

Contributions of infant word learning to language development

Daniel Swingley

Phil. Trans. R. Soc. B 2009 **364**, doi: 10.1098/rstb.2009.0107, published 16 November 2009

References

[This article cites 110 articles, 14 of which can be accessed free](#)

<http://rstb.royalsocietypublishing.org/content/364/1536/3617.full.html#ref-list-1>

[Article cited in:](#)

<http://rstb.royalsocietypublishing.org/content/364/1536/3617.full.html#related-urls>

Subject collections

Articles on similar topics can be found in the following collections

[cognition](#) (341 articles)

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click [here](#)

Review

Contributions of infant word learning to language development

Daniel Swingley*

*Department of Psychology, University of Pennsylvania, 3401 Walnut Street 302C,
Philadelphia, PA 19104, USA*

Infants learn the forms of words by listening to the speech they hear. Though little is known about the degree to which these forms are meaningful for young infants, the words still play a role in early language development. Words guide the infant to his or her first syntactic intuitions, aid in the development of the lexicon, and, it is proposed, may help infants learn phonetic categories.

Keywords: language acquisition; language development; word learning; speech perception; lexicon

1. INTRODUCTION

Infants begin learning their native language by discovering aspects of its sound structure. Precocious development of the auditory system, and innate sensitivity to acoustic variation along linguistically important dimensions, allow for rapid learning of the native language's consonants and vowels, in some cases even before children have attempted to say their first words. By the end of the first year, the average child has become attuned to his or her language with a facility that long-labouring adult second-language learners can only envy. These facts about speech-sound learning in infancy have, justifiably, captured the attention of many developmental psychologists and linguists, and have contributed to broad recognition of the importance of infant learning to language acquisition. That said, however, infants learn more than just the sounds of their language in the first year. They also learn words. This article reviews the evidence showing infants' word learning, and suggests ways in which this learning is an important contributor to the rapid pace of language acquisition in childhood (Jusczyk 1997). We begin by describing infant speech research up to the late 1980s, adopting a chronological perspective. Then we discuss subsequent work specifically exploring lexical knowledge, addressing the evidence that infants do learn words, describing how infants find those words in speech, and considering how lexical knowledge contributes to language development in infancy and beyond.

Observers as early as Taine (1876) and Darwin (1877) commented on the receptive language understanding of their infants, noting cases in which children responded appropriately to spoken words. But in contrast to these diarists' work, laboratory studies of infants have concentrated less on meaningful

interpretation of speech than on perceptual development in the categorization of speech sounds. The intellectual ancestors of experiments in this tradition are the classic studies of Eimas and his colleagues (Eimas *et al.* 1971), showing presumably innate biases in perceptual categorization, and the later studies of Werker & Tees (1983, 1984), and Kuhl (e.g. Kuhl *et al.* 1992), showing adaptation to the native language's phonology. This work, described in more detail below, used methods that were also being used in studies of categorization and sensory perception in non-linguistic domains. One consequence of this was that the modern era of controlled experimentation on infants' receptive language ability was more tightly linked to the speech science pioneered by researchers at the Haskins Laboratories (e.g. Lisker & Abramson 1967) than to the more ethological tradition of early diarists like Darwin or Leopold (e.g. Leopold 1939), a point discussed further below.

The first wave of studies of infant speech perception followed up on Eimas *et al.*'s studies of consonant discrimination. Eimas *et al.*'s experiments tested infants' abilities to tell apart the syllables [pa] and [ba], where the materials were synthesized to vary in voice onset time (VOT). VOT, the amount of time between the start of a consonant and the beginning of vocal vibrations, differentiates voiced consonants like [b], which have a short VOT, and unvoiced consonants like [p], which have a longer VOT. Adults find it easy to distinguish instances of [p] and [b] that differ only in VOT. It is much more difficult to tell apart variant realizations of [p] that are within the [p] category; likewise, different [b] sounds are hard to distinguish from one another. For example, VOTs of 0 and 20 ms both signal a [b], and are hard to tell apart; VOTs of 20 and 40 ms signal a [b] and a [p], and are easy to discriminate (e.g. Lisker 1975). In principle, this difference in detectability could be due to adults' long practice in categorizing speech sounds. But Eimas *et al.* found that 1- and 4-month-olds could discriminate between-category VOT changes (like 20 versus 40), but

*swingley@psych.upenn.edu

One contribution of 11 to a Theme Issue 'Word learning and lexical development across the lifespan'.

not within-category changes (like 0 versus 20). Infants' categories thus seemed to align with adults'. This suggested an innate basis to phonetic categorization.

Following Eimas *et al.*'s report, research from several laboratories confirmed and extended these results, using sounds drawn from a variety of languages. In most of these cases, the goal of the experiments was to probe for successful discrimination of pairs of speech sounds, and this was nearly always the result that was obtained (Aslin *et al.* 1998). Other studies revealed in infants context effects and cue-trading relations that had previously been shown in adults. For example, in the syllables [ba] and [wa] there is a characteristic pattern of phonetic changes in the transition from the consonant to the vowel. When these changes are quick, *ba* is perceived; when they are slow, *wa* is perceived. Among adults, what counts as 'quick' or 'slow' depends on the speaking rate. In fast speech, the transitions quick enough to indicate *ba* must be very speedy; in slow speech, the transitions for *ba* are of a medium duration, such that they would signal *wa* in slower speech. This *relative* interpretation of many acoustic cues is a common characteristic of speech perception. And, in several such cases, adult-like relative interpretation appears to be present in early infancy (e.g. Eimas & Miller 1980; for a review, Eimas *et al.* 1987).

Results of this sort suggested that innate predispositions could 'solve' much of the problem of speech perception for infants, with learning coming into play primarily to eliminate phonetic distinctions unused by the infant's particular language and to fine-tune the innate categories (Eimas *et al.* 1987; Kuhl 1995).

Through the 1980s, researchers' perspective shifted from this standpoint for two related reasons. One concerned the relevance of infant discrimination experiments for phonological categorization. The most interesting discovery of Eimas *et al.* (1971) was not that infants were sensitive to the distinction between [pa] and [ba]; it was that infants were apparently *only* sensitive to the distinction as implemented in many languages, including English, while failing to discriminate sounds that would count as linguistically equivalent. But many infant phonetic discrimination studies tested only phonologically relevant distinctions, without using within-category controls, and as a result could be viewed as demonstrating a generic sensitivity to acoustic variation rather than predispositions uniquely matched to the task of discovering linguistic structure. Similarly, the discovery of analogous categorization phenomena in non-human animals deflated hopes that a key to humans' unique language faculties had been found (e.g. Kuhl & Miller 1978; Kluender *et al.* 1987).

A second contributor to researchers' change in perspective came with demonstrations of infants' precocious adaptation to the native language's sound categories. Werker & Tees (1984) showed that between 6 and 12 months, English learners declined precipitously in their ability to differentiate the consonant pair [ɬi] and [ki] (glottalized uvular and glottalized velar stops from the Thompson Salish or Nthlakampx language), or the pair [t̪] and [ɖ] (dental and retroflex

stops from Hindi). For the sounds of both foreign languages, 6- to 8-month olds generally performed to criterion, while only about 60 per cent of 8–10-month olds succeeded, and very few 10–12-month olds succeeded. Changes in the perception of non-native vowels, as opposed to consonants, were shown in even younger infants (Kuhl *et al.* 1992; Polka & Werker 1994). These studies focused the research community's attention on the learning process, as opposed to the innate cognitive endowment. In the past 25 years or so these studies have been replicated and extended in several ways: identifying sources of individual variation (Kuhl *et al.* 2005; Liu *et al.* 2003), determining what aspects of language exposure are critical for category learning or maintenance (Kuhl *et al.* 2003), and examining bilingual development (Bosch & Sebastián-Gallés 2003; Burns *et al.* 2007), among others.

In the 1980s and early 1990s, several developmental psycholinguists began studying infants' interpretation of speech samples larger than syllables, a change in emphasis that led to more intensive study of word learning and precursors to syntax, as well as a greater concern for the natural ecology of the infant's speech environment. For example, Fernald (1985) tested whether 4-month-olds preferred to listen to speech delivered in the prototypical infant-directed register, with high pitch and exaggerated intonation, over speech in the relatively muted adult-directed register. They preferred the infant-directed register. This study turned out to be as influential for its methods as its conclusions. In what became known as the Head-turn Preference Procedure (HPP), infants were seated in a booth containing, on both the left and right sides, a loudspeaker with a light on it. On each trial, one of the side lights began to flash. When infants turned toward that side, speech was played from the associated loudspeaker. For each infant, either the left or right side was assigned to the infant-directed or adult-directed condition, so that more reliable orientation to a given side could be interpreted as a preference for that speech register. This procedure was modified by Hirsh-Pasek *et al.* (1987) (see also Colombo & Bundy 1981) to use listening times, rather than side preference, as the dependent measure. Continued presentation of the speech materials was made contingent on infants' continued orientation toward the active side, and trials of each sort were equally divided between the left and right sides. This modification has become standard in studies using the procedure. Results are reported in terms of infants' overall 'preference' for one stimulus type over another, averaging over trials.

The innovation of the HPP has led to an explosion of studies of infants' speech processing. In some cases, as in the Fernald (1985) study, infants are tested for preferences they already had when they were brought into the laboratory; in other cases, infants are first familiarized with speech samples, and then tested to see if aspects of this familiarization have an impact on preferences at test. Over the years, measurements of these preferences have dominated the study of infants' knowledge of linguistic structure above the level of the individual speech sound. The resulting

literature is rich and diverse, so what follows is only a brief and partial summary, focusing on early knowledge of words.

2. WORDS AS FORMS LEARNED IN INFANCY

Conventionally, ‘knowing a word’ requires knowing the sound form of the word (its sequence of consonants and vowels), and its denotation, including its semantic reference and its syntactic properties. In this sense, infants younger than 12 months or so may know very few words. But given infants’ natural attention to properties of the native language speech signal, it is reasonable to consider whether infants might learn the sound forms of words, even if their knowledge of denotation were fragmentary or absent. Evidence that infants know some word-forms has been shown most directly in studies comparing infants’ listening times for words and non-words. If infants can distinguish words and non-words, provided that the two sets are suitably matched phonologically, it suggests that they are familiar with the words from their experience with listening to speech. The first such study tested French-learning 11- and 12-month-olds. Using the HPP, Hallé and de Boysson-Bardies showed that infants listened longer to lists of words likely to be frequent in parental speech (such as *biberon*, [bibɛʁɔ̃], baby-bottle) than to words unlikely to be frequent (such as *busard*, [byzɑʁ], harrier).

This work was replicated and extended by Hallé & de Boysson-Bardies (1996) and Vihman *et al.* (2004). Both studies aimed to determine the degree to which infants retain the phonological details of the word-forms they know, by comparing familiar words, unfamiliar words and altered familiar words created by changing a word’s syllable-initial consonant into another wrong consonant. Just as words like *tummy* are likely to be more familiar to infants than words like *tenor*, and thus preferred, *tummy* should be preferred over e.g. *summy*. Likewise, if *summy* is not recognized as familiar, no preference between *summy* and an unfamiliar word is expected.

The results of these studies were complex, but overall the evidence supported three conclusions. First, infants were able to recognize words that had been produced using deviant forms, particularly when the alterations occurred in unstressed syllables (second syllables in English materials, first syllables in French materials), but infants did not recognize words consistently when the words’ stressed syllables were altered. Second, when unstressed syllables were altered, infants’ preference for deviant forms over unfamiliar words tended to appear only on later trials, suggesting that although deviant forms were recognizable in some cases, infants took longer to recognize them. These results suggested that infants do accurately retain the phonological features of the onset consonants in at least some words, though mispronunciation does not bar recognition. Swingley (2005a) extended these findings to Dutch-learning 11-month-olds, showing infant knowledge of both syllable-initial and syllable-final consonants in monosyllabic words. Third, Vihman *et al.* (2004) found that 9-month-olds did *not* show a

preference for canonically produced familiar words over unfamiliar words—a striking result given the very consistent performance of the 11-month-olds (e.g. in one experiment, 11 of 12 11-month-olds preferred familiar over unfamiliar words).

Why did 9-month-olds fail to respond differentially to potentially familiar and unfamiliar words in Vihman *et al.*’s study? One possibility is that 9-month-olds do not learn word forms. Another is that they do learn words but for some reason fail to reveal this knowledge through preferential listening. Yet, a third is that they are fully capable of learning word forms, but had not yet learned the particular words in the experiment’s stimulus set. Evidence favouring the last hypothesis is found in studies that use a training procedure. Jusczyk & Hohne (1997) visited 8-month-olds in their homes 10 times over two weeks. At each visit infants were played a 30-min recording of a woman reading three storybooks; over the 10 visits children heard the stories read by five different women. Two weeks after the last visit, children were brought into the laboratory and tested, using HPP, for their preferences for the most frequent of the content words in the stories (such as *jungle*, *python* and *sneeze*), versus foils with a similar English frequency of occurrence, and similar overall phonological shape, to the test words (e.g. *camel*, *lanterns* and *sloth*). Children showed a significant preference for the words from the stories. A control group of children who had not received prior exposure to the words showed no preference for story words over foils.

Thus, infants of 8 months can learn word forms that have no obvious relevance in the conversational interaction. These forms can be retained for at least two weeks, during which time the forms may be heard very infrequently if at all. No elaborate teaching process was required—just repetition, on average 13 times per researcher visit, in the context of stories.

In these studies showing preferences for familiar words, children heard multiple words on each trial (usually in quasirandom orders). As a result, there is no way to determine how many of the test words children knew. In principle, a preference for familiar words over unfamiliar words could be driven by only one or two words, which children seize upon each time they appear on a given trial. One way to make the procedure more informative about the number of words known is to present only subsets of the words on each trial. For example, in Swingley (2005a), half of the trials tested only animal words (the Dutch words for *bear*, *dog*, etc.) and half tested body words (*leg*, *mouth*...), and performance on each of these groups was equivalent. Thus, Dutch 11-month-olds appear to know enough animal words and body words to drive a preference for real-word lists in both categories. But it is still not clear how many words this must be. In principle, such effects could come from knowledge of only one animal word and one body word, though this sort of minimalist interpretation seems to be disfavoured by most researchers.

Some studies have avoided this problem, and also the concern raised by Vihman *et al.*’s (2004) failure to see evidence of word knowledge at 9 months, by testing only one or two words using stimuli calibrated

to the experience of each infant. For example, Mandel *et al.* (1995) presented 4.5-month-old infants with their own name, repeated several times by an unfamiliar talker, and also foil names with the same number of syllables and same stress pattern as their name. Infants preferred to listen to their own name. Although it is not clear from this demonstration how well-specified infants' knowledge of their name is phonologically, it certainly indicates a very young age at which infants appear to begin learning word forms—well before the earliest age at which children have been shown to adapt to their native language consonant or vowel categories (e.g. Polka & Werker 1994).

3. INFANTS' PROCEDURES FOR FINDING WORDS

If infants learn the sound forms of some words, which words are they, and how do infants find them? To answer these questions researchers have turned to a training version of HPP. In the first study to use this technique, Jusczyk & Aslin (1995) familiarized 7.5-month-olds with two words, and then tested whether infants would prefer little stories containing those familiarized words over stories containing other words. Familiarization was implemented as a prelude to the HPP test, using the HPP setup in which infants' orientation to the left or right side triggered presentation of a spoken list of different tokens of a word (such as *bike... bike...*). Infants accumulated 30 s of exposure to two such lists, and then passed directly to the preference test, where the auditory stimuli were 6-sentence stories. Two stories included the familiarized words (e.g. *His bike had big black wheels. The girl rode her big bike...*) and two included unfamiliarized words (which served as the familiarized words for other infants). Jusczyk and Aslin found that 7.5-month-olds listened longer to passages containing familiarized words than passages containing unfamiliarized words. The same effect was obtained when children were familiarized to the passages first and tested on the lists. This shows that children hearing full sentences can remember at least some of the frequently occurring words in those sentences, and recognize them again when hearing them in isolation. Additional studies have shown that under these testing conditions, infants do not respond to phonological variants of the tested words (e.g. familiarizing to *zeet* or *feek* and testing on the *feet* passage), showing that memory for the familiarized words is, at least to a first approximation, phonetically intact (Jusczyk & Aslin 1995; Tincoff & Jusczyk 1996).

Several factors influence the likelihood that infants will extract a given phonetic sequence from continuous speech. Infants are reluctant to consider as a single word any portions of the speech signal that straddle a prosodically marked phrase, clause or utterance boundary. Words whose edges align with such boundaries are easier for infants to detect. Seidl & Johnson (2006), for example, used Jusczyk & Aslin's (1995) procedure but systematically compared recognition of words appearing only in utterance-initial, utterance-medial, or utterance-final position, and found that

8-month-olds were better able to detect the match between words in passages (the familiarization) and in lists (test) when the words appeared initially or finally in the passages; they did not show recognition of sentence-medial words.

Other studies have demonstrated infants' use of more subtle prosodic boundary markers. Gout *et al.* (2004), using a different procedure called the Conditioned Headturn Procedure, trained 10- and 13-month-olds to look to the right to view an attractive animated display whenever they heard a target word, such as *paper* or *pay*. In this training, the target was presented as an isolated word, in the context of repetitions of another word. In a second session, children were tested for the headturn response when the target word was embedded in a sentence. Some sentences included a prosodic break amid the bisyllabic target (e.g. *The outstanding pay persuades him to go to France*), whereas other sentences included the bisyllabic target word without a break (*The scandalous paper sways him to tell the truth*).¹ Children trained on the bisyllabic word were, at both ages, much more likely to turn when hearing the second sort of sentence than when hearing the first; children trained on the monosyllable turned more upon hearing the first sort of sentence than the second, though this was only statistically significant in the 13-month-olds. Thus, very young children are sensitive to the way that prosodic phrasing 'packages' linguistic units. The same phenomenon has been documented extensively for larger units such as clauses and (some) phrases (e.g. Hirsh-Pasek *et al.* 1987; Jusczyk *et al.* 1992; Gerken *et al.* 1994; Nazzi *et al.* 2000; Seidl 2007).

Some of the intonational markers of clause boundaries are similar across languages and may be available to infants as unlearned perceptual grouping biases. But other indications of word boundaries vary considerably from one language to other, and must be learned in order to be used. The best-studied of these is the cue of lexical stress, which in some languages, including English, is a fairly good indicator of the initial syllable of a content word (Cutler & Carter 1987; Redford *et al.* 2004). In a series of studies, Cutler and her colleagues showed that English native speakers tended to interpret strong syllables as word onsets (e.g. Cutler & Norris 1988; Cutler & Butterfield, 1992). If infants were to do the same, in many cases it would help them avoid mis-segmentations, such as assuming that *a dog* is a word in *you got a dog*.² Indeed, English-learning 9-month-olds, though not 6-month-olds, prefer listening to lists of unfamiliar trochaic words (bisyllables with strong first syllables and weak second syllables) over lists of unfamiliar iambic words (bisyllables with a weak-strong pattern), as shown by Jusczyk *et al.* (1993). Such a preference suggests that 9-month-olds have established a simple model of how bisyllabic words are stressed in English. This pattern of favouring trochees is also evident, with some variation among children, in English learners' early word productions (e.g. McGregor & Johnson 1997) in contrast to French learners' largely iambic productions (Vihman *et al.* 1998), suggesting learning of language-specific production *templates* based on characteristic patterns in the words children hear.

The trochaic bias is also a parsing bias that guides English learners' word-finding in continuous speech by 8 months (Echols *et al.* 1997; Houston *et al.* 2000; Curtin *et al.* 2005; Morgan 1996; see also Polka & Sundara 2003; Nazzi *et al.* 2006, for results showing adaptation to French). English-learning 7.5-month-olds tested in the familiarization HPP procedure using bisyllables as targets are able to match isolated words to passages containing those words, and vice versa, when the bisyllables are trochaic but not when they are iambic (Jusczyk *et al.* 1999). When infants are familiarized with passages containing iambs like *guitar*, two interesting things happen in addition to the failure to match to *guitar* in isolation. First, infants do find the final syllable, preferring (e.g.) *tar* after hearing the *guitar* passage. Second, if the passages are rewritten so that a consistent unstressed word follows the iamb every time, infants seem to attach the iamb's stressed syllable to that word, forming a trochee in spite of the word boundary. In this case they do not show recognition of the stressed syllable in isolation (Jusczyk *et al.* 1999). This pattern of results is consistent with English-learning infants having a bias to extract trochees from continuous speech.

In addition to using lexical stress to guide word extraction, infants also appear to exploit knowledge of their language's consonantal patterns. Languages vary not only in which sounds they use, but also in how common each sound is, and how frequently different sounds co-occur. This *phonotactic* knowledge can be used in word segmentation. For example, the sequences [fh] and [ng] are extremely rare within words, but they do occur between words, as in the [fh] in *do it yourself honey*, or the [ng] in *you can go*. Likewise, [lg] and [ft] are relatively frequent within words like *kangaroo* and *lift* but are infrequent between words. Mattys & Jusczyk (2001; see also Mattys *et al.*, 1999) found that infants extracted *gaffe* in the context of [ng] and [fh], e.g. in the phrase *pine gaffe house ...*, but not in the context of the clusters that occur more often within words (e.g. in the phrase *strong gaffe tin ...*). Mattys and co-workers propose that infants not only know which sound sequences are most frequent in their language (e.g. Jusczyk *et al.* 1994); they also know something about how transitions from one sequence to another align with word boundaries. This may be possible only if children have antecedently extracted a 'database' of words over which to discover phonotactic generalizations, a point we will come back to below.

A third procedure infants use to discover words in speech is to group together syllables that tend to co-occur. If two syllables A and B occur together as the sequence AB most of the time that they occur at all, infants are likely to consider AB as a cohesive unit (Goodsitt *et al.* 1993). This has now been shown in numerous studies, most of which have used an 'artificial language' method. For example, Saffran *et al.* (1996) familiarized 8-month-olds to a 3-min auditory stimulus consisting of a quasirandom ordering of four trisyllabic nonsense words strung together to form a continuous sequence like *pabikututibopabikugolabu ...*. The nonsense words' syllables were each

unique to a single nonce word. The materials were 'spoken' by a speech synthesizer whose input did not distinguish word boundaries from other syllable boundaries. The positioning of words adjacent to one another resulted in some trisyllabic sequences whose syllables always occurred together (words), and other trisyllabic sequences whose syllables only sometimes occurred together, being made up of part of one word and part of another (part-words). After hearing such syllable streams, infants showed listening preferences for part-words over words. Infants could only distinguish these two stimulus types if they were able to track statistical properties of syllable co-occurrences. A follow-up study replicated this result with a modified design in which words and part-words were equally frequent in the familiarization (Aslin *et al.* 1998), showing that infants' lexical inferences were driven by *conditional probability* (A is usually followed by B, B is usually preceded by A) and not only *frequency* (AB is more common than other bisyllables). The same effect holds with real speech materials (e.g. Pelucchi *et al.* 2009).

These procedures, or processing biases, work together in guiding infants to a 'protollexicon' of word forms. Intonational grouping and statistical coherence may be useful to some degree in interpreting any language, whereas lexical prosody and phonotactic regularities are known to vary widely from one language to another. One question this naturally raises is whether the first two biases could provide the infant with sufficient information to derive the others (Thiessen & Saffran 2003). For example, do the syllable groupings that arise from statistical clustering support the generalization that English bisyllables are usually trochaic? Under some conditions, this turns out to be true (Swingley 2005b). The more frequent and statistically cohesive an English syllable pair is, the more likely it is to be a word, and also the more likely it is to be trochaic.³ An infant coming into the world with the ability to perceive and categorize syllables, and a bias to group together frequent sequences with mutually reinforcing co-occurrence statistics, could learn the English trochaic bias by constructing a lexicon first, and then reading the bias off the lexicon.

This account's developmental sequence appears to be correct: 6-month-olds presented with artificial-language materials in which stress and statistical coherence indicated different lexical groupings appeared to use the statistical coherence cue more heavily, whereas 8-month-olds clustered syllables using a trochaic strategy (Thiessen & Saffran 2003; see also Johnson & Jusczyk 2001). This suggests that in early speech segmentation, learning builds upon learning. Language-generic biases yield a database that reveals language-specific phonological patterns that are then exploited to find new words.

An additional example of this progressive learning phenomenon is shown in infants' use of learned words to detect additional ones. Words can serve as indicators of word boundaries (Brent & Cartwright 1996; Dahan & Brent 1999). If, for example, an infant knew *very* as a word form, then the sequence *they're not very tasty* could be divided into *they're not,*

very, and *tasty*. Of course, not all such segmentations would be correct or complete; *they're not* isn't a word, and knowledge of a word like *phone* could lead to misinterpretation of *tele* in *telephone*. Computational analyses of infant-directed speech corpora suggest that knowledge of word-forms would help in segmentation enough to be of significant benefit to the learner, at least in principle (e.g. Brent & Cartwright 1996; Swingley 2005b).

There is some evidence that this mechanism operates even in 6-month-olds. Bortfeld *et al.* (2005) used the HPP to test this idea in 6-month-olds, who generally fail in Jusczyk and Aslin's word-finding procedure. Bortfeld and co-workers recorded passages in which a novel word was preceded by the infant's own name, or another infant's name (e.g. *The boy played with Maggie's bike.*) Children were familiarized with both own-name and other-name passages. At test, infants showed a preference for listening to isolated repetitions of the novel word that had followed their own name, but not other children's names, relative to unfamiliarized novel words. The same effect was found using *Mommy* (or *Mama*) rather than the child's name. Children did not show any advantage in recognizing words following *Tommy* rather than *Mommy*, showing that the *Mommy* advantage could not be ascribed to local phonetic effects near the target word. (Children named *Tommy* were excluded.) This set of results suggests that words can be used as indicators of word boundaries, or that detection of words heightens infants' attention to the signal in the vicinity of those words (as in reports of the 'cocktail-party effect'). Either mechanism might apply to a broader range of words than just salient proper names, though perhaps not in children as young as 6 months of age.

A similar lexical segmentation process has been shown in studies of infants' interpretation of two-word noun phrases. French-learning 11-month-olds distinguish lists of article-noun phrases beginning with genuine articles (*le, la, les, un, une* and *des*) from phrases beginning with foils (e.g. */er/, /rœ/, /mã/...*), and furthermore only distinguish article-noun lists containing familiar and unfamiliar nouns when the preceding articles are real ones (Hallé *et al.* 2008). Similarly, phrases comprising a real English functor (*the, her, its...*) and a nonce noun (e.g. *tink*) are preferred over similar phrases with nonce functors by English-learning 11-month-olds and 13-month-olds, though not 8-month-olds (Shi *et al.* 2006b), showing developing knowledge of function words. Familiarization to phrases containing *the* and a nonce noun led 11-month-olds to prefer the familiarized nonce noun over an unfamiliarized noun, an effect that was not found upon familiarization with analogous functor-noun phrases using the pseudo-functor */kə/* (Shi *et al.* 2006a). Infants appear to treat common function words as separate from the following speech, facilitating interpretation of noun phrases containing novel nouns.

A common theme in these studies is that infants become increasingly adept at using regularities specific to their language to acquire word forms. How many

word forms might children know by the age of 12 months? Current experimental data provide little constraint on quantitative estimates of infants' lexical knowledge (Swingley 2005b). However, we can make some guesses by examining corpora of infant-directed speech and counting how many different words occur very frequently. It stands to reason that word forms infants hear most often are most likely to be learned, particularly if those forms appear in isolation (Brent & Siskind 2001), or in either utterance-initial or utterance-final position (Seidl & Johnson, 2006), though these are not prerequisites (e.g. Jusczyk & Aslin 1995). Figure 1 presents counts of frequently occurring words in an aggregated corpus of 14 mothers speaking to their 9–15-month-olds, from the Brent database. Each plot represents data from 207 h of recording. Within this period, there were about 35 000 single-word utterances, and about 245 000 word instances (tokens) appearing either utterance-initially or utterance-finally.

The lines on each graph show increases in the number of different words (types) occurring n or more times, for $n = 1, 5, 20, 50$, as the number of tokens in the corpus increases. For $n = 1$, the line is simply the type:token ratio. In the left plot, only types and tokens occurring in isolation are counted; in the right plot, only types and tokens occurring utterance-initially or -finally are counted. Suppose, for instance, a theory held that infants need to hear a word in isolation 20 times to learn it to some criterion. According to the left-hand plot, infants would know about 190 different words after interacting with their parents for about 200 h. According to the right-hand plot, if infants can learn a word after hearing it 20 times in either initial or final position, they would be expected to know about 715 words after about 100 h of parental interaction. (See Swingley (2007a) for the analogous plot assuming no sentence-position restrictions on encoding.)

These estimates of word counts per hour were derived directly from the time-annotated Brent and Siskind transcriptions (Brent & Siskind 2001). It is not clear how many hours in each day parents keep up their linguistic interaction at the levels recorded in this corpus, and even within this corpus different mothers varied substantially, with words per hour varying from about 1000 to 3200 (cf. Hart & Risley 1995). One study of words in speech recorded by 78 mothers measured monthly over a six-month period estimated average daily word exposure at about 11 300 words, with an s.d. among families at about 4200 words (Gilkerson & Richards 2007). The Brent data plotted on the graphs is drawn from about 140 000 utterances totalling 488 000 words; by Gilkerson and Richards' estimate, this token count might be achieved in about six weeks for the average parent.

Of course, as noted above, children's word segmentation cannot be perfect, and so it is to be expected that a substantial number of partial words and non-lexical portmanteaux would be included among the forms infants treat as familiar. In addition, these plots present counts over orthographic tokens, not words as uttered, perceived and identified, so some

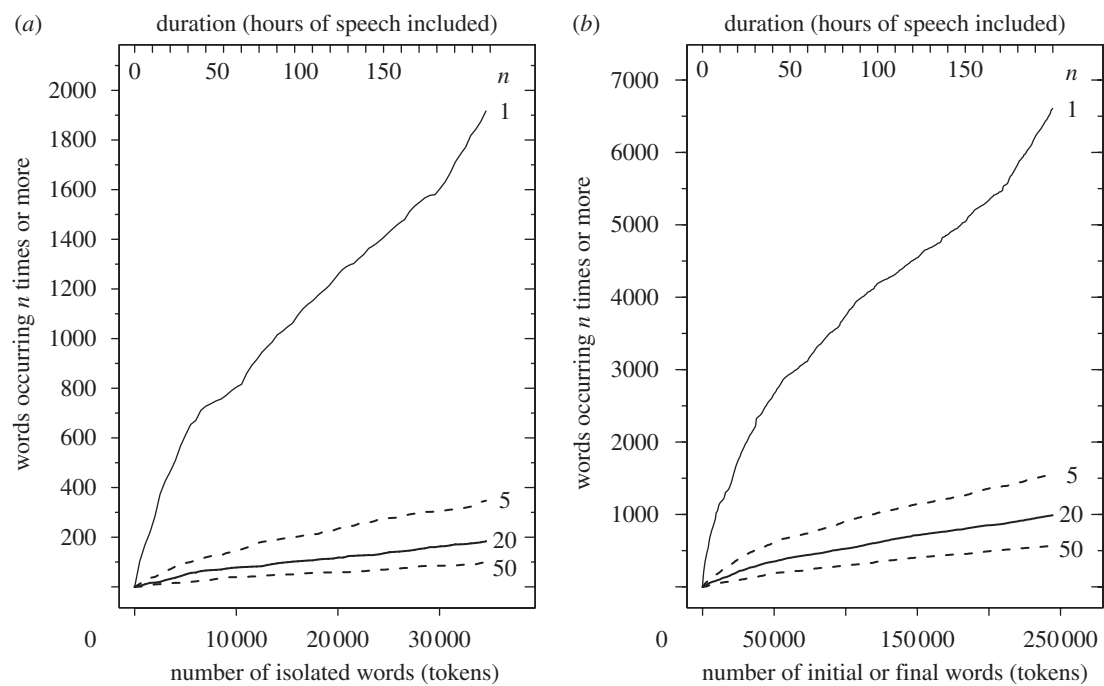


Figure 1. Counts of frequent word forms in the Brent infant-directed speech corpus. Plotted lines show the number of different words occurring $n = 1, 5, 20, 50$ or more times, as a function of how many word tokens the child has heard (bottom axis) or how many hours of speech in the corpus are included (top axis). (a) Plot includes only isolated words; (b) plot includes only utterance-initial and utterance-final words. Note that the scales on the axes are different in each graph. Both plots' data are measured from the same 207 h of recording time.

non-trivial proportion of words intended by parents are not in fact categorized accordingly by infants. Still, though the above estimates are necessarily very rough, there is nothing extreme about supposing that an infant of 10 months might know dozens or even hundreds of word forms as recognizable phonetic chunks.

4. THE USE OF WORD FORM KNOWLEDGE IN INFANCY AND BEYOND

As we have seen, infants learn the forms of words, and this learning leads to improvements in the discovery of additional word forms. It is natural to ask at this point whether all this learning does the child any good. When infants wake up on the morning of their first birthday, how does their past experience learning words help them interpret speech, learn word meanings or acquire grammar?

(a) *Experience-dependent phonological interpretation*

Infants show evidence of learning the phonetic categories of their native language during the first year (e.g. Kuhl *et al.* 1992; Polka & Werker 1994; Bosch & Sebastián-Gallés 2003). This perceptual tuning process has lifelong consequences in many cases (Sebastián-Gallés *et al.* 2005) and is already a factor in toddlers' representation of words: once infants learn to attend more to the phonetic distinctions that are relevant in their native language, they use this capacity in recognizing words throughout development. For example, toddlers are adept at recognizing words quickly and efficiently, as shown in eyetracking studies (Swingley *et al.* 1999). If

shown two pictures on a computer screen, and one picture is named, 18-month-olds shift their gaze from the un-named distracter picture toward the named target about 800 ms from the onset of the target word, on average (Fernald *et al.* 1998; Swingley 2009). If words are 'mispronounced' to children on some trials, children shift to the target less reliably and fixate the target less overall than when words are pronounced correctly (e.g. Swingley & Aslin 2000). This disruption in performance is not found for phonetic substitutions that toddlers had learned to ignore in early infancy (Ramon-Casas 2009). Thus, although toddlers are not always able to differentiate phonetically similar words as well as older children do (Stager & Werker 1997; Swingley & Aslin 2007), the phonetic category formation process that begins in infancy is continuous with the development of phonological interpretation skills in toddlers.

This category learning process takes place too early developmentally to be triggered by the child's detection of semantically contrasting minimal-pair words. That is, one might imagine that a child unsure of whether the vowels /e/ and /i/ were different might, upon learning the words *wheel* and *whale*, notice that references to wheels sound a bit different from references to whales, and conclude that the language separates the vowels. This is not a plausible acquisition story for phonetic category learning because children do not appear to know words that could play this role early enough. As a result, researchers assume that some form of perceptual distributional learning is responsible for infants' phonetic category learning (e.g. Kuhl *et al.* 2008; Maye *et al.* 2002; Werker *et al.* 2007). In essence, distributional learning refers to the discovery of categories by detecting clumps of instances in a

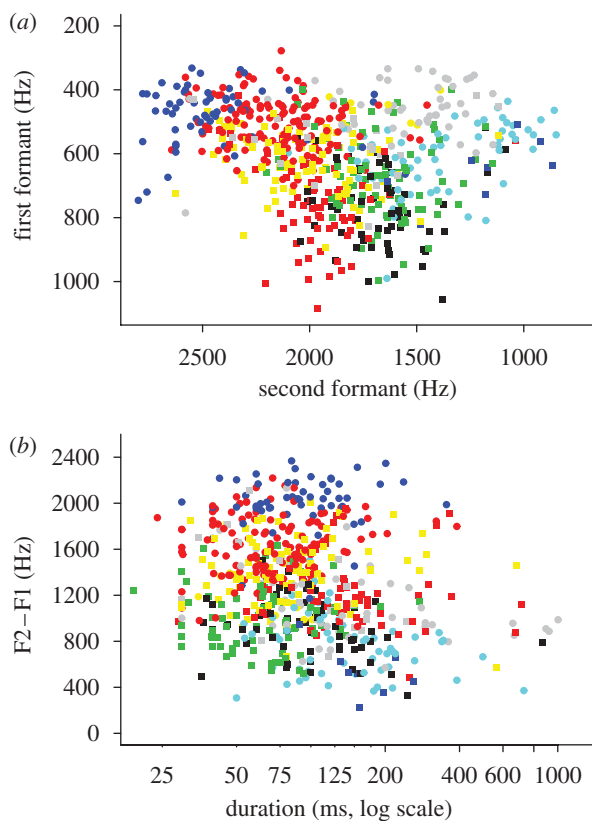


Figure 2. (a) First and second formants in about 700 vowels of one mother's speech to her infant. Each colour/shape combination indicates a different vowel. (b) Second formant minus first formant plotted against raw duration, for vowels of one mother's speech to her infant. Each colour/shape combination indicates a different vowel.

psychophysical space. If all [i] sounds are somewhat similar, and also distinct from instances of all other vowels, the infant learner could infer that [i] is a category. The same process could, in principle, lead to the formation of a language-appropriate set of phonetic categories.⁴

Adult and infant learners can discover auditory categories through distributional learning, even with fairly limited laboratory exposure (e.g. Holt & Lotto 2006; Francis *et al.* 2008; Maye *et al.* 2008; Goudbeek *et al.* in press). One of the questions that remains is whether the distributions that are available to infants in infant-directed speech are clear enough to support this form of learning as it is normally envisioned. Some studies of laboratory infant-directed speech suggest that they are (e.g. Vallabha *et al.* 2007), though the generalizability of such results needs to be assessed on the most naturalistic datasets possible. To see why the problem is a difficult one, consider figure 2, which shows measurements of several hundred vowels of infant-directed speech, spoken by one mother (coded as 'speaker f1') in the Brent & Siskind (2001) database. The speech recordings were entirely undirected and were made in a variety of environments.⁵ The graphs exclude schwa and some of the diphthongal vowels, leaving [i, ɪ, ɛ, æ, ɔ, ʌ, o, ʊ, u,] and [ɜ]. In figure 2a, the axes show the measured frequencies, in Hertz, of the first and second formants. Figure 2b shows the same data with the difference

between the formants (i.e. $F2 - F1$) on the ordinate and duration on the abscissa. It seems clear that the vowels do not cluster together distinctly, at least along the dimensions visible here. The mean values of each vowel category differ in predictable ways, and it would be easy to show that there are significant differences among the categories once the categories are identified. But of course infants do not know the categories in advance; they cannot see the colours on the graphs.

How can infants succeed at such a difficult problem? One possibility is that the apparent disarray of plots like figure 2 is overestimated. Perhaps with more dimensions, such as consonantal transition information, pitch, visible facial cues, additional formants, the dynamics of the formants, and so forth, categories would plainly emerge. This has yet to be demonstrated empirically, but it might be true (Davis & Lindblom 2001). Similarly, although formant values are good indicators of vowel identity, they might not map onto perceptual dimensions accurately, and this slippage might make the category learning process look harder than it is.

Another possibility that seems worth considering is that infants are aided by their knowledge of familiar word forms (Swingley 2007b). Infants do not hear many similar-sounding words, and among the words they hear very frequently, the number of phonological neighbours is smaller still. If an infant knew a few dozen word forms, some uncertainty about their component phonetic categories might not lead to many confusions; infants might recognize varied tokens of some familiar words while still in the process of determining where category boundaries are. Consider, for example, the vowel tokens plotted in figure 3. These are the subset of vowels of figure 2 that were transcribed as [i] (the filled blue dots) or as [ɪ] (the filled red squares). The [i] sounds tend toward higher second formant values and lower first formant values, but there is no clear break between the categories. An infant hearing these sounds might have little basis for determining whether there is one category or two in this dataset.

The four most common words that contain these vowels in this portion of the Brent corpus are *see*, *we*, *dillon* and *this*. Those vowel tokens that appeared in these four frequent words are marked with open symbols on the plot. Note that unlike the blue and red colours (which are 'invisible' to the infant's category finding mechanism), the open symbols are available, provided that the infant can recognize them as frequent word-forms. Just as the (inaudibly) blue dots tend toward the upper left of the plot, so do the (recognizable) instances of *see* and *we*. Likewise, instances of *dillon* and *this*, likely to be known to the child, tend toward the lower right. These tendencies are, statistically, extremely unlikely to occur by chance. For example, all but two of the 37 instances of *dillon* and *this* have $F2$ values less than 2250; all but one of the 19 instances of *see* and *we* have $F2$ values greater than 2250.

The proposal, then, is that words, which are identifiable by infants, might serve as rough indicators of where vowel category boundaries lie. Because speech

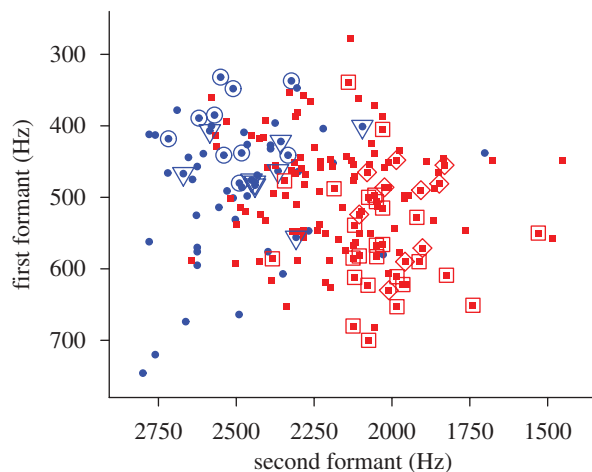


Figure 3. The vowels /i/ and /u/ in first- and second formant space, as spoken by one mother to her infant. The /i/ instances are plotted as blue circles, /u/ as red squares. Outlines around instances indicate tokens measured from the words *see* (open circles), *we* (open triangles), *dillon* (open squares), and *this* (open diamonds).

to infants does not contain very many frequent minimal pairs, infants whose phonological categorization abilities are not yet aligned with the mature, language-specific system may nonetheless identify distinct words, albeit with some uncertainty or error. These word forms may help suggest to infants the limits (or central tendencies) of distinct phonological categories. Thus, in contrast to the conventional notion that vowels are *segmented from their context* and placed in a multidimensional phonetic space for clustering, vowel tokens might retain information about their lexical origins, thereby helping to delimit the phonetic categories. Of course, this is not meant to suggest that the vowel tokens' intrinsic properties would not be used; rather, lexical categories may supplement the phonetics of individual tokens in defining the clusters of sounds that compose phonetic categories.

At present, this notion is speculative, and there are several ways it could break down. If phonetic context effects were to account for a substantial amount of the variability in vowel realizations, an infant might find herself with too many vowel categories—in the limit, one for each word, or perhaps one for each context. Inspection of the vowels in the present database suggests that relatively little of the variability among tokens is predictable from the immediate context; in figure 3, for example, the /i/s in *dillon* and *this* cover much of the same area despite their quite different following contexts, and the /i/s in *see* and *we* also overlap substantially. Another problem, which is equally fundamental, is that there is little evidence to suggest that infants are able to use markers like word identity to create categories that then admit members without the marking contexts. This remains an area for future research. If the proposal is correct, it would indicate a previously unappreciated role for the infant's acquisition of a 'protolexicon' of familiar word forms.

(b) Learning what words mean

As mentioned above, when we say we know a word, typically what we mean is that we know what a word sounds like (and how to say it), and what the word means. The research reviewed above indicates that infants are familiar with word forms, but do infants know what the words mean? Surprisingly little laboratory research has addressed this question, possibly because the origins of modern psycholinguistic work with infants lie in the domain of speech perception. Observational studies have remarked upon children's comprehension of words and have, in some cases, taken pains to ensure that children's responses to words really reflect some grasp of the words' meaning, as opposed to, for example, orienting to speech without understanding it. Such studies generally place the onset of word understanding at around 10 months, with some children not yielding consistent evidence until 11 or 12 months (e.g. [Huttenlocher 1974](#); [Benedict 1979](#)). Early experimental studies measuring children's looking at or touching named objects have placed the comprehension of first words closer to 13 months, although the procedures may have underestimated children's knowledge (e.g. [Thomas et al. 1981](#), used a four-alternative looking task that may have been too complex for the 11-month-olds who showed no recognition; and [Oviatt \(1980\)](#), used a word-teaching procedure rather than evaluating words infants may have already known).

Parents filling out vocabulary questionnaires like the MacArthur-Bates Communicative Development Inventory ([Fenson et al. 1994](#)) frequently report that their 8-month-olds understand a dozen or more words, a phenomenon that is viewed with some suspicion because most parents probably have much less stringent criteria for ascribing meaningful understanding than researchers do. On occasion parents who are also scientists have reported on their own children. [Darwin \(1877\)](#) writing of his son, recorded that 'when exactly 7 months old, he made the great step of associating his nurse with her name, so that if I called it out he would look round for her'. However, reports from trained parental observers are few in number.

The developmentally earliest empirical report of word understanding showed comprehension of *mommy* and *daddy*, or the analogous words used in the child's family, by 6-month-olds ([Tincoff & Jusczyk 1999](#)). Each parent was videotaped, and these films were presented to infants side by side. A synthesized voice produced several acoustically varying tokens of *mommy* and *daddy* while infants gazed at the screens. Relative to a baseline taken in silence, infants looked more at the named parent than the un-named parent. Infants in a second experiment did not look longer at matching films of unfamiliar parents, suggesting that the basis of children's matching fixations was a connection between the word and the named individual, and not a broader link between the words and all women and men.

Other studies have demonstrated trained associations between words and nonsense syllables. Infants can learn to link moving images and syllables and remember this connection over a substantial delay.

Gogate & Bahrick (2001) habituated 7-month-olds to consistent pairings of a film of two toys and two syllables (/a/ and /i/). For some children, each instance of the syllable was synchronized with movements of the toy; for other children, syllables and toy movements were not temporally correlated. Immediately following habituation, test trials measured infants' looking when a familiar vowel-toy pairing was shown and when a novel vowel-toy pairing was shown (e.g. calling the /a/ an /i/; Stager & Werker 1997). Infants looked longer on the novel, 'switched' trials—but only when words and movements were synchronized. This training and 'Switch' testing was repeated in a second experiment with just the synchronous condition, replicating that result. Four days later, infants were brought back to the laboratory and shown the films of the toys side by side while hearing one vowel or the other on successive trials. Infants' first looks were significantly more likely to be to the matching object than to the mismatch, suggesting long-term retention of the object-syllable links.

Parents do sometimes provide this kind of synchrony in labelling objects for their children (Matatyaho & Gogate 2008). Thus, infants from at least as young as 7 months may know some word forms that they connect associatively to objects or events. Opinion varies about whether the cognitive mechanisms underlying this learning are continuous with the mechanisms that underlie word learning in slightly older children (Werker & Patterson 2001). Even if the learning mechanisms are different, it seems most likely that the *knowledge* that infants acquire through early experience with words and objects develops into full-blown lexical knowledge of the sort anyone would credit (Golinkoff & Hirsh-Pasek 2006). Though infants associatively learning connections between a word and a specific object might not show either phonological or conceptual generalization patterns characteristic of the young toddler, such connections could still provide a foundation for later vocabulary development.

Even if infants learn many word forms to which they attach little or no semantic content, they might nevertheless benefit from knowing them as they develop their vocabulary. Studies of young toddlers support this notion. Werker and colleagues have shown that 14-month-olds can only learn pairs of similar-sounding words in a single session under certain conditions, and that these children's failure appears to be a consequence of a complex combination of weak cognitive capacities and immature pragmatic and phonological interpretation skills (Stager & Werker 1997; Fennell *et al.* 2007; Yoshida *et al.* 2009). Manipulations that enhance 14-month-olds' phonological processing of the words improve performance (e.g. Thiessen 2007; Rost & McMurray 2009), suggesting that children might learn the referents of pairs of similar-sounding words more readily if the words' forms were familiar.

Two recent studies support this notion, albeit in older children. Swingley (2007a) familiarized 19-month-olds with a novel word in the context of an animated story. During this familiarization, children heard the novel word 14 times, but were not

told or shown what the word meant. Immediately following the story, a novel object was displayed and named using the novel word. Learning was tested using a language-guided looking procedure. Children saw the named novel object and an alternative novel object on the screen, and the novel object was labelled in a sentence. Upon hearing the name for the object, children tended to look at it; but on trials when the name was pronounced in a distorted way (e.g. saying *kiebie* for the word *tiebie*), children fixated the target less, revealing their correct encoding of the initial consonant. A second group of children who heard a story that did not familiarize them with the form of the word to be learned did not show this decrement in fixation upon hearing the mispronunciation. This suggests that children can learn a word-form first, phonetically accurately, and then add meaningful content to this form as the situation allows.

Similarly, in another study, 17-month-olds familiarized with a word form in a complex statistical word-segmentation task (Saffran *et al.* 1996, described above) were then better able to demonstrate later learning of the word's meaning than were children who had not been familiarized with the word's form (Graf Estes *et al.* 2007).

Although these studies do not demonstrate toddlers' superior learning of words whose forms were familiar from infancy, they are consistent with this possibility. Thus, learning in the first several months of life may directly contribute to vocabulary acquisition in two ways: through learning of both words and aspects of their meaning, and through learning forms alone. This is an area that could benefit strongly from further research using existing techniques.

(c) *Grammatical structure*

Infants learn not only which sound sequences correspond to words, but also something about how words may be organized into functional categories or ordered in sentences. Much of the work on this issue has concerned infants' learning of functional elements (like determiners) and their typical positioning relative to content words (like nouns). Infants appear to learn at least some function words in the first year, as reviewed above. Computational modeling based on transcribed infant-directed speech suggests that information available to infants could in principle provide some purchase on the differentiation of open- and closed-class categories, nouns and verbs, and perhaps other categories as well (e.g. Mintz 2003; Swingley 2005b; Monaghan *et al.* 2007; Christiansen *et al.* 2009). There is no experimental evidence yet that infants have the category of (e.g.) *determiner* within the first year. However, there is evidence for some syntactic differentiation of categories after the first birthday (Booth & Waxman 2009), and experiments using training or artificial-language methods suggest that infants have the computational and representational wherewithal to form categories based on distributional information.

Production data suggest that children may have the category of determiners from the time of their earliest

production of them (Valian *et al.* 2009; for an opposing view, Pine & Martindale 1996), but this doesn't happen until well into the second year. One experimental approach to the question is to test whether infants make syntactically appropriate generalizations that transcend particular forms. For example, if a child heard a novel word following *the*, she would do well to assume that the word was a noun, given that the grammar of English places articles prenominally and given the rarity of interposed adjectives. She might therefore find the word to be appropriate in contexts that are suitable for nouns—even contexts without the *the* marker. Evidence for this sort of inference has been shown in German-learning 14–16-month-olds (though not 12–13 month-olds) by Höhle and co-workers (2004). They familiarized children with noun phrases like *ein glamm*, in which *ein* is an article and *glamm* is a nonce word. Following this familiarization, infants showed longer listening to 'verb-appropriate' passages like *Der Junge glamm immer auf dem Weg zur Schule* (The boy glammed always on his way to school) than to 'noun-appropriate' passages like *Den Kindern gefiel das wunderbare Glamm sehr gut* (The children loved the wonderful glamm very much). They did not prefer either sort of passage after familiarization with verbal phrases like *sie glamm*, where *sie* means 'she' and where *glamm* should be interpreted as a verb. Höhle *et al.* interpret the longer listening result as a novelty preference and ascribe the lack of a complementarity effect for the *sie* items to the substantially lower frequency of *sie* and its relatively smaller likelihood of immediately preceding a verb, as measured from a corpus of child-directed speech. If this is correct, it shows that German 15-month-olds know that *ein* marks words that belong in the contexts where nouns usually appear.

By 14 months (though not at 11 months), English learners distinguish between novel words presented as nouns and presented as adjectives (Waxman 1999; Waxman & Booth 2001, 2003). Providing labels for sets of objects alters infants' construal of the objects. For example, if shown a series of plush ducks in the context of a noun label (e.g. *This one is a blicket...*), infants then attend more to a plush spoon than a plush duck, as if weary of the procession of ducks; but they show no preference between a hard duck and a plush duck, unsated by the barrage of plushness. Shown the same ducks in the context of an adjective label (e.g. *This one is blickish...*), infants appear to focus on both the plushness and the properties that make the objects ducks, and exhibit preferences for both non-duck and non-plush objects. Exactly how children arrive at these generalizations about the reference of nouns or adjectives is not well understood, but the fact that children in the adjective conditions perform much like children in 'no-label' control conditions suggests that the noun category may have a special and more specific status for children.

Still younger children appear to know something about whether their language places functor elements (such as articles, prepositions and personal pronouns) before or after content words. Languages vary in this respect: some, like English and Italian, place heads of syntactic phrases before complements, a descriptive

fact that, in the world's languages, implies a likelihood that functor elements will tend to come at the beginning of syntactic phrases. Other languages, such as Japanese and Hindi, tend to order phrases the other way around. Because most functors are far more frequent than most content words, a language's placement of functors in phrase-initial position tends to result in utterances that start with a word drawn from a small, but high-frequency set, whereas a language's placement of functors phrase-finally results in utterances with high-frequency words utterance-finally. This is true of child-directed speech in Italian and Japanese, implying that if infants could keep track of highly frequent words and where they tend to appear at utterance boundaries, infants might gain language-appropriate intuitions about dividing multi-phrasal utterances at their phrase boundaries (Gervain *et al.* 2008).

To evaluate this possibility, Gervain *et al.* tested Italian and Japanese 8-month-olds' segmentation of continuous syllable streams. These were synthesized following the pattern *fiXgeY*, where **X** instances were selected from a set of nine syllables and **Y** from another set of nine. As a stream, the sequence had no clear beginning or ending, fading in and out at the start and finish; thus, infants were familiarized to something like ...*gefofibugedefikogepafimoge*.... An infant disposed to treat *ge* and *fi* as sequence onsets (just as frequent words tend to align with constituent onsets in Italian) might parse this stream as *gefo fibu gede fiko*...; conversely, the reverse disposition suggests the parse *ge fofi buge defi*.... At test, infants were offered four-syllable sequences consistent with the frequent-initial or frequent-final groupings. Japanese infants listened longer to frequent-final sequences, Italian infants to frequent-initial ones. Language-appropriate parsing biases driven by knowledge of how frequent words align with phrase boundaries could help young children rule out incorrect hypotheses about their grammar (e.g. Morgan 1986).

This case provides another example of the value of infants' learning of a word-form 'protolexicon'. Just as the trochaic bias cannot be simply 'read off' utterances, the frequent-functor directional bias cannot be inferred from distributions of syllables alone. In Italian infant-directed speech, for example, the set of utterance-final syllables contains a disproportionately large number of highly frequent syllables, whereas utterance-initial syllables are not particularly frequent—trends that are opposite to the pattern for words (J. Gervain 2009, personal communication). To discover the orientation of functors in Italian, infants must compute frequencies over larger units, perhaps using statistical co-occurrence properties of syllables to derive words.

Other experiments have tested infants' capacity for detecting word order regularities or for inducing categories of words from distributional patterns. These studies use training procedures with artificial languages or materials drawn from languages unfamiliar to the infants. For example, Gomez & Gerken (1999) familiarized 12-month-olds to nonsense syllable sequences whose ordering was determined using

a finite-state grammar. Infants showed listening preferences for novel sequences drawn from the grammar over novel sequences inconsistent with the grammar. This was true even in a test in which each specific syllable in the training was replaced, one to one, with a novel syllable, indicating that infants had learned something abstract about the sequences permitted in the grammar (see also Marcus *et al.* 1999). Gómez & Lakusta (2004) found that 12-month-olds presented with $a\mathbf{X}b\mathbf{Y}$ strings, with a and b being different sets containing two monosyllables, \mathbf{X} being a set of six bisyllables, and \mathbf{Y} being a set of six monosyllables, were able to learn that the a syllables always preceded bisyllables and the b always preceded monosyllables. This kind of association between one class of elements and a distinguishing feature of another class is argued to be an essential step in the distributional induction of form class categories (Braine 1987; Gerken *et al.* 2005). Studies using variations of these basic techniques have begun to elucidate which generalizations infants are able to make (e.g. Newport & Aslin 2004), what infants learn when the information presented supports multiple regularities (Gerken 2004, 2006), and how infants integrate the discovery of words and the detection of word-ordering regularities (Saffran & Wilson 2003; Swingley 2005b).

Given that infants do discover the forms of words early in life, it is assumed, quite reasonably, that the sorts of learning that infants demonstrate in artificial-language studies are also characteristic of natural language processing. That said, for the most part these studies are not yet able to predict or explain details of the timing of the emergence of grammatical knowledge in natural language. Ideally, one would want to use artificial-language learning experiments to establish which sources of information infants learn from, and under what conditions, and how best to characterize the learning process. Armed with this knowledge, computational models of the child could be applied to infant-directed speech data in various languages. The closer the model and the corpora are to the truth, the better the model should predict the emergence of linguistic structure in infants. To the extent that models fall short in explaining children's behaviour, this might be owing to poor characterization of the 'input' data to the child, or it might point to a need to constrain the model differently, including possibly adding innate biases of some sort.

At present, the body of experimental work on infants' artificial-language learning offers a number of novel points of further inquiry, and has led to significant changes in how researchers think about language learning. At the same time, this work does not constrain learning models much at the quantitative level. Partly this is because of the practical constraint that training experiments can only train infants a limited amount relative to the infant's natural language exposure, and as a result experiments have to present the regularity of interest in a highly simplified context. If infants can learn something simplified in 5 min, can they learn something complex in five months? It's not easy to find out. This is one reason why it is important to complement artificial-language studies with tests of what infants know of their own language.

As we have seen, the evidence suggests that infants' experience with words in the first year of life leads to some syntactic knowledge that is active early in the second year.

5. CONCLUSIONS

Parents have long considered the first spoken words to mark an important transition from the 'prelinguistic' infant to the verbal one. Just as infants' emerging social smiles in the second month are significant and accessible markers of development (e.g. Lavelli & Fogel 2005), first words signal the child's entry into linguistic dialogue, and are carefully noted in baby books the world over. Psychologists and linguists, too, have often taken first spoken words as the starting point for language acquisition, and have marvelled at how quickly children learn to use their native language. This perspective left young infants with a relatively minor contribution to make. Since the 1980s, we have known that infants begin to learn specific phonological characteristics of their language, and as a result, textbook treatments now tend to assign infants the task of phonetic category learning. Word learning on such accounts seems to be an enterprise that concerns the second and third years, as children become increasingly adept at using their developing knowledge of communicative intentions to discover the meanings of words and other linguistic expressions.

By contrast, the evidence we have reviewed suggests several ways in which infants' word-form learning contributes to language acquisition. In the first year, infants begin solving some of the problems that make language acquisition hard. It may be that the learning strategies available to infants are only a subset of those available to older children, and there might be fundamental differences between the ways in which linguistic knowledge is acquired in infancy and later in childhood. At the same time, the knowledge of language that infants acquire is continuous with the knowledge that they will build upon for the rest of their lives as they continue to use that language. Phonetic categories provide the foundation of phonology; spoken word forms give rise to the lexicon; and statistical patterns over these forms help point the child to his or her language's grammar.

This paper was written with the support of grants NIH R01 HD49681 to D.S., and NSF HSD 0433567 to Delphine Dahan and D.S. Thanks are due to Gareth Gaskell for suggesting that I write the paper, to Judit Gervain for analyses of Italian syllable frequencies and to Michael Brent for contributing his infant-directed speech corpus.

ENDNOTES

¹Sentences were matched in the region of the target syllables and were grammatical, if not always plausible, to aid naturalness in recording.

²In this paper, all examples of this sort were selected from transcriptions of infant-directed speech, primarily the Brent corpus (2001).

³These generalizations require both components—highly frequent syllable pairs are not particularly likely to be words because there are too many common pairs like *you want*, and infrequent but cohesive sequences are not especially likely to be words either, because

there are too many infrequent pairs of words made of uncommon syllables.

⁴By itself, even under ideal circumstances this process would not give infants the *phonology* of the language, which requires additional linguistic analysis. For example, the 'd' letter in the Dutch word *honden*, 'dogs', is pronounced [d]; the same letter in the singular *hond* 'dog', is pronounced [t], a phenomenon described as 'final devoicing'. The conventional linguistic analysis holds that both sounds are, at a linguistic level of representation, [d], but that the sound undergoes devoicing. This sort of process is not envisioned to occur only by clustering sounds in psychophysical space.

⁵This contrasts with the much neater phonetic data one sees in studies of 'phoneticese', the speech produced by careful talkers reading lists of minimal pairs into microphones in phonetics laboratories. Such laboratory studies are crucial for characterizing phonetic knowledge and revealing how linguistic variables interact in speech production, but they do not accurately portray the variability present in the phonetic data available to infants in their natural environments.

REFERENCES

- Aslin, R. N., Saffran, J. R. & Newport, E. L. 1998 Computation of conditional probability statistics by 8-month-old infants. *Psychol. Sci.* **9**, 321–324. (doi:10.1111/1467-9280.00063)
- Benedict, H. 1979 Early lexical development: comprehension and production. *J. Child Lang.* **6**, 183–200. (doi:10.1017/s0305000900002245)
- Booth, A. E. & Waxman, S. R. 2009 A horse of a different color: specifying with precision infants' mappings of novel nouns and adjectives. *Child Dev.* **80**, 15–22. (doi:10.1111/j.1467-8624.2008.01242.x)
- Bortfeld, H., Morgan, J. L., Golinkoff, R. M. & Rathbun, K. 2005 Mommy and me: familiar names help launch babies into speech-stream segmentation. *Psychol. Sci.* **16**, 298–304. (doi:10.1111/j.0956-7976.2005.01531.x)
- Bosch, L. & Sebastián-Gallés, N. 2003 Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Lang. Speech* **46**, 217–243. (doi:10.1177/00228309030460020801)
- Braine, M. D. S. 1987 What is learned in acquiring word classes—a step toward an acquisition theory. In *Mechanisms of language acquisition* (ed. B. MacWhinney), pp. 65–87. Hillsdale, NJ: Erlbaum.
- Brent, M. R. & Cartwright, T. A. 1996 Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition* **61**, 93–125. (doi:10.1016/s0010-0277(96)00719-6)
- Brent, M. R. & Siskind, J. M. 2001 The role of exposure to isolated words in early vocabulary development. *Cognition* **81**, B33–B44. (doi:10.1016/s0010-0277(01)00122-6)
- Burns, T. C., Yoshida, K. A., Hill, K. & Werker, J. F. 2007 The development of phonetic representation in bilingual and monolingual infants. *Appl. Psycholinguist.* **28**, 455–474. (doi:10.1017/s0142716407070257)
- Christiansen, M. H., Onnis, L. & Hockema, S. A. 2009 The secret is in the sound: from unsegmented speech to lexical categories. *Dev. Sci.* **12**, 388–395. (doi:10.1111/j.1467-7687.2009.00824.x)
- Colombo, J. A. & Bundy, R. S. 1981 A method for the measurement of infant auditory selectivity. *Infant Behav. Dev.* **4**, 219–223. (doi:10.1016/s0163-6383(81)800252)
- Curtin, S., Mintz, T. H. & Christiansen, M. H. 2005 Stress changes the representational landscape: evidence from word segmentation. *Cognition* **96**, 233–262. (doi:10.1016/j.cognition-2004.08.005)
- Cutler, A. & Butterfield, S. 1992 Rhythmic cues to speech segmentation: evidence from juncture misperception. *J. Mem. Lang.* **31**, 218–236. (doi:10.1016/0749-596x(92)90012-M)
- Cutler, A. & Carter, D. M. 1987 The predominance of strong initial syllables in the English vocabulary. *Computer Speech Lang.* **2**, 133–142. (doi:10.1016/0885-2308(87)90004-0)
- Cutler, A. & Norris, D. G. 1988 The role of strong syllables in segmentation for lexical access. *J. Exp. Psychol. Human Percept. Perform.* **14**, 113–121. (doi:10.1037/0096-1523.14.1.113)
- Dahan, D. & Brent, M. R. 1999 On the discovery of novel wordlike units from utterances: an artificial-language study with implications for native-language acquisition. *J. Exp. Psychol.: General* **128**, 165–185. (doi:10.1037/0096-3445.128.2.165)
- Darwin, C. 1877 A biographical sketch of an infant. *Mind* **2**, 285–294. (doi:10.1093/mind/03-2.7.285)
- Davis, B. L. & Lindblom, B. 2001 Phonetic variability in baby talk and development of vowel categories. In *Emerging cognitive abilities in early infancy* (eds F. Lacerda, C. von Hofsten & M. Heimann), pp. 135–171. Mahwah, NJ: Erlbaum.
- Echols, C. H., Crowhurst, M. J. & Childers, J. B. 1997 The perception of rhythmic units in speech by infants and adults. *J. Mem. Lang.* **36**, 202–225. (doi:10.1006/jmla.1996.2483)
- Eimas, P. D. & Miller, J. L. 1980 Contextual effects in infant speech perception. *Science* **209**, 1140–1141. (doi:10.1126/science.7403875)
- Eimas, P. D., Siqueland, E. R., Jusczyk, P. W. & Vigorito, J. 1971 Speech perception in infants. *Science* **171**, 303–306. (doi:10.1126/science.171.3968.303)
- Eimas, P. D., Miller, J. L. & Jusczyk, P. W. 1987 On infant speech perception and the acquisition of language. In *Categorical perception: the groundwork of cognition* (ed. S. Harnad), pp. 161–195. Cambridge, UK: Cambridge University Press.
- Fennell, C. T., Waxman, S. R. & Weisleder, A. 2007 With referential cues, infants successfully use phonetic detail in word learning. *Proc. 31st Annual Boston Univ. Conf. on Language Development*, pp. 206–217. Boston: Cascadia Press.
- Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J. & Pethick, S. J. 1994 Variability in early communicative development. *Monogr. Soc. Res. Child Dev.* **59**, Serial Number 242. (doi:10.2307/1166093)
- Fernald, A. 1985 Four-month-old infants prefer to listen to motherese. *Infant Behav. Dev.* **8**, 181–195. (doi:10.1016/S0163-6383(85)80005-9)
- Fernald, A., Pinto, J. P., Swingley, D., Weinberg, A. & McRoberts, G. W. 1998 Rapid gains in speed of verbal processing by infants in the second year. *Psychol. Sci.* **9**, 228–231. (doi:10.1111/1467-9280.00044)
- Francis, A. L., Kaganovich, N. & Driscoll, C. J. 2008 Cue-specific effects of categorization training on the relative weighting of acoustic cues to consonant voicing in English. *J. Acoust. Soc. Am.* **124**, 1234–1251. (doi:10.1121/1.2945161)
- Gerken, L. A. 2004 Nine-month-olds extract structural principles required for natural language. *Cognition* **93**, B89–B96. (doi:10.1016/j.cognition.2003.11.005)
- Gerken, L. A. 2006 Decisions, decisions: infant language learning when multiple generalizations are possible. *Cognition* **98**, B67–B74. (doi:10.1016/j.cognition.2005.03.003)
- Gerken, L. A., Jusczyk, P. W. & Mandel, D. R. 1994 When prosody fails to cue syntactic structure: nine month olds' sensitivity to phonological versus syntactic phrases. *Cognition* **51**, 257–265. (doi:10.1016/0010-0277(94)90055-8)

- Gerken, L. A., Wilson, R. & Lewis, W. 2005 Infants can use distributional cues to form syntactic categories. *J. Child Lang.* **32**, 249–268. (doi:10.1017/s0305000904006786)
- Gervain, J., Nespore, M., Mazuka, R., Horie, R. & Mehler, J. 2008 Bootstrapping word order in prelexical infants: a Japanese-Italian cross-linguistic study. *Cogn. Psychol.* **57**, 56–74. (doi:10.1016/j.cogpsych.2007.12.001)
- Gilkerson, J. & Richards, J. A. 2007 The Infoture natural language study, technical report itr-02-1. Technical Report. Boulder, CO: Infoture, Inc.
- Gogate, L. J. & Bahrick, L. E. 2001 Intersensory redundancy and 7-month-old infants' memory for arbitrary syllable-object relations. *Infancy* **2**, 219–231. (doi:10.1207/s153270781N0202_7)
- Golinkoff, R. M. & Hirsh-Pasek, K. 2006 Baby wordsmith: from associationist to social sophisticate. *Curr. Direct. Psychol. Sci.* **15**, 30–33. (doi:10.1111/j.0963-7214.2006.00401.x)
- Gomez, R. L. & Gerken, L. A. 1999 Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition* **70**, 109–135. (doi:10.1016/s0010-0277(99)00003-7)
- Gómez, R. L. & Lakusta, L. 2004 A first step in form-based category abstraction by 12-month-old infants. *Dev. Sci.* **7**, 567–580. (doi:10.1111/j.1467-7687.2004.00381.x)
- Goodsitt, J. V., Morgan, J. L. & Kuhl, P. K. 1993 Perceptual strategies in prelingual speech segmentation. *J. Child Lang.* **20**, 229–252. (doi:10.1017/s0305000900008266)
- Goudbeek, M., Smits, R. & Swingley, D. In press. Supervised and unsupervised learning of multidimensional auditory categories. *J. Exp. Psychol.: Human Percept. Perform.*
- Gout, A., Christophe, A. & Morgan, J. L. 2004 Phonological phrase boundaries constrain lexical access: II. Infant data. *J. Mem. Lang.* **51**, 548–567. (doi:10.1016/j.jml.2004.07.002)
- Graf Estes, K., Evans, J. L., Alibali, M. W. & Saffran, J. R. 2007 Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychol. Sci.* **18**, 254–260. (doi:10.1111/j.1467-9280.2007.01885.x)
- Hallé, P. A. & de Boysson-Bardies, B. 1996 The format of representation of recognized words in the infants' early receptive lexicon. *Infant Behav. Dev.* **19**, 463–481. (doi:10.1016/s0163-6383(96)90007-7)
- Hallé, P. A., Durand, C., Bardies, B. & de Boysson 2008 Do 11-month-old French infants process articles? *Lang. Speech* **51**, 23–44. (doi:10.1111/00238309080510010301)
- Hart, B. & Risley, T. R. 1995 *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Paul H. Brooks Publishing.
- Hirsh-Pasek, K., Kemler Nelson, D. G., Jusczyk, P. W., Cassidy, K. W., Druss, B. & Kennedy, L. 1987 Clauses are perceptual units for young infants. *Cognition* **26**, 269–286. (doi:10.1016/s0010-0277(87)80002-1)
- 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**, 341–353. (doi:10.1207/s15327078in0503_5)
- Holt, L. L. & Lotto, A. J. 2006 Cue weighting in auditory categorization: implications for first and second language acquisition. *J. Acoust. Soc. Am.* **119**, 3059–3071. (doi:10.1121/1.2188377)
- Houston, D. M., Jusczyk, P. W., Kuijpers, C., Coolen, R. & Cutler, A. 2000 Crosslanguage word segmentation by 9-month-olds. *Psychonom. Bull. Rev.* **7**, 504–509.
- Huttenlocher, J. 1974 The origins of language comprehension. In *Theories in cognitive psychology: the Loyola Symposium* (ed. R. L. Solso), pp. 331–368. Potomac, MD: Erlbaum.
- Johnson, E. K. & Jusczyk, P. W. 2001 Word segmentation by 8-month-olds: when speech cues count more than statistics. *J. Mem. Lang.* **44**, 548–567. (doi:10.1006/jmla.2000.2755)
- Jusczyk, P. W. 1997 *The discovery of spoken language*. Cambridge, MA: MIT Press.
- Jusczyk, P. W. & Aslin, R. N. 1995 Infants' detection of the sound patterns of words in fluent speech. *Cogn. Psychol.* **29**, 1–23. (doi:10.1006/cogp.1995.1010)
- Jusczyk, P. W. & Hohne, E. A. 1997 Infants' memory for spoken words. *Science* **277**, 1984–1986. (doi:10.1126/science.277.5334.1984)
- Jusczyk, P. W., Hirsh-Pasek, K., Kemler Nelson, D., Kennedy, L., Woodward, A. & Piwoz, J. 1992 Perception of acoustic correlates of major phrasal units by young infants. *Cogn. Psychol.* **24**, 252–293. (doi:10.1016/0010-0285(92)90009-Q)
- Jusczyk, P. W., Cutler, A. & Redanz, N. J. 1993 Infants' sensitivity to the predominant stress patterns of English words. *Child Dev.* **64**, 675–687. (doi:10.2307/1131210)
- Jusczyk, P. W., Luce, P. A. & Charles-Luce, J. 1994 Infants' sensitivity to phonotactic patterns in the native language. *J. Mem. Lang.* **33**, 630–645. (doi:10.1006/jmla.1994.1030)
- Jusczyk, P. W., Houston, D. M. & Newsome, M. 1999 The beginnings of word segmentation in English-learning infants. *Cogn. Psychol.* **39**, 159–207. (doi:10.1006/cogp.1999.0716)
- Kluender, K. R., Diehl, R. L. & Killeen, P. R. 1987 Japanese quail can learn phonetic categories. *Science* **237**, 1195–1197. (doi:10.1126/science.3629235)
- Kuhl, P. K. 1995 Mechanisms of developmental change in speech and language. In *Proc. XIIIth Int. Congress of Phonetic Sciences* (eds K. Elenius & P. Branderud), vol. 2, pp. 132–139. Stockholm: Stockholm University.
- Kuhl, P. K. & Miller, J. D. 1978 Speech perception in the chinchilla: identification functions for synthetic VOT stimuli. *J. Acoust. Soc. Am.* **63**, 905–917. (doi:10.1121/1.381770)
- Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N. & Lindblom, B. 1992 Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* **255**, 606–608. (doi:10.1126/science.1736364)
- Kuhl, P. K., Tsao, F. M. & Liu, H. M. 2003 Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. *Proc. Natl. Acad. Sci.* **100**, 9096–9091. (doi:10.1073/pnas.1532872100)
- Kuhl, P. K., Conboy, B. T., Padden, D., Nelson, T. & Pruitt, J. 2005 Early speech perception and later language development: implications for the 'critical period'. *Lang. Learn. Dev.* **1**, 237–264. (doi:10.1207/s15473341ld0103&4_2)
- Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M. & Nelson, T. 2008 Developmental phonetic perception: Native Language Magnet Theory Expanded (NLM-e). *Phil. Trans. R. Soc. B* **363**, 979–1000. (doi:10.1098/rstb.2007.2154)
- Lavelli, M. & Fogel, A. 2005 Developmental changes in the relationship between the infant's attention and emotion during early face-to-face communication: the 2-month transition. *Dev. Psychol.* **41**, 265–280. (doi:10.1037/0012-1649.41.1.265)
- Leopold, W. F. 1939 *Speech development of a bilingual child, a linguist's record*, vol. 1. Evanston, IL: Northwestern University Press.
- Lisker, L. 1975 Is it VOT or a first-formant transition detector? *J. Acoust. Soc. Am.* **57**, 1547–1551. (doi:10.1121/1.380602)
- Lisker, L. & Abramson, A. S. 1967 Some effects of context on voice onset time in English stops. *Lang. Speech* **10**, 1–28.

- Liu, H.-M., Kuhl, P. K. & Tsao, F.-M. 2003 An association between mothers' speech clarity and infants' speech discrimination skills. *Dev. Sci.* **6**, F1–F10. (doi:10.1111/1467-7687.00275)
- Mandel, D. R., Jusczyk, P. W. & Pisoni, D. B. 1995 Infants' recognition of the sound patterns of their own names. *Psychol. Sci.* **6**, 314–317. (doi:10.1111/j.1467-9280.1995.tb.00517.x)
- Marcus, G. F., Vijayan, S., Rao, S. B. & Vishton, P. M. 1999 Rule learning by seven-month-old infants. *Science* **283**, 77–80. (doi:10.1126/science.283.5398.77)
- Matatyaho, D. J. & Gogate, L. J. 2008 Type of maternal object motion during synchronous naming predicts preverbal infants' learning of word-object relations. *Infancy* **13**, 172–184. (doi:10.1080/15250000701795655)
- Mattys, S. L. & Jusczyk, P. W. 2001 Do infants segment words or recurring contiguous patterns? *J. Exp. Psychol.: Human Percept. Perform.* **27**, 644–655. (doi:10.1037/0096-1523.27.3.644)
- Mattys, S. L., Jusczyk, P. W., Luce, P. A. & Morgan, J. L. 1999 Phonotactic and prosodic effects on word segmentation in infants. *Cogn. Psychol.* **38**, 465–494. (doi:10.1006/cogp.1999.6721)
- Maye, J., Gerken, L. A. & Werker, J. F. 2002 Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* **82**, B101–B111. (doi:10.1016/S0010-0277(01)00157-3)
- Maye, J., Weiss, D. J. & Aslin, R. N. 2008 Statistical phonetic learning in infants: facilitation and feature generalization. *Dev. Sci.* **11**, 122–134.
- McGregor, K. K. & Johnson, A. C. 1997 Trochaic template use in early words and phrases. *J. Speech, Lang. Hearing Res.* **40**, 1220–1231.
- Mintz, T. H. 2003 Frequent frames as a cue for grammatical categories in child directed speech. *Cognition* **90**, 91–117. (doi:10.1016/S0010-0277(03)00140-9)
- Monaghan, P., Christiansen, M. H. & Chater, N. 2007 The phonological-distributional coherence hypothesis: cross-linguistic evidence in language acquisition. *Cogn. Psychol.* **55**, 259–305. (doi:10.1016/j.cogpsych.2006.12.001)
- Morgan, J. L. 1986 *From simple input to complex grammar*. Cambridge, MA: MIT Press.
- Morgan, J. L. 1996 A rhythmic bias in preverbal speech segmentation. *J. Mem. Lang.* **35**, 666–688. (doi:10.1006/jmla.1996.0035)
- Nazzi, T., Kemler Nelson, D. G., Jusczyk, P. W. & Jusczyk, A. M. 2000 Six month olds' detection of clauses embedded in continuous speech: effects of prosodic well-formedness. *Infancy* **1**, 123–147. (doi:10.1207/s15327078IN010111)
- Nazzi, T., Iakimova, G., Bertoncini, J., Ffledonie, S. & Alcantara, C. 2006 Early segmentation of fluent speech by infants acquiring French: emerging evidence for cross-linguistic differences. *J. Mem. Lang.* **54**, 283–299. (doi:10.1016/j.jml.2005.10.004)
- Newport, E. L. & Aslin, R. N. 2004 Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cogn. Psychol.* **48**, 127–162. (doi:10.1016/S0010-0285(03)00128-2)
- Oviatt, S. L. 1980 The emerging ability to comprehend language: an experimental approach. *Child Dev.* **51**, 97–106. (doi:10.2307/1129595)
- Pelucchi, B., Hay, J. F. & Saffran, J. R. 2009 Statistical learning in a natural language by 8-month-old infants. *Child Dev.* **80**, 674–685. (doi:10.1111/j.1467-8624.2009.01290.x)
- Pine, J. M. & Martindale, H. 1996 Syntactic categories in the speech of young children: the case of the determiner. *J. Child Lang.* **23**, 369–395. (doi:10.1017/S0305000900008849)
- Polka, L. & Sundara, M. 2003 Word segmentation in monolingual and bilingual infant learners of English and French. *Proc. 15th Int. Congress of Phonetic Sciences, Barcelona, Spain, August 2003* (eds M. J. Solé, D. Recasens & J. Romero), pp. 1021–1024.
- Polka, L. & Werker, J. F. 1994 Developmental changes in perception of nonnative vowel contrasts. *J. Exp. Psychol.: Human Percept. Perform.* **20**, 421–435. (doi:10.1037/0096-1523.20.2.421)
- Ramon-Casas, M., Swingley, D., Bosch, L. & Sebastián-Gallés, N. 2009 Vowel categorization during word recognition in bilingual toddlers. *Cogn. Psychol.* **59**, 96–121. (doi:10.1016/j.cogpsych.2009.02.002)
- Redford, M. A., Davis, B. L. & Miiikkulainen, R. 2004 Phonetic variability and prosodic structure in mothers. *Infant Behav. Dev.* **27**, 477–498. (doi:10.1016/j.infbeh.2004.05.001)
- Rost, G. & McMurray, B. 2009 Speaker variability augments phonological processing in early word learning. *Dev. Sci.* **12**, 339–349. (doi:10.1111/j.1467-7687.2008.00786.x)
- Saffran, J. R. & Wilson, D. P. 2003 From syllables to syntax: multilevel statistical learning by 12-month-old infants. *Infancy* **4**, 273–284. (doi:10.1207/s15327078IN0402_07)
- Saffran, J. R., Aslin, R. N. & Newport, E. L. 1996 Statistical learning by 8-month-olds. *Science* **274**, 1926–1928. (doi:10.1126/science.274.5294.1926)
- Sebastián-Gallés, N., Echeverría, S. & Bosch, L. 2005 The influence of initial exposure on lexical representation: comparing early and simultaneous bilinguals. *J. Mem. Lang.* **52**, 240–255. (doi:10.1016/j.jml.2004.11.001)
- Seidl, A. 2007 Infants' use and weighting of prosodic cues in clause segmentation. *J. Mem. Lang.* **57**, 24–48. (doi:10.1016/j.jml.2006.10.004)
- Seidl, A. & Johnson, E. K. 2006 Infant word segmentation revisited: edge alignment facilitates target extraction. *Dev. Sci.* **9**, 565–573. (doi:10.1111/j.1467-7687.2006.00534.x)
- Shi, R., Cutler, A., Werker, J. F. & Cruickshank, M. 2006a Frequency and form as determinants of functor sensitivity in English-acquiring infants. *JASA Exp. Lett.* **119**, EL61–EL67.
- Shi, R., Werker, J. F. & Cutler, A. 2006b Recognition and representation of function words in English-learning infants. *Infancy* **10**, 187–198. (doi:10.1207/s15327078in1002_5)
- Stager, C. L. & Werker, J. F. 1997 Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature* **388**, 381–382. (doi:10.1038/41102)
- Swingley, D. 2005a 11-month-olds' knowledge of how familiar words sound. *Dev. Sci.* **8**, 432–443. (doi:10.1111/j.1467-7687.2005.00432.x)
- Swingley, D. 2005b Statistical clustering and the contents of the infant vocabulary. *Cogn. Psychol.* **50**, 86–132. (doi:10.1016/j.cogpsych.2004.06.001)
- Swingley, D. 2007a Lexical exposure and word-form encoding in 1.5-year-olds. *Dev. Psychol.* **43**, 454–464. (doi:10.1037/0012-1649.43.2.454)
- Swingley, D. 2007b Phonetic categories: what statistics, and what generalizations? *Paper presented at the workshop Current issues in language acquisition: artificial languages and statistical learning, Calgary, Canada.*
- Swingley, D. 2009 Onsets and codas in 1.5-year-olds' word recognition. *J. Mem. Lang.* **60**, 252–269. (doi:10.1016/j.jml.2008.11.003)
- Swingley, D. & Aslin, R. N. 2000 Spoken word recognition and lexical representation in very young children. *Cognition* **76**, 147–166. (doi:10.1016/S0010-0277(00)00081.0)

- Swingley, D. & Aslin, R. N. 2007 Lexical competition in young children's word learning. *Cogn. Psychol.* **54**, 99–132. (doi:10.1016/j.cogpsych.2006.05.001)
- Swingley, D., Pinto, J. P. & Fernald, A. 1999 Continuous processing in word recognition at 24 months. *Cognition* **71**, 73–108. (doi:10.1016/s0010-0277(99)00021-9)
- Taine, H. 1876 Sur l'acquisition du langage chez les enfants et dans l'espèce humaine. *Revue Philosophique de la France et de l'Etranger* **1**, 5–23.
- Thiessen, E. D. 2007 The effect of distributional information on children's use of phonemic contrasts. *J. Mem. Lang.* **56**, 16–34. (doi:10.1016/j.jml.2006.07.002)
- Thiessen, E. D. & Saffran, J. R. 2003 When cues collide: use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Dev. Psychol.* **39**, 706–716. (doi:10.1037/0012-1649.39.4.706)
- Thomas, D. G., Campos, J. J., Shucard, D. W., Ramsay, D. S. & Shucard, J. 1981 Semantic comprehension in infancy: a signal detection analysis. *Child Dev.* **52**, 798–803. (doi:10.2307/1129079)
- Tincoff, R. & Jusczyk, P. W. 1996 Are word-final sounds perceptually salient for infants? *Poster presented at the Fifth Conf. on Laboratory Phonology, Evanston, IL, USA.*
- Tincoff, R. & Jusczyk, P. W. 1999 Some beginnings of word comprehension in 6-month-olds. *Psychol. Sci.* **10**, 172–175. (doi:10.1111/1467-9280.00127)
- Valian, V., Solt, S. & Stewart, J. 2009 Abstract categories or limited-scope formulae? The case of children's determiners. *J. Child Lang.* **36**, 743–778. (doi:10.1017/S0305000908009082)
- Vallabha, G. K., McClelland, J. L., Pons, F., Werker, J. F. & Amano, S. 2007 Unsupervised learning of vowel categories from infant-directed speech. *Proc. Natl Acad. Sci. USA* **104**, 16027–16031. (doi:10.1073/pnas.0705369104)
- Vihman, M. M., DePaolis, R. A. & Davis, B. L. 1998 Is there a trochaic bias in early word learning? evidence from infant production in English and French. *Child Dev.* **69**, 935–949. (doi:10.2307/1132354)
- Vihman, M. M., Nakai, S., DePaolis, R. A. & Hallé, P. 2004 The role of accentual pattern in early lexical representation. *J. Mem. Lang.* **50**, 336–353. (doi:10.1016/j.jml.2003.11.004)
- Waxman, S. R. 1999 Specifying the scope of 13-month-olds' expectations for novel words. *Cognition* **70**, B35–B50. (doi:10.1016/s0010-0277(99)00017-7)
- Waxman, S. R. & Booth, A. E. 2001 Seeing pink elephants: fourteen-month-olds' interpretations of novel nouns and adjectives. *Cogn. Psychol.* **43**, 217–242. (doi:10.1006/cogp.2001.0764)
- Waxman, S. & Booth, A. 2003 The origins and evolution of links between word learning and conceptual organization: new evidence from 11-month-olds. *Dev. Sci.* **6**, 128–135. (doi:10.1111/1467-7687.00262)
- Werker, J. F. & Patterson, M. L. 2001 Does perceptual learning really contribute to word learning? *Dev. Sci.* **4**, 26–28.
- Werker, J. F. & Tees, R. C. 1983 Developmental changes across childhood in the perception of non-native speech sounds. *Can. J. Psychol.* **37**, 278–286. (doi:10.1037/h0080725)
- Werker, J. F. & Tees, R. C. 1984 Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behav. Dev.* **7**, 49–63. (doi:10.1016/s0163-6383(84)80022-3)
- Werker, J. F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L. & Amano, S. 2007 Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition* **103**, 147–162. (doi:10.1016/j.cognition.2006.03.006)
- Yoshida, K., Fennell, C., Swingley, D. & Werker, J. F. 2009 14-month-olds learn similar-sounding words. *Dev. Sci.* **12**, 412–418. (doi:10.1111/j.1467-7687.2008.00789.x)