CHAPTER 9

Abstract linguistic representations and innateness

The development of determiners*

Virginia Valian
Hunter College and CUNY Graduate Center

This paper uses the syntactic category of determiner to address the issue of innateness in language acquisition. Reviewing data from infants and toddlers, I propose that categories are innate and that children show continuity in category acquisition. As development proceeds, children learn the individual words in each category in the target language and the specific syntactic properties of those words, but they do not construct the categories themselves.

1. Nativism and syntactic categories

This paper addresses the issue of innateness in language acquisition, using the syntactic category of determiner as the example. The nature-nurture controversy is alive and well, despite attempts to recast it or eliminate it (Elman, Bates, Johnson, Karmiloff-Smith, Parisi, & Plunkett 1996; Thelen & Smith 1994). It deserves to be alive and well because of the importance of the central question — whether and in what way the child's mind has content independent of experience (Spelke & Newport 1998). Since the focus of this paper is on syntax, when I speak of nativism and empiricism, I will be doing so with respect to syntax only. It is possible to be an empiricist in one domain, like syntax, but a nativist in another domain, such as semantics or cognition. Braine (1992), for example, proposed the absence of any specifically syntactic innate ideas but at the same time proposed the existence of innate ideas concerning the structure of logic, such as the notion of an argument.

* This work was supported in part by an award from the National Science Foundation to Hunter College (SBE-0123609). An early version of this paper was presented in 2006 at the Symposium on The role of function words in early language development at the International Conference on Infant Studies, Kyoto, Japan. I thank the contributors to that symposium for their papers and I thank the following contributors for their comments and suggestions about this manuscript: R. Shi, E. Johnson, P. Hallé, and A. Christophe. I also thank Sandeep Prasada and two anonymous readers for their comments.

The distinction between process or mechanism nativism and content nativism will help focus the discussion (Braine 1992). Process nativism claims that there are innate mechanisms for absorbing, learning, and remembering information. Humans are more sophisticated learners and users of information than any other species and may well use special processes unavailable to other species. Process nativism is not a locus of controversy and is not discussed here. Related to process nativism, and here subsumed underneath it, is what might be called architecture nativism (Braine 1992; Elman et al. 1996). By architecture I am not referring to how the brain is wired, but to how the mechanisms for learning and remembering are structured. Is the learning system, for example, an algorithm or is it connectionist? Architecture nativism is also not discussed here.

Both mechanisms and architecture require content. That is, the mechanisms for learning and remembering operate over content. The issue I address is whether any of that content is syntactic from the start. The controversy over nativism in syntax acquisition is thus not whether children innately have semantic or cognitive concepts, nor whether they have dispositions, nor whether they have one or another learning mechanism, nor whether language is used primarily for communicative purposes. The controversy is whether humans have innate ideas with specific syntactic content.

The nativist position with respect to syntax is that there are innate syntactic concepts; the empiricist position is that there are not. Thus, modern models that go by the names constructivism, interactionism, and emergentism are empiricist with respect to syntax. Claims to the contrary notwithstanding, such models do not stake out some neutral ground between nativism and empiricism (Elman et al, 1996; Thelen & Smith, 1994).

Categories are a good place to start in trying to choose between empiricism and nativism. That is because, on both positions, the final state includes syntactic categories and features, such as noun, verb, adjective/adverb, preposition (the lexical categories), and tense, person, determiner, and complementizer (among the functional categories and features). (Not all languages use articles like *the*, but all seem to use quantifiers, possessive pronouns, and demonstratives; they are functional categories that I am treating here as in the same class as articles.) For some areas of syntax, nativists and empiricists disagree about the nature of the final state; if there is disagreement about the final state, there is no point in comparing the theories with respect to the initial state. But, at least for English, there is agreement that the final state includes basic syntactic features and categories. Both positions also largely agree about the learning process: it is a form of pattern learning. The agreement about the endpoint and learning mechanism helps localize the areas of disagreement.

The argument concerns whether pattern learning operates over uncategorized words and yields categories as an outcome or whether pattern learning operates to match uncategorized words to categories. Empiricism takes the first position: categories are constructed. Nativism takes the second: words are mapped onto categories.

On syntactic empiricism, children create syntactic categories with no syntactic foundation: they use whatever non-syntactic concepts they possess plus the input they hear to create syntactic categories. *Lexical learning* — the term I will use for proposals that children have lexically-specific formulae (e.g., Pine & Lieven, 1997; Pine & Martindale, 1996), pivot-open grammars (Braine, 1963), or verb islands (Tomasello, 1992) — is an empiricist proposal that abjures an innate syntactic basis. According to *lexical learning*, the child has particular knowledge about particular words but each piece of knowledge is isolated from other similar pieces of knowledge and is tied to a particular lexical item. The child begins at the bottom with particulars and works her way up — at some point and in some manner — first to lexical categories, such as nouns, and after that to functional categories such as determiners. On this proposal, children around 24 months do not have any syntax. Their representations are category-free and consist of frames, such as *that's a*, into which words are slotted. The problem children face is not a mapping problem but a creation problem: creating categories from words.

On syntactic nativism, the child begins with an abstract representation of syntactic categories and has the task of mapping the input to those categories (Grimshaw 1981; Pinker 1984). Syntactic nativism is abetted by *semantic bootstrapping*. Semantic bootstrapping gives the child an entry point into the syntactic system which is already innately specified but is present at too abstract a level to allow direct mapping. The semantic correlates of words help the child map a word to an innate abstract syntactic characterization of a category. Children's first categorized words are likely to be those with clear external referents or semantic/cognitive correlates — such as nouns (Pinker 1984). Once the system has been cracked open, the child can use distributional regularities to map new words to categories and include functional categories like determiners, the category I will focus on in this chapter.

Nativism does not specify any particular set of syntactic features, though linguists in the generative tradition often assume that whatever features are part of universal grammar are innate. The nativism that I will propose here has two main claims: 1) syntactic categories are innate; 2) the general form of the grammar is innate. Although I am presupposing, for the sake of convenience, a generative type of grammar, what I am saying could be reformulated for any syntactic theory. With respect to determiners, the claims imply that the child innately knows

that a Determiner Phrase (DP) has a determiner as its head and takes a noun as its complement. Once the child has identified a noun — perhaps via semantic bootstrapping — she then has to identify items that can serve as the head of a DP. She knows that there will be heads, but not what they are.

Prosodic phrase boundaries and edges, which provide data about the beginnings and ends of phrases (Christophe, Guasti, Nespor, & van Ooyen 2003; Christophe, Millotte, Bernal, & Lidz 2008; Morgan 1996; Seidl & Johnson 2006), can help the child establish whether determiners precede or follow nouns. Not all important syntactic boundaries correspond to prosodic boundaries (Selkirk 1984), but prosodic boundaries do correspond to major syntactic boundaries and even newborns are sensitive to them (Christophe, Dupoux, Bertoncini, & Mehler 1994). Prosodic boundaries can help limit the processing of distributional regularities to clause-internal boundaries (Morgan 1996) and can help the child work out directionality patterns in her language (Christophe et al. 2003).

A lot is at stake in the controversy over whether children's early categories are syntactic. First, if categories are abstract from the earliest point at which we can measure, they are either innate (as I will argue) or acquired before child is stringing words together. Second, if categories are abstract, there is continuity in language acquisition, i.e., use of the same theoretical vocabulary in children's and adults' grammars. I will argue that there is no suggestion at any point in a child's development that there is a qualitative shift in her syntactic representations. Third, if categories are syntactic, evidence that has been claimed to show item-specific learning in early combinatorial speech is deceptive and does not really reveal lexical formulae. What counts as evidence for or against item-specific learning will have to be rethought.

How can we decide whether nativism or empiricism is a better account of category acquisition? The argument will take the form of an inference to the best explanation because there is no knock-down proof available for either position. To choose between the two positions we also need to agree on the domain of the theories, what they explain. Otherwise, we cannot compare them. I will assume that we want to explain how the child achieves the final state of category acquisition, and that that involves specifying the initial state, the role of input, and the mechanism. I will be taking a piece of that domain here: what can we infer about the initial state and the mechanism from the data available?

2. Determiners

I focus on determiners because they are less likely to benefit from semantic correlates or acoustic salience than lexical categories. Lexical categories — especially nouns, verbs, and prepositions — contribute to the determination of semantic roles. Semantic roles are related to the cognitive structure of events and might thus be more accessible to children at an early age. Of the four lexical categories, two of them — nouns and adjectives — often have referential correlates: you can point to a ball or to something that is red (though not, of course, to virtue or to something that is ineffable). Of the four lexical categories, three of them — nouns, verbs, and adjectives/adverbs — are acoustically salient thanks to a combination of stress, duration, vowel character, syllable structure, and position in the sentence (see, e.g., Monaghan, Chater, & Christensen 2005; Shi, Morgan, & Allopenna 1998). Even prepositions, the fourth lexical category, receive stress in some contexts, such as when used as a final particle: "pick this up". There is thus a weak correlation between referentiality and stress: many words with a referential correlate receive high stress, especially nouns. That cluster of properties should make lexical categories salient. For infants to both perceive and mentally represent determiners would, in contrast, be surprising: unlike lexical categories but like most functional categories, determiners have little meaning, seldom have clear referential correlates, do not participate directly in assigning thematic roles, and are "hard to hear". Compared to nouns, verbs, and adjectives, determiners have little acoustic substance (though they contribute to rhythm and coarticulation). Thus, if any categories should show late development, it is functional categories like determiners.

A key property of spoken English is the extremely high frequency of some determiners. For example, *the* and *a* account for almost 10 % of all word tokens in a sample of English texts (Kucera & Francis 1967) and in the British National Corpus, *the* is the most frequent word and *a* is the fifth most frequent word (forms of *be* are second, *of* and *and* are third and fourth). A frequency analysis of mothers' speech to 21 children from the Valian corpus (e.g., Valian 1991) shows that *you* is the most frequent morpheme, followed by *'s*, *what*, *the*, *that*, *it*, and *a* (together accounting for more than 20 % of all tokens).

Frequency matters: children attend more to high-frequency items than low-frequency items (see, e.g., the review by Werker & Yeung 2005). Adults in language experiments are notoriously sensitive to frequency (making psycholinguists' lives a nightmare). A beginning learner could, then, use determiners to establish word boundaries, to make predictions about the next word, identify new nouns, and distinguish nouns from verbs.

In an artificial language learning experiment with adults, Valian and Coulson (1988) proposed that high-frequency markers, akin to functional categories like determiners, serve as anchor points for distributional analysis. They help the learner tabulate co-occurrences (what occurs immediately after or before a marker), substitutions (what a given marker is in complementary distribution with),

and other distributional phenomena. In two versions of an identically-structured language, which differed only in how frequent a marker was relative to a content word, learners acquired the fine-grained dependency linking a marker type with a content word type *only* in the version with highly frequent markers. High frequency alone helps learners with structure. High frequency also helps whether a reference field is present or not: reference helps distinguish phrase types, but high frequency is independently helpful (Valian & Coulson). High frequency markers also help learners when they hear sentences rather than read them, and help whether intonation is list-like or phrasal (Valian & Levitt 1996).

When phonologically coherent cues for content words are combined with high- or low-frequency for content words, the two have separate but interdependent effects. Monaghan et al. (2005) adapted Valian and Coulson's (1988) language, manipulating phonological coherence of the two noun-like classes and manipulating lexical frequency within each class. In one version, each of the two classes of content words had a distinct group of phonological properties; in the other version, each class had mixed properties. In each version, each class of content words had some frequent tokens and some infrequent tokens. High frequency — now of content words rather than markers — aided learners, as did a phonologically coherent demarcation of content words, especially for low-frequency tokens. Adult learners can make use of multiple sources of information simultaneously, especially if they are in different domains (frequency, phonology, reference).

Word or frame frequency has also been manipulated in computer simulations of category acquisition (e.g., Mintz 2003; Monaghan et al 2005). Frequency has been moderately to very successful in accurately isolating at least some categories. Thus, frequency provides benefits to adults learning an artificial language and to computer simulations of category learning. This notion of frequency differs from transitional probability (though is often co-extensional with it), which infants are also responsive to (Saffran, Aslin, & Newport 1996). Frequency is not the be-all and end-all of processing, even in infancy: when speech cues to grouping conflict with transitional probability grouping cues, for example, 8-month-olds follow the speech cues, such as stress and co-articulation (Johnson & Jusczyk 2001). But highly frequent words — like determiners — can serve as anchor points around which structural analyses can take place.

3. Two-year-olds' knowledge of determiners

Some of the most skeptical claims about children's categories concern children between the ages of 2–4: they don't represent determiners, they don't represent verbs, they don't represent abstract syntactic structures like the passive. One reason for skepticism is that children frequently leave out functional categories like determiners and tense, at least in English. Even though young children's understanding of syntactic categories has been studied for decades, there is no agreement about when children have genuine categories rather than lexically-based formulae or local information about what precedes and follows individual words. The timing of the first evidence of genuine categories is an important aspect of the debate. The earlier in time that children provide evidence of categories, the better the evidence for nativism. The later, the better the evidence for empiricism.

But in neither case — early or late learning — is the evidence definitive. With respect to early learning, say age 2, there is the possibility that in an earlier period the child may have learned about categories, leaving nativism under-determined. With respect to later learning, say age 3, there is the possibility that the mapping problem for some categories is difficult and thus more extended in time, leaving empiricism under-determined. The overall pattern of the data will help make the inference to the best explanation.

On the basis of *distributional tests* of spontaneous speech from six 2-year-olds, whose Mean Length of Utterance (MLU) measured in morphemes ranged between 3–4, I have argued that 2-year-olds' grammars contain the categories determiner, adjective, noun, noun phrase, preposition, and prepositional phrase (Valian 1986). The evidence for determiners was of several sorts. One was that the children distinguished between determiners and adjectives. If children thought that determiners and adjectives were a single category — a modifier of some sort — they should use them in either order if they were both included in a phrase: "the red truck" and "red the truck" should be equally likely. But children always sequenced them correctly.

Similarly, adjectives can be repeated but (most) determiners cannot: "the green green green truck" (an actual child production) is permissible but "the the the truck" is not; "the tiny little truck" is permissible but "the my truck" is not (in English). Adjectives were infrequent, and sequenced adjectives were rarer still, but they occurred. In contrast, despite the high frequency of determiners, combinations of determiners never occurred. There was a possible exception, phrases like "this the truck". Such phrases were analyzed as missing a copula rather than as an illicit noun phrase with two determiners. The basis for the analysis was that such phrases never acted like phrases: they never occurred as either the subject of a sentence, the object of a verb, or the object of a preposition. Legitimate sequences, on the other hand, like "this truck" or "the truck" did occur in all possible positions. Although such reasoning is unexceptional linguistic reasoning, it is likely to strike the skeptic as excusing away errors. (My word-processing software puts a wavy green line under phrases like "this the truck"; it can only perform very superficial linguistic analyses.)

The children frequently failed to include a determiner when one was required. If one used a criterion like Brown's 90 % appearance in obligatory contexts as a necessary and sufficient condition for demonstrating mental representation of a category, two-year-olds would appear not to show evidence of determiners (Brown, 1973). I have argued that inconsistent use of a category or relation does not by itself argue for absence of a category (Valian 1991; Valian & Aubry 2005; Valian, Hoeffner, & Aubry 1996; Valian, Prasada, & Scarpa 2006). Instead, inconsistent use can be attributed to limitations in the child's planning and sequencing abilities, cognitive limitations, the information value of different constituents, and prosodic templates (Demuth 2001, Gerken 1996).

The child's utterances are the result of several different forces acting simultaneously (Valian & Eisenberg 1996). The child, like the adult, has a message to convey, a grammar she can use to give syntactic form to the message, and a limited set of resources. Resource allocation operates at the joint behest of message characteristics, grammar characteristics, and task characteristics. Children are, as Brown implicitly suggested, cognitive misers (Brown & Bellugi 1964; Brown & Fraser 1963). Depending on the message, the child's current grammar, and the discourse context, different parts of the message will be privileged.

The skeptic's view of the appeal to performance limitations is that they have not been demonstrated (Tomasello 2000). But the evidence does support the idea of performance limitations. For example, in an elicited imitation task, two-year-olds with low MLUs include subjects more often from short than long sentences, whereas two-year-olds with higher MLUs include subjects equally from short and long sentences (Valian et al. 1996). Longer sentences tax the child's sentence-processing abilities, leading to omission. As another example, children include more pronominal subjects in their imitations when they have a second opportunity to imitate a sentence (Valian & Aubry 2005). Having performed a partial analysis on the first attempt, children have more resources available on the second attempt, resources which are now directed toward less informative elements like pronominal subjects. As another example, children include more constituents when imitating sentences that have a predictable compared to an unpredictable direct object (Valian et al 2006). A predictable direct object requires less processing than an unpredictable one.

Are adults any different? Studies of adult processing suggest continuity: the adult processor, like the child's, is incremental and sensitive to a variety of sources of information. Message characteristics, grammar characteristics, real world knowledge, and task characteristics all play a role (e.g. MacDonald, Pearlmutter, & Seidenberg 1994; Trueswell & Tanenhaus 1994; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy 1995; Kamide, Altmann, & Haywood 2003). The limitations of the adult processor are less obvious during spontaneous speech than those of

the child's because there are few "super"-adults (such as Henry James) whose extended resources would make the average adult look limited.

Another objection on the skeptic's part is that claims about categories are interlocked: claims for determiners, say, involve claims about adjectives and nouns. The skeptic wants evidence for each category that is independent of the evidence used for any other category. From a linguistic perspective, this is impossible. The category system *is* an interlocking system and it *is* impossible to talk about one of its components in complete isolation from its other components. Similarly, knowledge of one element in the system cannot be independent of knowledge of every other element in the system. Part of understanding any individual element is understanding how that element fits into the system as a whole.

Subsequent observations have replicated Valian's (1986) category results. A longitudinal investigation of a child aged 27 months at the beginning of observations corroborated Valian's analysis, finding that early determiners were distributed across a variety of nouns and showed no semantic localization (Ihns & Leonard 1988). A second longitudinal study of the spontaneous speech of 17 children, beginning at 18 months, similarly found that children used determiners from the onset of combinatorial speech and made very few errors other than omissions (Abu-Akel, Bailey, & Thum 2004).

Elicited imitation data likewise suggest that very young children both attend to and understand determiners. Two-year-olds are more likely to repeat an English noun if it is preceded by an English determiner than if it is preceded by a short nonsense word (Gerken, Landau & Remez 1990: high MLU children, Experiment 1; low MLU children, Experiment 2). A child who hears "Pete push-o na car", for example, is less likely to repeat *car* than a child who hears "Pete pushes the car". Genuine determiners help in parsing unknown lexical categories. Children are also more likely to repeat nonsense words that occupy the determiner slot than to repeat English determiners, indicating that even very young children distinguish between English and nonsense determiners that have the same prosodic and segmental characteristics (Gerken et al. 1990). The fact that children preferentially omit genuine determiners demonstrates that they know what such words are and know that they are less important than other parts of the sentence.

Comprehension data similarly demonstrate an early sensitivity to determiners. Seventeen-month-old girls are able to choose one or another animal or doll depending on whether a determiner is used before a nonsense word to guide the child's choice (Katz, Baker & MacNamara 1974). In a later, better-controlled, experiment, children aged 31 months used the presence or absence of a determiner to guide their choice of a stuffed animal or block (Gelman & Taylor, 1984).

In sum, children who are just beginning to string words together seem to have grammars that include the abstract category determiner. But a skeptical rejoinder

is that children's achievements are more apparent than real and are due to lax criteria (Pine & Lieven, 1997; Pine & Martindale, 1996). If a child had lexically specific formulae, for example, she might appear to meet the distributional tests, but only because her repertoire consisted of stored frames extracted from the input, such as *the ball* and *a horse*. Pine and colleagues instead propose an overlap test for determiners. If the child is using frames, then *ball* will occur only with *the* and *horse* will occur only with *a*. If the child has the category determiner, *ball* should occur with both *the* and *a*, as should *horse*. The overlap test requires children to use more than one determiner (in their analyses, *the* and *a*) with the same noun. Without such overlap, they argue, the child is not operating with a category but only with lexically-specific formulae. In their analyses of children's speech, the children showed little or no overlap, in contrast to their mothers.

Valian, Solt, and Stewart (in press) agree with the overall reasoning. A child who knows that *a* and *the* are determiners should, *ceteris paribus*, use both of them before singular count nouns. But we think that Pine and colleagues' implementation of the test was faulty. Most importantly, the implementation did not take into account how often a child used a given noun with a determiner. Consider the limiting case, where a child uses a noun only once with a determiner: overlap is impossible by definition. With a small sample per child, or with children who produce many nouns only once or twice, overlap will be difficult for the observer to detect, purely on statistical grounds. We stratify the child's and parent's nouns according to how often each occurs with a determiner (and run additional tests as well). Across 21 children aged 22–32 months, and ranging in MLU from 1.53 to 4.38, there is no difference between the children and their mothers: both show overlap to the same degree.

Thus, two-year-olds, even those who are producing utterances that average less than two morphemes in length, have the determiner category. That finding undermines lexical learning — at least if the theory is offered, as it has been, as a theory for children up to the age of three. So far, nativism has the edge. In addition, the finding raises the possibility that other evidence thought to support lexical learning is also flawed, and that properly-implemented tests will reveal abstract categories across the board.

4. Infants' perception of determiners

If we accept the claim that 2-year-olds have the determiner category, we can use continuity — not simply as an argument in the abstract but empirically by demonstration. Continuity has always been a strong argument in principle. Continuity is the thesis that the child's grammar is commensurate with the adult's. As I am using the term, it means that the child's set of theoretical categories does not differ from the adult's in kind, only in degree: infants' categories are underspecified phonetically, morphologically, and syntactically. The infant begins knowing no language-specific facts about her language (other than those she acquired in utero). She has only abstract characterizations of categories and knowledge of what features they could include. She acquires language-specific information, such as which features actually are included in the language she is acquiring and what words are members of the category, after she is exposed to it. The category becomes fleshed out and mapped to lexical items but does not change its fundamental character. If the category is innate, continuity, rather than discontinuity, should be evident.

Discontinuities are costly in terms of extra theoretical apparatus: they require the introduction of different concepts and different mechanisms. If there is no good theory to account for the discontinuities, continuity is the best working assumption. To reiterate an earlier point, categories are a good domain in which to have an argument about the abstractness of children's representation because everyone agrees that children eventually do acquire categories. Eventually, determiners are part of an English speaker's grammar. In a discontinuous theory, you have to explain how they arise. Tomasello (2000) has suggested that lexically-specific learning evades the discontinuity problem, because the child goes not from one system to a different system but from no system to a system. Unfortunately, going from nothing to something is also a discontinuity. Without a mechanism in hand that succeeds in getting from one system to another, or from no system to a system, how to establish continuity remains an unsolved problem.

Although it is hard to say exactly how infants should perform if the determiner category is innate rather than constructed, we can make some educated guesses. If the child has abstract, underspecified categories, as I propose here, her errors should look like underspecification errors. If the child has lexically-specific categories, she should not make errors of underspecification because, by definition, she has encoded a specific form. The infant data I review support underspecification more than lexical specificity. Although high frequency items like determiners are, by hypothesis, underspecified phonetically and morphosyntactically, that should not interfere with their ability to serve as anchor points, as Valian & Coulson (1986) proposed. As long as they have few sound-alike competitors and are very frequent, determiners should help infants process speech, and that is what happens.

That any high-frequency item can help segment words is suggested by data on six-month-olds. They can use their own name or their mother's in possessive form to isolate nouns (Bortfeld, Morgan, Golinkoff, & Rathbun 2005). A child named Maggie who hears *Maggie's cup* is later more likely to recognize *cup* in isolation than if she had earlier heard *Hannah's cup*. Hearing *Mommy's cup* provides

a similar benefit. But whether the high-frequency items directly help parse the following word or whether the motivational importance of the child's and mother's names recruits more attention to the sentence, indirectly benefiting parsing, is not clear. Real determiners and similar sounding nonsense words provide more information.

For infants at 8 months, determiners are underspecified with respect to phonetic content. A detailed study compared 8- and 11-month-olds' ability to use real vs nonsense determiners to segment a nonsense noun from its preceding determiner, manipulating determiner frequency (Shi, Cutler, Werker, & Cruickshank 2006). Infants heard determiner-noun pairs half the time with a high-frequency real determiner (e.g., *the tink*) and half the time with a phonologically similar nonsense determiner (e.g., *kuh breek*). Other infants heard a low-frequency determiner, *her* vs *ler*.

If infants' first representations are tied to specific words, then they should be equally unable to segment words like *tink* and *breek* (i.e., equally unable to recognize them when they are presented in isolation), whether they are preceded by *the* or *kuh* during familiarization trials. Since they have never heard *tink* before, they have also never heard the sequence *the tink* before. Although the infants have heard *the* before, if *the* is tied in their representations only to nouns they have previously encountered, the sequence *the tink* should be perceived as a single two-syllable word. Since the children have never heard *kuh* before, they should similarly perceive *kuh tink* as a single two-syllable word.

Importantly, the 8-month-olds did not distinguish *the* and *kuh*, indicating that *the* is not fully specified phonetically. Eight-month olds *were* sensitive to frequency: infants who heard *the* or *kuh* were more likely to segment the accompanying nonsense noun (treating it as something they had heard before) than infants who heard *her* or *ler*. Thus, 8-month-olds can use the high-frequency determiner *the* to segment speech, but they represent it in an underspecified fashion that does not distinguish it from its phonologically similar mate *kuh*.

The 11-month-olds showed a different pattern. They distinguished between *the* and *kuh*, with the result that only *the* isolated the following nonsense noun. The other three items—*kuh*, *her*, and *ler*—did not help isolate the following noun. The step from 8 to 11 months was to develop a specific representation for the high-frequency determiner *the*, enough to distinguish it from *kuh*. The low-frequency determiner *her* was still undifferentiated from its phonologically similar nonsense mate.

Infants thus do not begin with a highly-specific representation. Instead, they have an underspecified representation of a very highly frequent form and can initially use that form to segment new words. By 11 months, the infant has phonetically specified *the* but not *her*. Infants appear to be working with the most highly frequent forms first. The 11-month-olds seem not to know just what items are included in the determiner category beyond its most frequently encountered member. They have yet to construct an equivalence class of elements that populate the determiner category. What they are missing, on this analysis, is not the category, but knowledge of all the specific elements that make up the category.

With a similar task but a different set of materials (Shi, Werker, & Cutler 2006), only 13-month-olds showed a reliable preference for a sequence of determiner-nonsense noun pairings that used real determiners (*the, his, her, their, its*) over nonsense determiners with similar sounds (*kuh, ris, ler, lier, ots*). Although it is possible that the 13-month-olds preferred the real determiners solely because they recognized *the*, that seems unlikely. Since 11-month-olds distinguish *the* and *kuh*, they would have shown the same performance as 13-month-olds if only recognition of *the* were driving the results. By 13 months, then, the infant has mapped out the phonetic form of more than one real determiner, but we do not know which ones. (English-speaking 11-month-olds demonstrate via brain responses that they distinguish a set of English words that play a functional role from phonologically similar nonsense words (Shafer, Shucard, Shucard, and Gerken 1998) but here it is possible that only *the* (out of a list consisting of *was, is, the, a, of, with*, or *that*) was responsible for the effect. The same caveat holds for preference data for 11-month-olds (Shady 1996).

Eleven-month-old French infants prefer real determiner-noun sequences over nonsense determiner-noun sequences when the noun is monosyllabic, whether the noun is familiar (*chat*) or unfamiliar (*dôme*) (Hallé, Durand, & de Boysson-Bardies 2008). When the noun is disyllabic, the infants prefer the sequence with real determiners for familiar (*canard*) but not unfamiliar (*soutard*) nouns. Since six different determiners were used, possible frequency differences among the determiners could not be determined. The overall pattern of data suggests that 11-month-olds recognize at least some familiar determiners and use them to segment monosyllabic nouns, data consistent with the reports for English that I reviewed earlier.

French 11-month-olds also prefer familiar nouns to unfamiliar nouns—but only if the noun is preceded by a real determiner, not if it is preceded by a nonsense determiner (Hallé et al. 2008). A preference for familiar nouns requires a mechanism to register frequency: a familiar noun cannot become familiar if each perception of it is erased. Infants are thus tabulating noun frequencies as well as determiner frequencies.

The acquisition sequence for specifying determiners phonologically seems, then, to have three steps: step 1) make use of very high frequency markers like determiners to segment them from content words, but (probably) use only the vowel to specify the marker; step 2) refine the representation of the marker to specify it

in more phonetic detail; step 3) proceed similarly with other markers that are still highly frequent relative to content words but less frequent than *the*.

That infants 14–16 months of age have some idea of what items constitute the class of determiners comes from a head-turning experiment with German-speaking infants. When children had familiarization trials with the indefinite article *ein* followed by a nonsense word (*glamm* or *pronk*), they subsequently showed a novelty effect, listening more to passages where the nonsense word was in a verb context rather than a noun context featuring a different determiner, such as the definite article *das* or the demonstrative *dieses* (Höhle, Weissenborn, Kiefer, Schulz & Schmitz 2004). Although several interpretations of this experiment are possible, it suggests that children recognize that *ein glamm* implies that *glamm* is a noun and accept other determiners, but not personal pronouns, as syntactically legitimate contexts.

Determiners receive a rich phonetic representation before they are fully specified morphosyntactically. Dutch-learning 24-month-olds, for example, cannot use the difference between *de* (common gender definite determiner) and *het* (neuter gender definite determiner) to guide eye gaze, indicating that gender is not well-specified (Johnson & Diks 2005). Fine-grained morphosyntactic detail concerning determiner features is thus a later occurrence than fine-grained phonetic detail. Again, the infant appears to be fleshing out a category rather than creating one *de novo*.

Dutch-learning infants at even 19 months, however, understand instructions better if they hear *de* or *het* rather than a nonsense determiner *se* (van Heugten 2006). Similar data exist for English (Gerken & McIntosh 1993; Kedar, Casasola, & Lust 2006; Zangl & Fernald 2007): 18-month-olds and older infants parse a speech stream better if they hear a genuine determiner than a nonsense form or function word from a different class (such as *and*), and, often, better than if they hear no determiner. Even though children at 18 months seldom produce determiners, their comprehension is improved when they hear real determiners, indicating that they have a determiner slot which they expect to be filled appropriately.

The final set of data suggesting underspecification comes from children who produce filler syllables, which are usually (though not always) syllables with reduced vowels (see, for example, Bottari, Cipriani, & Chilosi 1993/1994; Peters 2001; Tremblay 2005; Veneziano & Sinclair 2000). Not all children produce them, and not all children who produce them use them in exactly the same way, but there is a pattern. They appear to be positioned like syntactic markers, especially before nouns. The first function of these filler syllables may be completely prosodic — to make the child's output sound like the target language. But, later, around 19–22 months, the syllables before nouns appear to be serving a determiner-like syntactic function in European French (Veneziano & Sinclair 2000), Canadian French (Tremblay 2005), and Italian (Bottari, Cipriani, & Chilosi 1993/1994). The existence of filler syllables is easy to explain on an underspecification model, because the child has not mastered the specific knowledge about just which determiner precedes just which noun. An underspecified form meets the syntactic requirement of supplying a determiner without indicating features like number or gender. The contrasting model of lexical-specificity cannot account for filler syllables. If the child is learning, item by item, what combinations are possible, she should only produce combinations she has heard before. Filler syllables should not exist.

5. Conclusion

By age 2, there is clear evidence that children's grammars include the syntactic category determiner. Between early infancy and age 2, children's perception and comprehension of determiners and their production of filler syllables suggest that the determiner category is part of the child's grammar, but in unspecified form. The data on determiners now extend from no later than 11 months to 36 months. Determiners and their role in parsing have been more fully documented than any other functional category and as much as nouns. The data suggest that continuity is empirically true. Children never look as if they are abandoning one system for another. They never look as if they only have knowledge about individual, specific words and construct a category. Instead, they look as though they are fleshing out a category more and more fully, acquiring more and more specific details, mapping more and more items.

Where do determiners and nouns come from? This is where we came in. The inference to the best explanation is that categories are there from the beginning. The child's job is to find out what the members of those categories are in the target language she will become a speaker of and what the language-particular characteristics of those categories are. She accomplishes those tasks by age 2 for determiners.

References

Abu-Akel, A. Bailey, A. & Thum, Y.-M. 2004. Describing the acquisition of determiners in English: A growth modeling approach. *Journal of Psycholinguistic Research* 33:407–24.

Bortfeld, H. Morgan, J. Golinkoff, R. & Rathbun, K. 2005. Mommy and me: Familiar names help launch babies into speech-stream segmentation. *Psychological Science* 16:298–304.

Bottari, P. Cipriani, P. & Chilosi, A. 1993/1994. Protosyntactic devices in the acquisition of Italian free morphology. *Language Acquisition* 3:327–69.

Braine, M. D. S. 1963. The ontogeny of English phrase structure; the first phrase. *Language* 39: 1–13.

Braine, M. D. S. 1992. What sort of innate structure is needed to bootstrap into syntax? *Cognition* 45: 77–100.

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

Brown, R. & Bellugi, U. 1964. Three processes in the acquisition of syntax. *Harvard Educational Review* 34: 133–51.

Brown, R. & Fraser, C. 1963. The acquisition of syntax. In *Verbal Behavior and Learning: Problems and Processes*, C. Cofer & B. Musgrave (eds). New York NY: McGraw Hill.

Christophe, A. Dupoux, E. Bertoncini, J. & Mehler, J. 1994. Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition. *Journal of the Acoustical Society of America* 95: 1570–1580.

Christophe, A. Guasti, M. T Nespor, M. & van Ooyen, B. 2003. Prosodic structure and syntactic acquisition: The case of the head–complement parameter. *Developmental Science* 6: 213–22.

Christophe, A. Millotte, S. Bernal, S. & Lidz, J. 2008. Bootstrapping lexical and syntactic acquisition. *Language and Speech* 51: 61–75.

Demuth, K. 2001. Prosodic constraints on morphological development. In *Approaches to Bootstrapping: Phonological, Syntactic and Neurophysiological Aspects of Early Language Acquisition* [Language Acquisition and Language Disorders 24]. 3–21, J. Weissenborn & B. Höhle (eds). Amsterdam: John Benjamins.

Elman, J. Bates, E. Johnson, M. Karmiloff–Smith, A. Parisi, D. & Plunkett, K. 1996. *Rethinking Innateness: A Connectionist Perspective on Development*. Cambridge MA: The MIT Press.

Gelman, S. & Taylor, J. 1984. How two–year–old children interpret proper and common names for unfamiliar objects. *Child Development* 55: 153 –1540.

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

Gerken, L. Landau, B. & Remez, R. 1990. Function morphemes in young children's speech perception and production. *Developmental Psychology* 26: 204–16.

Gerken, L. & McIntosh, B. 1993. Interplay of function morphemes and prosody in early language. *Developmental Psychology* 29: 448–57.

Grimshaw, J. 1981. Form, function, and the language acquisition device. In *The Logical Problem of Language Acquisition*, C. Baker & J. J. McCarthy. Cambridge MA: The MIT Press.

Hallé, P. Durand, C. & de Boysson–Bardies, B. 2008. Do 11–month–old French infants process articles? *Language and Speech* 51: 23–44.

Höhle, B. Weissenborn, J. Kiefer, D. Schulz, A. & Schmitz, M. 2004. Functional elements in infants' speech processing: The role of determiners in the syntactic categorization of lexical elements. *Infancy* 5: 341–53.

Ihns, M. & Leonard, L. 1988. Syntactic categories in early child language: Some additional data. *Journal of Child Language* 15: 673–8.

Johnson, E. & Diks, M. 2005. On-line processing of grammatical gender in Dutch-learning toddlers. Paper presented at the Xth International Congress for the Study of Child Language, Berlin.

Johnson, E. & Jusczyk, P. 2001. Word segmentation by 8–month–olds: When speech cues count more than statistics. *Journal of Memory and Language* 44: 548–67.

Kamide, Y. Altmann, G. & Haywood, S. 2003. The time–course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language* 49: 133–56.

Katz, N. Baker, E. & Macnamara, J. 1974. What's in a name? A study of how children learn common and proper names. *Child Development* 45: 469–73.

Kedar, Y. Casasola, M. & Lust, B. 2006. Getting there faster: 18- and 24-month-old infants' use of function words to determine reference. *Child Development* 77: 325–38.

Kucera, H. & Francis, W. 1967. *Computational Analysis of Present–Day American English*. Providence RI: Brown University Press.

MacDonald, M. Pearlmutter, N. & Seidenberg, M. 1994. The lexical nature of syntactic ambiguity resolution. *Psychological Review* 101: 676–703.

Mintz, T. 2003. Frequent frames as a cue for grammatical categories in child directed speech. *Cognition* 90: 91–117.

Monaghan, P. Chater, N. & Christiansen, M. 2005. The differential role of phonological and distributional cues in grammatical categorization. *Cognition* 96: 143–82.

Morgan, J. 1996. Prosody and the roots of parsing. *Language and Cognitive Processes* 11: 69–106.

Peters, A. 2001. Filler syllables: What is their status in emerging grammar? *Journal of Child Language* 28: 229–42.

Pine, J. & Lieven, E. 1997. Slot and frame patterns and the development of the determiner category. *Applied Psycholinguistics* 18: 123–38.

Pine, J. & Martindale, H. 1996. Syntactic categories in the speech of young children: The case of the determiner. *Journal of Child Language* 23: 369–95.

Pinker, S. 1984. *Language Learnability and Language Development*. Cambridge MA: Harvard University Press.

Saffran, J. Newport, E. & Aslin, R. 1996. Word segmentation: The role of distributional cues. *Journal of Memory and Language* 35: 606–21.

Seidl, A. & Johnson, E. 2006. Infant word segmentation revisited: Edge alignment facilities target extraction. *Developmental Science* 9: 565–73.

Selkirk, E. 1984. *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge MA: The MIT Press.

Shady, M. 1996. Infants' Sensitivity to Function Morphemes. PhD Dissertation, SUNY-Buffalo.

Shafer, V. Shucard, D. Shucard, J. & Gerken, L. 1998. An electrophysiological study of infants' sensitivity to the sound patterns of English speech. *Journal of Speech, Language & Hearing Research* 41: 874–86.

Shi, R. Cutler, A. Werker J. & Cruickshank, M. 2006. Frequency and form as determinants of functor sensitivity in English–acquiring infants. *Journal of the Acoustical Society of America* 119: EL61–EL67.

Shi, R. Morgan, J. & Allopenna, P. 1998. Phonological and acoustic bases for earliest grammatical category assignment: A cross-linguistic perspective. *Journal of Child Language* 25: 169–201.

Shi, R. Werker, J. & Cutler, A. 2006. Recognition and representation of function words in English–learning infants. *Infancy* 10: 187–98.

Spelke, E. & Newport, E. 1998. Nativism, empiricism, and the development of knowledge. In *Handbook of Child Psychology*, 5th edn, Vol.1, *Theoretical Models of Human Development*, W. Damon & R. Lerner (eds), 275–340. New York NY: John Wiley.

Thelen, E. & Smith, L. 1994. *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge MA: The MIT Press.
Tomasello, M. 1992. *First Verbs: A Case Study of Early Grammatical Development*. Cambridge MA: Cambridge University Press.
Tomasello, M. 2000. Do young children have adult syntactic competence? *Cognition* 74: 209–53.
Tanenhaus, M. Spivey-Knowlton, M. Eberhard, K. & Sedivy, J. 1995. Integration of visual and linguistic information in spoken language comprehension. *Science* 268: 1632–1634.
Tremblay, A. 2005. On the status of determiner fillers in early French: What the child knows. In *Proceedings of the 29th Annual Boston University Conference on Language Development*, A. Brugos, M. Clark-Cotton & S. Ha, 604–15. Somerville MA: Cascadilla.
Trueswell, J. & Tanenhaus, M. 1994. Toward a lexicalist framework of constraint –based syntactic ambiguity resolution. In *Perspectives on Sentence Processing*, C. Clifton, L. Frazier & K. Rayner. Hillsdale NJ: Lawrence Erlbaum Associates.
Valian, V. 1986. Syntactic categories in the speech of young children. *Developmental Psychology* 22: 562–79.
Valian, V. 1991. Syntactic subjects in the early speech of American and Italian children. *Cognition* 40: 21–81.
Valian, V. & Aubry, S. 2005. When opportunity knocks twice: two–year–olds' repetition of sentence subjects. *Journal of Child Language* 32: 617 –641.
Valian, V. & Coulson, S. 1988. Anchor points in language learning: The role of marker frequency. *Journal of Memory and Language* 27: 71–86.
Valian, V. & Eisenberg, Z. 1996. Syntactic subjects in the speech of Portuguese-speaking children. *Journal of Child Language* 23: 103–28.
Valian, V. Hoeffner, J. & Aubry, S. 1996. Young children's imitation of sentence subjects: Evidence of processing limitations. *Developmental Psychology* 32: 153–64.
Valian, V. & Levitt, A. 1996. Prosody and adults' learning of syntactic structure. *Journal of Memory and Language* 35: 497–516.
Valian, V. Prasada, S. & Scarpa, J. 2006. Direct object predictability: Effects on young children's imitation of sentences. Journal of Child Language 33: 247–69.
Valian, V. Solt, S. & Stewart, J. In press. Abstract categories or limited — scope formulae? The case of children's determiners. *Journal of Child Language*.
van Heugten, M. 2006. Dutch-learners' use of definite articles in speech processing. MA Thesis, Radboud Universiteit, Nijmegen, The Netherlands.
Veneziano, E. & Sinclair, H. 2000. The changing status of 'filler syllables' on the way to grammatical morphemes. *Journal of Child Language* 27: 1–40.
Werker, J. & Yeung, H. 2005. Infant speech perception bootstraps word learning. *TRENDS in Cognitive Science* 9: 519–27.
Zangl, R. & Fernald, A. 2007. Increasing flexibility in children's online processing of grammatical and nonce determiners in fluent speech. *Language Learning and Development* 3: 199–231.

PART 3

Language as a formal system