph (1968) provides both answers to our question about the relationship between the child's stage / grammar and the adult's. The central and unifying claim of Jakobson's theory is that the speech of children from first words on is both *simpler* in a highly principled way and the *same* as the universal structure that underlies the language of their parents. Yet he also believed that there was a categorical difference between the babbling stage of the first year of life and the true language stage that begins in the second year of life. The hypothesized discontinuity – as striking as, say, a tadpole-frog discontinuity – was predicted to be accompanied in some cases by a silent period. On the first point, Jakobson's elegant theory is correct in substance if not always in detail.

With respect to the second point, however, there is considerable evidence that there is a close connection between the specific features of babbling and of first words and that babbling and word use are not discrete stages but overlap in time as well as in substance. These facts show that there is no major discontinuity between the so-called prelinguistic babbling period and the presumably phonological stages that begin in the second year. This has been taken by many to refute Jakobson's theories concerning phonology, as if by demonstrating the absence of a discontinuity we had proved there was no phonology, a view advanced by the “it's all tadpoles/phonetics” school. The pendulum has swung back somewhat in the direction of phonology with recent research that shows that, in the last six months of the first year, children's babbling takes on segmental and prosodic characteristics of the surrounding language and that their perceptual systems attune to the specific language of their communicative environment with the loss of the ability to discriminate noncontrastive differences by eight to ten months of age. These findings undermine the anti-Jakobsonian school's argument that the absence of a discontinuity between babbling and speech (early in the second year) shows the primacy of phonetics to the phonological systems of the second year. Rather, these findings show that specific learning about the child's own language is strongly affecting the supposedly “phonetics-only” stage of babbling and discrimination in the first year. Clearly, the label "prelinguistic" is an inaccurate description of the abilities of the first year. Equally clear is that there is a complex relationship between this stage and that of the second year, just as there is in general a complex relationship between phonetics and phonology in all aspects of speech. Let us return to the question of how the stages of language learning from the second year on are related to the end state.

1.1 The Strong Identity Thesis

Jakobson argued that one universal phonological system determines the structure of synchronic languages and the supposedly "extralinguistic" domains of acquisition, sound change and disordered speech. In advancing this strong identity relationship between the child and adult phonological capacity, Jakobson is one of many phonologists who have taken the same thesis to argue various issues in phonological theory – Schleicher, Paul, Ament, Grammont, Meillet, Jespersen, Saussure, Baudouin de Courtenay, Halle, Kiparsky, Stampe, and Ohala. Like Grammont (1902, 1933), many have viewed phonological acquisition as a microcosm of diachronic sound change. Ohala and others have observed the same acoustically-motivated rules in acquisition and sound change and taken this to show that both children and adults create such sound patterns independently because they possess the same physical phonetic apparatus (e.g., Greenlee and Ohala 1980). Generative theorists argue that the parallels between acquisition and diachrony are due to the same shared phonological system and that children may be the actual source of sound change (cf. 19th and early 20th century phonologists like Paul and Saussure). For example, Stampe (1969, 1972) argued there are universal processes and that change in a language occurs when processes are not correctly limited during acquisition. (See the discussion by Paul Kiparsky in chapter 21 of this volume for a related discussion.

In Jakobson's theory, features are the central unit of phonology: there is a small universal set of binary features that function to differentiate elements in natural language; the feature system is hierarchical and implicational; and this system constrains phonological inventories, systems, and rules. The same structural principles that determine this invariant hierarchy in phonemic systems also determine sound change and an invariant acquisition order of sound classes. The patterns of stratification, change, and acquisition derive ultimately from the principle of maximal contrast along acoustic axes of sonority and tonality: maximal contrasts are found in all languages and acquired first. Implicational relationships govern features, such that if Y occurs in a phonemic system, then X does too; and X is acquired by children before Y. The child's acquisition of feature oppositions proceeds through the universal feature hierarchy from the most general contrast to the finest, rarest contrasts. Furthermore, the relative frequency, combinatorial capacity, and assimilatory power of particular features once acquired, and the substitution patterns within each stage of the child's development are also determined by the priority relationships within the universal feature hierarchy (1968, p. 58).

To exemplify the specific proposals (since the theory is well – described in many places, e.g., Anderson 1985; Ingram 1989), we will look briefly at consonants. The stages should be (1) optimal consonant /p/ versus optimal vowel /a/, (2) / p:m/, (3) / p:t/ and (4) /m:n/, yielding the basic consonantal system /p, t, m, n/. Other predictions include that stop consonants are acquired before fricatives; front before back; voiceless unaspirated before voiced; fricatives before affricates; strident fricatives (/f, s/) before the corresponding mellow fricatives (/ø, θ); in early stages, fricatives and affricates are replaced by stops of the same place, voiced by voiceless unaspirated, continuants by noncontinuants; liquids as a class are acquired late with one, usually /1/, possibly early; dentals have a natural priority after stages (1)–(4).

Data from a wide variety of languages have been shown to conform to the general Jakobsonian outline of development – for example English (the classic study of Leopold 1947 and many others), French, Norwegian, Spanish, Greek; as well as Jakobson's source languages (then and since), Swedish, Danish, Russian, Serbo-Croatian, Polish, Czech, Bulgarian, and Zuni. Although many of the specific predictions have been shown to have exceptions in at least one child (particularly the precise order of stages (1) – (4) above), the general markedness relations – where [–voice] (voiceless and unaspirated), [–continuant], [+coronal][2] and "front" or [+anterior][3] are unmarked – hold for the great majority of children. In general, the unmarked consonants are acquired first (1), are most frequent in the child's lexicon (2), have the fewest restrictions on their distribution (3), and serve as replacements for the corresponding marked consonants during the stage when the constrast is neutralized (4), precisely as Jakobson predicted. As generalizations about simple inventories of segments in both children (e.g., Dinnsen 1992) and in languages of the world (e.g., Maddieson 1984), Jakobson's system of implicational relationships among features is overwhelmingly valid. While there is justifiable concern over the validity of statistical "universals" as opposed to absolute universals, the significance of this achievement should not be underestimated, thought it usually is today. From the earliest stages of learning the lexicon (and perhaps earlier), children are working within the same distinctive feature constraints that structure the phonological systems – inventories and segmental rules – of languages of the world.

Nevertheless, there are important problems with the specific predictions of the theory. The acquisition data do not clearly verify the notion that the underlying framework of oppositions is completely "contrastive" in the strict phonological sense (of phoneme minimal pair oppositions) or "acoustic" along Jakobson's sonority and tonality dimensions. While features play an important role, other units are equally (and in early stages more) important. Markedness relationships are more complicated than indicated, particularly in certain categories (e.g., "coronal/dental," where unmarked specifically means /t, s, n/ and later /1/ and not other members like /š, θ/; the category of glides is not included at all in the hierarchy or text; etc.) and in certain properties, like their supposedly greater assimilation power (Jakobson's fifth property of unmarked members): for [coronal] and possibly [voice], the more accurate observation (following Trubetzkoy's distinction) would be that the unmarked members are used by children in the earliest stages because they appear in neutralization contexts (cf. property (4) above) but, rather than having uniformly greater assimilatory power, they (may) undergo assimilation in later stages when both members of an opposition are represented; thus coronals serve as replacements for velars in the early stages but assimilate to velars in later stages (cf. 1968, p. 54). Other attested acquisition assimilatory relationships – like coronals to labials, labials to velars, and (less commonly) velars to labials – are not discussed by Jakobson. In contrast, the unmarked [-continuant] appears both to be used in neutralization contexts and to have greater assimilatory power throughout all stages. Finally, there are three basic types of unpredicted variation.

1.2 Variation

First, there are individual differences among children learning the same language. As we find in the study of early syntax (most children use single words during the one-word stage, yet some use large units only partially analyzed), so too in phonology: most children use feature-sized units fairly consistently but others work on larger, more global structures (typically the prosodic word, perhaps the syllable) and vary features considerably by prosodic position and context, where, e.g., the sequencing of place and manner features is linked to prosodically dominant (initial) and nondominant (medial or final) positions. Basically, any study of ten or more children acquiring the same language will be virtually a typological study of possible variations in structure and content. Yet there are no reports of normally developing children who produce forms outside the constraints of synchronic theory. Children with "phonological disorders" show delay and are sometimes highly idiosyncratic, but their rules too are phonetically and phonologically natural (see, e.g., the research and publications of Dinnsen, Edwards, Gandour, Grunwell, Kent, Leonard, Shriberg, Spencer, Stoel-Gammon). Children simply do not produce types of structures unattested in languages of the world.

The variation seen across different learners of the same language – like dialect variation – reveals the small number of particular parametric options within core grammar in a particularly clear way: since the basic system is in important respects the "same," potentially interacting variables are absent or inherently controlled and the system structure intrinsically clearer to the observing scientist. Acquisition data provide a particularly simple view of core phonology in part because there are few if any interactions with morphology.

Second, there are the consistent cross-linguistic differences (see, e.g., Pye, Ingram, and List 1987). For example, the palato-alveolar affricate is predicted by Jakobson to be acquired late, and this is the case in English. In Spanish, however, this affricate is acquired quite early, partly because it is more frequent in the corpus addressed to children by virtue of its high frequency in nicknames, diminutives, and sound-symbolic terms. Alternations between nasal and nonnasal stops is common in French acquisition but infrequent in other well studied languages. Laterals show substantial differences: [1] interacts with [j] in English (and retroflex [r] with [w]). In Spanish, [ð] (the spirant allophone of /d/, used intervocalically and in a few other contexts) frequently patterns with the liquids and is replaced by [1] and in some cases flap or trill [r]. The use of [1] for [ð] is common in the acquisition of Greek, but is extremely rare in the acquisition of English, the general replacement being [d] or [t]. In some languages, the acquisition alternation is [1] and [n] (e.g., Arabic, Yucatec Maya, and, with stage and learner variation, Spanish). In the case of [1], the acquisition differences are due to the intrinsic phonetic structure of [lateral], the phonetic properties of different [1]s in different languages and cross-linguistic phonological patterning differences – the same properties that underlie theoretical arguments about the feature geometric position of [lateral].

The third source of cross-linguistic acquistion differences comes from distributional regularities of the input. A consistent and originally surprising finding of the last ten years has been the at times close relationship between certain statistical properties of sounds and sound patterns in particular languages and the stages of children learning those languages. In Finnish, [d] – one of the easiest and earliest acquired obstruents in other languages – is acquired very late, which is very surprising; but phonologically and in the input, [d] is highly restricted in the adult language. In Spanish, the hierarchical dominance of labial in the adult language (Hooper 1976) underlies the predominance in Spanish acquisition of velar-to-labial harmony (which may be less common cross-linguistically than labial-to-velar harmony) and in the generalizations some children make.

The child's (and adult's) ability to extract distributional and statistical regularites must be accommodated in our theory by acknowledging that a phonological grammar has not only an abstract, symbolic algebraic system of the type proposed in current generative theories but also a statistical or stochastic component common to many connectionist and phonetic theories. To account for this and the other variation types, an acquisition model must recognize that the universal language acquisition device (LAD) is not so constrained as to result in invariant stages, as rigid order theories like Jakobson's would predict. Rather, the form of a possible phonology and a possible stage is universally constrained, but the learner has some freedom to work within this formal space to extract generalizations from the input. This constrained hypothesis formation or cognitive model (see Kiparsky and Menn 1977b and Macken and Ferguson 1983) incorporates the general acquisition patterns and universal structure envisioned by Jakobson and Chomsky, while recognizing the freedom the system must have to allow individual learners the creative flexibility they show in forming generalizations and inventing rules. Indeed, all rules may be invented or discovered by each learner. This degree of freedom may simply be a consequence of the freedom that the LAD must have in any case, since an interplay between learning and innate structure must take place to permit different languages to be acquired. The individual learner appears to have a similar degree of freedom to construct interim grammars. Why or how these grammars change over the stages of acquisition as the child's lexicon expands rapidly is a different though crucial question on which there is very little work.

1.3 Acquisition and Theory Results

The structure and variation in children's data fall within the same constraints found in the language of adults. This confirms the strong identity thesis that the same phonological system or grammar accounts for both synchronic language systems and diachronic and acquisitional change.[4] Thus, the theory of phonology can be applied to explain language acquisition data and data from children can, in principle, be used to change or confirm the theory. However, using acquistion evidence to modify theory or change theories has rarely been done. Menn (1980) cited three findings that she and other researchers thought were idiosyncratic phenomena in children (e.g., the prosodic word as a minimal unit), and yet each has independently found expression in later theoretical developments. The underlying unity of formally dissimilar rules effecting harmony and a CVCV canonical form could not be captured in the *SPE* framework of Smith (1973) (though recognized by Smith) – with the profound significance of their central role lost to the theory. Some of these problematic data can be handled by syllable theory (Spencer 1986) but could have been arguments for the syllable or other phonotactic structure template at the time. Similarly, early acquisition work identified lexical (phonotactic) patterns and associated constructs; what at the time seemed idiosyncratic to acquisition are, rather, instances of what are now called templates and planar segregation (Macken 1992a).

More often, acquisition data have been used to support (or disconfirm) aspects of phonological theory. Smith (1973) provides convincing arguments for the validity of segments, distinctive features, particular features (especially [coronal]), two levels of representation, realization rules, and some formal universals (especially rule ordering); and for the lack of evidence for other formal universals like particular abbreviatory devices (especially Greek letter variables) and *SPE* marking conventions.

Spencer (1986) reanalyzes Smith (1973) in a nonlinear framework, showing its advantages in several areas, notably in motivating underspecification and a third level of representation, in capturing the bidirectionality of lateral harmony, and in explaining the simultaneous changes in rules (e.g., labial assimilation and cluster rules) which were simply an accident in the original framework. The experimental work in Gordon (1985) provides supporting evidence for the innate status of grammatical levels and constraints on level ordering proposed in lexical phonology (Kiparsky 1982c). Bradley's work on the relationship between phonological acquisition and reading provides experimental evidence on the structure and role of subsyllabic constituents (e.g., Bradley's review article in Ferguson, Menn, and Stoel-Gammon 1992). Using acquisition data, Stemberger (1993) confirms the transparency of glottals to spreading rules and argues for the placelessness of glottals. Dresher and Kaye (1990) use learnability-theoretic computational modeling to provide a perspective on the formal properties of stress systems. Yet the richness of acquisition data for constructing phonological theory has been drawn on only minimally.

The strong identity thesis, however, does not imply that there cannot be differences between children and adults. The thesis is that there are no qualitative differences: phonological structure, features, levels, hierarchy, and constraints are available from the outset, universal then in the way basic syntactic categories and the binding principle may similarly be present. Phonology (and language generally) is like vision in these respects, where, similarly, the infant's first mechanisms for perceiving objects remains central to perception and thought, the fundamental capacities may be enhanced but not fundamentally changed, and the study of infants and children likewise helps reveal the core visual system (e.g., Spelke 1990 and Spelke et al. in press).

The rejection of a significant, qualitative stage theory of development is problematic to some theorists because it is at odds with the dominant theory of cognitive development - a significant problem if a language acquisition theory must integrate with theories of the mind and cognitive development. A discrepancy could mean several things - for example, that language is unlike other cognitive domains or, if language is like other cognitive domains, that the strong identity thesis is wrong for language or that stage theories are wrong for cognition. Let us turn, then, to Piaget, the premier stage theorist.

1.4 Stage Theory

In contrast to Kakobson and Chomsky, Piaget focused on fundamental discontinuities throughout development. His theory is that the thinking of children is qualitatively different from that of adults, that there are four stages (discontinuities) in reasoning during development, and that the reasoning anomalies at each stage will be across-the-board in all content domains. Piaget's experiments show that children of different ages “think” differently in that they give nonadult answers to questions. These task findings are robust: if a task is given as Piaget did to children from a culture like that of Piaget's children, then the results are as described. However, cross-cultural research has revealed significant differences between cultures which appear to be related to the experience children have in a particular area. For example, children from a pottery-making culture have considerably fewer problems at very young ages with conservation of matter tasks than French children do. More broadly, research of the last 15 years has challenged the foundations of Piaget's stage theory, by showing that after simplifying Piaget's tasks, children have many of the representational capacities that Piaget believed they lacked and by providing different explanations for children's performance. For example, if three years olds are drilled on different days on hundreds of pair-wise comparisons of balls of different sizes, they can then create linear orderings and make relevant deductive, transitive inferences over new comparisons where they previously could not. For both conservation and transitivity, there are differences in experience - and perhaps memory utilization differences - between children and adults, but children have the same representational capacity as adults. In general it appears that knowledge differences account for more “cognitive” differences between children and adults than do qualitative differences in, for example, representational capacity (e.g., Carey 1985). This shift to a quantitative model places cognitive development more in accord with Jakobsonian and Chomskian language development than under the traditional Piagetian view.

2 The Nature of Developmental Change

A theory of acquisition must include a theory of change, both describing the changes that occur, and then explaining them. We have found that, despite anatomical changes in the vocal tract during the second year, the words of even young children show the same basic system of features and featural relationships that structure adult languages of the world. Jakobson's original theory explained the feature hierarchy and the acquisition order in terms of phonological organization based ultimately on perception (the principle of maximal contrast along two acoustic dimensions), but the acoustic foundation to the theory has been the target of considerable criticism. Of the several possible explanations for acquisition stages, articulatory factors are primarily cited, but the evidence is usually indirect and the arguments often theory-rather than data-driven. To explore these issues, we will look at the earliest stage, where words are either one or two syllables in length, characteristically CV (CV) or CVC structures composed from a small inventory of segments and prosodically restricted to initial stress or level tones.

Even the simplest cases present analytic and theoretical issues. The child who says [du] for *juice*, as most English-speaking children are reported to say, is drawing on the basic Jakobsonian inventory and implications relationships. But we also must ask: Does the child know her word differs from the adult form? Does she know segments (as represented in the adult's transcription)? Is her word two segments long? Why two? Why CV? Why a voiced [d]? Why is the vowel more accurate than the consonant(s)? What significance does this simplification in form and content have for adult systems? Many explanations are compatible with the surface form: the simplification can be due to random or idiosyncratic effects, context effects, memory limits, articulatory inability, perceptual confusions, phonological organization or simplification, or a perceptual miscategorization or expectancy bias in the adult observer/transcriber.

2.1 Methodological Issues

Acquisition research deals with the methodological problem by obtaining many tokens of each word spontaneously produced in different contexts (with detailed verification of form, meaning, and intent) and collecting such data for the child's complete lexicon weekly or biweekly for many months. Complete lexicon weekly or biweekly for many months. Complete lexicon and longitudinal data are crucial. Children are acquiring systematic rules (or operating under systematic constraints), and it is systems of rules (constraints) that are changing. Acquisition data can seem chaotic when taken out of the context of a given child's system, and many of the differences between children are due to differences between types of systems. Only longitudinal data can show the complete nature of structure. All developmental stages and diachronic changes yield possible synchronic states. But any “stage” can be the observer's arbitrary cuts along the time dimension, and any cross-sectional observation may contain odd elements. Some are unassimilated or idiosyncratic elements (e.g., family conventionalized forms). But most synchronically “odd” elements are temporary residues of earlier stages, atypical only with respect to the current primary system. Given the variation, similar data must be collected from several children matched for all variables.

As to the problem of perception and bias, we as adults listen with highly phonemically (and psychophysically) categorized ears, and as observers and theoreticians we look and “see” with other types of preexisting ideas. An example is the way in which we hear the voiced stop in [du]. Cross-linguistic acoustic analyses of the speech of 18 to 24(+) month old children show they initially produce all stops in the short-lag voice-onset time (VOT) range – voiceless and unaspirated, as Jakobson predicted – and that English-learning children go through a second stage in which their short-lag stops show a significant difference in mean VOT between stops that correspond to adult voiced versus voiceless phonemes (Macken and Barton 1980). From the perspective of the adult observer, the voiceless stops in both stages fall within the perceptual boundary category of voiced phonemes /b,d,g/ for adult English speakers. This accounts for the strong tendency for English-speaking writers to (mis-)transcribe [du] and (erroneously) discuss the (context-free) “voicing” of all stops in the speech of very young children.[5] The significant difference in mean VOT shows that the voicing contrast exists in the child's underlying representation at least by stage two. While acoustic evidence is rarely available, the listener's perceptual bias problem can be dealt with in a number of ways. The problem of projecting our theories of the world onto our observations is tougher. Assuming for the moment that forms like [du] (or[tu]) are systematic and accurate realizations of their targets and representative of the child's lexicon and stage of development, we can turn to the major explanatory theories – qualitative capacity constraints, motor development, perceptual development, and phonological systematization.

2.2 Capacity

Let us consider the canonical length of words during this stage. This restriction is of particular interest, because the next stages are not similarly constrained: there are no three-syllable or four-syllable stages, etc. Do these characteristically short utterances verify the universality and core status of the CV syllable and the disyllabic foot – a conclusion that assumes that the explanation lies with innate, end-state phonological constraints? Are there other possible explanations? As for nonlinguistic capacity constraints, there is little direct evidence that bears on a possible memory or general cognitive, neurological or biological base to this restriction of lexical forms to one or two syllables.[6] But there are data on a similar string restriction in digit and letter span, arguably an analogical domain. It is well established that there are great differences between children and adults when asked to repeat strings of letters or numbers. Adults can remember and reproduce seven (plus or minus two) items, while the four year old generally can repeat only three. If this difference is due to a qualitative difference, say in memory capacity, we might expect that adults have more "slots" in short-term memory. Similarly, the two-year old child who produces [du] for *juice* or [nana] for *banana* may have fewer "slots" for lexical performance. If this difference were due to such a fundamental capacity difference, we would not expect any necessary relationship between how children and adults "lexicalize" digit spans or word strings. The fact that a two-year-old child might "lexicalize" a string "485791" as "91" would have no particular significance for adult cognition and the status of binary units. Similarly for words, forms like [du] might show the same type of string or slot constraint and be of equal nonsignficance for adult representation.

If, however, this difference in performance is due to basic differences in knowledge or experience (quantitative differences between children and adults), we would expect that because children are less familiar with numbers or words, they are less able to use their memory capacity, while adults who have much greater knowledge of numbers can use that knowledge and familiarity to increase their efficiency (where, e.g., noticing even-odd patterns or ascending descending patterns simplifies the recall of "485791"). For digit and letter span, experimental work has been done that distinguishes these alternative explanations. If stimuli of equal familiarity are presented to adults and children, the marked developmental difference is nearly wiped out: on strings of very high frequency lexical items or equally unfamiliar nonsense items, the adult-to-child advantage goes from over 2 to 1 to only 1.3 to 1 (Chi 1976). Thus for number, letter, and word span, memory capacity does not change much over the course of development.[7]

In the case of acquisition data on word production, we have further evidence that memory or some other capacity constraint is not the explanation. Complexity of the string matchings makes this constraint questionable. For example, the full range of a given child's forms will not be explainable by either a simple left-right or right-left filter (e.g.,*banana* [bana], *balloon* [bun], *juice* [du]). More importantly, the reductions correspond to constituent structure. To know that [du] for [dzus] is a segmented, syllable-type constituent structure and not the result of recency slot mapping, we must know that the child segments the word and "recognizes" in some way the last string element [s] and that [d] corresponds to [dz]. Indeed, other evidence does show such knowledge. In general, children know considerably more than their production behavior would suggest. To show this brings us to the fundamental issue of the child's mental representation and the remaining major theories – production, perception and /or phonological systematization – proposed for explaining developmental change.

We can infer the structure of the internal grammar by looking at the linguistic descriptions of the child's performance – which leads usually to indeterminacy – or by examining the order in which particular structures are acquired, much as Jakobson did. A different window on the mind can be found by looking at the properties of change itself. If the characteristics of diachronic change differentiate linguistic synchronic structure, the nature of structure will become clear during change in the same way that objects in a complex visual field – objects that appear interlocked when still and perceptually fused – separate into distinct forms when one of them moves. By looking for explanations in the properties of change in this way, we can identify three basic types of rules, each associated with a hypothesis about the child's underlying lexical representation and one aspect of development (Macken 1992b).

2.3 Acoustic Constraints

For some words, we find that they are produced correctly (in certain respects) at one stage and incorrectly at a next stage. Consider the following pairs of words, where the first member of a pair shows this incorrect second stage: (a) *chalk* [tšak] T1 (Time 1), [trak] T2 (Time 2); *train* [tšen] T1, [tren] T2; (b) *bran* [bræn] T1, [bænd] T2; *hand* [hæn] T1, [hænd] T2. For this type of change, Type 1, (i) two segments x and y are neutralized as y, (ii) a phonological change x' spreads slowly through the learner's lexicon (over a period of months) (iii) word-by-word, with some consistent lexical exceptions, and (iv) appears in the correct x and incorrect y environments, resulting in errors on y. The simplest explanation of this kind of change is that the child's underlying representations at T1 are the same as the surface representations and, in other words, there is no actual phonological rule of neutralization operating during this stage: e.g., adult /tr/ and /tš/ are both underlyingly /tš/ for the child. Actual rules for Type 1 phenomena operate at Time 2 when the lexical representations are changed: e.g., at Time 2, a rule of [tš] → [tr] is applied piecemeal to (correctly) construct underlying representations in words like *train* but creating errors on words like *chalk*. What appears to be going on in these cases is that the child does not completely preceive the relevant adult contrast at Time 1. Thus, even after age two, some perceptual learning takes place (cf. also Ingram 1974 and Waterson 1987) – typically in words where the segmental structure is complex and the acoustic cues for the constituent segmental contrasts are difficult to perceive. Type 1 change appears to characterize a small number of the rules found between the ages of two and four.

2.4 Articulatory Constraints

Type 2 change is considerably more common than Type 1 between the ages of two and four. In this type, (1) two segments x and y are neutralized as y, (2) a phonological change x' spreads rapidly through the learner's lexicon (in a matter of days usually) (3) in all and only the correct environments (4) with no errors on y. Examples of Type 2 change include the acquisition of fricatives and [r] shown in the first word of the following pairs of words: (a) *bus* [bət] T1, [bəs] T2; *but* [bət] T1, [bət] T2; (b) *pretty* [bIdi] T1, [prIdi] T2; *pip* [bIp] T1, [pIp] T2; (c) *rain* [dein] T1, [rein] T2;*den* [dɛn] T1, [dɛn] T2. The across-the-board nature of these changes, in all and only the correct environments, suggest that the child's underlying representations are accurate at T1 and at T2 and what has been learned at T2 is how to say particular segments that have been perceived and represented correctly for some period of time (Smith 1973; Stampe 1969, p. 146). For example, *bus* is /bəs/ at both T1 and T2; and /s/ is changed at the surface at T2. If so, articulatory constraints account for much of the child's development. Many Type 2 rules are, however, instances of Jakobson's implicational rules (e.g., fricatives and affricates are replaced by stops of the same place of articulation). Thus, although fricatives may have to be "learned" motorically, the phonological relationship between fricatives and stops (which explains why the fricative replacement is a stop) may be based on acoustic principles of contrast within phonological systems.

Type 1 relationships at Time 1 show the existence of a perceptual filter; Type 1 rules proper at Time 2 operate between the child's underlying representation and the child's lexical (or phonemic) representation, while Type 2 rules operate between the latter and the child's surface phonetic form. In terms of properties like slow, word-by-word spread, Type 1 rules are a developmental analogue of lexical diffusion sound change; rapid across-the-board changes show Type 2 rules to be the analog of classical Neogrammarian change. This parallel suggests that substantive phonetic content of features (acoustic versus articulatory) may play a role in these two different types of historical change as well. The two rule types also resemble the distinction in lexical phonology between, respectively, lexical rules which may have exceptions and post-lexical rules which do not, and thus may be an early reflex of the grammar organization. There are other types of variation not covered by Type 1 and Type 2 rules. First, because children are unskilled, there is much phonetic variation (especially in fricatives) as children acquire adult-like control; this variation is not evidence of a slow spread of a Type 1 rule across the lexicon. Second, occasionally a word can become an isolated lexical exception and remain so for a long period of time; such words have been relexicalized or "restructured" (e.g., *take* which remained [keik] until 3 years, 1 month, and 15 days of age [abbreviated henceforth in the style 3; 1.15], long after the velar harmony rule disappeared (2; 8.4), Smith 1973). Such restructured exceptions are not necessarily evidence of a Type 1 rule change.

2.5 Lexical Generalizations

Type 3 rules are more complicated than either Type 1 or Type 2, in that the underlying mechanism appears to be the imposing of pure, systemic phonological organization itself. In Type 3 change, (1) two or more segments x and y (, z) are distinct, (2) then a phonological change x′ spreads through the learner's lexicon in the correct environments and (3), some time later (typically, a couple of weeks), spreads rapidly and across the board to the incorrect y (, z) environment(s), (4) resulting in errors on y (, z). For example, at T1 (from 1;6 to 1;10), a child reduced initial stop + /r/ clusters by deleting /r/: *pretty* [pIti]; *tree* [ti], *drink* [tInk]; *cradle* [kedəl]. At T2 (1;11.0), initial /tr/ and /dr/ change to [f], *tree* [fi], *drinking* [finkin], etc. The coalecence of frication and rounding in [f] is probably due to the strong aspiration in this cluster in English and the labiality of English /r/. To this point, the rule is a typical Type 1 or Type 2 rule: if the former, then we would expect later some errors on true /f/-initial words when the rule is "unlearned", but no errors if a Type 2 rule. The next development however, at T3 (2;0), is one where the rule is generalized, rapidly spreading across the board to all voiceless stop clusters, wiping out the previous contrasts at all three places of articulation: e.g., *pretty* [flti], *tree* [fi], *cradle* [fedəl], and so on. The child here has made a generalization in terms of phonological (onset) categories and a major phonetic step backwards (or regression) at the level of correct [p, t, k] segment production.

That the generalization to / pr / and / kr / takes place very fast (within days) and in an across-the-board manner tells us, as with Type 2 rules, that the change is operating on underlying, well-defined categories. Thus, we can conclude that the child has distinct underlying representations for /pr/ and / kr / initial clusters (versus initial / p / and / k / words). This fact, with the delay of several weeks between the origin of the [f]-rule and its subsequent spread to labial and velar clusters, tells us that all three places of articulation are underlyingly distinct and that the rule is not a Type 1 rule (cf. also harmony regressions, Macken and Ferguson 1983, p. 269). Is it an articulatorily based Type 2 rule that just happens to have been independently generated in each of the three clusters? Not likely. First, the properties of change are not completely the same as with Type 2 rules (cf. the delay). Second, although all acquisition rules are phonetically natural, the articulatory motivation for each case is not equally convincing: [f] for [tr] is phonetically natural and common among children; the phonetic case for [f] from [pr] is weaker, and this alternation is rare among children; the phonetic argument for independently deriving [f] from [kr] is even weaker, and there are no other cases in what is a very large literature of a child using [f] for only [kr]. Finally, it is extremely unlikely that three such formally similar rules would spontaneously arise independently at the same time. What these Type 3 rules show is the child actively forming generalizations over classes of segments and subsets of the lexicon, in effect constructing interim, relatively abstract, autonomous phonological rule systems. These cases show, sometimes dramatically, that a significant part of development involves cases of getting better *phonologically*, by getting worse phonetically. This kind of (individual rule) nonlinearity in development shows most clearly the phonological aspects of acquisition and demonstrates that some aspects of acquisition fall outside the explanatory range of phonetic (segment-centered) theories.

Type 3 rules generally simplify lexical representations by creating symmetry along some abstract dimension of phonological organization as in the above example and in Amahl's "acquisition" of a velar fricative which filled out a symmetric system of /bdg, mnn, wlγ/ (see Smith 1973, pp. 109, 179; cf. also examples cited in Macken and Ferguson 1983). These rules operate between the underlying representation and the child's lexical representation on perceptually and phonologically distinct categories, and hence may be psychologically real in a way that Type 2 rules do not seem to be. Type 3 rules appear to be discovered or invented by the child: the general pattern is (1) isolated accuracy (i.e., "progressive idioms"); (2) a period of experimentation; (3) construction of a rule; (4) overgeneralization, which causes the loss of accuracy or regression in some forms; (5) construction of a new, more general (and ultimately accurate) rule (Macken and Ferguson 1983). To take a single word as an example, *pretty*, Hildegard Leopold's first permanent word, was pronounced with near perfect phonetic accuracy for a year (a "progressive idiom") and then was systematically pronounced as [pIti] and still later in a third stage as [bIdi], a month later (Leopold 1947). The two regressions correspond to the times when rules of consonant cluster reduction and consonant voicing (respectively) appeared in her system. Both these rules show the properties of change associated with Type 2 rules and are typically assumed in the literature to be articulatorily-motivated rules; however, the first year accuracy shows that the child could literally articulate both clusters and voiceless stops – which raises a question about the underlying explanation for the rules. While Type 2 rules most likely do have a basic articulatory component and Type 3 rules are most clearly nonautomatic, organizational generalizations, it may be, as suggested by Kiparsky and Menn (1977b), that *all* actual rules are discovered or invented by the child. These three types of developmental change suggest that there are three basic aspects to acquisition – perception, articulation, and phonological generalization.

2.6 Phonological Units

The Jakobsonian data and examples like the generalization underlying the velar fricative in Smith (1973) show that features are integral to acquisition and can be the unit of generalization. For the types of questions raised at the beginning of this section, we can reasonably be certain that the surface [d] is distinct from the affricate /dz/ and that the surface omission of /s/ is likely to be a similar Type 2, articulatorily-based rule. As rules change, systematically affecting all lexical instances of, say, /d/ or /s/, we reasonably infer that the child "segments" the input and represents words in terms of a linear sequence of segment-sized or equivalent-length units. As to how early segmentation takes place prior to this stage, we only can say that from the earliest stages of word use, in most cases, children behave in accordance with the hypothesis of segmented underlying representations; and we acknowledge that, alternatively, the underlying representations may be in a nonsegmented, Gestalt format at an early point during the first stage (e.g., Waterson 1987). The format of these beginning representations cannot be decided at this point. This raises other issues and is a problem to which we return below.

The acquisition evidence strongly suggests that the CV syllable, disyllabic foot, and prosodic word are basic to the universal core of phonological systems. We can exclude general memory constraints as an explanation here. In the case of Type 3 rules, children form interim phonological generalizations over the subsets of the lexicon known at each stage. Humans are powerful pattern recognizers, and this process of rule discovery or invention (Type 3 rules and possibly Type 2 rules) is a central part of phonological acquistion. What is the purpose of creating generalizations only to discard them at the next stage? Perhaps the construction of these rules is a way to learn about and systematize the lexicon. These cross-lexicon patterns then improve memory for the sound system and increase performance accuracy similar to the way that learning about relationships between numbers increases the efficiency of memory and improves performance on tasks like digit span. Thus memory would play the same indirect role in the quantitative changes that occur in phonological development as in cognitive development (transitivity, conservation, etc.). In both, change is tied to learning knowledge about specific domains. For phonology, while the content of the specific rules changes as the subset of the lexicon known changes at different stages, the form of rules, the capacity and fundamental structure do not change over development. It is equally important to recognize that it is not the lexicon per se driving the change or rules: for the same lexical subset, different children will use/invent different rules; and children will change rules when there has been no corresponding structural change in the child's lexicon. Rules are not passive, emergent properties of the lexicon, but active, creative constructions of the user.

3 Differences between Child and Adult Phonology – Problems for the Theory?

A corollary, though not a necessary one, of the strong-identity thesis is the view advanced by Halle (1962), and Kiparsky (e.g., 1968b, 1971, 1988), among others, that children may be one source of historical change in individual languages, either through imperfect learning (Kiparsky; cf. also Paul 1886) and/or through an alternative grammar compatible with the set of utterances learned at a particular stage (Halle). Children are a highly unlikely source of the particular rules of language change, because first and foremost, children do ultimately get it right: they learn to sound identical to their (native speaker) parents and presumably have the same underlying representations and the same or an equivalent grammar by age five or six generally (e.g., the nearly complete adult competence at 4; 0 in Smith 1973).[8] Second, the general universality of their stages would hardly be compatible with the diversification of languages or of a particular language, and conversely, where children do differ (from other children), it is precisely the idiosyncracy of those individual children's interim rules that limits their "adaption" by others and the potential explanatory value of such rules for typical sound changes. Third, given cross-cultural child-rearing practices, children's interim acquisition rules have absolutely no social support for transmission.[9] Finally, and of particular interest for this paper, there are certain cases where the content of acquisition rules and of sound change is different and, likewise, where acquistion and synchrony – the resulting states of sound change – differ. Do these differences create problems for the strong-identity thesis?

At the most specific, we can find highly unusual interim rules. Consider, for example, the rule system reported in Priestly (1977), where a child generalized a *CV$_j$VC* template to cover most disyllabic words: *cracker* [kajak], *breadman* [bijan], *records* [rejas], etc. The form of this rule at the CV template level is typical of children who construct templatic generalizations (though the use of both left-to-right and edge-in association is not typical); however, the content of the rule, particularly the medial glide, is unique in acquisition as far as we know, and unattested in known templatic languages of the world. Classical generative theory, which distinguished formal and substantive universals, would not be concerned with the singularity of a particular rule (or statistical facts of rule distribution, see below) provided a rule was a possible rule on formal grounds, which this rule is. However, more restrictive theories that require that constraints on form and content be universal may encounter difficulty with such data (e.g., McCarthy and Prince, chapter 9 of this volume). For such theories, these cases may demonstrate that some acquisition data fall outside the domain of phonological theory; or there can be a synchronic grammar with such a constraint, and the theory must generate both the child and adult structures.

3.1 Asymmetries

More interesting are two types of asymmetries between acquistion on the one hand and diachrony/synchrony on the other: (1) phonological processes with one typical directionality in acquistion and typically the inverse in diachrony – e.g., fricatives frequently replaced by stops in acquistion versus the more common sound change where stops change to fricatives; and (2) inverse distributions – consonant harmony common in acquisition, rare in synchrony, while vowel harmony is common synchronically in languages of the world yet infrequent in acquistion (and rare in recorded histories of sound change for that matter). For both (1) and (2), the issue is not one of mutually exclusive occurrence but rather statistical asymmetry – a fact about the world that under classical generative theory would again not necessarily concern the formal theory but that may be the province of Praguean marking theory and nonlinear theories.

The (1) cases seem to stem largely from what is, I believe, one central but not previously discussed difference between acquistion and diachrony: many historical sound changes and synchronic alternations are due to contextual effects between string adjacent consonants and vowels (e.g., palatalization), while none of the primary rules of acquisition and few of the other attested rules in the first year or two (ages one to three) show interactions between consonants and vowels. To the extent that sound changes like stops to fricatives begin in and/or are limited to positions after particular vowels or strictly string-based post-vocalic/intervocalic weakening processes (e.g., Spanish spirantization of medial voiced stops), they would not occur in early acquisition. The nearest acquisition analogues are the context-sensitive, word-based or syllable-sensitive processes like devoicing in final position (voicing in initial and medial positions) and fricatives acquired first in final (or medial) position, and low-level phonetic effects like stops showing somewhat greater closure variability in medial position. A full-scale acquisition analogue of Grimm's Law is unattested (and thus the kind of case where children are clearly unlikely sources of sound change); and the typical, across-the-board acquisition pattern of fricatives to stops is only diachronically common for the interdental fricatives (e.g., Germanic). Not enough is known about the rule types of diachrony and acquisition to pursue this question much further for other rules (cf. Dressler 1974 for other suggestive examples).

3.2 Harmony

Though the data are incomplete also for the (2) consonant-vowel asymmetries, these are more widespread and raise interesting questions for nonlinear theories of spereading, adjacency, and locality. The basic facts to explain appear to be an early acquistion of vowels as a set, frequency of consonant harmony, infrequency of vowel harmony, the protracted stage of consonant harmony, and patterns of individual differences. Typically, vowel harmony (where documented) characterizes only the stags before age two, while consonant harmony is more widespread in a child's system and lasts well into the third year (e.g., Smith 1973).

As to the individual differences, some children are reported to have a minimal consonant inventory and extensive vowel contrasts (e.g., Braine 1974), while the reverse (many consonantal constrasts, small number of vowels) is reported for other children acquiring the same language (e.g., Velten 1943); such a difference presumably could be correlated with consonant harmony versus vowel harmony, respectively. To some extent, observer categories and/or psychophysical constraints may contribute to the reported greater diversity among vowels: consonants are highly categorically perceived, while vowels are perceived in a more continuous manner; an adult then will tend to perceive greater variation and contrast in the child's vowels and fewer differences in the consonants.

Finally, consonant harmony though common is not universal among children (e.g., Vihman 1978). Some children show basically no consonant harmony (and no melody rules), such as G reported in Stemberger (1988) and E in Moskowitz (1970). For those children who use harmony, a wide variety of place and manner harmony rules will be found. In contrast, there are children who have long distance rules but who do not or only rarely use harmony. These children show constraints of a very different kind, namely where place and manner features are linearly ordered – a “melody” grammar, as opposed to the other or “harmony” grammar: thus, for example, while coronal consonants frequently undergo harmony in the former type systems – *sopa* “soup” [popa] – melody grammars will metathesize or otherwise sequence coronal consonants to the right of a noncoronal consonant – *sopa* [pota] (Macken 1992a).

In synchronic languages, in contrast, vowel harmony is common and consonant harmony is infrequent. Furthermore, synchronic consonant harmony almost always involves coronal or laryngeal features only (e.g., Chumash sibilant harmony or Rendaku/Lyman's Law in Japanese, respectively), and place of articulation harmony is rare (reported for Eskimo Inupiaq dialects). Nonlinear theory treats assimilation as spreading between a specified element and an underspecified element (e.g., coronals), both of which are adjacent at some level of representation, as for example on an autosegmental tier or between adjacent nodes in the feature geometry. One motivation for the geometric separation of vowel features from consonant features is to capture is to capture the natural assimilation of vowels over intervening consonants and the rarety or nonexistence of noncoronal place assimilation of consonants over vowels (e.g., Archangeli and Pulleyblank in press). Such theories cannot in principle account for the naturalness of acquistion harmony rules nor any acuisition place harmony over an intervening vowel that shares the same place feature as the trigger consonant (e.g., *doggie* [gogi]), a violation of the no-line crossing constraint. Alternatively, treating harmony as V-to-C assmilation (to handle [gogi]) fails with the equally frequent velar harmony over high front vowels (e.g., *drink* [gInk]). The consonant-vowel metathesis rules of melody grammars similarly violate the no-line crossing constraint in feature-geometric theories of locality. The observed differences between adult and child harmony rules, then, would place acquistion data outside the explanatory framework of the synchronic (adult) theory.

3.3 Templates

However, planar segregation provides a third approach to locality, one that permits consonants and vowels to be on separate planes under specified morphological or phonological conditions (McCarthy 1981a, 1989b). With planar segregation, acquisition harmony templates and the templatic, melody V/C metachesis representations are well-formed, if in the latter cases, the medial consonant is a default consonant: precisely as would be expected, the medial consonant in such melody templates is typically a voiceless coronal stop.[10] In addition, acquisition melody and harmony grammers fit all the McCarthy (1989b) diagnostics for phonological planar segregation: rigid consonant-consonant constraints in CVC(V) words (cf. harmony), V/C metathesis, and highly restricted prosodic structure constraints (e.g., CV syllables), as found for example in the Mayan and Oceanic languages. With planar segregation, then, the frequency of consonant harmony is related to the predominance of highly rigid, simple prosodic structure; the rarity in synchronic languages is a function of the rarity of the enabling conditions. Moreover, the same representational structure – planar segregation –provides a uniform account of what had been though to be two unrelated acquisition facts (the harmony versus melody grammars) and shows the underlying unity between the adult and child data. In general, the consonant-consonant constraints in acquisition strongly resemble, not distant assimilation, but (1) string-adjacent assimilation rules in sound change (e.g., Pali intervocalic consonant assimilatory changes from Sanskrit) and synchronic grammars (e.g., place assimilation in Spanish nasal clusters) and (2) string adjacent consonant constraints that permit only geminates or homorganic consonant clusters in languages like Diola Fogny (Sapir 1965), those that permit place to vary in consonant clusters only if the second consonant is coronal as in languages like Attic Greek (Steriade 1982), or the less common constraints where CC sequences are ordered left-right according to point of articulation as in Georgian where C1 must be further front than C2 (Tschenkéli 1958). The apparent difference between child and adult harmony and melody rules is, rather, a function of the same constraints operating on adjacent consonants - string adjacent consonants in adult languages and consonants adjacent at a planar level of representation in child templatic harmony and melody systems (Macken 1992a).

3.4 C-V Interactions

If, then, consonants and vowels are segregated on different planes, we would expect in acquisition that there would not be assimilatory changes between vowels and consonants, which is the general case, as previously noted. There is, however, a counterexample reported sporadically for the very first stage of word use for some children acquiring English: for these children, there is a statistical tendency for coronal consonants to occur before high front vowels (e.g., Braine 1974; Fudge 1969) and possibly labials before back round vowels (e.g., MacNeilage and Davis 1993). If these representations are segmented (cf. the discussion above) and the data are not related to extraneous factors like the structure of the adult models, reduplication or the (English) diminutive, this distribution would suggest consonant-vowel interactions (C-to-V assimilation). If in addition, these children are among the children who use harmony or melody processes, then planar segregation is not the solution; it remains to be seen whether in such a case other locality mechanisms (like feature geometry or C-to-V assimilation for the harmony cases) would uniformly work.

It appears, however, that many cases of C-V co-occurrences are not due to active vowel-to-consonant assimilation rules in the child's system but are cases where the C-V distributions are a direct reflection of characteristics of the lexicon and further that some of the children who show the C-V co-occurrence restrictions do not produce active harmony or melody forms (e.g., Braine 1974; the first stage in Fudge 1969, sect. 3.1). It may, however, be that some of these cases indicate that another option available to children is the syllable as the primary organizing unit (cf. Moskowitz 1973) although there is little data that would support the centrality of syllables as a general stage. Most likely, these early representations are rather in a gestalt, nonsegmented form and nonrepresentative of those children's representations in later stages. For example, in Fudge (1969), at the second stage (sect. 3.2) when vowel harmony and velar harmony appear, the evidence for active C-V harmony (tenuous even in the first stage) is sparse at best. For the majority of English-learning children, there is little or no evidence of C-to-V assimilation, of consonant-vowel co-occurrence restrictions, or of the types of developmental progressions that would be expected if consonants and vowels were on the same representational plane or if the primary phonological unit were the syllable and children used an inventory of syllable types. In contrast, there is considerable evidence that vowels and consonants are representationally and developmentally autonomous (during the time period in question). Recent cross-linguistic research has not found C-V co-occurrences except where the child data reflects the adult language distribution (e.g., Boysson-Bardies 1993). Thus, this would be one Aristotelian case where the structure is not in the child but in the world.

While phonological theory has not settled all locality and adjacency issues and confirming evidence must come for the acquisition issues as well, current nonlinear theory provides a unifying account of the major case of differences between children and adults - the (2) type consonant-vowel asymmetries - and brings greater insight into the variation found in the acquisition literature as well, turning an apparent exception into verification of the strong-identity thesis.

4 Conclusions

For many issues concerning underlying representations and for many different phonological domains - particularly prosodic structure, tone, the interaction with morphology and syntax - much additional research is needed. But for the well-studied phonology of features, segmental rules, templates, and the prosody of the CV syllable and the foot, the evidence from children shows the same structure and variation as found in adult systems. Given the validity of the strong-identity thesis, acquistition data not only can be used to confirm and change phonological theory, but each relevant theoretical proposal must be able to incorporate acquisition data.

Ultimately, phonological theory will explain the stages, constraints on variation, and the form of rules found across all learners either as options available in universal grammar or through interactions between subdomains. Yet, as striking as the differences are between harmony and melody learners or between children who appear to construct rule systems based on quite different core units, there is no evidence at all that subsequent stages show the effects of earlier rule types. Thus, structure of language does exist separately from the particular acquisitional paths traveled. The absence of any residue of the content of particular interim grammars is a significant problem for all empiricist learning theories: stochastic learning is cumulative and where paths differ, outcomes differ. The acquistion data confirm that language is, rather, a formal problem space. It does not matter to the learner the different points of entry or the particular trajectories through this space. The outcome is the same – a shared, fully adult phonological grammar and competence by six years of age. We can look at the variation across children as a mix of "correct" rules (in the sense of being steps that can be considered, with miminal adjustments, as toward the target language) and "incorrect" rules (where no miminal, local adjustments could reorient toward the target). Given the sheer number of children in the latter category, it is not unremarkable that those children get to end state at all. That they do so in virtually the same amount of time would approach miraculous in a stochastic, indeed in any empiricist, world.

1 The term "discontinuity" is used in two senses in the literature. In the most common, Jakobsonian sense, the theorized discontinuity between babbling and speech is supposed to indicate a complete break between the abilities and knowledge that characterize the two stages. The term is also used to describe the nonlinearity in development seen during phonological acquisition when the acquisition of a new rule or reorganization of structures causes a temporary, surface loss of ability or other anomaly; in this case, the change is within one part of a single system and does not mean the two temporal stages show two fundamentally different systems.

2 The Jakobson term is [dental], a category that excludes the true palatals, as did [coronal] in the original features (Chomsky and Halle 1968, hereafter cited as *SPE*). For the dentals, Jakobson explicitly characterizes their corresponding marked place of articulation as the velars.

3 In general, "front" is a valid category for acquisition order and developmental and cross-linguistic inventories. This is in opposition to the general lack of evidence for the corresponding Chomsky/Halle feature [+anterior]as a natural class.

4 Most acquistion researchers agree tacitly or explicitly in the strong identity relationship but disagree among themselves what the "system" is that is shared: formal phonology (Dresher 1981; Dresher and Kaye 1990), both an abstract, formal phonology and a stochastic component (Macken 1987), only a stochastic (connectionist) system (Stemberger 1992a), only a concrete, phonetic system (Lindblom 1992; MacNeilage and Davis 1993).

5 This same kind of problem arises in cross-linguistic typological surveys and all areas of phonology.

6 Or, similarly, of sentences to two words during the first stage of syntax acquistion.

7 This is the same kind of result noted earlier for Piagetian conservation and transitivity. Children differ from adults in quantitative knowledge and experience but, it now appears, not in basic or core capacity.

8 This age is based on the existing acquistion studies. However, little if any work has been done on languages with complex phrasal phonology (e.g., the interaction of tone and phrasal position in Chaga) or languages with highly complex, productive morphology (e.g., the many thousand possible forms of verbs in Shona); children learning such languages may take longer to reach adult competence. If there is any doubt that phonological acquisition is the learning of rules rather than the memorization of word forms, data from these children will no doubt resolve this debate resoundingly on the side of rules.

9 These arguments apply to primary phonological acquistion – the child learning her native language. This of course leaves open whether adolescents who recreate language for peer solidarity can be a source of sound change.

10 Some melody template systems metathesize labial-velar consonants (e.g., Grammont 1902).

Cite this article

MACKEN, MARLYS A. "Phonological Acquisition." *The Handbook of Phonological Theory*. Glodsmith, John A. Blackwell Publishing, 1996. Blackwell Reference Online. 31 December 2007 <http://www.blackwellreference.com/subscriber/tocnode?id=g9780631201267_chunk_g978063120126724>

Bibliographic Details

The Handbook of Phonological Theory

Edited by: John A. Glodsmith
eISBN: 9780631201267
Print publication date: 1996