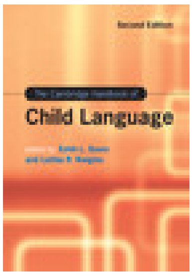


The Cambridge Handbook of Child Language

Edited by Edith L. Bavin, Letitia R. Naigles

Book DOI: http://dx.doi.org/10.1017/CBO9781316095829

Online ISBN: 9781316095829

Hardback ISBN: 9781107087323

Paperback ISBN: 9781107455504

Chapter

11 - The acquisition of prosodic phonology and morphology pp. 230-249

Chapter DOI: http://dx.doi.org/10.1017/CBO9781316095829.011

Cambridge University Press

11

The acquisition of prosodic phonology and morphology

Katherine Demuth

11.1 Introduction

Much of the early work on the acquisition of phonology has traditionally focused on the acquisition of segments, and the transition from babbling to first words (see Vihman 1996 and Chapter 10). However, since the mid-1990s, research on children's phonological acquisition has increasingly begun to examine phonological development at higher levels of prosodic structure (e.g. at the levels of the syllable, the prosodic word and the phonological phrase), and the implications for understanding why children produce the particular early word shapes and utterances they do. This shifting focus has been stimulated in part by new approaches to phonological theory (e.g. Optimality Theory (Prince & Smolensky 2004)). Optimality Theory (or OT, as it is often called) was originally designed to better understand how phonological systems, and morphophonology, work. The groundbreaking insight here was that phonological systems contain a series of constraints, and what qualifies as 'grammatical' in a given language is merely the best way to satisfy the most important (or the most highly ranked) of those constraints. For example, nicknames in English can be made very short – e.g. *Elizabeth* > *Liz, Joseph* > *Joe*. Both result in a bimoraic foot – i.e. they contain either a coda consonant or a long vowel. However, truncating *Liz* to *Li* [lɪ], would be 'ungrammatical' (see Section 11.2). This is because English has a constraint that all open class words, including names, must contain a certain amount of phonological structure – i.e. they can't be *too* short. Such a view is ideal for understanding how children's phonological systems develop. Optimality Theory has therefore provided the theoretical and practical tools needed for investigating children's early language productions as a series of competing constraints rather than rules, whereby simple (unmarked) structures are predicted to appear earlier than those that are phonologically (and articulatorily) more complex.

At the same time, there has been an increase in the availability of longitudinal, phonetically transcribed corpora of child speech between the ages of 1 and 3 from an ever-growing variety of languages. Much of this can be found on the CHILDES database (see MacWhinney 2000). Some of these data also provide information about the language input (child-directed speech) children receive. Researchers are therefore now able to investigate both frequency and markedness factors in making within-language and crosslinguistic predictions about the course of phonological development. Further developments, such as increasing attention to the relationship between perception and production abilities, acoustic analysis of child speech, and the rise of articulatory phonology (Browman & Goldstein 1986, 1988), have begun to lay the groundwork for thinking of children's early speech productions in terms of a *developmental model of speech planning and production* (Demuth 2014).

This chapter first reviews some of the structures that are important to the study of prosodic development. It then highlights some of the classic and more recent findings regarding prosodic development, interactions with the perception and production of morphosyntax, the various methods used, and the implications for understanding the mechanisms underlying early speech planning and production. It concludes by identifying areas for further research.

11.2 Prosodic structures

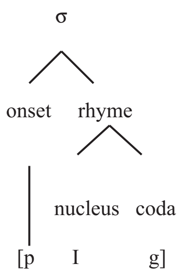
To investigate the structure of children's early syllables, words and morphemes it is useful to consider the Prosodic Hierarchy in (1) (Nespor & Vogel 1986, Selkirk 1984, 1996). Of particular interest are Prosodic Words (PWs) (also called Phonological Words), which are composed of feet (metrical units) and syllables. These PWs may also be embedded in higher-level phonological phrases (PPs), phonological utterances and intonational phrases.

(1) The Prosodic Hierarchy

Utt	(Phonological Utterance)	*I saw the man give the kitty the banana*
|		
IP	(Intonational Phrase)	*I saw the man*
|		
PP	(Phonological Phrase)	*the man*
|		
PW	(Prosodic Word)	*banana*
|		
Ft	(Foot)	*man/kitty*
|		
Σ	(Syllable)	*man*
|		
μ	(Mora)	*ma*

Syllables in turn are composed of an onset consonant and a rhyme, as in (2). The rhyme consists of an obligatory nucleus, and an optional coda. These subsyllabic units are called moras. Thus, monomoraic syllables contain only a nucleus, whereas bimoraic syllables may contain either a vowel plus coda consonant (*dog*), a diphthong (*play*), or a long/tense vowel (*see*).

(2) Basic Syllable Structure



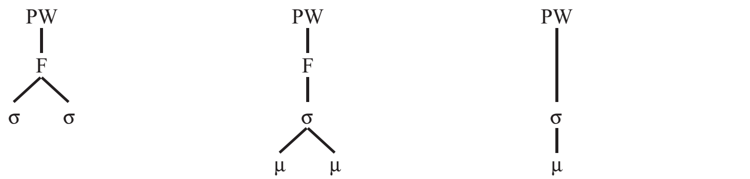
Some languages also permit complex (branching) onsets and codas. These are realized as consonant clusters. The permissible consonant clusters vary depending on the language. However, most consonant clusters obey the Sonority Sequencing Principle (SSP), where sonority is greatest in the nucleus, and decreases toward the edges of the syllable (Clements 1990, Selkirk 1984). This is captured by the Sonority Hierarchy in (3), where each sound can be categorized in terms of one of seven manners of articulation (Ladefoged 1993). More sonorous segments tend to fill the nucleus of the syllable, and less sonorous segments tend to fill onset and coda positions. In the case of a consonant cluster, sonority typically decreases from the nucleus outward. For example, in the word *blend* /blɛnd/, /ɛ/ is a vowel, /b/ and /d/ are stops; /l/ and /n/ are a liquid and nasal, which are both less sonorant that a stop, but more sonorous than a vowel.

(3) The Sonority Hierarchy
Stops > Affricates > Fricatives > Nasals > Liquids > Glides > Vowels
least sonorant ——————————→ most sonorant

Languages differ in the types of syllable structures, foot structures, and PW structures permitted. Children must therefore learn what types of prosodic structures their target language allows. Moras play an important role in languages such as English and Dutch, where stress assignment is sensitive to the syllable weight (how many moras it contains), and where stress generally falls on heavy syllables (i.e. those containing two moras of structure). Foot structure also differs from language to language. Languages such as English and Dutch permit one-syllable bimoraic feet such as in *dog*, whereas Bantu languages like Sesotho permit only monomoraic syllables, and therefore have disyllabic feet, as in *nama* ‘meat’. Languages also differ in the directionality of feet, many exhibiting Strong–(weak) trochaic feet (e.g.

kitchen – English, Dutch), but some exhibiting binary or longer (w)(w)(S) iambic feet (e.g. *chapeau* 'hat' – French, K'iche'). Binary feet can be disyllabic (4a) or monosyllabic (bimoraic) (4b). They therefore constitute well-formed minimal words (McCarthy & Prince 1994). Some languages also permit words containing only a light (monomoraic) syllable, or a subminimal word. Subminimal words are generally considered to be marked and unusual since they are PWs that do not contain a foot. However, words of this type are permitted in Romance languages and Japanese.

(4) Prosodic words composed of a foot (a,b), and a subminimal word (c)



(a) Disyllabic foot (*kitty*) (b) Bimoraic foot (*dog*) (c) Monomoraic subminimal word

The frequency of different PW shapes varies from language to language. Although both English and Spanish permit four-syllable PWs containing 2 feet (5a), as well as a foot plus an initial unfooted syllable (5b), both are much more frequent in Spanish. In contrast, English and Dutch contain many monosyllabic and disyllabic PWs like those in (4a) and (4b).

(5) Prosodic words composed of more than a foot

(a) Two feet (e.g. *alligator*) (b) One foot plus an initial unfooted syllable (e.g. *banana*)

With these structural preliminaries, we can now consider how children learn these various prosodic structures. We first review early findings in the field, and then discuss more recent research.

11.3 Prosodic development

Although much of the early research on the acquisition of phonology focused on *segments*, European researchers focused on the *word* as an important unit in children's early phonological organization. Drawing on insights from Firth (1948), Waterson (1971, 1987) proposed that children's early phonologies could best be characterized as consisting of holistic, non-segmental prosodic units. Allen and Hawkins (1978, 1980) subsequently proposed that English-speaking children's early words tended to take the form of disyllabic trochaic (strong–weak) feet (e.g. *kitty*). They observed that children's early words are often augmented (*cup* > *cupy*) or truncated (e.g. *banana* > *nana*), both resulting in a disyllabic trochaic foot. They further proposed that such early word shapes might be universal, representing the default, or *unmarked* form of early words.

Following research on the prosody–syntax interface (Selkirk 1984), Matthei (1989) investigated across-word processes in children's early speech. Consistent with Allen and Hawkins (1978, 1980), he found that some lexical items were augmented to a disyllabic trochaic foot when produced in isolation (6a–b). However, when the two are combined into a larger phonological phrase, both were reduced to yielding a disyllabic trochaic foot (6c).

(6)		Child	Adult target		
	(a)	[ˈbebi]	/ˈbebi/	'baby'	(1;5)
	(b)	[ˈbʊkɔ]	/ˈbʊk/	'book'	
	(c)	[ˈbebʊ]	/ˈbebiz ˈbʌk/	'baby's book'	

Macken (1978, 1979) also found that some children exhibited templatic patterns in their early words. That is, they went through a period of development where their early words exhibited certain distributions of consonants, such as only labial consonants word-initially, and only coronal consonants word-medially. Thus, words such as Spanish *Fernando* were realized as [mano], and *libro* 'book' as [pito]. Such findings led to proposals that children had both a perception and a production representation (Kiparsky & Menn 1977, Menn 1983, Menn & Matthei 1992), though others disagree (Smolensky 1996). The early findings began to lay the groundwork for thinking of children's early phonologies in terms of *output constraints*.

Having encountered limitations in rule-based, segmental accounts of children's early productions (e.g. Smith 1973), acquisition researchers also began to explore autosegmental (above the level of the segment) approaches to understanding the nature of early phonological systems. This was motivated by much work on tonal systems, where tones and segments appear to move independently of each segment (see Gussenhoven 2004). Using this approach, Demuth (1993) found that 2-year-old Sesotho-speaking children had no problem learning lexical tone, but only acquired grammatical tone melodies (tone sandhi) around the age of 3. Other researchers used similar approaches to understand aspects of phonological development in both first and second language acquisition (e.g. Archibald 1995, Yavas 1994).

11.4 The emergence of unmarked prosodic structures

Jakobson (1941) proposed that children begin the process of language acquisition by initially producing only 'unmarked' consonants (i.e. those that are easy to produce, and widely found amongst the world's languages). Although this proposal has never been verified at the segmental level, phonologically simple structures, such as stop consonants (e.g. /p/, /t/, /k/) and simple CV syllable structures (such as /ba/) do tend to be acquired early. Fee (1995) and Demuth and Fee (1995) suggested that both weak initial-syllable truncation (*banana* > *nana*) and reduplication/vowel epenthesis (e.g. *dog* > *dada*) could be understood in terms of markedness. Drawing on developments in prosodic phonology (Nespor & Vogel 1986, Selkirk 1984, 1996), they proposed that children's early productions will tend to avoid more 'marked' prosodic structures such as syllable-final coda consonants and initial weak (unstressed) syllables, at least in English. Observing that the same types of constraints could also account for early word-shapes in Dutch (see Fikkert 1994), they proposed that perhaps children learning all languages would exhibit a similar stage of early development, where prosodic words were both minimally and maximally a binary foot, or 'minimal word' (Demuth 1995).

Similarly, Gnanadesikan (2004) proposed that the 'emergence of the unmarked' could help account for the fact that children tend to preserve the least sonorant consonant in cases of consonant cluster reduction at the beginnings of words (e.g. *tree* > *tee, stop* > *top*). Pater (1997) integrated these proposals, showing that children's early word truncations could be understood in terms of markedness constraints at both the level of the syllable and prosodic word. Thus, *banana* is often truncated to *bana*, preserving the least sonorant (least marked) consonant in the syllable/word onset. Note that such truncations also indicate that children have perceived at least the onset of the weak, unstressed syllable, even though they have not fully produced it. This notion of 'markedness' characterizing early stages of acquisition now helps to inform much current research on children's early language productions (see below), including recent findings on the acquisition of morphology (see Demuth 2014 for an overview).

11.5 The acquisition of syllable structures

Clements and Keyser's (1983) work on the relevance of the syllable as a phonological unit highlighted the importance of the sonority hierarchy and the sonority sequencing principle for understanding some of the crosslinguistic restrictions on syllable structures (see (2), (3) and (4) above). This set the stage for examining how and when different types of syllable structures are acquired, both within and across languages. Thus, although there are individual differences in the timing of syllable structure acquisition within a given language, there are also robust crosslinguistic differences that can be understood in terms of the relative frequency of different syllable structures in the language children hear, and the prosodic context in which these appear.

11.5.1 Coda consonant acquisition

Many children's earliest syllable structures consist of simple CV structures, with coda consonants omitted (*dog* [dɑ]). Over time, children develop the ability to produce coda consonants (*dog* [dɑg]), and other, more marked, complex syllable structures (*desk* [dɛsk]). Interestingly, coda consonants tend to appear earlier in languages where codas and coda clusters are common. Lleó (2003) reported that some German-speaking children begin to use coda consonants while still babbling. In contrast, she found that Spanish-speaking children's first use of coda consonants is much more delayed, with many coda consonants still being omitted after the age of 2 (e.g. *espejo* [eˈpeho] 'mirror'). In contrast, Demuth and McCullough (2009a) found that French-speaking children produce most coda consonants around 1;8 years (*canne* [kan] 'stick'). In addition to possible segmental factors, these crosslinguistic differences in the timing of coda consonant acquisition can be explained by the interaction of at least two factors: the overall frequency of coda consonants in the ambient language, and the prosodic position in which they occur within the word. For example, using an elicited imitation task with novel words, Kirk and Demuth (2006) found that English-speaking children were much more likely to produce coda consonants in stressed or word-final syllables, as compared with unstressed and/or word-medial syllables. They suggest that this is due to the fact that both stressed and final syllables, in English and many other languages, tend to be longer in duration than medial or unstressed syllables. This may provide young language learners with more time to articulate more complexity within the syllable. It is perhaps not surprising, then, that coda consonants are acquired later in Spanish (around 2;3 years), since many of these occur in unstressed and/or word-medial position (e.g. *es̲pejo* [eˈpeho] 'mirror'). In contrast, Prieto (2006) shows that coda consonants are used a year earlier (1;2 years) in closely related Catalan, in part due to the higher frequency of (weak)–Strong (CV)ˈCVC words ending in a coda (e.g. *carn* [tan] 'meat'), even if all the segments are not yet in place. Some of the within-speaker variability in the production of coda consonants may therefore be a function of the prosodic contexts in which these appear. Thus, both frequency and prosodic context play a role in the determining when coda consonants may emerge. This may also help explain some of the crosslinguistic differences in the timing of coda consonant acquisition.

These findings do not address the types of consonants that are first acquired in the coda. On markedness grounds it might be expected that more sonorous consonants would be acquired in the coda first. However, in a corpus study of English child-directed speech, Stites, Demuth and Kirk (2004) found that alveolar stops are the most frequent coda consonants in English. In a longitudinal study of child speech they also found that most English-speaking children's first coda consonants are alveolar stops rather than the less frequent, phonologically less marked sonorant coda consonants. Kehoe and Stoel-Gammon (2001), in a larger cross-sectional study, confirmed this finding, showing that /t/ was the first coda consonant acquired by most children, followed quickly by /d/. Thus, although frequency and markedness typically pattern together, children may show a preference for frequency over markedness effects in their early productions, all else being equal. This raises questions about the notion of markedness as a whole, and its relationship to frequency for learners of a particular language. It also raises the question of which linguistic units learners are using for calculating 'frequency'. For example, Zamuner, Gerken and Hammond (2004) showed that coda consonant production is a function of neighbourhood density. That is, it is the frequency of the rhyme + coda, rather than simply the coda consonant itself, that is the best predictor of accuracy in coda consonant production, at least for English. On the other hand, /ʁ/ is one of the most frequent consonants in French, yet several studies have found that at least some French-speaking children have persistent problems with the production of /ʁ/ (e.g. Demuth & McCullough 2009a, dos Santos 2007, Rose 2000). This may be due to the articulatory challenges experienced in trying to produce this uvular fricative, or due to its variable realization in the input children hear.

11.5.2 Consonant cluster acquisition

Research on the structure of the syllable has provided a framework for examining the acquisition of consonant clusters as well. Some of the early research focused on consonant cluster reduction in children with phonological delay, where various explanations were given for why clusters are simplified the way they are (e.g. Chin & Dinnsen 1992, Gierut 1999) (for review see Bernhardt & Stemberger 1998). Following Pater (1997), some researchers proposed that children typically preserve the least marked onset, i.e. the least sonorant segment of the cluster (e.g. Barlow 1997, Ohala 1996, 1999). Thus, in a word like *stop*, the obstruent /t/ would be preserved, but in a word like *sleep*, the /s/ would be preserved. Others noted the limitations of the sonority account (e.g. Barlow 1997, 2001). Goad and Rose (2004) proposed that children preserve the consonant that is the head of the syllable (e.g. *plate* > *pate; slate* >*late)*. However, Pater and Barlow (2003) show that some children simplify *sneeze* to *neeze*, but *sleep* to *seep*. Jongstra (2003) therefore proposed that when the sonority distance is close, the segment contiguous with the nucleus will be preserved (*sneeze* > *neeze*), whereas when the sonority distance is sufficiently far, the least sonorous segment will be preserved (*sleep* > *seep*). However, a recent study of cluster simplification calls all the above into question, noting that features from both consonants often remain in cluster reduction (e.g. *spin* > *fin*) (Kirk 2008). This suggests that articulatory factors, involving place of articulation, may help explain some of these phenomena. Most of these studies have been carried out in Germanic languages; research on other languages will contribute to our understanding of how articulatory factors influence early language production.

The studies mentioned above all examine word- and syllable-onset clusters. Only a few studies have investigated the acquisition of word- and syllable-final clusters. One might predict these to be acquired later since codas are more marked than onsets. However, Lleó and Prinz (1996) found that final clusters were acquired several months earlier than word-initial clusters in a longitudinal study of German-speaking 1–2-year-olds. Levelt, Schiller and Levelt (2000) also found that the majority of the children in the Dutch CLPF corpus acquired word-final before word-initial consonant clusters, though both patterns occur, probably due to equal frequency in child-directed speech. Kirk and Demuth (2005) found that English-speaking 2-year-olds were more accurate at producing word-final as opposed to word-initial consonant clusters. In English, coda clusters are more frequent than onset clusters. Interestingly, the English-speaking children in their study also exhibited better production of final nasal+*s* and stop+*s* clusters than final nasal+stop and *s*+stop clusters. Furthermore, children often metathesized the *s*+stop clusters (*wasp* > *waps*), suggesting that phonotactic probability or frequency factors may be involved. Note also that the most accurately produced clusters are those that typically occur with morphologically complex forms, suggesting that morphology may provide a further perceptual or production advantage for these coda clusters.

To explore these issues further, Demuth and Kehoe (2006) examined the acquisition of consonant clusters in French. They found that 2-year-olds were more accurate at producing onset rather than word-final clusters in picture identification tasks, a finding confirmed in a subsequent longitudinal study (Demuth & McCullough 2009a). Some researchers have proposed that some word-final consonants in French (and other languages) prosodify as onsets to empty-headed syllables (e.g. *partir* 'to leave' /paʁ.ti.ʁØ/) (Charette 1991). It is possible that this structure is more marked, and therefore later acquired (though Goad and Brannen (2003) claim that such structures are universal at early stages of acquisition). Rose (2000) noted, however, that one child from his longitudinal study of two children learning Canadian French had acquired all but /ʁ/ in word-final position, but had /ʁ/ as a coda word-internally. He therefore proposed that this child had a coda representation for /ʁ/ in all positions. However, others have also noted that the acoustic and articulatory characteristics of French /ʁ/ are extremely variable, both within and between speakers (for review see Demuth & McCullough 2009a). Little is known about the acquisition of segments that are variably realized in the input, or where the syllabic representation is ambiguous (see discussion in Rose 2000, Kehoe, Hilaire-Debove, Demuth & Lleó 2008). However, in a recent study of the acquisition of the alveolar stops /t,d/ Song, Shattuck-Hufnagel and Demuth (in press) show that children aged 1;6–2;6 are much more conservative than their mothers, producing few flaps, glottal stops and unreleased stops. This suggests that, despite variability in the input children hear, they are able to extract information about that target phoneme form, and tend to be conservative, slowly increasing their use of these phonetic variants over time.

11.5.3 Articulatory aspects of phonological acquisition

Much of traditional phonology explored the featural aspects of phonemes. Yet many of these features, such as place (high, front, back, labial, velar, etc.) and manner (sonorant, fricative, stop, etc.) actually refer to how different segments are articulated. Articulatory Phonology (Browman & Goldstein 1986, 1988) was therefore developed to explore these issues in more depth, helping to explain how the realization of different segments depends in part on the phonological environment in which a segment occurs. This has led to the adaptation of ultrasound imaging methods for examining movements of the tongue. It is now possible to explore aspects of children's articulation during word production in a non-invasive manner. Researchers have explored how children make various articulatory contrasts, and how this might explain some of the reported delays in the acquisition of certain phonemes. For example, it has long been observed that English-speaking children exhibit a protracted period of acquisition with liquids such as 'r' and 'l'. Both are composed with multiple tongue gestures, whereas children often use only one. Our recent research with English-speaking 4–7-year-olds explores the acquisition of word-initial onset /l/ (as in *lake*) and word-final coda /l/ (as in *pail*), showing that the acquisition of adult-like articulatory patterns takes some time to develop (Lin & Demuth 2013, 2015). As expected, there was a tendency for only one articulatory gesture to be used in coda position, and this slightly increased with age. Unexpectedly, however, only one gesture appeared in onset position as well, despite the fact that the percept sounded adult-like. More research will be needed to determine exactly when and how children acquire adult-like articulatory gestures for this segment. Other research on children's late acquisition of 'r' has shown that visual input from watching ultrasound videos of their tongue can help children learn to produce the appropriate articulatory gestures (e.g. Bernhardt, Gick, Bacsfalvi & Ashdown 2003).

Recent studies have also explored the articulatory gestures underlying children's production of final clusters, comparing the production of monomorphemic words like *box* and bimorphemic words like *rocks*, both with the same word-final /ks/ cluster. Researchers often wonder if children's early grammatical morphemes are really 'productive', or simply learned as unanalysed, 'frozen' forms. If the latter were true, we might expect both types of word-final /ks/ clusters to be produced in the same way. However, if the plural morpheme is productive, and must be retrieved from the child's mental lexicon independently from the lexical word itself, then we might expect to find evidence for this in the articulatory gestures used, as has been found for adult speech (e.g. Cho 2001). Song, Demuth, Shattuck-Hufnagel and Ménard (2013) found that 2;6-year-olds do show morphemic effects, showing differences in tongue height when producing the morphemic compared to the non-morphemic /ks/ clusters. This suggests that, even at the age of 2, children show both cognitive and articulatory planning mechanisms that are surprisingly adult-like. Such results provide a framework for much further research exploring the nature of children's lexical and higher-level prosodic and morpho-phonological representations, and how these develop over time (Sections 11.6 and 11.7).

11.6 The acquisition of prosodic word structure

Initial research on the acquisition of PW structure (Demuth 1995, Pater 1997) suggested that children had an early awareness of word-minimality effects, and that this could be captured in terms of constraint interactions. Using acoustic evidence, Ota (1999) also showed that Japanese learners exhibit compensatory lengthening of the vowel when a coda is omitted, thereby preserving moraic (and minimimal word) structure. But Japanese is a mora-timed language. What about word-minimality effects in a syllable-timed language like French, where CV subminimal words are also permitted? Demuth and Johnson (2003) examined this issue in longitudinal data from one French-speaking child. They found that her earliest words (1;3–1;5) were all target or reduplicated CVCV forms. As in other languages, her early grammar showed a highly ranked constraint against word-final (coda) consonants, resulting in either reduplicated CVCV repairs, or truncated CV outputs. Interestingly, she also reduced some disyllabic CVCV words to monosyllabic CV form. Further analysis showed that segmental constraints against fricatives, velar stops and clusters were more highly ranked than faithfulness to syllable preservation and/or word minimality (see dos Santos (2007) for similar observations from another child who does have velar consonants). Demuth and Johnson (2003) showed that CV subminimal words account for 20 per cent of all words French-speaking children hear. They suggested that learners are sensitive to the high-frequency phonological structures of the target language, and quickly begin to adjust their grammars (constraint ranking) to accommodate such forms. Note that such a perspective on the development of early grammars minimizes the role of universal markedness. Rather, higher-frequency phonological forms become the 'unmarked' structures on a language-specific basis.

This issue has been subsequently pursued in several other studies. For example, Goad and Buckley (2006) proposed that one Canadian French-speaking child did show early word-minimality effects through compensatory vowel lengthening (CVC > CV:), though no acoustic analysis was provided. However, analysis of two French children showed no systematic lengthening of the vowel when the word-final consonant was missing (Demuth & Tremblay 2008). The number of subjects examined in all these studies is small, suggesting that further study with more children at the early stages of acquisition (1–2 years) is required to resolve this issue. Returning to English, Demuth, Culbertson and Alter (2006) examined word-minimality in four children between the ages of 1 and 3. Although some children showed apparent compensatory vowel lengthening, this occurred on both monosyllabic and disyllabic words, and on both long/tense as well as short/lax vowels. If learners were using compensatory lengthening to preserve word-minimality, one would expect it to be restricted to monosyllabic words with short/lax vowels, where a second mora of structure is required to preserve a bimoraic foot, or minimal word. Further acoustic analysis of three children's compensatory processes found that two of the children exhibited compensatory lengthening for missing codas (in monosyllabic words) with all vowels, whereas only one (older) child showed compensatory lengthening only for target words with a short/lax vowel (Song & Demuth 2008). This suggests that English-speaking children may initially compensate for omitted coda segments, and only later (around the age of 2) come to realize that English has word-minimality constraints. The English findings contrast with those of Ota (1999) for Japanese. However, since coda consonants are always moraic in Japanese, it is possible that compensatory lengthening is due to segmental factors here as well. Alternatively, perhaps children become more aware of moraic structure and its consequences for PW structure earlier in a mora-timed language. Recent findings, however, suggest that even for English, there may be some early awareness of both moraic structure and word-minimality effects. Using an elicited imitation task with CVC and target words, Miles, Cox, Yuen and Demuth (in press) found that 2-year-olds were more likely to preserve the coda consonant when it followed a short/lax vowel (e.g. *lid*) compared to when followed a long/tense vowel (e.g. *seed*). This suggests that there is some early sensitivity to word minimality by at least 2;3 years. It would be interesting to explore this crosslinguistically to see how and when this constraint is learned for other languages.

Roark and Demuth (2000) proposed that the frequency of syllable and prosodic word shapes in the input children hear may help determine the PW structures children use in their early utterances. In a corpus study of child-directed speech they showed that most words in English are monosyllabic, whereas Spanish has many more trisyllabic and quadrasyllabic words. They suggested that these word-shape characteristics may account for English-speaking children's tendency to truncate words like *banana* until around 2;6 years (Pater 1997). In contrast Spanish-speaking children permit larger PWs much earlier (see also Lleó 2006). Further support for a frequency-based account comes from studies of European Portuguese

(Vigário, Freitas & Frota 2006). However, Prieto (2006) suggested that the relative frequency of foot shape, rather than PW shape, helps explain why Catalan learners (but not Spanish learners) exhibit a stage of development where they truncate disyllabic (w)S PWs (e.g. *aquí* > [t^hi] 'here'). Finally, Ota (2006) suggested that lexical frequency effects best account for the few cases of truncation found in child Japanese. Thus, frequency effects at different levels of prosodic structure may help determine the relative ranking of constraints in the grammars of children learning different languages, resulting in different truncation patterns in early PW development.

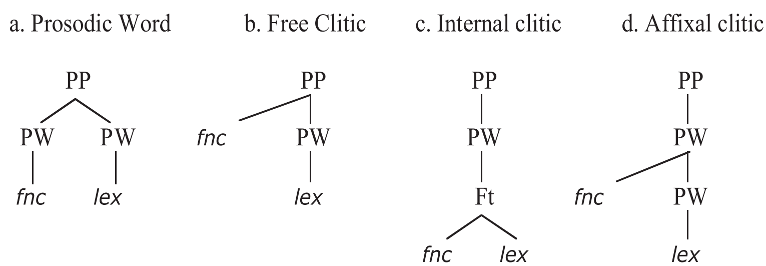
Critically, these patterns of truncation appear to be due to phonological, not perceptual or articulatory constraints. For example, Carter and Gerken (2004) found that children left a prosodic 'trace' of the missing syllable (realized as a silent duration) when they omitted the initial unstressed syllable of a three-syllable word. This suggests that, in some cases, children have 'planned' for the syllable, even though no segmental content is realized. Such 'covert contrasts' in children's early speech are often missed in traditional phonetic transcription. This points to the need for finer-grained acoustic analysis of children's speech, where measures of duration and other aspects of both the spectrogram and waveform can be consulted to determine if phonemic representations might be present despite that lack of an adult percept (see Munson, Edwards, Schellinger, Beckman & Meyer 2010, Scobbie, Gibbon, Hardcastle & Fletcher 2000, Theodore, Demuth & Shattuck-Hufnagel 2012). This also raises questions about the extent to which other 'omissions' in child speech may be realized at some level of analysis, perhaps present in the child's phonological or morphological representation, but simply not produced. This suggests that a *developmental model of speech planning and production* will be needed to better understand the nature of children's early (variable) productions. This would involve looking above the level of the prosodic word (PW) to consider the higher levels of the phonological phrase (PP), the intonational phrase (IP) and the phonological utterance (PhU).

11.7 The acquisition of prosodic morphology

Drawing on insights from the Prosodic Hierarchy, researchers in the 1990s began to examine children's acquisition of grammatical morphemes (see Demuth 1994, Gerken 1996). Since many grammatical morphemes are variably produced for a certain period in development, syntacticians have often claimed that children's morphosyntactic representations take time to be fully acquired. However, researchers have also found that some of the variability in children's production of grammatical morphemes is not random, but predictably constrained by aspects of children's developing prosodic representations. That is, there may be phonological (as well as syntactic and semantic) restrictions on children's use of grammatical morphemes. For example, researchers of Bantu languages such as Sesotho reported that children tend to produce noun class prefixes with monosyllabic stems before consistently producing them with disyllabic stems (Connelly 1984). Demuth (1994) suggested that children first produce noun class prefixes that constitute part of a disyllabic foot (*mo-tho* 'person'), and tend to omit those that are unfooted (*mo-sadi* > [sadi] 'woman'). Demuth and Ellis (2009) have shown that this tendency holds until the age of 2;3.

Selkirk (1996) shows that different languages prosodify grammatical function items at different levels of structure (7). She also suggests that unfooted grammatical morphemes that were prosodified at the level of the Phonological Phrases (PP) (7b) violate constraints on well-formed prosodic structure, where each level of the Prosodic Hierarchy is immediately dominated by the next higher level (e.g. Syllable > Foot > PW). Thus, grammatical morphemes that are prosodified as free clitics (7b) (e.g. French) require the child to produce a marked type of structure. This is also the case with the affixal clitics in (7d) (e.g. Spanish). In contrast, grammatical morphemes that can be prosodified as an internal clitic as part of a foot (7c) should be the easiest and earliest acquired. We hypothesized that this is the form that the earliest noun class prefixes assume in Sesotho. Finally, those grammatical morphemes that themselves constitute a PW (7a) (as in German) would require the child to produce yet another 'word'.

(7) The prosodic structure of grammatical function items



Gerken and colleagues (Gerken 1994, Gerken & McIntosh 1993) also found that English learners were more likely to produce grammatical morphemes such as pronouns and determiners (such as articles) when these could be prosodified as part of a foot (e.g. *Tom [hit the]*$_{Ft}$ *pig* vs *Tom [wanted]*$_{Ft}$ *the pig*). Gerken (1996) then showed that this could also be captured in terms of Selkirk's (1996) markedness constraints. Thus, children's variable omission of grammatical function items could be understood in terms of prosodic constraints, where those that could be prosodified as part of a foot were more likely to be produced at a certain stage of acquisition.

Lleó (1996) had long noted that Spanish-speaking children (unlike German-speaking children) exhibit the use of (proto)determiners from the beginning of their speech. This was explained in terms of the high frequency of Spanish three-syllable words, which required a monomorphemic structure like that in (7d). This then provides Spanish-speaking children with the prosodic structure (i.e. a three-syllable window) needed for the inclusion of articles in their early output forms (e.g. *la mesa* 'the table') (Demuth 2001, Lleó 2001, Lleó & Demuth 1999). Further support for this finding came from the fact that three-syllable words that are truncated to two syllables are nonetheless accompanied by a (proto)determiner (e.g. *la muñeca* 'the doll' > [aˈmeka]) (Demuth 2001, Demuth, Patrolia, Song & Masapollo 2012). This suggests that Spanish-speaking children can use the prosodic structure in (7d) at this point in development, and can fill the initial prosodic slot with either lexical or functional material. This is the main insight of the *Prosodic Licensing Hypothesis*, which predicts that grammatical morphemes will be included in children's early productions to the extent that these can be realized as part of a simple, unmarked prosodic structure – i.e. a foot, or a phonotactic and/or higher-level prosodic structure they already have in their grammar (Demuth 2014, Demuth & McCullough 2009b).

Research on other languages similarly shows that young children are more likely to produce grammatical morphemes that are 'prosodically licensed' than those that are not. For example, Demuth and Tremblay (2008) showed that French-speaking children consistently use determiners with monosyllabic words around 1;10 years, whereas consistent use of determiners with disyllabic and trisyllabic words lags by two and four months, respectively. This suggests that the early determiners are prosodified as part of a foot, and that determiner use with two- and three-syllable words appears only once these can be prosodified at the level of the PP (7b). Similarly, Demuth and McCullough (2009b) found that English-speaking children had significantly higher use of articles when these could be prosodified as part of a foot with the *preceding* word. In contrast, children tended to omit articles that remained unfooted (those prosodified at the level of the PP) (e.g. *Tom [hit the]*$_{FT}$ *ball* vs *Tom [wanted]*$_{FT}$ *(the) ball*). This pattern persisted for four to five months, disappearing as the children approached 2–2;6 years. Note that this is about the same time that children begin to more reliably produce the initial unstressed syllables of lexical items like *banana* (Pater 1997).

The prosodic licensing of grammatical morphemes appears to occur at the level of the syllable as well, where some children exhibit syllable structure (phonotactic) restrictions on the acquisition of English third-person *-s* (e.g. Stemberger & Bernhardt 1997). Thus, in both spontaneous speech and in more controlled elicited imitation tasks, 2-year-olds are much more likely to produce this grammatical morpheme when it occurs as a simple coda consonant than when it forms part of a consonant cluster (e.g. *sees* vs *hits*) (Song, Sundara & Demuth 2009), and similar findings are reported for plurals and possessive morphemes (see Mealings & Demuth 2014a, Theodore, Demuth, & Shattuck-Hufnagel 2011, 2012). These effects are particularly prominent *utterance-medially* compared to when the word occurs in utterance-final position, where phrase-final lengthening provides more time to completely produce all segments in the final syllable (e.g. *The dogs barked* vs *I like dogs*). Note also that the articulatory content is more complex utterance-medially, where the three consonants /gzb/ must be produced in rapid succession. The shorter duration of the morpheme utterance-medially, plus possible masking from the consonant of the following word, also renders the morpheme less perceptually salient (see Sundara, Demuth & Kuhl 2011). Thus, both lower perceptual salience and articulatory challenges may contribute to lower production of utterance-medial grammatical morphemes. Given that English is an SVO language, where verbs occur utterance-medially about 75 per cent of the time, it is perhaps no wonder that English-speaking children take some time to fully acquire third-person singular *-s* (see Hsieh, Leonard & Swanson 1999, Song *et al.* 2009). Mealings, Cox and Demuth (2013) extended these findings to syllabic *-es* morphemes as well, showing that articulatory problems persist with the production of the plural on words like *buses*, even when compared to disyllabic plural production on disyllabic words like *farmers* (which in non-rhotic Australian English ends in the same schwa+/z/). This suggests that the articulatory challenges of producing such forms, plus their lower frequency, may help explain their later acquisition. However, the fact that these morphemes can and are produced in utterance-final position suggests that at least some aspects of the morphological representation are present, even if the morpheme is not consistently realized in phonologically challenged environments.

The production of these inflectional morphemes also appears to be influenced by lexical and processing factors, where lower performance is found with *longer utterances* (e.g. Mealings & Demuth 2014b). Interestingly, Theodore and colleagues have found that making the utterance-medial context easier, by having the following word begin with a vowel, can enhance morpheme production (e.g. *the dogs eat* vs *the dogs bark*). This suggests that there is still much to be discovered about the phonology–syntax interface in children's developing grammars, where constraints on prosodic representations, combined with lexical and processing factors, may account for much of the variable production of grammatical morphemes. All these factors will play a role in understanding the development of speech planning and production.

These findings suggest that children's acquisition of grammatical morphemes is closely tied to the development of prosodic representations. Given that many grammatical morphemes are unstressed prosodic clitics, their acquisition is dependent on the development of higher-level prosodic structures. The *Prosodic Licensing Hypothesis* therefore provides a framework for exploring the development of higher-level prosodic representations, and how this changes over time. It also provides a principled means for making predictions about the course of grammatical morpheme development within and across languages. As shown in the case of Spanish determiner acquisition, these developments are *also* closely tied to the prosodic properties of the lexicon. And the recent research on English shows the close connection with processing factors as well.

11.8 The future of phonological acquisition

11.8.1 Theoretical developments

The field of phonological acquisition has been significantly influenced by the developments in phonological theory, including the prosodic issues outlined above. Many other developments have implications for our understanding of children's phonological systems as well, and this will continue to develop in years to come. For example, the constraint-based approaches to the study of phonological systems (e.g. Prince & Smolensky 2004) provide a framework for investigating interactions between different types of constraints in the developing system, and for viewing phonological acquisition as a constraint-satisfaction problem, as illustrated above by many of the prosodic word and morphological characteristics of early child speech. Thus, this approach provides a much-needed vocabulary for understanding what constraints change over time.

11.8.2 Frequency versus prosodic factors

There is still the problem of understanding the mechanisms underlying phonological development. Researchers have long known that lexical frequency plays an important role in psycholinguistic processing (e.g. MacDonald, Pearlmutter & Seidenberg 1994), and infant speech perception studies show that infants are also sensitive to the frequency of the segments and prosodic structures they hear (e.g. Anderson, Morgan & White 2003). It has also long been known that 3–5-year-olds' representation of familiar, high-frequency words is more robust in both perception and production than that of novel and low-frequency words (Edwards, Beckman & Munson 2004). And, as noted above, researchers have found frequency effects on children's production of syllable and prosodic word structures.

One of the challenges of studies of frequency effects is in identifying what to count. Demuth (2001) suggests that language learners may be keeping track of the statistics of structures at all levels of the prosodic hierarchy, as well as the segmental interactions therein. For example, much of the research on lexical acquisition finds that children's accuracy in the production of lexical items is closely related to neighbourhood density (Edwards *et al.* 2004, Storkel 2004). Thus, some of the variability found in the acquisition of syllable structures, as well as words and morphemes, may be explained by the frequency with which these occur in the lexicon. However, as mentioned above, there are also limits to the frequency accounts. Across different prosodic contexts, articulatory planning phenomena may better account for some of the variable production found. For example, the position within the word or within the phonological utterance (Hsieh *et al.* 1999), as well as the presence or absence of stress, may also play an important role in determining the nature of children's early syllable, word and morpheme productions. Such issues are not currently incorporated into models of early acquisition. Controlling for such prosodic factors may provide a clearer understanding of children's phonological competence and the factors that contribute to variability in production.

11.8.3 Articulatory and acoustic factors

Given the complexities of language production, there may also be acoustic and/or articulatory (see Section 11.5.3) evidence that children are actually approximating certain contrasts that are not heard by the listener/transcriber (e.g. Scobbie *et al.* 2000, Song & Demuth 2008). The renewed interest in investigating such 'covert contrasts' provides acoustic and articulatory evidence for children's developing phonological representations. For example, Stoel-Gammon and Buder (2002) showed that most English-speaking children control extrinsic vowel lengthening before voiced/voiceless consonants by the age of 2, and Yuen, Cox and Demuth (2014) have shown that 2-year-olds can make the distinction between phonemically long and short vowels even as a function of contrastive stress and phrase-final lengthening. There is much to be learned about the prosodic organization of children's early productions, and how this interacts with both articulatory factors and the development of language planning and production abilities.

11.8.4 Sources of data

Another challenge to the field has been the lack of longitudinal phonetically transcribed data from multiple children between the ages of 1 and 2. This type of data is particularly important since children are actively acquiring the phonology of their language during this time – a point at which it is often difficult to conduct elicited production experiments. New language acquisition corpora are continually becoming available on the CHILDES database (MacWhinney 2000). Many of these include interactions with parents, providing important information about the input children hear. Some also contain linked acoustic and video files and/or phonetic transcription, allowing for the acoustic/phonetic analysis of both child and adult speech. Phonological and phonetic analysis tools (e.g. PHON tools – see CHILDES (Rose *et al.* 2006), Praat tools (Boersma & Weenink 2005)) are also increasingly available to facilitate phonological and acoustic analysis. As mentioned above, it is now also possible to explore aspects of children's speech planning and production using ultrasound imaging of tongue movements for the examination of both segmental, syllabic and morphological organization in early speech (e.g. Lin & Demuth 2013, 2015, Song *et al.* 2013). Together with online comprehension methods that provide a complement to the production research, such as eye-tracking and EEG, as well as crosslinguistic evidence, a more detailed picture of children's emerging phonological and morphological abilities is emerging.

Most of the above research has been carried out with monolingual, typically developing children. However, it is also well known that children with Specific Language Impairment (SLI) experience protracted problems with grammatical morphology. Recent research suggests that this is not merely a problem with tense morphemes, but may also interact with phonological, frequency and processing issues for both nominal and verbal inflection, where syllabic morphemes like *buses* persist in being problematic (e.g. Tomas, Demuth, Smith-Lock, & Petocz in press). Furthermore, it has long been known that children developing bilingually exhibit certain challenges, especially when the two languages differ phonologically, morphologically and syntactically. Recent research with Mandarin-speaking preschoolers who are beginning to learn English suggests that coda clusters and inflectional morphemes may present a particular challenge, much as they do for Mandarin-speaking adults learning English as an L2 later in life (see Xu Rattansone & Demuth 2014). This suggests that many of the issues that are relevant for understanding the course of typical monolingual phonological, prosodic and morphological development will also be relevant from many other populations of children who experience various challenges with learning language.

11.9 Conclusion

The field of phonological acquisition has grown significantly since the 1990s in exploring interactions between the acquisition of segments and higher-level prosodic structures more systematically. This has been due to several developments in phonological theory, as well as the increasing availability of early, phonologically transcribed longitudinal language acquisition data. Both have allowed researchers to more thoroughly explore the nature of the constraints on children's early phonologies, and how these change over time. This in turn has allowed the field to make testable predictions about the factors that influence the process of phonological development. These advances now begin to provide a clearer picture of how phonological systems are acquired in normally developing individuals, with implications for better understanding the nature of language delay.

Suggestions for further reading

Demuth, K. (ed.) (2006). Crosslinguistic perspectives on the development of prosodic words. *Language and Speech* (special issue), 49, 129–297.

Goad, H., & Rose, Y. (eds.) (2003). *Segmental-Prosodic Interaction in Phonological Development: A Comparative Investigation*, special issue of Canadian Journal of Linguistics, 48, 139–452.

Kager, R., Pater, J. & Zonneveld, W. (2004). *Constraints in Phonological Acquisition*. Cambridge University Press.