than do French and English. Unfortunately, phoneme discrimination studies with bilingual infants are not abundant, especially those including sensitive methods such as tracking eye movements or measuring brain responses. Research is yet to demonstrate the full extent of the differences and similarities in how phoneme discrimination develops in bilingual versus monolingual children.

Paola Escudero
*University of Western Sydney*

**See Also:** Bilingual Language Development; Bilingualism: Interactions Between Languages; Early Second Language Acquisition; Early Word Learning; Electrophysiology Methods in Language Development; Phonological Development; Speech Perception and Categorization; Statistical Learning.

**Further Readings**


**Phonological and Prosodic Bootstrapping**

Around their first birthday, most infants start producing words. However, it is typically not until children have acquired some 50 words in their productive vocabulary that they start combining words into utterances. Early theories of syntactic development—based primarily on corpus work—were consequently developed to account for children’s early (syntactically deprived) utterances. By contrast, more recent work has started to focus on children’s comprehension of grammatical structures at ages that precede productive evidence of syntactic development. Such work has suggested that, within their first two to three years of life, infants gain sensitivity to many aspects of the syntactic structure of their native language. This sensitivity is thus in place even before children have acquired a rich vocabulary that may help them tune into the complexities of the ambient language.

The relatively early sensitivity to grammatical structure has raised the question of what cues enable young, preverbal children to acquire grammar. According to one account, prosodic information, and phonological cues more broadly, immediately perceivable from the speech signal, may be the key ingredient providing infants’ initial bootstrap into the language. The central tenet of this hypothesis, typically referred to as the *phonological bootstrapping hypothesis*, or bootstrapping from the signal, is based on the idea that crude structural properties of syntax are signaled by their phonological correlates. That is, children can learn certain aspects of the structure of their language through an analysis of the surface form of the incoming speech signal. Because this approach to bootstrapping into one’s native language requires no prior linguistic knowledge, it can potentially explain the earliest steps of language acquisition.

The notion that prosodic information in the speech stream contains cues to syntactic structure was first introduced in the 1980s by researchers such as Lila Gleitman, James Morgan, Elissa Newport, Ann Peters, and Eric Warren and forms the basis of the current phonological bootstrapping hypothesis. Originally called the *prosodic bootstrapping hypothesis*, a term introduced by Steven Pinker, the hypothesis mainly focused on infants’ use of prosody (i.e., information such as pitch modulation and rhythmic variation) to bootstrap into the structural properties...
of the language. Later accounts also have started to incorporate nonprosodic information available from the speech signal, such as phonetic and phono-tactic cues. To integrate this broader view, Morgan and Katherine Demuth referred to the phonological bootstrapping hypothesis, which is currently the most commonly used term.

The idea that a purely perceptual analysis of the speech signal may reveal grammatical structure rests on a few critical assumptions: that (1) structural properties of sentences are reliably correlated with prosodic or phonetic features; (2) infants are sensitive to the acoustic correlates of these prosodic features; and (3) infants are able to rely on these acoustic correlates of prosodic features during speech perception early in life. We here address each of these assumptions and discuss the plausibility of the phonological bootstrapping hypothesis in more detail.

Relationship Between Phonology and Syntax

In order for phonology to cue syntax, prosodic and syntactic phrases need to reliably coincide, and phrasal prosody should provide language learners with cues that help them divide sentences into rudimentary prosodic–syntactic units. Prosodic units should furthermore be organized hierarchically according to the depth of their acoustic markers. Evidence suggests that this is indeed the case. Specifically, phonological words are grouped into phonological phrases, which in turn constitute intonational phrases. Generally speaking, each of these levels correlates with syntactic levels of utterances, such that the prosodic structure might provide a natural bracketing of speech into syntactically relevant phrases and clauses. For example, intonational phrases—corresponding to clauses or propositions within a sentence—tend to be universally marked by phonological cues such as phrase-initial strengthening, phrase-final lengthening, pitch declination, and pauses. Phonological phrases—intermediate prosodic phrases, corresponding to syntactic phrases typically consisting of a few function and content words—are indicated by similar, but reduced, acoustic correlates. Although the exact set of prosodic cues that serve to delimit clauses and phrases differs across languages, a reliable correspondence between the prosodic and the syntactic organization is systematically found in both adult- and child-directed speech across various languages. For example, corpus work suggests that, in Japanese, much like in English, prosodic cues correctly identify clause boundaries approximately 88 percent of the time.

Of course, a correlation between the prosodic information and syntactic structure does not imply that the two are perfectly aligned. Perhaps one of the most frequently raised criticisms of the prosodic bootstrapping hypothesis concerns the caveat that there is no one-to-one correspondence between prosodic and syntactic phrases. For example, although prosodic boundaries typically correspond to syntactic boundaries, the reverse is not necessarily the case. More specifically, while clause boundaries are fairly reliably associated with prosodic cues in child-directed speech, intermediate phrasal structure has much weaker prosodic prominence. A sentence such as He is eating, for instance, is encapsulated in a single prosodic unit even though the pronoun is a clear syntactic constituent on its own, leaving the major syntactic boundary separating the subject from the verb phrase prosodically unmarked. This suggests that not all syntactic boundaries may be learnable from the phonological structure. Nonetheless, the phonological information can provide children with the first step to decode the signal and can, as such, be used as a first proxy to syntactic analyses. That is, even if some syntactic boundaries will have no clear prosodic marker, and will hence not be recovered on first pass, a phonological analysis of the speech signal would provide learners with at least some reliable cues that may guide their early syntactic analysis of sentences. Once bootstrapped, such initial syntactic information can subsequently be used to analyze the input in a more fine-grained fashion.

Prosodic phrase boundaries are not the only source of information infants could access in the speech signal to cue the syntactic structure. Within phonological phrases, the position of prosodic prominence is correlated with word order and could thus provide a perceptually available surface cue to the typology of the language. For example, prosodic prominence is phrase initial in languages where the object follows the verb (e.g., [they] ride, bikes,) and the noun follows the preposition (e.g., in, stores,), such as English or French, and is marked by a pitch accent on the initial word. By contrast, prosodic prominence is phrase final in languages where the object precedes the verb (e.g., bikes, ride,) and postpositions follow nouns (e.g., stores, in,), such as Turkish, Japanese, or Hindi, and is marked by a longer duration of the final word. Sensitivity to phrasal prosodic units may
thus help infants deduce the syntactic organization of their native language.

A second prosodic cue to the word order of a language involves the acoustic salience of individual words. Across a variety of different languages, function words (frequently occurring words with little lexical meaning, such as determiners, auxiliaries, and pronouns) tend to be shorter in duration, less stressed, phonologically reduced, and produced with lower intensity compared to content words (words that are rich in meaning, such as nouns, verbs, and adjectives). As function words are situated in initial or final position within a syntactic phrase according to the word order of the language (phrase initial for verb–object languages and phrase final for object–verb languages), this acoustic distinction between function and content words can help infants derive the basic word order of their native language.

**Sensitivity to Prosodic Features**

The mere alignment of phonological and syntactic cues, while necessary, is not sufficient for the phonological bootstrapping hypothesis to hold. In order for phonological cues to facilitate the acquisition of syntactic patterns, children should also be sensitive to the phonological manifestation of these cues. It is no surprise that adults, with years of experience of listening to their language, are able to identify prosodic boundaries with great ease. For example, naïve adults can perceptually rank prosodic boundary strengths even when the lexical content of an utterance is not accessible. Thus, even in the absence of recognizable words, adults can use acoustic cues to determine the prosodic structure of a sentence. But do infants, who have much less experience listening to language, display a similar sensitivity?

Children are known to process phrasal prosody from very early on. In fact, infants gain sensitivity to different aspects of prosody with remarkable speed. For example, immediately after birth, children discriminate disyllabic strings that differ only in the presence or the absence of a phonological phrase boundary. *Latí* extracted from *gorila tísico* (with a phonological phrase boundary) is thus perceived as being functionally different from *lati* originating from *gelatina* (without the phonological phrase boundary). Furthermore, by 2 months of age, infants are sensitive to pitch changes, and a few months later, infants begin to display sensitivity to the position of prosodic markers in fluent speech, exhibiting a preference for passages containing artificial pauses inserted at clause boundaries (e.g., *in a great big house / but it was sort of dark*) over passages containing artificial pauses inserted at other, less natural, positions within the clause (e.g., *in a great big house but it was / sort of dark*). A preference for passages containing pauses at the (somewhat less reliable) phrase boundary level (e.g., *That / looks great*) over passages containing pauses at other positions in the clause (e.g., *That looks / great*) emerges by 9 months of age.

With regard to the sensitivity to phonological correlates of word order, very young infants—between 6 and 12 weeks of age—can discriminate sentences originating from languages such as French and Turkish that differ in word order (and its associated prosodic correlates) but that are otherwise similar in their phonological properties. Young infants, even newborns, also experience no difficulty discriminating lists of function words from lists of content words, and this ability to tell apart word categories develops into a preference for content words by 6 months of age, suggesting that the ability to employ this information is in place in time to be used for phonological bootstrapping. Moreover, even before their first birthday, infants are sensitive to the frequency distribution of function and content words in their native language, leading to expectations regarding the relative position of these items in the input. Such expectations are modulated by prosodic cues. Specifically, children learning a verb–object order such as French or English expect to hear comparatively infrequent words (i.e., the content words) realized with phrase-final prominence (i.e., lengthening) at the end of utterances, whereas children learning an object–verb language such as Japanese or Hindi expect to hear such infrequent items realized with phrase-initial prominence (i.e., high pitch) at the onset of utterances. Taken together, these findings of both lines of research make a strong case for the attunement to prosodic information during early language acquisition.

**Use of Prosodic Features for Language Development**

Infants are thus sensitive to the prosodic features that correlate with the structural properties of sentences. This suggests that the first two assumptions of the prosodic bootstrapping hypothesis are met. Recent work has further started to address the third assumption, that infants use the prosodic information they have access to during sentence processing. Adults use
this information to constrain syntactic analysis. For example, when hearing a homophone such as *mort/*mord* (dead/bite), which could either be an adjective (e.g., *[le petit chien mort]*) the little dead dog or a verb (e.g., *[le petit chien][mord]*) the little dog bites), adult listeners immediately assign the appropriate syntactic category to a word, suggesting that the prosodic organization of the sentence helps parse sentences into syntactic categories.

How does this ability to utilize prosodic information online develop during the course of language acquisition? Infants have been shown to rapidly employ their early sensitivity to phonological information. By 2 months of age, infants' memory of words is enhanced when they are part of a single clause (e.g., *Cats like park benches*) as opposed to when the words occur in list form without sentence intonation (e.g., *Cats. Like. Park. Benches.*) or in separate clauses (what *cats like. Park benches are*). Slightly older infants parse and store syntactic phrases with much greater ease when they occur in prosodically well-formed phrases than when they cross clause or phonological phrase boundaries. And as early as by 13 months of age, infants, much like adults, use phonological phrases to constrain lexical access. That is, children trained to recognize the word form *paper* recognize this word when it has previously occurred as such within a phonological phrase but not when it straddles a phonological boundary (e.g., *[men with the highest pay]*). In the latter case, children recognize the word *pay* instead. Moreover, preschoolers use prosodic information to resolve syntactic ambiguities (e.g., an ambiguous prepositional-attachment sentence such as *feel / the frog with the feather* versus *feel the frog / with the feather*) online. This demonstrates that children readily employ the prosodic features of syntactic structure during sentence processing and, hence, that all three conditions underlying the plausibility of the phonological bootstrapping hypothesis have been met.

**Syntactic Skeleton**

Although the discussion of phonological bootstrapping has thus far focused on the plausibility of phonological bootstrapping on its own, this does not preclude the possibility that phonological or prosodic cues are processed in tandem with other cues in the signal. In fact, while the acoustic correlates of syntactic structure may boost sensitivity to the grammar of the language, on their own, they do not provide direct clues to the syntactic labels of constituents (e.g., noun phrase or verb phrase). To derive this information, syntactic indices such as function words need to be incorporated. This joint integration of prosodic information and frequently occurring, and acoustically distinct, function words is central to the latest theorized model of phonological bootstrapping. Anne Christophe and colleagues have proposed that infants could combine these two cues to build a partial syntactic representation of sentences: a syntactic skeleton. In such model, the prosodic boundaries, signaling the syntactic constituent boundaries, would not only aid syntactic processing but also facilitate the recognition of the function words that are situated at the edges of phonological units.

Function words are crucial in this model because they both promote access to the neighboring content word and permit one to categorize the units (verb phrase or noun phrase). For example, a sentence such as *The little dog is eating a big bone* could be analyzed as *[The X][is X-ing][VP]* where the prosodic boundaries would be indicative of the syntactic constituent boundaries, the determiner the (typically occurring before nouns but not before verbs) would label the unit as a noun phrase, while the auxiliary is (typically occurring before verbs but not before nouns) would label the unit as a verb phrase. Even without knowledge of any of the content words in the previous sentence, it is possible to compute this partial syntactic representation. Adult studies using jabberwocky sentences where all content words are replaced by invented words (e.g., *[the moop]*) show that the syntactic skeleton allows adults to determine that *moop* is a noun, while *blick* is a verb, in less than 500 milliseconds. Studies examining how young children deal with such newly created words are currently underway.

**Conclusion**

To what extent can phonological bootstrapping explain the language acquisition process? The findings reviewed here suggest that (1) structural properties of sentences are reliably correlated with prosodic features, that (2) infants are sensitive to the acoustic correlates of these prosodic features, and that (3) infants are able to rely on these acoustic correlates of prosodic features during online speech perception early in life. At the same time, syntactic information is not always reliably aligned with prosodic information, and by solely attending to the
Phonological information, infants will not be able to discover the complete syntactic system of their language. Nonetheless, phonological bootstrapping may allow the language learner to compute a first, rudimentary grouping of the input into smaller units, roughly corresponding to clauses. Subsequently, both phonological cues and other information from the speech signal, such as the presence of function words, may further refine these rudimentary groupings into phrase-like units and help detect the basic word order within those units. Other heuristics, such as syntactic, semantic, or frequency-based bootstrapping, may also play roles at this stage.

Taken together, the speech signal presents children with a window into the syntax of their language. Children rapidly learn to exploit this information and use it to acquire various aspects of the structural properties of their native language.

Marieke van Heugten
Isabelle Dautriche
Anne Christophe
Ecole Normale Supérieure

See Also: Distributional Knowledge and Language Learning; Grammatical Categories; Lexical Bootstrapping; Multiple Cues in Language Acquisition; Neonatal Speech Perception; Parsing/Sentence Processing; Speech Prosody; Syntactic Bootstrapping.

Further Readings

Phonological Awareness

Phonological awareness is the explicit understanding that spoken language comprises discrete linguistic units of sound structure, such as words, syllables, and phonemes. Phonological awareness encompasses a wide variety of skills that indicate an understanding that spoken language can be analyzed and manipulated based on the sound structure of words alone, independent of word meaning. Phonological awareness primarily develops during the preschool years and is considered one of the strongest indicators for later reading success. Phonological awareness is measured by tasks along a hierarchy of both linguistic complexity (e.g., words, syllables, and phonemes) and cognitive demand (e.g., recognition or manipulation).

Spoken language does not naturally provide acoustic cues to distinguish the beginnings and endings of words or individual phonemes in words. Early in development, infants and toddlers process spoken language holistically with little awareness that spoken language comprises discrete words, syllables, and phonemes. Vocabulary expansion is linked to an increased sensitivity to individual sounds because of the need to distinguish between different words with similar sounds. Children's awareness of the sound structure of language develops along a predictable continuum from phonological sensitivity to larger units of spoken language (e.g., words) to phonological sensitivity to the smaller units of sound in language (i.e., phonemes).

In the early preschool years, children begin to recognize that words comprise individual syllables, and as such, they are able to clap or tap out the number of syllables in various words. This is an important step in phonological awareness as it involves parsing words into segments that, on their own, do not necessarily carry meaning. For example, the word computer comprises three separate syllables: com, pu, and ter. The next progression in phonological awareness is development of sensitivity for intrasyllabic word characteristics, such as onset and rime. Onset refers to the beginning consonant (or cluster of consonants) that come before a vowel in a given syllable. Rime is the portion of the syllable following the onset, which typically comprises a vowel and any subsequent consonants. For example, in the word boat, /b/ is the onset, and /ot/ is the rime. Similarly, in the word spoon, /s/ is the onset, and /un/ is the rime. Multisyllabic words have the same number of rimes as syllables. Finally, children develop an awareness of the individual