CHAPTER 29

THE DEVELOPMENT OF PROSODIC PHONOLOGY

KATHERINE DEMUTH

29.1 Introduction

How do children become adult-like speakers of a language? Anyone who has listened to a two-year-old knows that they are young not only due to the high pitch of their voice, but also due to the word-by-word like structure of their early utterances. This is often called "telegraphese" (Brown, 1973), where a target sentence such as "I want to go to the store" might only be produced as "want go store."

Even by the age of three, children's speech begins to sound more "fluent," with an emerging ability to string together longer sequences of words and phrases. They also begin to include more of the grammatical morphemes (e.g., articles, pronouns, prepositions, auxiliaries) that were omitted before, and start to use more complex, multisyllabic words (*banana* instead of truncated form like "nana") and more complex sentences. Some of this emerges along with their growing lexicon and an increasing knowledge of grammar (syntax and semantics). But part of sounding more "adult-like" is due to development at the level of phonology as well.

This raises the question of what phonology is. In general, *phonology* is defined as the "sound system" of language. Much attention has focused at the segmental level of phonology—or the "sounds" that make up a language, how these are "contrastive" or "phonemic" (i.e., lead to creating different meanings in words, such as *cat*/kæt/ vs. *pat*/pæt/). Thus, it has long been common to think of the acquisition of phonology as being about the acquisition of sounds, or segments.

However, linguists have also long realized that there is much more to the sound system of language than segments. Many languages make tonal contrasts, for example, that can change the meaning of a word. This includes many of the languages found in Asia, but also much of the rest of the world, especially Africa and Latin America. But even in a language like English, the placement of stress on one syllable or another can change the meaning of a word, as in *re*cord vs. re*cord*. In this case, the shift in stress from the first syllable to the second syllable results in a noun versus a verb. So, even for English, the placement of stress on one syllable or another can signal a difference in meaning. Researchers have thus begun

to explore how and when children master some of these other aspects of the phonological system that can occur above the level of the segment.

The use of tone and stress are part of a much larger class of phonological phenomena that children must learn to become a competent speaker of a language. They must also learn how to produce complex syllables with consonant clusters (e.g., *street*/stɹit/) (cf. Kirk & Demuth, 2005; Kehoe, Hilaire-Debove, Demuth, & Lleó, 2008), long words with both stressed and unstressed syllables (*elephant*/ˈɛləfənt/) (Kehoe & Stoel-Gammon, 1997) and grammatical morphemes (e.g., articles, auxiliary verbs, and so on; see Brown, 1973). They will also need to learn that intonation can be used to signal different meanings. For example, the two utterances *Where are YOU?* vs. *Where ARE you?* have slightly different meanings that involve contrastive focus, which must be mastered to become a competent speaker of English, and many other languages. Finally, children must also learn aspects of prosodic phonology, such as the fact that, in English, unstressed syllables are typically shorter in duration than stressed syllables, and that at the end of a phrase or utterance, the pitch of the utterance will fall in a statement, signaling a phrasal boundary. The last syllable of the phrase/utterance will also be longer in duration than that of the other syllables in the utterance (see next).

Thus, learning the "phonology" of a language entails much more than merely learning the segmental inventory (or phonemes) of a language, and much more than learning just words. Even learning to produce words in an adult-like fashion requires much of what is called "prosodic phonology." One of the ways in which this can be captured is by appealing to the Prosodic Hierarchy in (1) next (Nespor & Vogel, 1986; Selkirk, 1984), where a Prosodic Word (PW) is composed of a foot which contains one or more syllables, and the PW itself is part of a larger Phonological Phrase (PP), Intonational Phrase (IP), and Phonological Utterance (Utt).

(1) The Prosodic Hierarchy

Utt	(Phonological Utterance)	*I hope we find some bananas*
|		
IP	(Intonational Phrase)	*I like bananas*
|		
PP	(Phonological Phrase)	*like bananas*
|		
PW	(Phonological/Prosodic Word)	*bananas*
|		
Ft	(Foot)	*nanas*
|		
σ	(Syllable)	*nas*
|		
μ	(Mora)	*na*

The rest of this chapter will briefly summarize what we know about the acquisition of PWs, what we are beginning to know about the acquisition of higher-level PP and IP structures, how this provides a framework for understanding within-speaker variability in early production, and how this is relevant for understanding the emergence of grammatical morphemes across languages.

29.2 The acquisition of prosodic words

Researchers and parents alike have long been intrigued by how and when children begin to learn their first words. Although this is often thought to occur when a child first says something like "baba," and a parent thinks they have (obviously) intended "mama" or "papa," the child probably first recognizes their own name around the age of six months (Bortfeld, Morgan, Golinkoff, & Rathbun, 2005), and may have many words in their receptive lexicon before they actually produce their first words, typically between 11 to 16 months. Jakobson (1941) proposed that the period of canonical consonant-vowel (CV) babbling was followed by a silent period before the onset of first words, but this is now no longer thought to be the case. Rather, both babbling and children's first words overlap for several months, with words typically beginning to dominate and babbling diminishing by around one to six years. But even those first words may bear little resemblance to the actual adult form. Consider, for example, the target word *rice*, which one child at one to two years produced as [ˈwʌki] for several months (cf. Demuth, Culbertson, & Alter, 2006). Some precocious children, who produce their first words from around 11 months of age, have been observed to go through this brief stage of development where early monosyllabic target words are produced with an epenthetic vowel, possibly to form a disyllabic "word" (e.g., *clean* [ˈklinə]; one to three years: Demuth et al., 2006, p. 174; see also Vihman & Velleman, 2000). This occurs primarily with those target words that end in a voiced consonant. It is possible that this apparently epenthetic form may be a reflex of voicing articulatory release, ensuring that the final consonant is clearly produced. If one considers that young children have a large tongue and fairly small vocal tract, perhaps it is not surprising that such forms, as well as many well-known early "phonological processes" just as backing (*pack* produced as *pat*), are common around or before the age of two or three (Kent, 1976).

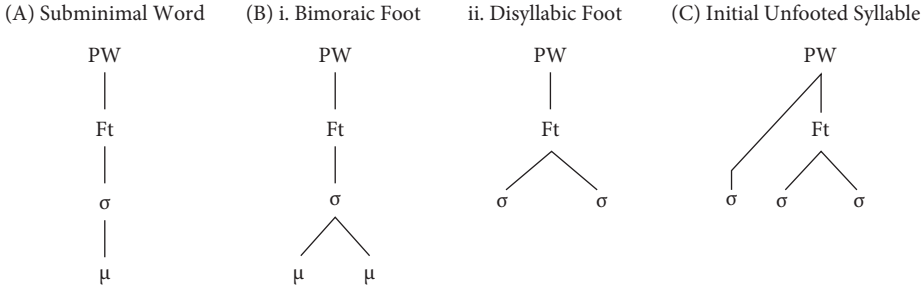
Another way in which young children's word shapes may differ from that of the target form is in the truncation of word-initial unstressed syllables (e.g., *banana* [ˈnænə]) (cf. Demuth, 1996), and the reduction of word medial unstressed syllables (e.g., *telephone* [ˈtɛfon]) (cf., Kehoe & Stoel-Gammon, 1997). This happens not only in English, but also in Dutch (Fikkert, 1994), Spanish (Gennari & Demuth, 1997), and many other languages. Note that the reductions in many of these cases also lead to a disyllabic output form. This led Demuth (1996) to propose that children's early PWs were initially composed of a (disyllabic) Foot. As shown in the following section, this is a common early PW form that many children exhibit.

29.3 Minimal words

Well-formed prosodic words in English must contain two moras of structure, that is, either a long/tense vowel or diphthong, or a short vowel and a coda consonant if a disyllable (2b), or be disyllabic (2c). Words with just a short vowel (2a) in English are ill-formed. Nicknames are a productive way to show this: one can shorten Philip to Phil /fil/ (containing a short vowel plus coda consonant), and Susan to Sue /su/ (containing a long vowel), but not /fi/ or /sʌ/, respectively—both with only a short vowel. But how and when do children learn that English

words must have a certain amount of phonological "weight," containing at least a bimoraic or disyllabic Foot? Many other languages have such a constraint as well. For example, Bantu languages like Sesotho (Doke & Mofokeng, 1957), spoken in southern Africa, add another mora to monomoraic words, so that ill-formed **ja!* "eat!" becomes *eja!* or *jaa!* But this is not universal: languages like French have no such constraint, permitting words that contain only one mora of structure, like *o* "water." This gives rise to the various PW structures presented in (2) next.

(2) Prosodic word structures



Demuth (2006) proposed that English-speaking children's early words might take the form of phonologically well-formed Minimal Words. This entailed that early words would contain at least two moras of prosodic structure—that is, either a long vowel or diphthong, or a short vowel and a coda consonant—or a Foot. Evidence from children's spontaneous speech productions appeared to suggest that children's early target words containing a short vowel tended to more reliably occur with a coda consonant than those target words containing a long vowel. Thus, learners might have an early sensitivity to this language-specific minimal word constraint.

However, children variably omit coda consonants under the age of 1;6. This also raised the possibility that children might be lengthening a short vowel when a coda was omitted, thus preserving word minimality. If so this would provide additional evidence for an early awareness of this constraint. Song and Demuth (2008) therefore conducted an acoustic study of children's early speech productions, and found that vowels were indeed lengthened in the case of missing codas, but that this occurred for both long and short vowels. Thus, it appeared that processes of compensatory lengthening took place to compensate for the missing consonant, not for mora preservation.

However, all the children in these studies were under the age of two. In a final attempt to address this issue in a more controlled fashion, Miles, Yuen, Cox, and Demuth (2016) conducted an elicited imitation experiment with children aged two to three years. They confirmed that children were more likely to produce coda consonants in the words with short vowels, and to drop codas more often in words with long vowels. This then suggests that, at least for children learning English, coda consonants are more likely to be produced in those prosodic contexts that require them for phonological well-formedness. These English findings then raise many questions about the course of prosodic word development in other languages. We turn now to the case of subminimal words in French.

29.4 Subminimal words

As we have already mentioned, many languages have a word minimality constraint, but this is not universal. Rather, it must be learned. The previous discussion suggests that learners of English have an emerging awareness of this constraint before the age of two, and that it appears fairly stable by two to three years. But what about learners of a language like French, which has submininal words such as *o*, meaning "water"? Do French-learning children produce these as monomoraic, subminimal words? Or do they augment them, as required in Sesotho, adding an epenthetic vowel, or another syllable?

Children often have a limited segmental inventory under the age of two (cf. Smit (1993) for developmental norms for English). Thus, young children may often preserve a syllable, but change the sounds to those they can more easily produce. This often has the effect of replacing late acquired target sounds, such as fricatives, affricates, and liquids, to earlier acquired (and easier to produce) segments, such as stops. French-learning children start by producing PWs that take the shape of disyllabic feet, though these are iambic (weak-strong, final stress), with the more prominent (longer) syllable being the final one (Demuth & Tremblay, 2008). Some of these, as in English and other languages, are reduplicated early on (e.g., *chapeau* "hat" [popo]). But then, at least for some children, there is a stage of development where they not only produce subminimal words like *o* "water," but also truncate longer words to produce subminimal forms (e.g., *chapeau* "hat" [po]) (cf. Demuth & Johnson, 2003). This is quite surprising! Why would a child who is producing disyllabic word forms suddenly start producing them as truncated monosyllables? Such forms then appear to be "subminimal" (monomoraic) words, and are widely attested in other French-learning children under the age of two (cf. Archibald & Carson, 2000).

Why might such truncated forms appear? Why would a child not simply persist with the disyllabic reduplicated form of the word until they can produce the articulatorily more challenging consonants? Demuth and Johnson (2003) suggest that this can be understood in terms of the child's shifting priorities (or, from the perspective of Optimality Theory (Prince & Smolensky, 2004), a "reranking of constraints"). Thus, early on, the child tries to be "faithful" to the number of syllables in the word, producing the same number of syllables as the target word, even at the cost of producing non-target segments. Thus, the two-syllable word *chausson*/ʃosɔ̃/ "slipper" is produced as [tɔtɔ̃], even though the segments of the first syllable are not target-like. At this point in development, this type of approximation is the best the child can do. But a month later, she decides this is "not good enough," that she should produce syllables that are more target-like, with the appropriate consonants and vowels. Thus, at one to five months, though she still cannot produce the fricative /ʃ/, rather than omitting or modifying the consonant, she omits the entire first syllable! This may suggest that the onset consonant and vowel are closely linked in her early phonological representations such that if the consonant is omitted, the vowel must go as well. In this way, children's early phonological "errors" can provide evidence for the nature of their emerging phonological (and lexical) representations. In the phonological grammar of this French child, we can understand her sudden production of monosyllabic forms not as a "regression," but rather the next stage in her developing phonological representation of words. At this point she is no longer prepared to produce segments (and the syllables that contain them) that do not match the target form.

This shows an early sophisticated awareness of segment-specific content in production that is similar, in some ways, to the feature-specific sensitivity to segmental contrasts in infant speech perception mispronunciation tasks (cf. White & Morgan, 2008).

Further confirmation that a developing phonological sensitivity may underlie this type of behavior comes from other studies of infant speech perception as well. In particular, it has been shown that, at one to two years, infants treat a segmental change as mapping onto a different word (Swingley & Aslin, 2000). This type of "switch task" can thus be used to probe children's early phonological representations, showing that they can perceive a language-specific change in phonological features (e.g., place of articulation) as having a different lexical referent. Perhaps then, the French child mentioned here has similar perceptual abilities which then influence her early productions. This suggests that our understanding of children's production of early words, and how this changes over time, may well be much more closely related to developments in perception than often thought. Exploration of both perception and production in the same children, and the role of both with respect to the developing lexicon, is thus an area of research that needs much more investigation.

Most of the word forms just discussed have been very simple, monosyllabic, or disyllable word shapes. In the following section we discuss what happens to longer words, and how these are realized in children's early phonological grammars.

29.5 Feet and the emergence of unmarked PW structures

We have seen that early learners of English have an emerging awareness that a well-formed PW must contain at least two moras of structure, as either a heavy monosyllable (with a long vowel or coda consonant), or as a disyllable. In contrast, we have also seen that learners of a language like French, where subminimal, monomoraic PWs are permitted, go through a stage of development where they truncated disyllables extensively, something not typically reported for a language like English. Thus, although children learning English may produce monomoraic CV word forms for target consonant-vowel-consonant (CVC) words when they cannot yet produce coda consonants, it is highly unusual to find disyllables truncated to a monosyllable unless the first syllable is unstressed. So, for example, a word like *giraffe* may be realized as "raf." Note that this is true in early perception for English-learners as well (cf. Jusczyk, Cutler, & Redanz, 2002), where the truncated form "raf" is treated as a word. This is probably due to the fact that the majority of words in English begin with a stressed syllable, and this is even more the case in infant-directed speech.

This raises the question of how young children produce longer words like *banana*? Anyone who has listened to a two-year-old will be well aware of some of the segmental modifications just mentioned, but also that children reduce longer words as well. This results in words such as *banana* being truncated to a disyllabic trochaic foot [ˈnænə]. Since most of the vocabulary young English-speaking children are exposed to contain a trochaic foot (i.e., stress on the first syllable), this tends to also be the form that is most frequently used in their early speech productions. This is the case for other closely related Germanic languages as well, such as Dutch (cf. Fikkert, 1994). Interestingly, Fikkert (1994) provides examples of some children actually shifting stress to the initial syllable at 1;10 years of age, so that a word like *balloon* (ballon

/bɑˈlon/ is produced with initial stress, as [ˈbo:mi]. Thus, the output that these Dutch children produce is similar in form to English children's truncations of words like *banana*: both result in a PW shape that can be characterized as a (disyllabic) trochaic Foot (see (1)).

This led Demuth (1996) to suggest that children's early words were prosodically constrained to take the form of strong-weak (stressed-unstressed) trochaic feet, such as words like *dolly*. This appears to capture the structure of many children's early word productions before the age of two, at least for a language like English, where longer words are truncated to take this form. Thus, words like *banana* are often first realized as *nana*. Around two to six years, children begin to produce unstressed (unfooted) word-initial syllables, finally producing full word forms like *banana*.

Interestingly, some children preserve the word-initial consonant /b/, and produce truncated forms like *bana*, instead of *nana*. This can be thought of in terms of competing "constraints," where the segments of the target word can be mapped into the output in a variety of ways. Thus, *banana* can be realized as either [ˈbænə] or [ˈnænə]. Both are prosodic words that take the shape of a trochaic foot, but the segments that make up that foot are drawn from different parts of the original target word.

Note that the mapping of the initial /b/ of the word *banana* into the output form provides strong evidence that the initial unstressed syllable is actually parsed, and that its omission is not due to a lack of perceiving it. Rather, the disyllabic output seems to be a phonological constraint on output form. Such constraints have been interpreted as the child producing early "unmarked" structures. Thus, early words like *nana* can be thought of as instances of the early "emergence of the unmarked" in child speech (Prince & Smolensky, 2004).

This tendency to truncate words to a disyllable lasts for several months in English and Dutch. But in languages where many more words are longer, as in Spanish or Italian (Roark & Demuth, 2000), these early unmarked forms quickly disappear. Thus, although Spanish-speaking children truncate early words such as *muñeca* "doll" to [ˈmɛka] around one to six years of age (replacing the more challenging nasal with the word-initial /m/), longer words begin to be produced before the age of two (Gennari & Demuth, 1997).

Thus, children's early word shapes gradually become more complex. Words such as *elephant*, with primary stress on the first syllable and secondary stress on the last syllable, tend to be produced by preserving both types of syllables by two years, resulting in word forms such as [ˈɛ.fɑn], in both English (Kehoe & Stoel-Gammon, 1997) and Dutch (Fikkert, 1994). Sometimes such word forms may actually be composed of two feet, e.g., [ˈɛˈfɑn]. Thus, one begins to find an "unfolding" of prosodic structure (cf. Frota, 2012; Frota, Cruz, Matos, & Vigário, 2016), both at the level of the simple PW, but also at higher levels of prosodic structure. To explore this development in more detail, we need to consider the emergence of morphologically complex prosodic words; another fascinating chapter in the development of children's phonological grammar.

29.6 The acquisition of higher levels of prosodic structure

It has long been known that some children produce what have been called "filler syllables" (e.g., Peters, 1983; Veneziano & Sinclair, 2000). These typically take the form of a reduced

vowel that tends to occur where a grammatical morpheme (e.g., article, pronoun, auxiliary verb, and so on) would be expected. Thus, early attempts at multiword utterances might take the form of something like "a go now" for "I'm going now." Such early stages of development suggest that some children attempt forms that they cannot completely produce, and that this may be an area with much individual variation. Critically, the acquisition of these higher levels of prosodic structure involve longer, morphologically, and syntactically more complex utterances. It has therefore often been treated as the domain of morpho-syntax. However, our research reveals that much of the reported within-speaker variability in the apparently gradual emergence of grammatical morphemes is due to principled interactions at the phonology-morphology interface (Demuth, 2015). Thus, as a child's increased complexity at the level of the PW unfolds, so does complexity at the higher levels of the PP and the IP. Different developmental paths are found cross-linguistically because of language-specific PW and higher-level prosodic structures, allowing for predictions to be made about how and when certain structures will be acquired. This can be captured in terms of the Prosodic Licensing Hypotheses (Demuth, 2014), as explained next (see also Lleó, 2003).

Much of this research began in the early 1990s, where researchers began to move beyond the simple word to explore morphologically more complex structures (e.g., Selkirk, 1996). This then was employed to examine how and when children began to produce grammatical function words such as articles, pronouns, and so on. Many of these were either omitted and/or produced as "filler syllables" prior to being produced in their full forms. Yet even filler syllables tend to appear in certain prosodic contexts. Typically, they occur in contexts where they can be prosodified as part of a disyllabic foot. For example, in the southern Bantu language Sesotho, where nouns take a singular/plural noun class prefix, children often truncate a word like *ba-sadi* "women" to *sadi*, appearing to omit the noun class prefix *ba-* (Demuth, 2003). This suggested that this might simply be a problem of knowing which of several noun class prefixes the noun belonged to. However, follow-up corpus analysis found that Sesotho-speaking children did produce noun class prefixes when these occurred on monosyllabic lexical stems (e.g., *ba-tho* "people" was produced in its full form; cf. Demuth, 2001; Demuth & Ellis, 2009). This suggested that children have no problem with noun class prefixes per se, but are merely truncating morphological complex prosodic words to a disyllabic foot—much in the way that English *banana* was reduced to *nana*. This was further confirmed by the fact that nominal agreement was appropriately marked on postnominal modifiers, even when the prefix was omitted on the noun itself (e.g., *ba-sadi ba-ne* produced as [sadi *ba*ne] "those women") (Demuth & Weschler, 2012). These findings raise many questions about the nature of phonological and morphological representations, and whether omitted (weak) syllables/morphemes are actually perceived. That is, perhaps they are omitted due to a lack of being perceived. However, the Sesotho data just mentioned again suggest that the morpheme is perceived even though it is not produced, as agreement errors do not occur.

Interestingly, Gerken and McIntosh (1993) showed that English-speaking children under the age of two years show different behavior when presented with a nonce word in a position where "the" should be, even if they consistently omit "the" in their speech productions. This strongly suggests that these function word omissions are due to constraints on production rather than perception. To probe the nature of these limitations further, Gerken (1994, 1996) showed that children are more likely to produce articles before an object when the article is proceeded by a monosyllabic verb rather than a disyllabic verb. That is, when an

article can be prosodified as part of a Foot with the proceeding word, as in *Tom [saw the]$_{FT}$ [chickens]$_{FT}$*, it was more likely to be produced than when the article was unfooted, as in *Tom's [washing]$_{FT}$ the [chickens]$_{FT}$.* Unfooted articles then only appear once children have access to higher levels of prosodic structure, such as that of the IP. However, these studies were conducted in an elicited imitation task with two- to three-year-olds, raising questions about what happens in everyday spontaneous speech. Demuth and McCullough (2009), using spontaneous speech productions from children aged one to three in the Providence Corpus (Demuth et al., 2006), then showed exactly the same results over developmental time.

These are very interesting results, for a variety of reasons. Like the results from Sesotho, they point to the fact that much of the variable production of many grammatical morphemes may be prosodically conditioned. That is, grammatical morphemes will be more likely to appear during early stages of development in precisely those contexts where they can be prosodified as part of a Foot. The Prosodic Licensing Hypothesis (Demuth, 2014) therefore provides a principled means of making predictions about the course of early morphological production across languages. If one knows about how different grammatical morphemes are prosodified in the target language, it is then possible to make explicit predictions about which morphemes in which prosodic contexts will be more likely to be acquired first, and which are more likely to appear later, when the syntax and semantics are held constant.

The Prosodic Licensing Hypothesis thus provides a principled explanation for why, once a child has some knowledge of the syntax and the semantic environment in which a grammatical morpheme must be used, the actual use of the morpheme is still highly variable. It also provides a framework for making predictions across speakers, suggesting that similar patterns of use and/or omission will appear.

This can then be tested cross-linguistically. For example, how are articles (and determiners more generally) acquired in Romantic languages such as French, Spanish, and Italian? To test this hypothesis, we need either an experimental paradigm like that used by Gerken (1996), and/or longitudinal data of children's spontaneous speech productions. We are fortunate, as part of the CHILDES database (cf. MacWhinney, 2000), to have French longitudinal data in the form of the Lyon Corpus (Demuth & Tremblay, 2008) and Spanish longitudinal corpora from several sources (cf. Demuth, Patrolia, Song, & Masapollo, 2012). Both studies have explored this issue in detail. As in the Sesotho case, French-speaking 1;10-year-olds were more likely to use articles with monosyllabic (or truncated) nouns (e.g., *la couronne*/laku^{l}ʁɔn/ > [la^{l}ʁɔn] "the (fem.sg.) crown"), and use articles with disyllabic nouns by around the age of two years (e.g., *les poubelles*/lepu^{l}bɛl/ > [lepu^{l}bel] "the (pl.) garbage"). Thus, despite the fact that French has a very different prosodic system, with phrase-final lengthening (vs. phrase penultimate lengthening in Sesotho (Doke & Mofokeng, 1957)), French- and Sesotho-speaking children showed the same patterns of early inclusion of prenominal morphology with monosyllabic words, and only later inclusion of grammatical morphemes with disyllabic words.

But what about Spanish, with a tendency toward penultimate stress? As mentioned, children under the age of two tend to truncate Spanish words to a disyllable [meka] for *muñeca* "doll"). How and when do articles begin to appear? About a third of the words that a Spanish-speaking child hears contain three or four syllables (Roark & Demuth, 2000). For these longer words, Spanish-speaking children take one of two paths to prosodically incorporate articles. Some add determiners to these truncated forms, producing morphologically complex trisyllabic prosodic words like [a^{l}meka] for *la muñeca* "the doll." (Similar findings

are reported for Italian; cf. Giusti & Gozzi, 2006.) However, other Spanish-speaking children prefer not to truncate the lexical item, producing full trisyllabic words like *muñeca* before eventually producing four-syllable morphologically complex prosodic words like *la muñeca* (Demuth et al., 2012; Gennari & Demuth, 1997; Lleó & Demuth, 1999). However, these studies of Spanish have only looked at a few children, as these were the only longitudinal phonemically transcribed data available at the time that allowed for a full assessment of these issues, and the Italian study only looked at one child. It would therefore be very interesting to explore these findings more fully, either with other children with newly available longitudinal data, and/or by conducting cross-sectional experiments similar to that used in Gerken (1996). This would provide further information regarding the nature of children's developing phonology representations, both at the level of the PW, as well as at the higher level of the PP and the IP.

These issues also become highly relevant in considering the variable production of inflectional morphemes as the ends of words, such as English plurals, third person singular, and past tense (e.g., *dogs, hits, washed*). Inspired by research with children with SLI showing worse performance on past tense morphemes with increasingly complex consonant clusters (e.g., *sewed*/sod/, *whipped*/wɪpt/, *danced*/dænst/; see Marshall & van der Lely, 2007), a large body of research has now shown that much of the within-speaker variability found in the production of inflectional morphemes is due to the prosodic context in which they appear. Thus, in addition to being better at producing inflectional morphemes appearing in simple rather than complex coda consonants (e.g., *sees* vs. *hits*), young children are more likely to produce these morphemes utterance-finally compared to utterance medially. This is because the final syllable of the phrase in English is lengthened (phrase-final lengthening), providing ample opportunity to articulate all the consonants (e.g., Kirk & Demuth, 2006) and the inflectional morphemes at the end of the word. Furthermore, children have acquired phrase-final lengthening by the age of two (e.g., Snow, 1994). In contrast, producing these inflectional morphemes is much more challenging utterance medially, where no such lengthening occurs (cf. Song, Sundara, & Demuth, 2009; see also Hsieh, Leonard, & Swanson, 1999). This results in utterance-medial inflectional morphemes being more often omitted during the toddler years, when lexical and morphological representations are still developing (cf. Mealings & Demuth, 2014; Theodore, Demuth, & Shattuck-Hufnagel, 2015). They are also subject to greater omission as function of increased utterance length, so omissions continue to occur even at three years of age, as children's syntax, semantics, and utterance lengths increase (e.g., Song et al., 2009; Valian, 1991).

These results have again raised many questions about how these grammatical morphemes might be perceived. Studies now also show that infants' perception of third person singular –s is much better utterance-finally compared to utterance medially, where the fricative is about half the duration (e.g., *She cries now* vs. *Now she cries*) (Sundara, Demuth, & Kuhl, 2011). This has now also been replicated with adults using online neurological measures employing EEG (Dube et al., 2016). This raises the possibility that increased perceptual salience may play a more important role in understanding both the development of early grammars and variability in production than often thought. For example, recent findings from children's eye-tracking shows that two-year-olds are sensitive to plural –s (e.g., *cats*), but not plural –z (e.g., *dogs*), in a task using novel words (Davies, Xu Rattanasone, & Demuth, 2017). Plural /s/ is not more frequent, but it is longer

in duration than /z/, suggesting that this enhances its perceptual salience. However, frequency can also play a role, as in the case of the later acquired –es allomorph in words like *bus-es* (Tomas, Demuth, & Petocz, 2017). Thus, although this allomorph might be thought to be more perceptually salient by virtue of its being an entire syllable (rather than merely a segment), its overall low frequency in the English lexicon (only 5% of the plural input young children hear) appears to contribute to its later acquisition, in both perception and production.

29.7 Discussion

We have outlined in this chapter the course of prosodic word development, and discussed the fact that grammatical morphemes (and unfooted syllables more generally) are often omitted in children's early speech, then slowly begin to appear as children's prosodic representations become more complex. We have also shown how prosodic factors, such as *where* in the utterance grammatical morphemes appear, can contribute to the morpheme's greater perceptual salience, thereby increasing the likelihood that it will be perceived and produced. Appealing to the Prosodic Licensing Hypothesis, as well as frequency factors and perceptual salience, helps to explain much of the within-speaker and cross-linguistic variability in the use of grammatical morphemes that has long been attested, not only in typically developing children's speech, but also in children with language disorders such as SLI (cf. Tomas et al., 2017). Note that much of this research involves not only careful control of the phonological and prosodic context in which early words and grammatical morphemes occur, but also acoustic analysis of the forms themselves in these different prosodic environments (cf. Theodore, Demuth, & Shattuck-Hufnagel, 2012).

What is still not clear is how and when the higher levels of the PP and IP are acquired, especially with respect to the use of prosodic clitics, both within and across languages. These issues have begun to be explored in European Portuguese (Frota et al., 2016), with much more to be done for English and other languages (see Leonard, 2016, for a discussion of nominal morphology).

This research then also raises many questions about how and when other populations (early L2 learners, bilinguals, children with hearing loss) develop perceptual sensitivity to grammatical morphemes, and use them in early speech. A better understanding of prosodic phonology more generally will be needed to inform the processes of speech planning and production, and how these develop over time. It is hoped that this chapter will provide a framework for exploring many of the factors that play a role in acquisition at the prosody/morphology interface, both across populations and languages. The findings will be both theoretically and methodologically important for the field of language development, with important implications for clinical intervention.

Acknowledgments

This work was funded in part by the following grants: ARC FL130100014, ARC CE110001021.

References

Archibald, J., & Carson, J. (2000). The acquisition of Québec French stress. Paper presented at the Annual Meeting of the Canadian Linguistic Association, University of Alberta.

Bortfeld, H., Morgan, J. L., Golinkoff, R. M., & Rathbun, K. (2005). Mommy and me: Familiar names help launch babies into speech stream segmentation. *Psychological Science*, *16*, 298–307.

Brown, R. (1973). *A First Language*. Harvard University Press, Cambridge, MA.

Davies, B., Xu Rattanasone, N., & Demuth, K. (2017). Two-year-olds' sensitivity to inflectional plural morphology: Allomorphic effects. *Language, Learning and Development*, *13*, 38–53.

Demuth, K. (1996). Alignment, stress and parsing in early phonological words. In: Bernhardt, B., Gilbert, J., & Ingram, D. (Eds.), *Proceedings of the International Conference on Phonological Acquisition* (pp. 113–24). Cascadilla Press, Somerville, MA.

Demuth, K. (2001). Prosodic constraints on morphological development. In: Weissenborn, J., & Höhle, B. (Eds.), *Approaches to Bootstrapping: Phonological, Syntactic and Neurophysiological Aspects of Early Language Acquisition* (pp. 3–21). John Benjamins, Amsterdam.

Demuth, K. (2003). The acquisition of Bantu languages. In: Nurse, D., & Phillipson, G. (Eds.), *The Bantu Languages* (pp. 209–22). Routledge, London.

Demuth, K. (2006). Cross-linguistic perspectives on the development of prosodic words. Guest Editor, Special Issue, *Language and Speech*, *49*(2), 129–297.

Demuth, K. (2014). Prosodic licensing and the development of phonological and morphological representations. In: Farris-Trimble, A., & Barlow, J. (Eds.), *Perspectives on Phonological Theory and Acquisition: Papers in Honor of Daniel A. Dinnsen* (pp. 11–24). Language Acquisition and Language Disorders Series, 56. John Benjamins, Amsterdam.

Demuth, K. (2015). The acquisition of prosodic phonology and morphology. In: Bavin, E., & Naigles, L. (Eds.), *Cambridge Handbook on Child Language* (pp. 230–49). Cambridge University Press, Cambridge.

Demuth, K., Culbertson, J., & Alter, J. (2006). Word-minimality, epenthesis, and coda licensing in the acquisition of English. *Language & Speech*, *49*, 137–74.

Demuth, K., & Ellis, D. (2009). Revisiting the acquisition of Sesotho noun class prefixes. In: Guo, J., Lieven, E., Budwig, N., Ervin-Tripp, S., Nakamura, K., Ozçalikan, S. (Eds.), *Crosslinguistic Approaches to the Psychology of Language: Festschrift for Dan Slobin* (pp. 93–104). Psychology Press, New York, NY.

Demuth, K., & Johnson, M. (2003). Truncation to subminimal words in early French. *Canadian Journal of Linguistics*, *48*, 211–41.

Demuth, K., & McCullough, E. (2009). The prosodic (re)organization of children's early English articles. *Journal of Child Language*, *36*, 173–200.

Demuth, K., Patrolia, M., Song, J. Y., & Masapollo, M. (2012). The development of articles in children's early Spanish: Prosodic interactions between lexical and grammatical form. *Linguistic Interfaces and Language Acquisition in Childhood*. Rothman, J., & Guijarro-Fuentes, P. (Eds.), *First Language*, *32*, 17–37.

Demuth, K., & Tremblay, A. (2008). Prosodically-conditioned variability in children's production of French determiners. *Journal of Child Language*, *35*, 99–127.

Demuth, K., & Weschler, S. (2012). The acquisition of Sesotho nominal agreement. *Morphology*, *21*, 67–88.

Doke, C., & Mofokeng, S. (1957). *Textbook of Southern Sotho Grammar*. Longman, Cape Town.

Dube, S., Kung, C., Peter, V., Brock, J., & Demuth, K. (2016). Effects of sentence position and type of violation on the auditory processing of subject-verb agreement: An auditory ERP study. *Frontiers in Psychology*, *7*, 1276.

Fikkert, P. (1994). *On the Acquisition of Prosodic Structure* [Dissertation]. University of Leiden, Leiden.

Frota, S. (2012). Prosodic structure, constituents and their representations. In: Cohn, A., Fougeron, C., & Huffman, M. (Eds.), *The Oxford Handbook of Laboratory Phonology* (pp. 255–65). Oxford University Press, Oxford.

Frota, S., Cruz, M., Matos, N., & Vigário, M. (2016). Early prosodic development: Emerging intonation and phrasing in European Portuguese. In: Armstrong, M., Henriksen, N. C., & Vanrell, M. M. (Eds.), *Intonational Grammar in Ibero-Romance: Approaches Across Linguistic Subfields* (pp. 295–324). John Benjamins, Philadelphia, PA.

Gennari, S., & Demuth, K. (1997). Syllable omission in Spanish. In: Hughes, E. M., & Green, A. (Eds.), *Proceedings of the 21st Annual Boston University Conference on Language Development* (pp. 182–93). Cascadilla Press, Somerville, MA.

Gerken, L. A. (1994). Young children's representation of prosodic structure: Evidence from English-speakers' weak syllable omissions. *Journal of Memory and Language*, *33*, 19–38.

Gerken, L. (1996). Prosodic structure in young children's language production. *Language*, *72*, 683–712.

Gerken, L., & McIntosh, B. (1993). The interplay of function morphemes and prosody in early language. *Developmental Psychology*, *29*, 448–57.

Giusti, G., & Gozzi, R. (2006). The acquisition of determiners: Evidence for the full competence hypothesis. In: Belletti, A., Bennati, E., Chesi, C., Di Domenico, E., & Ferrari, I. (Eds.), *Language Acquisition and Development: Proceedings of GALA 2005* (pp. 232–8). Cambridge Scholars Press, Newcastle upon Tyne, UK.

Hsieh, L., Leonard, L. B., & Swanson, L. A. (1999). Some differences between English plural noun inflections and third singular verb inflections in the input: The contribution of frequency, sentence position, and duration. *Journal of Child Language*, *26*, 31–543.

Jakobson, R. (1941). *Kindersprache, aphasie und allgemeine lautgesetze*. Almqvist & Wiksells Boktryckeri, Uppsala.

Jusczyk, P. W., Cutler, A., & Redanz, N. J. (2002). Infants' preference for the predominant stress patterns of English words. *Child Development*, *64*, 675–87.

Kehoe, M., & Stoel-Gammon, C. (1997). Truncation patterns in English-speaking children's word productions. *Journal of Speech and Hearing Research*, *40*, 526–41.

Kehoe, M., Hilaire-Debove, G., Demuth, K., & Lleó, C. (2008). The structure of branching onsets and rising diphthongs: Evidence from the acquisition of French and Spanish. *Language Acquisition*, *15*, 5–57.

Kent, R. D. (1976). Anatomical and neuromuscular maturation of the speech mechanism: Evidence from acoustic studies. *Journal of Speech and Hearing Research*, *19*, 421–47.

Kirk, C., & Demuth, K. (2005). Asymmetries in the acquisition of word-initial and word-final consonant clusters. *Journal of Child Language*, *32*, 709–34.

Kirk, C., & Demuth, K. (2006). Accounting for variability in 2-year-olds' production of coda consonants. *Language Learning and Development*, *2*, 97–118.

Leonard, L. B. (2016). Noun-related morphosyntactic difficulties in specific language impairment across languages. *First Language*, *6*, 3–29.

Lleó, C. (2003). Prosodic licensing of codas. *Probus*, *15*, 257–81.

Lleó, C., & Demuth, K. (1999). Prosodic constraints on the emergence of grammatical morphemes: Crosslinguistic evidence from Germanic and Romance languages. In: Greenhill, A., Littlefield, H., & Tano, C. (Eds.), *Proceedings of the 23rd Annual Boston University Conference on Language Development* (pp. 407–18). Cascadilla Press, Somerville, MA.

Marshall, C., & van der Lely, H. (2007). The impact of phonological complexity on past tense inflection in children with Grammatical-SLI. *Advances in Speech Language Pathology*, *9*, 191–203.

MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk*. Lawrence Erlbaum, Mahwah, NJ.

Mealings, K., & Demuth, K. (2014). The role of utterance length and position in three-year-olds' production of third person singular *-s*. *Journal of Speech, Language, and Hearing Research*, *57*, 484–94.

Miles, K., Yuen, I., Cox, F., & Demuth, K. (2016). The prosodic licensing of coda consonants in early speech: interactions with vowel length. *Journal of Child Language*, *43*, 265–83.

Nespor, M., & Vogel, I. (1986). *Prosodic Phonology*. Foris, Dordrecht/Riverton, NJ.

Peters, A. (1983). *The Units of Language Acquisition, Monographs in Applied Psycholinguistics*. Cambridge University Press, Cambridge.

Prince, A., & Smolensky, P. (2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Blackwell, Malden, MA.

Roark, B., & Demuth, K. (2000). Prosodic constraints and the learner's environment: A corpus study. In: Howell, S. C., Fish, S. A., & Keith-Lucas, T. (Eds.), *Proceedings of the 24th Annual Boston University Conference on Language Development* (pp. 597–608). Cascadilla Press, Somerville, MA.

Selkirk, E. O. (1984). *Phonology and Syntax: The Relation Between Sound and Structure*. MIT Press, Cambridge, MA.

Selkirk, E. O. (1996). The prosodic structure of function words. In: Morgan, J. L., & Demuth, K. (Eds.), *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition* (pp. 187–213). Lawrence Erlbaum, Mahwah, NJ.

Smit, A. B. (1993). Phonologic error distributions in the Iowa-Nebraska Articulation Norms Project: Consonant singletons. *Journal of Speech and Hearing Research*, *36*, 533–47.

Snow, D. (1994). Phrase-final syllable lengthening and intonation in early child speech. *Journal of Speech, Language, and Hearing Research*, *37*, 831–40.

Song, J. Y., & Demuth, K. (2008). Compensatory vowel lengthening for omitted coda consonants: A phonetic investigation of children's early representations of prosodic words. *Language & Speech*, *51*, 382–99.

Song, J. Y., Sundara, M., & Demuth, K. (2009). Phonological constraints on children's production of English third person singular *-s*. *Journal of Speech, Language, and Hearing Research*, *52*, 623–42.

Sundara, M., Demuth, K., & Kuhl, P. (2011). Sentence-position effects on children's perception and production of English 3rd person singular *–s*. *Journal of Speech, Language, and Hearing Research*, *54*, 55–71.

Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, *76*, 147–66.

Theodore, R., Demuth, K., & Shattuck-Hufnagel, S. (2012). Segmental and positional effects on children's coda production: Comparing evidence from perceptual judgments and acoustic analysis. *Clinical Linguistics & Phonetics*, *26*, 755–73.

Theodore, R., Demuth, K., & Shattuck-Hufnagel, S. (2015). Examination of the locus of positional effects on children's production of plural *–s*: Considerations from local and global speech planning. *Journal of Speech, Language, and Hearing Research*, *58*, 946–53.

Tomas, E., Demuth, K., & Petocz, P. (2017). The role of frequency in learning morphophonological alternations: Implications for children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *60*, 1316–29.

Valian, V. (1991). Syntactic subjects in the early speech of American and Italian children. *Cognition*, *40*, 21–81.

Veneziano, E., & Sinclair, H. (2000). The changing status of "filler syllables" on the way to grammatical morphemes. *Journal of Child Language*, *27*, 461–500.

Vihman, M. M., & Velleman, S. L. (2000). The construction of a first phonology, *Phonetica*, *57*, 255–66.

White, K. S., & Morgan, J. L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language*, *59*, 114–32.