Global and detailed speech representations in early language acquisition

PIERRE HALLÉ AND ALEJANDRINA CHRISTIA

Abstract: We review data and hypotheses dealing with the mental representations for perceived and produced speech that infants build and use over the course of learning a language. In the early stages of speech perception and vocal production, before the emergence of a receptive or a productive lexicon, the dominant picture emerging from the literature suggests rather non-analytic representations based on units of the size of the syllable: Young children seem to parse speech into syllable-sized units in spite of their ability to detect sound equivalence based on shared phonetic features. Once a productive lexicon has emerged, word form representations are initially rather underspecified phonetically but gradually become more specified with lexical growth, up to the phoneme level. The situation is different for the receptive lexicon, in which phonetic specification for consonants and vowels seem to follow different developmental paths. Consonants in stressed syllables are somewhat well specified already at the first signs of a receptive lexicon, and become even better specified with lexical growth. Vowels seem to follow a different developmental path, with increasing flexibility throughout lexical development. Thus, children come to exhibit a consonant-vowel asymmetry in lexical representations, which is clear in adult representations.

1 Introduction

To begin with, what do we mean by speech representations? We simply refer to the mental representations that speakers/listeners of a given language have built during acquisition and use to produce and understand spoken utterances of their language. We adopt the generativist view according to which production and perception of speech is accomplished via the manipulation -in production or perception- of basic speech units that combine into higher order units through the application of grammatical rules. Words are combined units with respect to, for example, phonemes but are basic units with respect to multi-word sentences. In that sense, words play a pivotal role at the interface between basic sound

units and sentences. Importantly, the units we are talking about are not just useful to describe languages and how languages work but are meant to have a psychological reality in language users' minds, following the goals of modern linguistics, as summarized in the following passage:

"There has always been a tension between two ways of understanding linguistics: On one view [...] (which was dominant in the first part of this century), [language] has a structure that can be explored independently of any efforts to figure out what particular speakers may do or think [...]. On the other view of linguistics (a view that has come to be relatively dominant in the past several decades), *the goal of linguistics is to model what it is that goes inside a speaker's head.*" (Goldsmith (1999), p.5, our stress).

The discussion in this chapter will therefore revolve around the mental representations speakers/listeners use in processing speech. We restrict ourselves to prelexical and lexical units and will not cover the issue of how the rules that combine units are themselves represented. Although many other questions could be posed, this chapter mainly focuses on the following questions: How detailed, in terms of phonetic specification, are the speech representations used by children acquiring their mother tongue? Do representations change throughout development? The literature on language acquisition, from Ferguson and Farwell (1975) onwards converges to suggest that overall children follow a holistic-to-analytic progression in the way they code words (from whole-word units to decomposed representations), at least in production. The motivation of such a progression is clear: The need to adopt systematic strategies to code words increases with vocabulary size, eventually leading to analytic representations into unit combinations. Work in the last 20 years has led to a refinement of this proposal by exploring lexical specification in both production and perception. In addition, some of this work suggests that there is an asymmetry in terms of the level of specification of consonants and vowels. In this chapter, we present a global view of both classical and recent results bearing on this proposal. This chapter is organized as follows: Sections 2 and 3 deal with prelexical infants' speech representations in production, beginning with the babbling; and in perception, beginning with newborns' speech

perception. The following two sections address the lexical representations that emerge around 10-11 months and often grow dramatically around 18-20 months. Section 4 briefly surveys the "child phonology" literature on early word representations in production; section 5 summarizes recent findings on early lexical representations in the receptive lexicon and on how these representations develop as a function of language acquisition. The final section (section 6) surveys adult perception and linguistic data indicating that consonants and vowels serve somewhat different linguistic functions and reviews recent child data bearing on a consonant-vowel asymmetry in lexical representations.

2   Speech representations for production in pre-lexical children

Before they produce (or are discovered to produce) their first words, young children normally go through different stages of vocal productions (see, among others, Oller (1980, 2000); Stark (1980); Kent and Murray (1982)). Among these, vocalizations and babbling undoubtedly are intentional, voluntary productions. The early vocalizations of young children usually are long, sustained vowels modulated in pitch and intensity, with the occasional occurrence of consonant-like onsets, thereby forming "proto-syllables" (Oller, 1980). Vocalizations thus are certainly not quite rich in terms of phonetic detail. The picture changes notably with babbling. Babbling is followed by the child's first words and is characterized by the production of syllables roughly conforming with the syllables of adult speech in terms of timing. These syllables are often reduplicated a few times (canonical babbling) but possibly also differ from one another with respect to vowel or consonant (variegated babbling). Although the dominant opinion is that variegated and canonical babbling appear more or less simultaneously (MacNeilage and Davis, 2002), this issue (Vihman et al., 1985) is still debated. The disagreement may be due to the difficulty of defining phonetic variation vs. phonetic constancy when it comes to describing the phonetic content of babbling productions. In other words, the phonetic substance of babbling productions might not be well defined in terms of phonetic categories. This suggests phonetically underspecified rather than detailed mental rep-

resentations in production, although it could be the case that phonetic variation in babbling is in part explained by an incomplete maturation of the vocal production system. At any rate, babbling productions are more readily described in terms of syllables than of consonants and vowels, as the defining characteristic of babbling itself suggests: Babbling consists of adult-like syllables.

In the frame-then-content (henceforth, F-then-C) view of children's babbling and first word productions promoted by MacNeilage and Davis (Davis and MacNeilage, 1990, 1995; MacNeilage, 1997), the syllable indeed explicitly appears as the basic unit of production: One syllable corresponds to precisely one cycle of mandibular oscillation, that is, to one "frame". The F-then-C account proposes that frames are initially underspecified segmentally. Early frames are "pure frames", only specified by a cyclic closing-opening movement of the jaw superimposed on laryngeal voice excitation. The result is heard as CVs whose Cs and Vs mechanically reflect unintentional, targetless positioning of tongue (and lips) riding passively on the jaw oscillatory cycle (typically a labial obstruent and a central vowel): Thus, these frames have no "content" (see Hodge (1989), for a similar idea). Content appears when voluntary maneuvers of the articulators are superimposed on the closing-opening cycle. At this stage, three more elaborate frames emerge, namely "front frames", "back frames", and "nasal frames", in which just one single articulatory parameter is set: Tongue position for front vs. back frames and velum opening for nasal frames (Matyear et al., 1998). Such minimal specifications are thought to prevail until rather late in language development, that is, until around 16-18 months (MacNeilage, 1996, 1997). In particular, the three basic frames "pure", "front", and "back" seem to explain most of the consonant-vowel cooccurrence data, although the dominant patterns of cooccurrence are somewhat debated (for a review, see Chen and Kent (2005)). The F-then-C account contends that the CV cooccurrence patterns observed for children during speech acquisition reflect a universal trend (MacNeilage and Davis, 2000). Whether or not such a universal trend exists (see Whalen et al. (ress), for a discussion), babbling and early word productions are largely underspecified according to the F-then-C account: Only four types of frames, hence four classes of syllables make up the building bricks of intended utterances and each class is defined by a single parameter.

The articulatory phonology approach (Browman and Goldstein, 1989, 1992) similarly holds that the syllable is the time frame wherein oscillatory systems are synched to produce consonant and vowel gestures. However, instead of positing syllabic gestures specified by a single parameter for both consonants and vowels, this approach considers that consonant and vowel gestures, although initially achieved with great imprecision, are *intended* separately. That certain CV cooccurrences are favored over others (e.g., front vowels follow alveolar rather than velar consonants) and are more noticeable in child than adult speech is attributable to gestural overlap. Young children still do not control well phasing relationships and durations - which are at the heart of articulatory phonology - thereby producing variable and unwanted gestural overlap. Articulatory phonology thus proposes a similar account for early speech CV cooccurrences and for assimilation processes: Gestural overlap. Browman and Goldstein (1992) suggest that children's early speech productions reduce to a few "dynamically stable patterns", wherein C and V gestures remain undifferentiated and are not accurately phased together. Children then progressively learn to differentiate these patterns into separate C and V gestures, eventually acquiring CV combinations specific to the language they learn (de Boysson-Bardies, 1993). The gestural approach thus also describes early speech production as initially underspecified for consonants and vowels and implicitly suggests that later emerging CV specification is still constrained by the syllabic time frame.

To sum up, children's early speech productions seem to be underspecified in terms of consonants and vowels and, rather, to be specified in terms of syllables as whole units. The F-then-C account holds that syllable-based speech productions - mostly specified by place only - are still the rule at the stage of early words and that children gradually escape this pattern through C and V variegation. Articulatory phonology assumes a less constrained development, which tends toward children's learning of phasing relationships within CV syllables.

3   Speech representations for perception in pre-lexical children

In the preceding section, the nature of speech representations in production was inferred from the extent to which vowels and consonants are independently controlled in children's productions. Speech representations in perception can be inferred from several other sources of evidence. Yet, the relevant data seem to converge toward a similar conclusion: Young children code speech in terms of syllables. One source of evidence is provided by the capacity of newborns to "count" syllables rather than phonemes within simple utterances. This finding was first reported by Bijeljac-Babic et al. (1993). These authors used a habituation-dishabituation paradigm based on the classic High Amplitude Sucking (henceforth, HAS) procedure (Eimas et al., 1971). In one experiment, four-day-old infants habituated to a set of disyllabic CVCV speech stimuli, such as {*rifu, kepa...*}, dishabituated when presented with a set of trisyllabic CVCVCV stimuli, e.g. {*mazopu, rekiva...*}, and vice-versa. Discrimination was not based on overall stimulus duration but on stimulus syllabic structure, as shown by another experiment in which the distributions of stimulus durations for the two sets were made to overlap by expanding and compressing the stimuli. This manipulation preserved the discrimination between the two- and three-syllable sets. An alternative explanation for these results simply is that infants discriminate between four- and six-phoneme utterances (CVCVs and CVCVCVs). Bijeljac-Babic et al. tested this possibility in comparing stimuli with four vs. six phonemes but all with two syllables (e.g., {*rifu, iblo*} vs. {*treklu, suldri*}). Infants did not discriminate 4-phoneme from 6-phoneme disyllabic items. Altogether, these data therefore strongly suggest that infants are more sensitive to syllabic rather than to phonemic units in the input speech. Another study by the same group further suggested that French infants count syllables rather than moras (Bertoncini et al., 1995): Amongst disyllabic items, infants did not discriminate 2-mora items (e.g., {*kago, mika, seki, buke*}) from 3-mora ones (e.g., {*kaNgo, mikaN, seQki, buuke*}).
Another source of evidence, in line with the idea that infants perceive speech as a string of syllables, is provided by Bertoncini and Mehler

(1981). The same acoustic contrast may be discriminated or not by two-month-olds (French-learning) according to whether the stimuli sound like speech, with a salient syllabic structure, or not. French two-month-olds do not discriminate [tʃp]-[pʃt] but do discriminate [utʃpu]-[upʃtu]. Indeed, syllabic structure is presumably much more salient in the latter than the former contrast, at least for French-learning infants. A more salient syllabic structure would allow infants to process the stimuli as speech and parse them into syllables they can discriminate.

A third line of evidence is provided by experiments that addressed the level of detail infants code within the syllable using the standard HAS procedure. For example, in (Jusczyk, 1987), American two-month-olds were habituated to a set of CVs sharing C ([bi, bo, bɚ, ba]). After habituation (post-shift phase), a new CV stimulus with a new C ([di]), a new V ([bu]), or both a new C and V ([du]) was added to the initial set of stimuli. The assumption was that infants would dishabituate with the [d] but not the [b] new stimuli, had they been able to extract and code [b] as a "sound" common to all the habituation CVs. But infants equally dishabituated for [d] and [b]. There was thus no evidence that two-month-olds can identify a syllable-initial consonant as a property shared by several syllables with different vowels. In other words, infants that age do not seem to extract and code a consonant within a syllable. Bertoncini et al. (1988) replicated Jusczyk and Derrah's (1987) results, using the same stimuli and procedure, with French two-month-olds. They also tested newborns who differed from two-month-olds in that they dishabituated only when the new stimulus introduced a new vowel (i.e., for [bu] and [du], not [da]). In a second experiment, infants were tested with an habituation set of CVs sharing V ([bi, si, li, mi]), and post-shift new CV stimuli with a new V ([ba]), a new C ([di]), or both a new V and C ([da]). The very same pattern of results obtained: Two-month-olds dishabituated for any new CV stimulus, whereas newborns dishabituated only when the new stimulus introduced a new vowel (i.e., for [ba] and [da], not [di]). These data were taken to suggest that very young infants code CV syllables as whole units, and that newborns are insensitive to consonant variation. The newborns' reduced sensitivity to consonants compared to vowels may reflect their experience in utero, which filters much of the high-frequency energy relevant for consonantal contrasts while relatively preserving vowels. Do older prelexical in-

fants come to represent consonants and vowels separately? The answer to this question may come from a more recent line of research, where young infants' encoding of sound sequences has been investigated. In this work, infants hear a large number of wordforms (50-100), all of which bear the same abstract sequence; for example, nasal vowels are always followed by fricative consonants, and oral vowels by stops. At test, infants are presented with new wordforms, some of which follow the abstract pattern (i.e., a new nasal vowel is followed by a fricative) while others do not (a new nasal vowel is followed by a stop). Prelexical infants, some tested as young as 4 months, exhibit stable preferences, typically for the more novel-sounding illegal patterns (see a summary in Cristia et al. (2011). This appears to indicate that prelexical infants come to represent consonants and vowels specifically. However, a recent study (Cristia and Peperkamp, 2011) suggests that such results may be best accommodated through acoustic, whole-word representations. In this study, six-month-olds were first familiarized with word forms sharing onset voicing. They were then tested on new word forms in which onset voicing and novelty (from the familiarization set or not) were manipulated. They preferred the new voicing and this was not due to novelty only since they showed no preference for novel over familiarization word forms when voicing was kept constant. At first sight, the infants' behavior could be explained by their reliance on the [voice] feature value in onsets. Yet, the results better fitted acoustic distances between word forms computed on entire items than on item onsets only. This suggested that the infants' behavior actually reflected acoustic- or auditory-based comparison between whole-word forms rather than feature-based analytic representations.

To summarize, prelexical infants appear to rely on syllabic representations, or perhaps on whole-word forms, in which consonants and vowels are bound together. Now what about the lexical stage? Do young children who are starting a productive or a receptive lexicon represent words (be they production targets or recognized spoken items) as composed of syllables or of, for example, consonants and vowels?

4 Lexical representations in production: Children's early words

Most evidence suggests that children first go through a whole-word stage during which the word is the basic unit (the "prosodic word": Macken (1978, 1979); see also Vihman (1996)). Following that stage, children gradually develop more adult-like phonological representations, that is, rule-based representations gradually leading to principled segmental and featural units. Some children, usually during the second year, develop a few templates (i.e., word patterns) consisting of a stable skeleton of consonants - or of consonants and vowels - that constrains all their attempted words. This is well illustrated in Macken's classic longitudinal study (Macken, 1978) from 1;6 to 2;5 years of the words produced by "Si", a child raised in a Mexican Spanish environment. From 1;7 to 1;9, virtually all the words attempted by this child followed a "labial-dental" disyllabic word pattern, whether this pattern reflected the adult model or not. For example, she produced *zapato* ('shoe') as [pwatːo], closely following the adult model, but also *sopa* ('soup') as [pwæta], reversing the dental-labial order of the adult model, or even *reloj* ('clock') as [budːo], although the adult model had no labial consonant. These single word patterns have been interpreted as reflecting non-analytic representations of produced speech into "whole-word" units. The rigid single pattern followed in virtually all produced words indeed suggests nondecomposed representations in terms of consonants or in terms of syllables. Neither the consonant skeleton nor the vowel pattern of the adult model are usually preserved in the productions of the children who follow the path of a single template whole-word stage, suggesting these children do not analyze words into segments. However, children are quite variable with respect to the observable path they follow in acquiring words. This is illustrated in Table 1, drawn from de Boysson-Bardies (1996): At comparable chronological age and/or estimated productive vocabulary size, early words are close to their adult model in some children (suggesting analytic coding), but rather underspecified in other children. Despite such individual variation, children's first words often are rather holistic approximations of the adult word forms. De Boysson-Bardies' data (e.g., Table 1) suggest that this holis-

Table 1: First words produced by two 14-month-old French children (with a productive vocabulary estimated at 30-40 words), from de Boysson-Bardies (1996); adult glosses in italics. Marie's productions may be described as "analytic" and those of Émilie as "holistic".

Émilie		Marie	
[ba]	balle	[ættæ]	attend
[bø]	bouton	[hatø]	bateau
[bebe]	bébé	[bebe]	bébé
[poe]	pomme	[dodo]	dodo q
[pɔ]	chapeau	[təbɔ]	c'est beau
[popo]	petit pot	[ebotsa]	c'est beau ça
[ka]	canard	[taːtinn]	tartine
[ke]	clef	[papɨdʒɐ]	papillon
[kkɪ]	cuillère	[voaʃy]	voiture
[kχ]	Mickey	[hemjetsɐ]	mimichat
[qa]	sac	[popi]	poupée

tic stage is difficult to notice in some children. Holistic approximations may be underlain by whole-word rather than analytic representations, as was proposed in the classic paper by Ferguson and Farwell (1975). Throughout development, however, word form representations become more and more clearly organized in terms of segments in that their consonants and vowels do not depart from those of adult forms in an erratic way but, rather, in a progressively more systematic way (see, for example, Vihman and Greenlee (1987)). This increasing systematicity is quite appealing for phonologists and a huge "Child Phonology" literature has been devoted to describing in terms of phonological processes how child forms differ from adult forms. We will not expand on that aspect and simply note that most of the processes described in this literature involve consonants rather than vowels (Vihman and Greenlee (1987), but see Pollock and Berni (2003) on both normal and disordered acquisition of English vowels). This might be an indication that consonants are more important than vowels in children's early lexical representations or, possibly, that vocalic variations in children's early words are less easily noticed by adult hearers and therefore get underreported (see section 6 on adults' greater sensitivity to consonant than vowel vari-

ation). To sum up, a gross analogy is observed between prelexical and lexical representations for production in that both become increasingly detailed throughout development.

5 Children's receptive lexical representations in word-learning and word-recognition

Children recognize word forms earlier than they can produce them intentionally. A further, and logically more difficult accomplishment is to consistently associate word forms with meanings. Many studies have used a variety of "word-learning" tasks, whereby children learn made-up associations between novel words and novel objects, and have directly addressed the issue of how much phonetic detail children are able to code for newly learned words. A somewhat different line of research has focused on how children code the words they have learned either during experimental training or from natural exposure to the language spoken in their linguistic environment, and has addressed that issue with "word recognition" tasks. We will see that the two approaches yield somewhat different pictures of children's representations of word forms. Six-eight months is about the youngest age investigated by word recognition studies, such as those conducted by Peter Jusczyk's group. Those studies show that 8-month-old American infants can "segment", or pull out monosyllabic words out of continuous speech and retain them at least for the duration of an experimental session. For example, Jusczyk and Aslin (1995) trained 6- and 7.5-month-olds with two words (*cup* and *dog* for half the infants, *foot* and *bike* for the others) appearing repeatedly in a few sentences during a familiarization phase, then tested them on their preference for trained over untrained words, using the now classical Headturn Preference Procedure (henceforth, HPP). They found that 7.5-month-olds but not 6-month olds preferred listening to the two words they had been trained on over the two other, untrained words. However, this preference, suggesting word recognition, did not resist a change in word-initial consonant: Infants trained with *cup* and *dog* did not prefer *tup* and *bawg* over untrained words. Hence, word form representations at this age seem phonetically detailed, at least with respect to the onset consonant.

Older infants (at about 11 months) have been shown to recognize words presumably familiar to them through natural, not experimental exposure (Hallé and de Boysson-Bardies, 1994; Vihman et al., 2004; Swingley, 2005). For example, 11-month-old French infants prefer listening to *ballon* ('balloon') than *félin* ('feline': presumably not a familiar word for children) right away, that is, without experimental training on *ballon* (Hallé and de Boysson-Bardies, 1994). This suggested 11-month-olds have coded some familiar words in long term memory, in an early receptive lexicon. Further studies examined how strictly infants might code word forms: How detailed are word form representations in infants' early receptive lexicon? Hallé and de Boysson-Bardies (1996) used mispronunciations of familiar words to address that issue. They found that 11-month-old French infants still preferred mispronounced familiar over unfamiliar words, when the mispronunciation affected the word-initial consonant (e.g., *poupée* ('doll') > *boupée* or *foupée*). Moreover, they did not prefer unaltered over mispronounced familiar word forms. The preference for mispronounced familiar over unfamiliar words tended to fade away when the mispronunciation affected the word-medial consonant (e.g., *poupée* > *poufée*). Omission of the word-initial consonant resulted in no preference at all for mispronounced familiar words (e.g., *poupée* > *oupée*). As proposed by Hallé and de Boysson-Bardies (1996), these data thus show a relative elasticity of 11-month-olds' wordform representations. It might seem surprising that infants tolerate word-initial but not word-medial consonant mispronunciation. More recent work by Vihman et al. (2004) provides a clue to this question. Vihman et al. (2004) tested 11-month-old British infants with familiar vs. unfamiliar disyllabic English words or noun phrases (so as to manipulate stress placement). The results they obtained were similar to those in Hallé and de Boysson-Bardies (1996) but with an important difference. The British infants tolerated mispronunciations of word medial consonant but not of word initial consonant (e.g., *dirty* > *dirny* tolerated; *dirty* > *nirty* not tolerated). This pattern is the opposite of that found for French. Since English and French words have opposite dominant metric patterns (trochaic and iambic, respectively), a sensible interpretation is that 11-month-olds rather strictly code stressed syllables and less strictly so unstressed syllables in their early receptive lexicon (Figure 1). This is compatible with the earlier finding by Jusczyk and Aslin (1995) that 7.5-

month-olds fail to recognize *tup* and *bawg* after they have been trained on *cup* and *dog*, since monosyllabic content words are stressed. Similarly, Swingley (2005) replicated the preference for familiar over novel words in Dutch 11-month-olds, using monosyllabic words. In line with our explanation, this preference was abolished if the words were mispronounced (e.g., *mont* ('mouth') > *nont* or *monk*) and infants systematically preferred unaltered over mispronounced familiar words. Further data
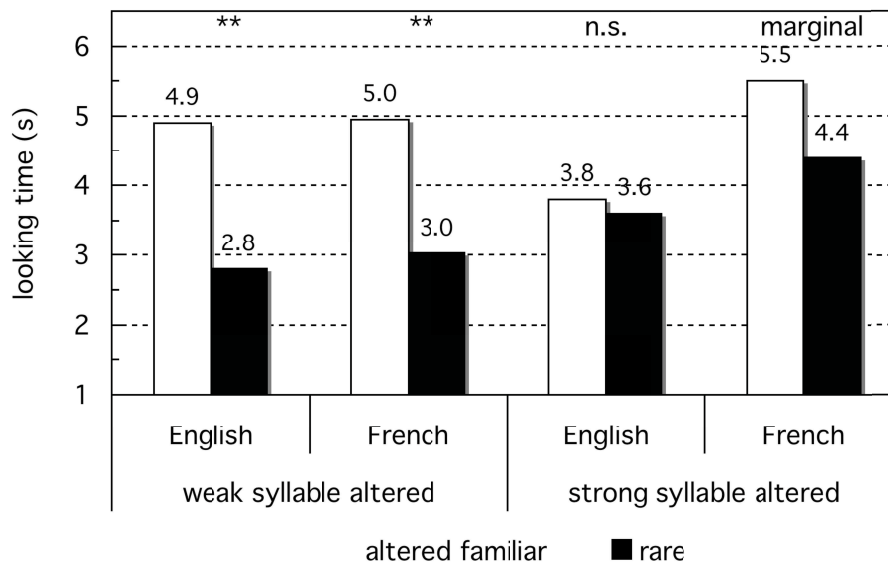
Figure 1: Looking times to disyllabic unfamiliar vs. familiar words, place-altered on the initial consonant of the weak vs. strong syllable. Recognition of familiar words obtains for weak but not strong syllable alteration (from Hallé and de Boysson-Bardies (1996) and Vihman et al. (2004).

with slightly older children confirm the view that infants' word form representations in early receptive lexicon are rather detailed phonetically. These data come from studies using the preferential looking procedure, whereby infants are presented with two pictures on a screen (one target, one distracter), then prompted with a sentence such as "Where is the [target]?" Longer looking times to the target than the distracter are assumed to indicate word recognition (together with knowledge of word-picture association). Swingley and Aslin (2002) found that 14-month-olds recognize both *dog* and *tog* as referring to the picture of a dog rather than that of a shoe, but look longer to that picture when presented with *dog* than *tog*. Bailey and Plunkett (2002) obtained an even more radical advantage for *dog* over *tog*, with 14-month-olds recogniz-

ing *dog* but not *tog*. When children learn new, arbitrary word-object associations, they seem to show some difficulty at coding words with full phonetic detail.

At 14 months, American children, when tested on a word-learning paradigm known as the "Switch" procedure (an habituation-dishabituation paradigm on word-object associations Werker et al. (1998)), do not distinguish *bih* from *dih* (Stager and Werker, 1997; Werker et al., 2002). Children do not succeed in this task before 17 months (Werker et al., 2002). However, when the contrasted words are familiar minimal pairs, such as *ball* vs. *doll* (presumably well known by young children), even 14-month-olds succeed in the Switch task (Fennell and Werker, 2003). But this is not the whole story, as recent data bearing on the "bih-dih issue" suggest: Children's success at associating *bih* and *dih* (or similar minimal pairs) to different referent objects depends on a variety of non-phonetic factors, such as the testing procedure - new minimal pairs can be learned at 14 months with the preferential looking procedure - (Ballem and Plunkett, 2005; Fennell and Waxman, 2010), the distracter items used (Thiessen, 2007), or the simple fact that referent objects move in synchrony with target speech items or not (see Lakshmi Gogate's research on 8-month-olds: e.g., Gogate (2010)). In other words, the picture of phonetic detail in word-learning is quite heterogeneous. If we limit our purpose to drawing a meaningful developmental trajectory, the data obtained with the Switch procedure might be the most telling, since they clearly reveal infants' improving ability to code phonetic detail in newly learned words. We therefore rely on these data so as to draw a coherent picture of children's sensitivity to phonetic detail in newly learned words.

Altogether, then, current research on lexical development, especially that using the Switch procedure, suggests that when children "learn" words, they code word forms in a less phonetically specified way than for the long term representations they have built in their early receptive lexicon over the course of normal language acquisition. Yet, this difference is probably quantitative rather than qualitative: Coding word forms during a word-learning task, after a limited exposure to arbitrary word-object pairs (typically presented only 3-9 times), is logically more demanding than detecting and memorizing recurring word forms from natural exposure to speech over the course of weeks or months. We

therefore propose that children use similar word form representations in both word-learning and in familiar word-recognition. The observed delay in forming phonetically detailed representations in word-learning tasks therefore suggests a developmental trend from initially somewhat underspecified toward fully specified word form representations.

The literature reviewed so far examined the issue of phonetic detail in children's word form representations for consonants (e.g., *ball* vs. *doll*) but not for vowels, or non-systematically so, as in Swingley and Aslin (2000) whose materials include *apple* vs. *opple*. In the next section, we review the recent data obtained for vowel compared to consonant variation.

6   Vowels versus consonants in children's lexical representations

During the first decade or so of its existence, the research on children's word form representations in reception focused exclusively on consonants, from Jusczyck and Aslin's onwards (Jusczyk and Aslin, 1995). This is in itself revealing of an implicit assumption made by researchers that consonants matter more than vowels for lexical forms. There are indeed reasons to believe so. Recently, the logical arguments for a consonant-vowel dissociation in terms of their functional roles were neatly laid out in an often cited paper by Nespor et al. (2003). In this work, the argument is made that lexical forms are mainly specified by consonants; in contrast, vowels carry prosodic, syntactic, and morphosyntactic information, and are less important for lexical identity.

To test these claims, Mehler's group at SISSA ("Scuola Internazionale Superiore di Studi Avanzati" in Trieste) has conducted numerous artificial grammar studies which suggest that adults segment out word forms from a continuous stream of syllables more readily when "words" are defined by consonant rather than vowel patterns (Bonatti et al., 2005). In contrast, other artificial grammar studies suggest that abstract regularities (e.g., structural sequencing such as reduplication, or simple patterns such as AAB, ABA, etc.) can be extracted by listeners when they are defined on vowels but not when they are defined on consonants (adults: Peña et al. (2002); Toro et al. (2008), infants: Pons and Toro (2010)), in line

with the idea that consonants specify words and vowels specify rules (Nespor et al., 2003). Evidence that is more immediately relevant to our question of whether consonants and vowels are coded equally in word representations comes from an older line of work. Word-reconstruction studies provide more convincing evidence for the lexical motivation of consonants because they bear on natural word form representations in long-term lexicon. In these experiments, listeners are asked to produce a word form closest to a pseudoword derived from a word by either vowel or consonant change. For example, *keebra* is derived from *cobra* by a vowel change or from *zebra* by a consonant change. Listeners more often and more easily "reconstruct" *cobra* than *zebra* from *keebra*, thereby preserving consonant rather than vowel information (van Ooijen (1996); see Sharp et al. (2005) for recent behavioral and brain imagery data). Note that this result obtains regardless of phonological inventory. That is, one might imagine that the observed consonant-vowel asymmetry only holds for languages such as English, in which there are many vowels and vowels vary with regional accent. In fact, the asymmetry also obtains in Spanish or Japanese (which have only a handful of vowels) as clearly as in English (Cutler et al., 2000; Cutler and Otake, 2002). Listeners tend to provide words that match the consonantal frame more frequently than those that match the vocalic frame, suggesting that word form representations are universally based on consonants more than vowels. Now is there evidence for the lexical coding C-V asymmetry in child data? Nazzi's study (Nazzi, 2005) marks a turning point in the children's lexical representation research in that it showed, for the first time, that consonants weigh more than vowels in 20-month-olds' word form representations. Nazzi (2005) used the "name-based categorization" paradigm (henceforth, NBC): In this paradigm, children are first taught three name-object pairings, in which two different objects share the same label (e.g., /pize/) and the third object has a different label (e.g., /tize/). In the test phase after each such triplet, the child's task is to put together the objects that share the same label; their success in the task indicates, among other things, their ability to learn two different labels (Nazzi and Gopnik, 2001). Of interest here is how the two labels differ. Nazzi (2005) manipulated this difference in various ways. In particular, he compared differences in consonant with differences in vowel (e.g., /pize/-/tize/ vs. /pize/-/pyze/); he found that 20-month-old

French children performed better for consonant than vowel differences, showing a "consonantal bias" in word-learning. Further follow-up experiments using the NBC procedure demonstrated that this bias was not confounded with a positional bias, that is, a possible advantage for syllable-initial phonemes (Nazzi and Bertoncini, 2009). The consonantal bias was also found for younger children, at 16 months, using a simplified version of the NBC task (Havy and Nazzi, 2009). Using variants of the NBC paradigm, Nazzi and colleagues (Nazzi et al., 2009) found that both French- and English-learning 30-month-olds are able to learn a single feature vowel change (/pize/-/pyze/) but still rely more on consonants when asked to match a mispronounced form with a learned label (e.g., they match /pide/ with /pyde/ rather than /tide/[1]). Altogether, this line of research suggests a developmental trend toward more attention paid to vowel variation (at 30 months), but still with an advantage for consonants over vowels. We return to this developmental issue in the general discussion.

A recent study by the SISSA's group suggests this consonant-vowel functional distinction also holds for 12-month-olds. Using a variant of the preferential looking paradigm, whereby an auditory word predicts the apparition of an associated picture at one side of the screen (Kovács and Mehler, 2009; Hochmann, 2010; Hochmann et al., 2011) showed that, when encoding words, 12-month-olds give more weight to consonants than vowels: After they learn to associate *keke* with side A and *dudu* with side B, they orient in anticipation to side A rather than B when presented with *kuku*, and vice versa with *dede*.[2] That is, the consonant code prevails over the vowel code for word recognition.

In an other experiment, Hochman showed that, when extracting a second order regularity, 12-month-olds give more weight to vowels than consonants: They do learn to associate vowel but not consonant reduplication with one side of the screen. Altogether, then, these children behaved just like adults: They rely on consonants to code lexical forms and on vowels to extract rules.

---

1 This experimental design is thus strongly reminiscent of the word-reconstruction paradigm used with adults. Note that it allows for directly comparing tolerance to consonant vs. vowel variation.

2 This experimental design is a constant target design, in contrast to most other studies, thus avoiding possible confounding factors in the consonant-vowel comparison.

Another line of research, conducted by Plunkett's group and using the preferential looking procedure, investigated the consonantal bias in word-recognition for well-known words. Mani and Plunkett (2007) found that English-learning children, aged 15, 18, and 24 months, were sensitive to both consonant and vowel mispronunciations. They did not look at the image of a *bib* (a well known word for children this age) when hearing either *bab* or *dib* (vowel or consonant mispronunciation, respectively), although 15-month-olds non-significantly tended to look at the *bib* upon hearing *bab*. Mani and Plunkett (2010) also found that British 12-month-olds were as sensitive to vowel as to consonant changes in familiar words (e.g., *cups* mispronounced *keps* or *tups*). Do these results contradict those obtained in Nazzi's group? A possible reason for the observed discrepancies might be the difference in word form representation for word-learning and word-recognition, as we discussed earlier. However, Mani and Plunkett (2008) obtained vowel-mispronunciation effects using novel words learned by 14-month-olds (e.g., learned padge mispronounced *poudge*). Likewise, Curtin et al. (2009) (see also Dietrich et al. (2007)), using the Switch procedure, found that English-learning 14-month-olds successfully learned a *deet-dit* pair (although not a *deet-doot* pair or a *dit-doot* pair). Note that these negative results with respect to the consonantal bias run contrary to the current findings with adult listeners from various languages mentioned above.

To sum up, the consonantal bias in children's word form representations is still a matter of debate. The discrepancies found across the various studies just reviewed may be due to a variety of factors: Children's target language, children'age, type of lexical representation (for word-learning vs. word-recognition), type of word form (e.g., mono- vs. disyllabic), or experimental procedure. Yet, on theoretical grounds and on the basis of data consistency, the dominant picture is an advantage of consonants over vowels for lexical coding. Interestingly, we note that the consonant bias in adults, which is rather well established, may have consequences in their report of children's early produced word forms: Child phonology studies have mainly focused on the regular or quasi-regular consonant variation in children's early words. Alternatively, it might be the case that vowel variation in these productions is random rather than regular. This point certainly deserves systematic research.

7   General discussion

In this review of prelexical and lexical representations in children, we came across rather different issues, addressed by various research groups. Yet, one aspect seems to characterize the entire spectrum of data: Children start with somewhat vague, global, holistic, underspecified representations of sound sequences and later of words, in both perception and production; they then orient toward more precise, analytic, phonetically specified representations.

Representations certainly cover different realities for perception and production, for prelexical sound sequences and for words. We focus in this discussion on *word form* representations but occasionally refer to prelexical representations. While representations for production definitely develop from whole-word to analytic representations into phoneme-like elements at around 2 years of age (Macken, 1978, 1979), those for perception follow a less clear, and less agreed-on developmental trajectory. Difference between word-learning and word-recognition might help understand this trajectory. As we argued, the kinds of representation that infants use in word-learning tasks and for coding/recognizing words in long term memory (in their receptive lexicon) are conceivably similar and should follow parallel, although somewhat time-delayed developmental trajectories. Whereas little change is observed for word recognition representations in terms of phonetic detail, a change from holistic to analytic is more apparent for word-learning between about 14 and 18 months (section 5). Such a change is also predictable on theoretical grounds: According to Metsala, "...representations of lexical items may become increasingly segmented (phonemic) with development from the pressure of an increasing vocabulary size." Metsala (1997): 161, (see also Metsala and Walley (1998); Walley (1993); Walley and Metsala (1990)). We therefore surmise that word form representations in children's early receptive lexicons gradually shift from holistic to analytic formats during the second year of life.

Now, this holistic-to-analytic shift is not the whole story. Indeed, as the recent research highlights, consonants and vowels play different roles and it seems that consonants are eventually more strictly coded than vowels in the receptive lexicon. Thus, the holistic-to analytic picture needs some qualification. The current literature on consonant-vowel

asymmetries suggests that consonants are strictly coded very early on: Even 7.5 month-olds recognize *cup* but not *tup* after they are familiarized with *cup* (Jusczyk and Aslin, 1995). But some confusion exists in the currently available data with respect to vowels. Prelexical data suggest that vowels initially have more weight that consonants (Bertoncini et al., 1988), which is also in line with the classical findings that vowels are perceived in a language specific way earlier than consonants (vowels: around 5-6 months (Kuhl et al., 1992); consonants: around 8-10 months (Werker and Tees, 1984)). For lexical coding, Mani and Plunkett's data (Mani and Plunkett, 2007) suggest a more elastic coding of vowels at 15 than 18 or 24 months. Nazzi et al.'s data (Nazzi et al., 2009) suggest that sensitivity to vowel variation increases between 20 and 30 months. Yet it seems that these data also depend on which consonant or vowel contrast is tested (consonants: Pomiechowska (2011); vowels: Curtin et al. (2009)). It is therefore clear that more research is needed in this domain. A study by Best and colleagues (Best et al., 2009) might serendipitously shed light on that issue. Although it was primarily aimed at detecting the emergence of "phonological constancy", this study suggests that 15-month-old American children are sensitive to vowel quality for the familiar words in their receptive lexicon. Indeed, 15-month-olds did not recognize familiar words[3] when spoken in Jamaican English, although they did recognize them when spoken in their natice Connecticut American English. Importantly, vowel quality is the main difference between these two English dialects.[4] At 19 months, children recognized familiar words, whether spoken in American or Jamaican English. That is, somewhere between 15 and 19 months, children develop the ability to ignore irrelevant vowel variation in lexical items. In our perspective, vowel information in known words becomes coded in a more and more flexible way with linguistic experience. Such a developmental trajectory seems more cogent with a general developmental trend from infancy

---

3 Best et al. (2009) used a conditioned-fixation version of Hallé and de Boysson-Bardies's (1994) familiar-word preference task.

4 Jamaican and American English mainly differ on vowels: different locations in vowel space (Wassink, 2006), different diphthongs, and different tense-lax distinctions. Jamaican English also tends to avoid vowel reduction, hence to be more syllable-timed than American English. The two dialects differ on some consonants (e.g., Jamaican English merges /f/ and /θ/, and sometimes drops /h/ or /r/, etc.); they also may differ on lexical stress patterns (Patrick, 1999).

to adulthood. In the adult state, vowels indeed have a lesser weight than consonants for coding lexical information. As we mentioned earlier, the consonant-vowel asymmetry in adults is demonstrated in word-reconstruction studies (Cutler et al., 2000; Sharp et al., 2005; van Ooijen, 1996), as well as in artificial grammar studies (Bonatti et al., 2005). Note that the perception data reviewed in section 3 for newborns vs. 2-month-olds (Bertoncini et al., 1988) may suggest a precursor of the change toward the C-V asymmetry observed in adults: Newborns appear more sensitive to vowel than consonant variation, whereas 2-month-olds are equally sensitive to both and seem to only attend to CV whole syllables. To conclude, whereas we agree with the general notion that lexical coding becomes more and more specified overall throughout the development of both productive and receptive lexicons - logically under the pressure of lexical growth -, we believe this notion needs an important qualification in the case of the receptive lexicon. With respect to consonants, high sensitivity to phonetic detail in word form representations can be observed very early on, at least in stressed syllables, and may only slightly increase in older children. Children's sensitivity to vowels seems to follow an opposite trajectory. We propose that, contrary to the general holistic-analytic developmental trend, children's lexical representations become more and more flexible for vowels with increased experience with the language they learn. Only this developmental path is compatible with the adult state of affairs, whereby the consonant-vowel asymmetry is clearly established.

References

Bailey, T. and Plunkett, K. (2002). Phonological specificity in early words. *Cognitive Development*, 17:1265–1282.

Ballem, K. and Plunkett, K. (2005). Phonological specificity in children at 1;2. *Journal of Child Language*, 32:159–173.

Bertoncini, J., Bijeljac-Babic, R., Jusczyk, P., Kennedy, L., and Mehler, J. (1988). An investigation of young infants' perceptual representations of speech sounds. *Journal of Experimental Psychology: General*, 117:21–33.

Bertoncini, J., Floccia, C., Nazzi, T., and Mehler, J. (1995). Morae and syllables: Rhythmical basis of speech representations in neonates. *Language and Speech*, 38:311–329.

Bertoncini, J. and Mehler, J. (1981). Syllables as units in infant speech perception. *Infant Behavior and Development*, 4:247–260.

Best, C., Tyler, M., Gooding, T., Orlando, C., and Quann, C. (2009). Development of phonological constancy: Toddlers' perception of native- and Jamaican-accented words. *Psychological Science*, 20:539–542.

Bijeljac-Babic, R., Bertoncini, J., and Mehler, J. (1993). How do four-day-old infants categorize multisyllabic utterances. *Developmental Psychology*, 29:711–721.

Bonatti, L., Peña, M., Nespor, M., and Mehler, J. (2005). Linguistic constraints on statistical computations: The role of consonants and vowels in continuous speech processing. *Psychological Science*, 16:451–459.

Browman, C. and Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology*, 6:201–251.

Browman, C. and Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, 49:155–180.

Chen, L.-M. and Kent, R. (2005). Consonant-vowel co-occurrence patterns in mandarin-learning infants. *Journal of Child Language*, 32:507–534.

Cristia, A. and Peperkamp, S. (2011). Generalizing without encoding specifics: Infants infer phonotactic patterns on sound classes. In *Communication at the 36th annual Boston University Conference on Language Development*.

Cristia, A., Seidl, A., and Francis, A. (2011). Phonological features in infancy. In Clements, G. and Ridouane, R., editors, *Where do Phonological Contrasts Come from? Cognitive, Physical and Developmental Bases of Phonological Features*, pages 303–326. Benjamins, Amsterdam.

Curtin, S., Fennell, C., and Escudero, P. (2009). Weighting of vowel cues explains patterns of word-object associative learning. *Developmental Science*, 12:725–731.

Cutler, A. and Otake, T. (2002). Rhythmic categories in spoken-word recognition. *Journal of Memory and Language*, 46:296–322.

Cutler, A., Sebastian-Galles, N., Soler-Vilageliu, O., and van Ooijen, B. (2000). Constraints of vowels and consonants on lexical selection: Cross-linguistic comparisons. *Memory & Cognition*, 28:746–755.

Davis, B. and MacNeilage, P. (1990). Acquisition of correct vowel production: a quantitative case study. *Journal of Speech and Hearing Research*, 33:16–27.

Davis, B. and MacNeilage, P. (1995). The articulatory basis of babbling. *Journal of Speech and Hearing Research*, 38:1199–1211.

de Boysson-Bardies, B. (1993). Ontogeny of language-specific syllabic productions. In de Boysson-Bardies, B., de Schonen, S., Jusczyk, P., MacNeilage, P., and Morton, J., editors, *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*. Kluwer, Dordrecht.

de Boysson-Bardies, B. (1996). *Comment la Parole Vient aux Enfants.* Odile Jacob, Paris.

Dietrich, C., Swingley, D., and Werker, J. (2007). One-year- olds' language-specific phonological categorization in word learning: a cross-linguistic study. *PNAS*, 104(41):16027–16031.

Eimas, P., Siqueland, E.and Jusczyk, P., and Vigorito, J. (1971). Speech perception in infants. *Science*, 171:303–306.

Fennell, C. and Waxman, S. (2010). What paradox? Referential cues allow for infant use of phonetic detail in word learning. *Child Development*, 81:1376–1383.

Fennell, C. and Werker, J. (2003). Early word learners' ability to access phonetic detail in well-known words. *Language and Speech*, 46:245–264.

Ferguson, C. and Farwell, C. (1975). Words and sounds in early language acquisition. *Language*, 51:419–439.

Gogate, L. (2010). Learning of syllable-object relations by preverbal infants: The role of temporal synchrony and syllable distinctiveness. *Journal of Experimental Child Psychology*, 105:178–197.

Goldsmith, J. (1999). Introduction. In Goldsmith, J., editor, *Phonological Theory: The Essential Readings*, pages 1–16. Blackwell, Malden MA.

Hallé, P. and de Boysson-Bardies, B. (1994). Emergence of an early lexicon: Infants' recognition of words. *Infant Behavior and Development*, 17:119–129.

Hallé, P. and de Boysson-Bardies, B. (1996). The format of representation of recognized words in infants' early receptive lexicon. *Infant Behavior and Development*, 19:465–483.

Havy, M. and Nazzi, T. (2009). Better processing of consonantal over vocalic information in word learning at 16 months of age. *Infancy*, 14:439–456.

Hochmann, J. (2010). *Categories, words and rules in language acquisition.* PhD thesis, SISSA, Trieste, Italy.

Hochmann, J.-R., Benavides-Varela, S., Nespor, M., and Mehler, J. (2011). Consonants and vowels: Different roles in early language acquisition. *Developmental Science*, 14:1445–1458.

Hodge, M. (1989). *A comparison of spectral-temporal measures across speaker age: implications for an acoustic characterization of speech maturation.* PhD thesis, University of Wisconsin-Madison.

Jusczyk, P. W. and.Derrah, C. (1987). A study of phonetic categorization in 2-month-olds. *Developmental Psychology*, 23:648–654.

Jusczyk, P. and Aslin, R. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, 29:1–23.

Kent, R. D. and Murray, A. (1982). Acoustic features of infant vocalic utterances at 3, 6 and 9 months. *Journal of the Acoustic Society of America*, 72:353–365.

Kovács, A. and Mehler (2009). Flexible learning of multiple speech structures in bilingual infants. *Science*, 325:611–612.

Kuhl, P., Williams, K., Lacerda, F., Stevens, K., and Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, 255:606–608.

Macken, M. (1978). Permitted complexity in phonological development: One child's acquisition of spanish consonants. *Lingua*, 44:219–253.

Macken, M. (1979). Developmental reorganization of phonology: A hierarchy of basic units of acquisition. *Lingua*, 49:11–49.

MacNeilage, P. (1996). Formal versus functional approaches to the acquisition of speech production. In *Paper presented at the Symposium on Research in Child Language Disorders.*

MacNeilage, P. (1997). Acquisition of speech. In Hardcastle, W. and Laver, J., editors, *The Handbook of Phonetic Sciences*. Blackwell, Cambridge MA.

MacNeilage, P. and Davis, B. (2000). On the origin of internal structure of word forms. *Science*, 288:527–531.

MacNeilage, P. and Davis, B. (2002). On the origins of intersyllabic complexity. In Malle, B., editor, *The Rise of Language out of Pre-language*, pages 118–136. John Benjamins, Amsterdam.

Mani, N. and Plunkett, K. (2007). Phonological specificity of vowels and consonants in early lexical representations. *Journal of Memory and Language*, 57:252–272.

Mani, N. and Plunkett, K. (2008). Fourteen-month-olds pay attention to vowels in novel words. *Developmental Science*, 11:53–59.

Mani, N. and Plunkett, K. (2010). Twelve-month-olds know their 'cups' from their 'keps' and 'tups'. *Infancy*, 15:445–470.

Matyear, C., MacNeilage, P., and Davis, B. (1998). Nasalization of vowels in nasal environments in babbling: Evidence for frame dominance. *Phonetica*, 55:1–17.

Metsala, J. (1997). Spoken word recognition in reading disabled children. *Journal of Education Psychology*, 89:159 –169.

Metsala, J. and Walley, A. (1998). Spoken vocabulary growth and the segmental re-structuring of lexical representations: Precursors to phonemic awareness and early reading ability. In Metsala, J. and Ehri, L., editors, *Word Recognition in Beginning Literacy*, pages 89–120. Erlbaum, Mahwah NJ.

Nazzi, T. (2005). Use of phonetic specificity during the acquisition of new words: dif-ferences between consonants and vowels. *Cognition*, 98:13–30.

Nazzi, T. and Bertoncini, J. (2009). Consonant specificity in onset and coda positions in early lexical acquisition. *Language & Speech*, 52:463–480.

Nazzi, T., Floccia, C., Moquet, B., and Butler, J. (2009). Bias for consonantal information over vocalic information in 30-month-olds: Cross-linguistic evidence from French. *Journal of Experimental Child Psychology*, 102:522–537.

Nazzi, T. and Gopnik, A. (2001). Linguistic and cognitive abilities in infancy: When does language become a tool for categorization? *Cognition*, 80:B11–B20.

Nespor, M., Peña, M., and Mehler, J. (2003). On the different roles of vowels and conso-nants in speech processing and language acquisition. *Lingue e Linguaggio*, 2:221–247.

Oller, D. (1980). The emergence of the sounds of speech in infancy. In Yeni-Komshian, G., Kavanagh, C., and Ferguson, C., editors, *Child Phonology 1: Production*, pages 93–112. Academic Press, New York.

Oller, D. (2000). *The Emergence of the Speech Capacity.* Lawrence Erlbaum Associates, Mahwah N.J.

Patrick, P. (1999). *Urban Jamaican Creole: Variation in the Mesolect.* Benjamins: Amster-dam.

Peña, M., Bonatti, L., Nespor, M., and Mehler, J. (2002). Signal-driven computations in speech processing. *Science*, 298:604–607.

Pollock, K. and Berni, M. (2003). Incidence of non-rhotic vowel errors in children: Data from the Memphis Vowel Project. *Clinical Linguistics and Phonetics*, 17(3-4):393–401.

Pomiechowska, B. (2011). Consonantal bias in word learning at 3 years of age: A crosslinguistic study. Master's thesis, LPP Paris.

Pons, F. and Toro, J. (2010). Structural generalizations over consonants and vowels in 11-month-old infants. *Cognition*, 116:361–367.

Sharp, D., Scott, S., Cutler, A., and Wise, R. (2005). Lexical retrieval constrained by sound structure: The role of the left inferior frontal gyrus. *Brain and Language*, 92:309–319.

Stager, C. and Werker, J. (1997). Infants listen for more phonetic detail in speech perception than in word learning tasks. *Nature*, 388:381–382.

Stark, R. (1980). Stages of speech development in the first year of life. In Yeni-Komshian, G., Kavanagh, J., and Ferguson, C., editors, *Child Phonology 1: Production*, pages 73–90. Academic Press, New York.

Swingley, D. (2005). 11-month-olds' knowledge of how familiar words sound. *Developmental Science*, 8:432–443.

Swingley, D. and Aslin, R. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, 76:147–166.

Swingley, D. and Aslin, R. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, 13:480–484.

Thiessen, E. (2007). The effect of distributional information on children's use of phonemic contrasts. *Journal of Memory and Language*, 56:16–34.

Toro, J., Bonatti, L., Nespor, M., and Mehler, J. (2008). Finding words and rules in a speech stream: Functional differences between vowels and consonants. *Psychological Science*, 19:137–144.

van Ooijen, B. (1996). Vowel mutability and lexical selection in english: Evidence from a word reconstruction task. *Memory & Cognition*, 24:573–583.

Vihman, M. (1996). *Phonological Development: The Origins of Language in the Child.* Blackwell, Oxford.

Vihman, M., dePaolis, R., Nakai, S., and Hallé, P. (2004). The role of accentual pattern in early lexical representation. *Journal of Memory and Language*, 50:336–353.

Vihman, M. and Greenlee, M. (1987). Individual differences in phonological development: Ages one and three years. *Journal of Speech and Hearing Research*, 30:503–521.

Vihman, M., Macken, M., Miller, R., Simmons, H., and Miller, J. (1985). From babbling to speech: A re-assessment of the continuity issue. *Language*, 51:397–445.

Walley, A. (1993). The role of vocabulary development in children's spoken word recognition and segmentation ability. *Developmental Review*, 13:286–350.

Walley, A. and Metsala, J. (1990). The growth of lexical constraints on spoken word recognition. *Perception and Psychophysics*, 47:267–280.

Wassink, A. (2006). A geometric representation of spectral and temporal vowel features: Quantification of vowel overlap in three linguistic varieties. *Journal of the Acoustical Society of America*, 119:2334–2350.

Werker, J., Cohen, L., Lloyd, V., Casasola, M., and Stager, C. (1998). Acquisition of word-object associations by 14-month-old infants. *Developmental Psychology*, 34:1289–1309.

Werker, J., Fennell, C., Corcoran, K., and Stager, C. (2002). Infants' ability to learn phonetically similar words: Effects of age and vocabulary size. *Infancy*, 3:1–30.

Werker, J. and Tees, R. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7:49–63.

Whalen, D., Giulivi, S., Nam, H., Levitt, A., Hallé, P., and Goldstein, L. (in press). Biomechanically preferred consonant-vowel combinations fail to appear in adult spoken corpora. *Language and Speech*.