ABSTRACT—In this article, I discuss the acquisition of functional morphemes during the first years of life. Infants begin to process functional items at birth. They start encoding functional elements and their structural relations in phrases and sentences long before they can produce these items. Functional items also assist various language acquisition tasks. These findings demonstrate the initial acquisition of functional morphemes and early grammatical knowledge, challenging the view that no syntactic structures are represented before the end of the toddler years. The findings have important implications for theories of language acquisition and for the debate concerning nature versus nurture in development.

KEYWORDS—language acquisition; infant speech perception; functional morphemes; grammatical relations; linguistic structures; word learning; syntax; nature versus nurture

Functional morphemes are elements such as determiners, auxiliaries, and tense endings. Lexical morphemes (including nouns, verbs, adjectives, and adverbs) carry the main weight of meaning and are often supported by observable conceptual attributes in the world, but the meaning of functional morphemes is subtle. Functional elements, crucial for indicating grammatical structures, are thus a major focus in research on language acquisition.

Across languages, children’s speech lacks functional morphemes from the one-word stage to the initial multiword stage. This characteristic is consistent with the theory that children do not know about functional morphemes. Even when functional items emerge in speech production from age 2, their representation is the subject of debate: Some (e.g., those who subscribe to item-specific models) consider them part of memorized item-specific formulae rather than abstract grammatical categories (Pine & Lieven, 1997). According to this view, syntactic representations are created gradually at a later age, after children store many specific exemplars (e.g., Pine & Lieven, 1997; Tomasello, 2000). Others (e.g., Valian, Solt, & Stewart, 2009; Yang, 2013) suggest that children’s early multiword speech demonstrates an abstract grammar with categories including determiners. They argue that acquisition begins with innate syntactic structures.

Studies of functor omission suggest that young children represent functional morphemes. In an imitation task (Gerken, Landa, & Remez, 1990), 2-year-olds omitted functors (e.g., -es in Pete bounces the ball), but not prosodically matched nonsense functors (e.g., -a in Pete pusha ko truck), suggesting that the -es omission was intentional and reflected knowledge about the suffix. In natural production, 2-year-olds omit functors due to phonological reasons or processing limitations such as sentence complexity and memory (e.g., Valian & Aubry, 2005). For instance, French-learning children omitted determiners more before disyllabic nouns than before monosyllabic nouns because the determiner and the noun formed a cohesive unit (i.e., prosodic foot) in the latter case (e.g., du lait), but not in the former (Demuth & Tremblay, 2008). Thus, the lack of functors in production does not necessarily imply lack of representation.

Production data are limited. They offer little information about functor acquisition during the initial 2 years of life. Due to performance factors (e.g., motor limitations), production develops much later than perception. Moreover, deciding whether children’s natural production represents abstract syntax or simply item-specific chunks is challenging. One indication of an
abstract grammatical category is the overlapping coproduction of its members with a member of another category (e.g., various determiners with the same noun: *the dog*), but the degree of overlap may be low due to other factors such as corpus size, word frequency, or sampling method (Valian et al., 2009).

The prosody–functor bootstrapping model proposes the mechanism underlying initial acquisition, focusing on perception data from early infancy. This theory predicts that functional morphemes are present from the onset of acquisition. Perceptual experiments can test whether infants have an abstract grammar. For example, they can assess whether abstract grammatical categories exist by presenting overlapping use of category members (e.g., determiners) and measuring infants’ listening responses. According to this model, functional categories bootstrap language acquisition long before they emerge in production. Specifically, input contains language-general distributional and sound properties supporting the distinction between lexical and functional words, and infants can innately use those properties to perceive this distinction. Subsequently, infants discover and encode individual functors of their ambient language. The encoded forms of functional elements bootstrap early acquisition tasks, including segmentation, word learning, and grammatical categorization of lexical words. Prosody (e.g., rhythmic or intonational patterns in utterances) may work with functional elements to bootstrap syntactic learning. (For earlier versions of this model, see Christophe, Guasti, & Nespor, 1997; Christophe, Millotte, Bernal, & Lidz, 2008; Morgan, Shi, & Alloppena, 1996; Shi, 2005.) Next, I discuss empirical research that bears on the claims of this model, focusing on perceptual studies of infants demonstrating the early learning of functional morphemes and their roles in bootstrapping language acquisition.

**FUNCTIONAL AND LEXICAL MORPHEMES: LEARNING THE LANGUAGE-UNIVERSAL FUNDAMENTAL DISTINCTION FROM BIRTH**

The prosody–functor bootstrapping model highlights the most basic distinction between functional and lexical categories. This bifurcation is language-universal and superordinate over finer grammatical categories. The two syntactic categories, corresponding largely to the closed-class versus open-class distinction in psychology, differ in distributional and sound properties. Functional morphemes are a small set of items, each extremely frequent. In contrast, lexical categories contain many items (nearly the entire dictionary), each occurring infrequently. These drastic distributional differences are characteristic of children’s input speech across languages (e.g., Gervain, Nespor, Mazuka, Horie, & Mehler, 2008; Monaghan, Christiansen, & Chater, 2007; Shi, 1996; Shi, Morgan, & Alloppena, 1998). In terms of sound properties, across languages, functional elements are reduced in their spoken form. They have fewer syllables per word and fewer consonants/vowels per syllable than lexical elements, and they tend to have weak prosody (e.g., weaker in intensity, shorter in duration, less prominent in pitch; Shi, 1996; Shi et al., 1998).

The distributional, phonological, and acoustical cues in the input may jointly support the categorization of lexical versus functional items, as shown in connectionist simulations of category learning (Monaghan et al., 2007; Shi, 1996; Shi et al., 1998). Indeed, in high-amplitude sucking experiments (Shi, Werker, & Morgan, 1999), these cues guided 1- to 3-day-olds to categorize functional versus lexical items. Moreover, newborns whose mothers spoke English and those whose mothers spoke typologically different languages (but not English) categorized in the same fashion the two classes of items in English, suggesting that newborns respond innately to language-general cues.

Similarly, in preferential looking experiments, both English- and Chinese-learning 6-month-olds discriminated English lexical and functional items (Shi & Werker, 2001, 2003). Moreover, when frequency differences were removed, newborns and 6-month-olds still distinguished the two categories, indicating that language-general prosodic and phonetic cues sufficed for the initial discrimination. These findings suggest that babies are endowed with the universal mechanism to distinguish the sound forms of functional versus lexical categories.

**DISCOVERING FUNCTIONAL MORPHEMES IN THE NATIVE LANGUAGE**

Acquiring functional categories requires infants to learn individual words in the ambient language that belong to the categories. In two studies, German-learning 8-month-olds (Hohle & Weisssenborn, 2003) and French-learning 6-month-olds (Shi, Marquis, & Gauthier, 2006) segmented function word forms. After being familiarized with isolated functors (e.g., German *von, French des*), infants segmented them from longer utterances. Infants also store individual functors of their native language at an early age (e.g., Hallé, Durant, & de Boysson-Bardies, 2008; Shi, Werker, & Cutler, 2006), starting with most frequent forms. English-learning 11-month-olds’ brain responses (ERP) to English functors differed from their responses to nonsense syllables that replaced the functors in sentences (Shafer, Shucard, Shucard, & Gerken, 1998). Similarly, 11-month-olds preferred listening to utterances containing frequent functors over those containing nonsense functors (e.g., Shi, Werker, & Cutler, 2006)—that is, infants recognized familiar function words in their native language.

Infants’ initial function words are underspecified. They are specified only on the vowel. French-learning 6-month-olds stored phonetically similar functors (e.g., French *la, ta*) the same, and distinguished dissimilar ones (*la, des*; Shi Marquis, & Gauthier, 2006). The frequent functor *the* and its mispronunciation *ke* are equally familiar to French-learning 8-month-olds, more familiar than the less frequent *her* (and its mispronounced *ler*) (Shi, Cutler, Werker, & Cruickshank, 2006). Later, the forms become well specified in the representation. For example, by 11 months, the mispronounced *ke* (for *the*) is no longer
acceptable in English (Shi, Cutler, et al., 2006). In comparison to English-learning infants, French-learning infants begin to recognize and specify the forms of frequent functors earlier: At 8 months, they recognized the frequent functor *des* over the nonsense *kes* and over the less frequent functor *ros* (Shi & Lepage, 2000). Thus, infants first have a general underspecified functor class containing a few underspecified items and gradually specify more individual items within the class, working from more frequent to less frequent functors.

Infants also recognize bound functional morphemes (e.g., *-ing*), although these items are likely harder to segment than free functional morphemes (e.g., *the*). English-learning infants segment the *-ing* suffix from the stem by 15 months (Mintz, 2004). Even preverbal infants segmented bound functional morphemes (Marquis & Shi, 2012): French-learning 11-month-olds segmented the suffix *-le* based on its high frequency and understood the morphological alternations associated with this morpheme. They treated root forms (e.g., *tridel*) and their suffixed alternations (*tridel*) as related forms, even though the stimuli were nonsense words, indicating that infants’ morphological representation is generalized. This frequency-based learning of morphological alternations in preverbal infants challenges the standard view that this ability is acquired at a much later age with the help of word meaning.

**FUNCTIONAL MORPHEME BooTSTRAP LANGUAGE ACQUISITION**

Once some functors are stored in the representation, infants begin learning their structures and using them for other acquisition tasks. As I discuss later, besides learning the groupings and relations among functors, which are important components of syntactic acquisition, infants use functors as anchors for segmenting neighboring lexical words and discovering their syntactic/semantic properties.

**Learning Grammatical Relations of Functors**

Shortly after their first birthdays, infants start grouping individual functional morphemes into refined categories, suggesting that they attend to their relations with neighboring words. French-learning 14-month-olds perceived different determiners (e.g., *ton, des, le*) as a common class and distinguished them from functors of another class such as pronouns (Shi & Melançon, 2010). By 20 months, infants divided determiners into subclasses of grammatical genders, distinguishing the masculine *un* and *le* from the feminine *une* and *la* (Cyr & Shi, 2013). The function words in those studies co-occurred with nonsense nouns, so the phrases were not exemplars memorized previously. Rather, infants’ perception of functors reflected abstract categorical representation, as further shown by their responses to overlapping use of functors within the same class. For example, after hearing a novel noun *ravole* following *un*, infants perceived the new phrase *le ravole* as grammatical and the new phrase *la ravole* as ungrammatical.

Infants also analyze the grammatical relations between functional morphemes, as shown in studies using functors as well as real and nonsense lexical words. English-learning 16-month-olds perceived that sentences such as *meep are good* were ungrammatical (Soderstrom, White, Conwell, & Morgan, 2007). Upon hearing misused functors such as *‘Can you see and ball’*, 18-month-olds’ noun recognition was impeded (Kedar, Casasola, & Lust, 2006). Eighteen-month-olds preferred listening to sentences containing nonadjacent grammatical dependencies such as *is ing (e.g., … is starting …)* over the ungrammatical *can ing (… *can starting …; Santelmann & Jusczyk, 1998).* German-learning 19-month-olds had similar preferences (Höhle, Schmitz, Santelmann, & Weissenborn, 2006). Moreover, French-learning 17-month-olds perceived the functor dependencies even when the words between nonadjacent functors were nonsense lexical words (e.g., *le nonword va; Van Heugten & Shi, 2010*). Thus, their preference for grammatical dependencies reflects generalized structural learning, not word sequences previously memorized.

**Segmentation**

Word segmentation is a prerequisite for vocabulary learning. Preverbal 8- to 11-month-olds use familiar functors to segment adjacent lexical words. In one study (Shi & Lepage, 2006), French-learning infants were first presented a determiner preceding a novel noun (e.g., *des preuves*) versus a nonsense syllable preceding another noun (e.g., *kes sangles*). Infants were tested with the nouns in isolation (*preuves vs. sangles*). They preferred the noun from the determiner context, suggesting that the determiner assisted the segmentation of the adjacent word. English-learning infants responded similarly (Shi, Cutler, et al., 2006). Moreover, 11-month-olds can use their knowledge of functional suffixes (e.g., *-le* in French) to segment lexical word roots (Marquis & Shi, 2012).

Segmenting larger grammatical units (e.g., phrases) is important for syntactic acquisition. Functional morphemes tend to occur at edges of these units in input speech (e.g., Shi et al., 1998). Whether they occur at the beginning or the end of units varies according to language. Indeed, infants use functors to place the phrasal boundary consistent with their ambient language (e.g., functor at the beginning for Italian and at the end for Japanese; Gervain et al., 2008). Phrases are often produced with prosodic cues (e.g., lengthening and distinct intonation at the end of phrases) and headed by functors (e.g., *the in the little boy*). Correspondingly, infants use these correlated cues to segment phrases (Bernard & Gervain, 2012).

**Learning the Meaning of Words**

Functional morphemes affect word learning. A novel word (e.g., *larp*) in the context of functors supporting the noun category (e.g., “He is waving a *larp*”) is interpreted by toddlers as referring to the object in a scene (e.g., scene: man waving object). In contrast, if the novel word is surrounded by functors indicating
its use as a verb (e.g., “He is larping a ball”), toddlers interpret the word as referring to the action of the same scene (Waxman, Lidz, Braun, & Lavin, 2009). English-learning toddlers used neighboring functors to interpret a novel word as referring to transitive versus intransitive actions (e.g., Naigles, 1996; Yuan & Fisher, 2009). Morphosyntactic cues involving functors (e.g., a _ is _-ing) also support toddlers’ learning of the meanings of nouns and verbs (i.e., object vs. action) in other languages (e.g., French: Bernal, Lidz, Millote, & Christophe, 2007; Japanese: Oshima-Takane, Ariyama, Kobayashi, Katerelos, & Poulin-Dubois, 2011). Infants understand that functors are more structural than semantic; 17-month-olds were familiarized with a foreign language, and then they heard a functor and a lexical word from that language while viewing a novel object (Hochmann, Endress, & Mehler, 2010). They associated the object to the lexical word, not to the functor. That is, they preferred to assign meaning (i.e., semantic properties) to the lexical word rather than to the functor.

**Grammatical Categorization and Online Comprehension**

Assigning words to grammatical categories is a fundamental step in syntactic acquisition, and functors stored during early infancy bootstrap this categorization. In one study (Höhle, Weissenborn, Kiefer, Schulz, & Schmütz, 2004), German-learning infants were familiarized with pseudowords preceded by a determiner or by a pronoun, then tested with the pseudowords in novel sentences. Fourteen- to 16-month-olds used determiners to categorize nouns. French-learning 14-month-olds also categorized nouns using determiners (Shi & Melançon, 2010). After familiarization with novel words each following a determiner (e.g., ton mige; des miges; ton crale; des crales), infants categorized the words as nouns in the context of another determiner (e.g., le mige; le crale). After hearing a novel word within function word frames supporting the noun category (e.g., the _ in), English-learning 12-month-olds discriminated the use of the novel word in other noun frames versus in verb frames (Mintz, 2006). Infants also used function words to categorize novel words into refined categories such as noun grammatical genders (Cyr & Shi, 2013). Finally, in experiments using an artificial (Gómez & Lakusta, 2004) and a natural language unknown to infants (Gerken, Wilson, & Lewis, 2005), functor-like elements combined with phonological cues led to successful categorization of lexical word classes. Collectively, these studies make a compelling case for the role of function words in helping infants across languages assign words to grammatical categories. Their ability to do so for novel lexical words demonstrates abstract grammatical representation.

Functional elements not only affect the assignment of grammatical categories but also influence infants’ activation of this knowledge during language comprehension. Online comprehension studies demonstrate that toddlers use grammatical cues of function words for predicting subsequent speech (Bernal, Dehaene-Lambertz, Millote, & Christophe, 2010). In tasks that measure eye responses to a named object, the gender feature in determiners facilitated toddlers’ comprehension of the upcoming gender-agreeing noun (Johnson, 2005; Lew-Williams & Fernald, 2007; Van Heugten & Shi, 2009). The same effect was found even with novel nouns (Melançon & Shi, 2011): When presented the pairing of a novel object with a novel noun in the context of a gender-marked determiner, toddlers assigned the determiner gender to the noun. In subsequent test trials, they recognized the noun better when they heard the noun following another gender-agreeing determiner than when it followed a gender-disagreeing determiner. The word–object association had just been introduced and the test utterance had never been encountered before, indicating that toddlers automatically activated the abstract syntactic representation of the determiner and used it to interpret the upcoming noun referent. These results demonstrate that infants have a productive grammar during early acquisition.

**Functional Morphemes and Early Grammatical Knowledge**

These empirical findings tell a coherent story. Infants acquire functional elements from birth, relying on frequency and sound properties (Shi et al., 1999). They start with a broad lexical versus functional distinction. From 6 months, they begin segmenting and storing individual functional morphemes in their native language, starting with the most frequent ones. Shortly afterward, they organize functors into abstract subcategories and represent their codependent grammatical relations. Starting from the earliest stage, infants use functors to segment lexical words, learn meaning, assign lexical words to refined grammatical categories, and assist online comprehension. These findings follow the predictions of the prosody–functor bootstrapping model. The perception and use of functors at such a young age reveals grammatical knowledge in infancy that is far more sophisticated than understood previously.

These findings bear on the issue of how children acquire grammar. Acquisition theories assuming the generative grammar framework claim that learning involves mapping input to innate abstract structures. Constructivists hold the view that no grammatical knowledge is innately available and all structures are created from the input.

Functional morphemes are central to this debate. The lack of these items in early speech production could support the constructivist views. However, the same observation can be considered as the instantiation of innate knowledge later in development. The remarkable perception of functional items from birth to toddlerhood contributes crucial data. Functional morphemes emerge early in children’s grammar and affect acquisition and processing, long before children can demonstrate knowledge of functors in their speech. This is significant because functional morphemes are abstract and little environmental support is available for their meaning; it suggests that acquisition likely starts with an innate base and is not driven completely by input. The findings have implications for the nature versus nurture debate that is central to development.

In conclusion, research on infants’ acquisition of functional morphemes is valuable not only in identifying the developmental
path for an understudied element of language but also in distinguishing different theories of language acquisition and providing insights into the contributions of nature and nurture to development.

REFERENCES


