Language Learning and Development, 00: 1–26, 2013

ISSN: 1547-5441 print / 1547-3341 online
DOI: 10.1080/15475441.2013.829740

Routledge
Taylor & Francis Group

The Relationship Between Infants' Production Experience and Their Processing of Speech

Marinella Majorano
Department of Philosophy, Education and Psychology, University of Verona

Marilyn M. Vihman
Department of Language and Linguistic Science, University of York

Rory A. DePaolis
Department of Communication Science and Disorders, James Madison University

The early relationship between children's emerging articulatory abilities and their capacity to process speech input was investigated, following recent studies with English-learning infants. Twenty-six monolingual Italian-learning infants were tested at 6 months (no consistent and stable use of consonants, or vocal motor schemes [VMS]) and at the age at which they displayed use of at least one VMS. Perceptual testing was based on lists of nonwords containing one of three categories of sounds each: produced by infant (own VMS), not yet produced but typical of that age (other VMS), or not typically produced by infants at that age (non-VMS). In addition, size of expressive lexicon at 12 months and 18 months was assessed using an Italian version of the MacArthur-Bates Communicative Development Inventory (CDI). The results confirmed a relation between infant preverbal production and attentional response to VMS and also between age at first VMS and 12-month vocabulary. Maternal input is shown not to be a specific determinant of individual infant production preferences. A comparison between the English and Italian experimental findings shows a stronger attentional response to VMS in isolated words as compared to sentences. These results confirm the existence of an interaction between perception and production that helps to shape the way that language develops.

INTRODUCTION

Evidence as to the course of early lexical and phonological development has steadily accumulated since the appearance of the first studies of babbling (e.g., Oller, Wieman, Doyle, & Ross, 1976), the first-word production study to go beyond diary data (Ferguson & Farwell, 1975), and the first experimental study of infant speech perception (Eimas, Sequeland, Jusczyk, & Vigorito, 1971). Recently, the link between children's emerging articulatory abilities and their capacity to process speech input has begun to be seen as a crucial issue for a deeper and more satisfactory

Correspondence should be addressed to Marinella Majorano, Department of Philosophy, Education and Psychology, Via San Francesco, 22, 37129 Verona, Italy. E-mail: marinella.majorano@univr.it

understanding of phonological development (e.g., Kuhl et al., 2008). Although a few studies have reported an early connection between an infant's ability to organize perceptual categories and to produce language-specific patterns (e.g., Kuhl & Meltzoff, 1996), the mechanism for the development of this relationship remains unclear. As part of the effort to identify such a mechanism this study replicates and expands, with Italian-learning children, previous work carried out with English-learning children that showed an effect of production patterns on the processing of novel word forms.

Two contrasting approaches have independently investigated the issue of the link between speech perception and production. A first line of research suggests that infants must construct the link, and thus acquire phonological units from the linguistic environment, on the basis of perceptual experience alone (e.g., Jusczyk, 1992, 1993, 1997; Kuhl, 2004; Werker & Curtin, 2005; Werker & Yeung, 2005). Many studies have investigated early perceptual development, particularly the decline in discrimination of non-native contrasts over the first year (e.g., Werker & Tees, 1984; Best, 1994) or what has come to be seen more broadly as "perceptual narrowing" (Lewkowicz, 2011). These studies have shown that speech perception is affected by nonlinguistic domain-general factors such as attention (Werker & Lalonde, 1988; Werker & Pegg, 1992), distributional learning (Maye, Werker, & Gerken, 2002; Anderson, Morgan, & White, 2003), referential mapping (Yeung & Werker, 2009), and the articulatory gestures that shape linguistic units (Best & McRoberts, 2003).

Within this approach, studies using different methodologies (e.g., event-related potentials, the conditioned head-turn procedure), different perception indicators, and different theoretical models have also indicated a relationship between early infant perception and later language acquisition. For example, according to the Native Language Magnet model-expanded (NLM-e, Kuhl et al., 2008), a combination of stronger perceptual responses to native phonetic categories and a weaker perceptual response to nonnative phonetic units predicts faster lexical advance. Similarly but based on different indices, rapid auditory processing in the first year of life (at a mean age of 7.5 months) is claimed to be a unique precursor of later language development (comprehension and production at 24 and 36 months) and an early marker of language delay in populations at risk for language disorders (e.g., Benasich & Tallal, 2002; Benasich, Thomas, Choudhury, & Leppänen, 2002). In short, according to the first line of research, perception is understood as playing the primary role in language development.

A second line of research has emphasized the role of production. The motor theory of speech perception was the first to postulate a role for the motor system in speech perception (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Liberman & Mattingly, 1985; Liberman & Whalen, 2000). This theory assumes that both perception and production are based on a repertoire of motor primitives responsible for the construction of articulatory gestures, which are not distinguished from phonological representations. Since this theory was first proposed it has been the subject of a great deal of criticism. Specifically, Galantucci, Fowler, and Turvey (2006) reviewed several decades of studies based on the motor theory and found support only for a weak version. However, the current view is that gestural analysis may influence speech processing, even if it does not constitute the sole pathway (see, for example, the direct realist approach described in Best, 1995).

At the same time, the discovery of the mirror-neuron system has lent support to the idea of a motor origin for language (Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). This observation/execution matching system is thought to be implicated in many important human cognitive processes, such as imitation, a fundamental process for social and linguistic development. As applied to the study of language, this proposal assumes that perceiving speech implies the activation of motor-articulatory representations in the brain that provide a mapping for perception (Rizzolatti & Craighero, 2004), similarly to what was originally hypothesized by the motor theory. This assumption is supported by empirical studies of adults using transcranial magnetic stimulation (Fadiga, Craighero, Buccino, & Rizzolatti, 2002) and functional magnetic resonance imaging (Wilson, Saygin, Sereno, & Iacoboni, 2004). Although these authors do not specifically discuss the role of the mirror system in the emergence of language (but see the theoretical account in Vihman, 2002), recent studies using brain imaging with the whole-head MEG system have found a perceptuomotor link as early as 3 months (Dehaene-Lambertz et al., 2006). Furthermore, Imada et al. (2006) document increasingly strong links at 6 and 12 months, when most infants are just beginning to produce adult-like vocalizations; this study uncovers early developmental synchronization of motor and auditory brain areas at the end of the first year of life, which requires experience with preverbal production.

A good deal of empirical evidence on the effect of infants' early production on later language development can also be found in observational studies describing infants' preverbal production. Specifically, some authors have focused on babbling and early words and their relation to later language acquisition, based on typically developing children or late-talkers (Paul & Jennings, 2002; Stoel-Gammon, 1985, 1989; Vihman, Keren-Portnoy, Whitaker, Bidgood, & McGillion, 2013). The first longitudinal studies indicated that children who are more advanced in their preverbal production (i.e., who produced more babbling vocalizations at an early age or who displayed more stable sound production and greater syllabic complexity) show more rapid lexical advance (e.g., Keren-Portnoy, Majorano, & Vihman, 2009; McCune & Vihman, 2001; Stoel-Gammon, 1989) and better phonological memory some months later (Keren-Portnoy, Vihman, DePaolis, Whitaker, & Williams, 2010).

According to these studies, babbling constitutes the key to the construction of the lexicon for at least two reasons. On the one hand, it represents an opportunity, on the basis of universal, biologically grounded patterns (Davis & MacNeilage, 2000; Lindblom, 2000; Locke, 2001), to construct sensorimotor representations with the help of the "auditory feedback loop" (Fry, 1966, p. 189): This enables the child to master articulatory routines matched by auditory representations through implicit or procedural learning (Squire & Zola, 1996). On the other hand, babble increases the attention of adults interacting with the child, since they value it more highly than other infant vocal behaviors (McCune, 1992). Thus babble is usually associated with an immediate caregiver response (typically a repetition or a recast) in interactive "proto-conversations" (Bloom, 1975; Bloom, Russell, & Wassenberg, 1987; Jones, 2007; Gros-Louis, West, Goldstein, & King, 2006; Veneziano, 1981). At the same time, children broadly modify their production in accordance with their caregiver's phonological patterns (Goldstein & Schwade, 2008), and this too is predictive of language advance (see Chapman, 2000, for a review).

Among the studies relating perception and production that have focused on the role of early infant production in later phonological development some have emphasized articulatory schemes. Vihman (1991, 1993, 1996) postulated the existence of an "articulatory filter" constructed on the basis of well-practiced vocal routines or "vocal motor schemes" (defined as frequent and stably used consonants in infants' emerging production; for an operational definition see Methods, below). These motoric routines, which emerge during the preverbal period, can be distinguished from the phones of canonical babbling in two ways. First, they are defined in terms of not only "frequency" but also "stability" across longitudinal observations. The co-occurrence of both frequency and stability suffices to establish that the child has constructed relatively robust sensorimotor patterns that should consequently be available for use in phonological memory. Second, these schemas, which resemble the constructs that Piaget (1936) identified in the sensorimotor stage of cognitive development, are the product of an (implicitly) perceived match between output and input (see also Thelen, Corbetta, & Spencer, 1996, who draw on similar constructs to account for the development of reaching).

Vocal motor schemes (VMS) can be considered to be early integrated representations available for use as building blocks for language acquisition. They constitute "schemes" since their production involves both an internal effect (internal loop) and an effect on the environment. According to the articulatory filter hypothesis these schemes serve to "spotlight" similar patterns in input speech, establishing a robust association between perception and production. Children who are more advanced in preverbal production (i.e., children who have developed more than one VMS) can be expected to show more rapid lexical development because their articulatory representations permit them not only to exercise their articulatory schemes more effectively in production but also to pick out and retain more word forms from input speech and consequently to learn words more quickly. This predicted effect was first demonstrated by McCune and Vihman (2001), who found a positive relationship between early articulatory skills (age of acquisition of VMS) and the acquisition of referential words.

This idea of an early link between perception and production, as conceptualized in the articulatory filter hypothesis, differs from the assumptions of the NLM-e and, more generally, the first line of research discussed above, since those theoretical frameworks ascribe the primary role in word learning to perception. The articulatory filter view is theoretically related to the broader Dynamic Systems developmental model (Smith & Thelen, 2003; Thelen & Smith, 1994), which emphasizes the developmental importance of action-perception links, without assuming a causal direction. In particular, the articulatory filter hypothesis identifies the period immediately preceding first word production as crucial to the formation of a specific learned mapping between perception and production. Thus although this approach, like the perception-based theories discussed above, argues for an early link between perception and production, it emphasizes the role of production in the construction of such a link, an aspect disregarded by the other models.

Recently, Vihman and her colleagues have established a contribution of language production experience to phonological working memory and early speech perception in three experimental studies. In the first, Keren-Portnoy et al. (2010) identified a relation between experience in consonant production (VMS) and phonological memory in 12 English-speaking 2-year-olds. In the second, DePaolis, Vihman, and Keren-Portnoy (2011) reported a role for infant's vocal experience in the segmentation of nonwords that do or do not contain a consonant strongly represented in the infant's babbling (as measured using VMS). Specifically, a role for articulatory practice in new word learning (i.e., in attention to novel word patterns) was suggested, since children producing at least two VMS displayed longer attention in response to nonwords containing consonants that they were *not* yet producing with any consistency than to nonwords featuring consonants that they *were* producing. Children producing only one VMS showed no such preference. In the third study, DePaolis, Vihman, and Nakai (2013) found that infants' preference for nonwords featuring alveolar stops /t,d/ over nonwords with the alveolar fricative /s/ was negatively correlated with the number of alveolar stops that the infant produced. Since the contrasted consonants (/t,d/ vs. /s/) occur with equal frequency in input speech, and since none of the infants are producing many alveolar fricatives, the authors interpreted this finding as reflecting a shift from a familiarity to a novelty preference based upon the individual infant's production.

Like the studies of Vihman and her colleagues, the present study is designed to assess the effect of preverbal production on the perceptual processing of nonwords at an age corresponding to the onset of VMS production; it also adopts the criteria for identifying VMS provided in DePaolis et al (2011). There are several differences between this study and that of DePaolis et al. (2011), however. First, the present research uses a longitudinal design, assessing speech perception first in the absence of either VMS or lexical production (at 6 months of age) and then immediately after infants display use of at least one VMS. This makes it possible to establish whether the link between perception and production (i.e., the difference in looking time in response to stimuli with or without the child's most practiced consonant) emerges toward the end of the first year, as reported in other studies, or whether preferential attention to stimuli containing a specific consonant predates babbling practice.

A second, related difference between the present study and that of DePaolis et al. (2011) is that lists of isolated words are used instead of words embedded in passages. Isolated words can be assumed to constitute less challenging stimuli, since familiar words presented as word lists are recognized at a younger age than are words embedded in passages (DePaolis, Vihman, & Keren-Portnoy, 2013); these stimuli should therefore be better suited to the younger infants.

A third important difference between this study and that of DePaolis et al. is that the participants were Italian, not English infants. Cross-linguistic comparison makes it possible to test the sensorimotor model on the basis of a language with different phonological characteristics. Italian children produce preverbal utterances with higher percentages of labials than do children exposed to English (e.g., Keren-Portnoy et al., 2009; Majorano & D'Odorico, 2011). In addition, based on the Italian version of the CDI (Caselli & Casadio, 1995), 43.6% of the words reported as used by 75% or more of the children (aged 8 to 17 months) have bilabials in initial position and 38.5% have stops (Zmarich et al., 2011). The first words produced are typically bi- or trisyllabic, consisting mostly of CV and, medially, CCV syllable types (e.g. CVCV, CVCCV, CVCVCV). Although no detailed phonetic description of Italian infant-directed speech has been carried out to date, these characteristics can be supposed to apply to maternal input to 1- and 2-year-old children. Note that alveolars are the most frequent consonants in adult Italian (Batinti, 1993; Stella & Job, 2001).

Two additional points are tested in this study. First, as also observed by DePaolis et al. (2011), specific maternal input (i.e., the most frequent sounds produced) could affect babbling preferences and thus potentially also listening preferences. Accordingly, this study examines the closeness of fit between individual mother and child phonetic profiles, thereby also providing new information regarding the nature of infant-directed speech in Italian. Second, in this study vocabulary size was assessed at 12 and 18 months, based on parental report, in order to evaluate the predictive relation between preverbal production ability and lexical development.

In summary, this study is intended to test the two lines of research outlined above. If perception is a key factor in shaping production late in the first year of life, it should be possible to identify individual perceptual preferences for one of the nonword lists in the first head turn preference procedure (HPP) experiment, with a possible later effect of that preference on VMS production. In contrast, if production shapes perception, a difference in infant response to nonword forms should be seen only in the second HPP, when the first frequent and stably used consonants have been established in production (VMS), and the timing of VMS establishment should also predict advances in subsequent word learning. Furthermore, the production-based response seen in English, the language to which the infant participants in DePaolis et al. (2011) were exposed, should also be evident in Italian, with its differing phonological characteristics. Finally, the hypothesized effect of production on speech processing should be a stronger predictor of infant perceptual preferences than individual maternal input speech characteristics if it is to be taken as a general mechanism for language development.

In short, the study tests the following hypotheses:

1. A preference for specific consonants emerges in relation to preferred babbling patterns (VMS) only after the latter have developed; that is, there is no contributory preference for those consonants in the prebabbling period.
2. Production practice affects speech processing. Specifically, we expect to see, as reported in DePaolis et al. (2011), a difference in looking time in response to nonwords that feature the child's VMS vs. VMS used by other infants.
3. The child's preference for producing certain sounds and for attending to certain perceptual stimuli is independent of specific maternal input characteristics. Specifically, we expect to find that children's individual differences in production are not systematically related to the sounds most often produced by the mother.
4. The association between early perception and production will be seen in Italian as well as in English infants, despite the differences in the phonological characteristics of the input. We assess this hypothesis by comparing the findings for Italian and English infants.
5. Preverbal production will predict vocabulary size at 12 and 18 months. Specifically, children with earlier achievement of VMS are expected to show earlier lexical advance, because vocal practice should enable them to retain a larger number of word forms at a younger age.

METHOD

Participants

Thirty Italian monolingual infants (14 males) recruited from the local maternity ward participated in the study. All of the infants were full-term and passed a screening test for hearing. The children's families were middle class, with maternal education ranging from high school to university; parents were informed in advance about the research procedure and general aims. Of the initial sample three children did not complete the study due to illness, and for one child no VMS were identified. The final sample included 26 children (12 males).

Overall Experimental Design

The study included three different data collection procedures: 1) two experimental sessions (Head-Turn Paradigm procedures, or HPP) to assess infants' speech processing at different developmental points, 2) several longitudinal observational sessions to assess infants' vocal production, and 3) maternal reports on vocabulary size at 12 and 18 months of age. The first experimental session was conducted when infants were not yet producing any consonants consistently and the second when at least one VMS had been produced to the criteria reported below. The number of observational sessions varied across infants in relation to their production; all were conducted at home. The data collection sequence was as follows: 1) first HPP in the absence of consistent consonant production; 2) longitudinal observational sessions; 3) second HPP when at least one VMS had been identified; 4) vocabulary assessment based on maternal report at 12 and 18 months of age. To simplify the presentation of the methods, we begin with the observational sessions, followed by the stimuli and the HPP. We then describe the methods used for coding infant-directed speech (IDS) and conclude the Methods with a description of the vocabulary assessment procedure.

Observational Sessions: Procedure and Analysis

Each family was contacted for the first time when the infant was around 3 months of age in order to explain the research and the criteria for the observational and experimental sessions. After two months, each mother was contacted again to plan the first head-turn experiment, to take place before the child had begun to produce identifiable babbling (at around 6 months, range 0;5.23 -0;6.9).

Following the procedure outlined by DePaolis et al. (2011), for production data collection mothers were instructed to contact the experimenter when the child began to produce identifiable consonant sounds. To help them to do this, before the six-month session mothers attended a presentation at the University of Parma that illustrated children's consonant production. In addition, the mothers were contacted one month after the first HPP (at around seven months) and again every two weeks thereafter to monitor VMS production and to plan the first home visit, if the mother had not yet contacted the experimenter to report the onset of babbling in speech-like syllables. Infants were then video-recorded biweekly in their home for 30 minutes of free play, using a standard set of toys, until VMS were identified. The audio and video material was recorded using a Sony digital Handycam DCR-PC 105; a Sony ICD-P17 microphone was attached to a cloth vest worn by the infants. The observations were supplemented by a maternal questionnaire about prenatal and perinatal indices and first sensory-motor behavior and about the characteristics of the child's environment (e.g., time spent at the nursery, time spent with mother/father or grandparents). The infants' ages at the first home visit ranged between 0;7.08 and 0;10.28.

Two trained transcribers, working independently of one another and using broad phonetic transcription, transcribed each child's vocal or verbal production, noting duration. In the case of particularly difficult vocalizations a third transcriber was included and, if there was no consensus after four joint listenings, the utterance was excluded. The guidelines for the transcription process were similar to those used by other authors (e.g., McCune & Vihman, 2001; Olswang, Stoel-Gammon, Coggins, & Carpenter, 1987).

Infant productions of each consonant were computed in each session, and the total number of occurrences was calculated for each infant. Two criteria were used to identify VMS, following McCune and Vihman (2001) and DePaolis et al. (2011). According to the first criterion, a consonant had to be produced at least 10 times in each of three out of four successive sessions. The second criterion was 50 occurrences of a consonant within one to three sessions; this criterion made it possible to bring a child in for testing more rapidly, once high use of a particular VMS had been identified. Voicing was not considered distinctive, since it is generally not well-controlled in this developmental period (Macken, 1980). Infants were classified into two categories according to VMS production: (i) single-VMS producers and (ii) multiple-VMS producers (two or more VMS produced).

Following McCune and Vihman (2001), broad phonetic-transcription reliability was evaluated for point-to-point agreement between the two transcribers, based on approximately 10 minutes' recording from each participant for VMS identification (/p,b/, /t,d/, /k,g/, /m/, /n/). Utterances that could not be identified after four listenings and with the aid of a third transcriber were excluded. Inter-judge agreement was .85, consistent with other studies of consonants produced in the prelinguistic period (e.g., Davis & MacNeilage, 1995; DePaolis et al., 2011).

Stimuli for the Head-turn Preference Procedure

A female native Italian speaker recorded four lists of 12 nonwords. All stimuli were recorded in a sound treated room using a RODE NT2A condenser microphone connected to a Metrichalo soundcard sampling at 48 KHz-24 bit. The stimuli had a speech peak within $\pm$ 1.5 dB and an ISI of 750 ms.

Three contrasting lists of nonwords were presented: one featuring the infant's identified VMS (own-VMS), one featuring VMS typical for that age but not yet acquired by the infant tested (other-VMS; both own- and other-VMS are stops grouped regardless of voicing: /p,b/, /t,d/, and /k,g/), and a third featuring consonants not often produced in babbling (/f,v/). Each of the consonants was combined with vowels according to the most frequent C-V associations in infant vocalizations, following Davis and MacNeilage (1995); nasals were excluded, following DePaolis et al. (2011). The phonotactic patterns of the stimulus words conformed to adult Italian. Although /ʌ/ is not present in Italian adult speech, this sound is reported for babbling and the first words of Italian children (Zmarich & Miotti, 2003); accordingly, it was included here in the lists of nonwords, to vary the phonetic shapes of the stimuli (Table 1). The word-final vowels conformed to adult Italian.

TABLE 1
Nonwords Used in the Head-Turn Experiment

p/b	*t/d*	*k/g*	*f/v*
pabʌ	tidɛ	kɔgu	favʌ
pʌba	dɛti	kugu	fʌva
bapʌ	ditɛ	gɔko	vafʌ
bʌpa	tɛdi	guku	vʌfa
papʌ	titɛ	kɔku	fafʌ
pʌpa	tɛti	kuku	fʌfa
babʌ	didɛ	gɔgo	vavʌ
bʌba	dɛdi	gugu	vʌva
papa	titi	kɔkɔ	fava
pʌpʌ	tɛtɛ	gogo	fʌvʌ
baba	didi	gɔku	vafa
bʌbʌ	dɛdɛ	gukɔ	vʌfʌ

In contrast with DePaolis et al. (2011) but following DePaolis et al. (2013b), lists of isolated nonwords were presented in this study instead of nonwords embedded in passages.

Head-turn Preference Procedure

For perception assessment each child was tested twice using the HPP, first at 6 months and then again at the age at which the criteria for VMS production were observed (each child participated in the head-turn experiment within a week after the relevant observation session). At the start of the first experimental session each mother was asked about the infant's vocal production, using a specially designed questionnaire; a short observation period was also included to establish whether the infant was babbling in syllables.

The experimental procedure was set up in accordance with Kemler-Nelson et al. (1995) and DePaolis et al. (2011). The mother was seated with the infant on her lap in a three-sided test booth. A camera and a red light were positioned at the center, and a blue light and a speaker were mounted on the right and left side panels. The observer controlled stimulus presentation using a video monitor and PC located in a separate room. Each experiment was composed of a familiarization and a test phase. During the familiarization phase the non-word lists were all presented (all four lists at 6 months and own-VMS, other-VMS and non-VMS in the second HPP), counterbalanced for order. This phase was designed to acquaint the infants with the speech stimuli. During the test phase the three stop-pair lists (/p,b/, /t,d/, /k,g/) and the one fricative list (/f,v/) were presented in the 6-month session while lists with own-VMS, other-VMS, and non-VMS (/f,v/) were presented in the second session.

Each trial started only when the experimenter judged that the infant's gaze was directed toward the flashing red light. At that point, one of the blue lights to one side of the infant would begin to flash and, once the infant looked toward the flashing light, one of the randomly selected lists of 12 nonwords would begin to play after a half-second delay. The infant's looking time was recorded by the experimenter pressing a button on the computer when the infant changed their gaze direction, turning their head at least 30° toward the side light. The side of presentation was pseudo-randomized by the program. Each trial ended when the infant failed to turn his/her head after the initial stimuli or turned away for more than two seconds. In the first session, at 6 months of age, the four lists were presented four times (16 trials); in the second session, the three lists were presented four times (12 trials). Participants were tested in the second head-turn procedure between the ages of 0;7.28 and 0;11.08. Only the own- and other-VMS pairs identified before the second HPP were included in the analysis of the first experimental session.

To assess reliability, a second experimenter coded eight infants' looking time off-line for each experimental session; interjudge agreement was .88 (Cohen's index).

Coding of Infant-Directed Speech

The IDS of a sample of six mothers was analyzed. As in DePaolis et al. (2011), mothers of infants who produced different VMS were selected for comparative analysis; both the pre-VMS session and the session in which VMS was identified were then transcribed. Following Vihman, Kay, Boysson-Bardies, Durand, and Sundberg (1994), consonants were tallied in (i) all child-directed running speech (content and function words), (ii) content words only, and (iii) onset position in content words. Geminate consonants in content words were counted as well, because this category, together with the initial consonant, is reported to have special salience for children (Savinainen-Makkonen, 2007; Vihman & Velleman, 2000).

To assess reliability, 20 minutes of maternal speech from each session were transcribed by a second transcriber; agreement of .89 (Cohen's index) was reached in the sounds corresponding to VMS identification in the maternal speech samples.

Vocabulary Assessment

The size of the children's comprehension and production vocabulary was assessed at 12 months; production only was assessed at 18 months, based on the Italian version of the MacArthur-Bates Communicative Development Inventory (CDI; Caselli & Casadio, 1995).

RESULTS

Infants' Production

Table 2 reports the occurrences of each sound produced, total occurrences and the number of VMS. Twelve children (mean age at first observational session 0;8.21) produced a single VMS, while 14 children (mean age at first observational session 0;9.01) produced multiple VMS; 8/12 single-VMS children produced p/b to criterion, 3/12 produced t/d, and one infant produced k/g. Most children in the multiple-VMS group produced two VMS (9/14), generally the pairs p/b and t/d, but 5/14 produced three. The single-VMS infants produced a mean of 85.33 consonants altogether ($SD = 12.84$), while the multiple-VMS infants produced a mean of 157 ($SD = 35.28$).

Infant Speech Perception-Production Relationship

Table 3 reports the VMS identified, the consonant pairs tested for each child in the second HPP and the age at first observational visit and at the second test.

To test possible differences in infants' preferences for consonantal patterns in relation to their production, a *k-means* cluster analysis was conducted on the total phoneme occurrences, using R free software programming language for statistical computing (version 2.15.0). This nonhierarchical clustering procedure makes it possible to group participants into subgroups with similar characteristics according to a specific variable (MacQueen, 1967). Using this analysis, a *k* number of clusters (specified a priori) are constructed in order to test the structure of the data set. Then centroids for a set of trial clusters are calculated, each participant is placed in the cluster with the nearest centroid, and the centroids are recalculated and the participants reallocated. The cluster analysis made it possible to identify infants based on production. The infants were then separated into groups (i.e., into the optimal number of clusters) based upon the Euclidean distance between each infant's production total and the three mean production values that minimized the respective distances.

TABLE 2
Each Child's VMS Production (Ordered by Total Number of Occurrences During the Sessions)

				Occurrences of					
	Age at first session	*N sessions*	*VMS*	*p/b*	*t/d*	*k/g*	*m*	*n*	*Total occurrences*
M.B.	8.08	2	p,b	52	9	1		2	64
A.A.	8.15	1	p,b	51	6		15		72
C.D	8.11	1	p,b	58	8	3	6		75
F.V.	10.28	1	p,b	56	4		3	12	75
A.G.	8.02	4	p,b	53	12		3	8	76
L.B.	8.15	2	p,b	56	10		16		82
C.A.	7.08	2	t,d	9	57	9	13		88
M.G.	8.10	2	p,b	58	10		22	5	95
V.S.	10.16	2	t,d	10	64		12	10	96
D.C.	9.04	2	k,g	11	5	65	10	7	98
M.B.1	8.03	4	t,d	4	66		20	11	101
M.B2	9.02	3	p,b	62	8		20	12	102
M.C.	8.25	2	p,b; t,d	50	56	4	6	2	118
A.R.	9.18	2	p,b; t,d	54	51		12	3	120
G.F.	10.13	1	p,b; m	60	4		53	6	123
H.B.	9.10	4	p,b; m	56	8		54	11	129
A.B.	9.03	4	p,b; t,d	68	57	2	9		136
F.V.	7.28	2	p,b; t,d	60	62		11	6	139
G.R.	9.02	2	p,b; k,g	61	2	51	22	8	144
M.T.	9.28	3	t,d; k,g	8	52	54	26	5	145
P.A.	8.27	2	p,b; t,d	57	56	10	21	5	149
F.M.	9.29	2	p,b; t,d; m	55	64		52	5	176
R.D.	8.18	3	p,b; t,d; m	69	58	6	52	10	195
L.G.	8.09	3	p,b; t,d; m	68	61	8	52	8	197
A.P.	9.27	3	p,b; k,g; m	68	8	60	61	2	199
A.L.	7.15	3	p,b; t,d; m	72	70	6	57	23	228

As can be seen in Figure 1, the analysis yielded three clusters: group 1: low production, composed of 12 infants (M = 85.33; SD = 12.39, group 2: middle production, 9 infants (M = 133.67; SD = 11.58) and group 3: high production, 5 infants (M = 199.00; SD = 18.64) ($p < .0001$). Comparing the composition of group 1 (Figure 1) and the number of VMS produced by each infant (Tables 2 and 3), it can be noted that the low production group corresponds exactly to the group of children identified as producing a single VMS. Following DePaolis et al. (2011), the infants' preference ratios (P-ratios) were computed for looking time (LT) in response to 'own' consonants over 'own+other' consonants (P-ratio = LT(own)/LT(own+other). Individual infant preference ratios for Own-VMS are reported for each of the three groups in each experiment.

Figures 2a and 2b indicate that in the second HPP only infants in group 1 looked longer in response to own-VMS lists than to other-VMS lists. On the other hand, children who produced more consonants (groups 2 and 3) looked longer in response to other- than to own-VMS list. Both figures show that at 6 months individual P-ratios are less differentiated by group, while in the second HPP most of group 1 is concentrated above the middle line (.50% level) whereas the

TABLE 3
Age at First Visit, Number of Sessions Needed to Identify VMS for Each Infant, VMS Identified, Age at Second HPP and Own- and Other-VMS Used in HPP

	Age at first visit	*N sessions*	*VMS identified*	*Age at 2nd HPP*	*Own-VMS*	*Other-VMS*
A.G.	8.02	4	p,b	9.20	p,b	t,d
C.A.	7.08	2	t,d	7.28	t,d	p,b
M.B.	8.08	2	p,b	8.26	p,b	t,d
M.B.1	8.03	4	t,d	9.20	t,d	k,g
L.B.	8.15	2	p,b	9.02	p,b	t,d
M.B2	9.02	3	p,b	9.29	p,b	t,d
D.C.	9.04	2	k,g	9.22	k,g	p,b
C.D	8.11	1	p,b	8.11	p,b	t,d
M.G.	8.10	2	p,b	8.27	p,b	k,g
A.A.	8.15	1	p,b	8.16	p,b	t,d
V.S.	10.16	2	t,d	11.06	t,d	p,b
F.V.	10.28	1	p,b	10.28	p,b	t,d
F.M.	9.29	2	p,b; t,d m	10.18	t,d	k,g
H.B.	9.10	4	p,b; m	10.27	p,b	t,d
A.B.	9.03	4	p,b; t,d	10.19	t,d	k,g
P.A.	8.27	2	p,b; t,d	9.15	p,b	k,g
R.D.	8.18	3	p,b; t,d; m	9.28	t,d	k,g
G.F.	10.13	1	p,b; m	10.18	p,b	k,g
L.G.	8.09	3	p,b; t,d; m	9.09	p,b	k,g
A.L.	7.15	3	p,b; t,d; m	8.23	t,d	k,g
F.V.	7.28	2	p,b; t,d	8.16	p,b	k,g
A.P.	9.27	3	p,b; k,g; m	10.26	k,g	t,d
G.R.	9.02	2	p,b; k,g	9.20	p,b	t,d
A.R.	9.18	2	p,b; t,d	10.11	t,d	k,g
M.T.	9.28	3	t,d; k,g	11.08	k,g	p,b
M.C.	8.25	2	p,b; t,d	9.12	t,d	k,g

other two groups fall in the lower part of the plot (Figure 2a) for own-VMS lists; the pattern is reversed for other-VMS lists (Figure 2b). As shown, the preference for own-VMS increases from the 6-month HPP to the second HPP only for group 1 and decreases for the other two groups. Looking time to the non-VMS lists was quite similar in the three groups (Figure 3).

A mixed 3 × 3 × 2 (VMS category [own-VMS, other-VMS, non-VMS] × infant production [low, middle and high production] × experimental session [6 months and age of VMS attainment]) Analysis of Covariance (ANCOVA) was performed on the test trials, using looking time as the dependent variable. Since the first observation of the multi-VMS group (and, consequently, the second HPP) was carried out later than the first observation of the single-VMS group (9.01 vs. 8.21 mos. and 10.02 vs. 9.12 mos.) and since this might reflect a later recognition of babbling onset on the parents' part, age at second HPP was added as covariate.

Analysis indicated no main effect for the single factors (VMS category: $F(2, 48) = 2.11$; $p = .13$, $\eta^2 = .08$; infant production: $F(1, 24) = .64$; $p = .54$, $\eta^2 = .05$; experimental session: $F(1, 24) = .07$; $p = .78$, $\eta^2 = .01$) but revealed a significant interaction between VMS categories, infant production, and experimental session ($F(2, 48) = 3.34$; $p = .05$, $\eta^2 = .18$). Posthoc analysis

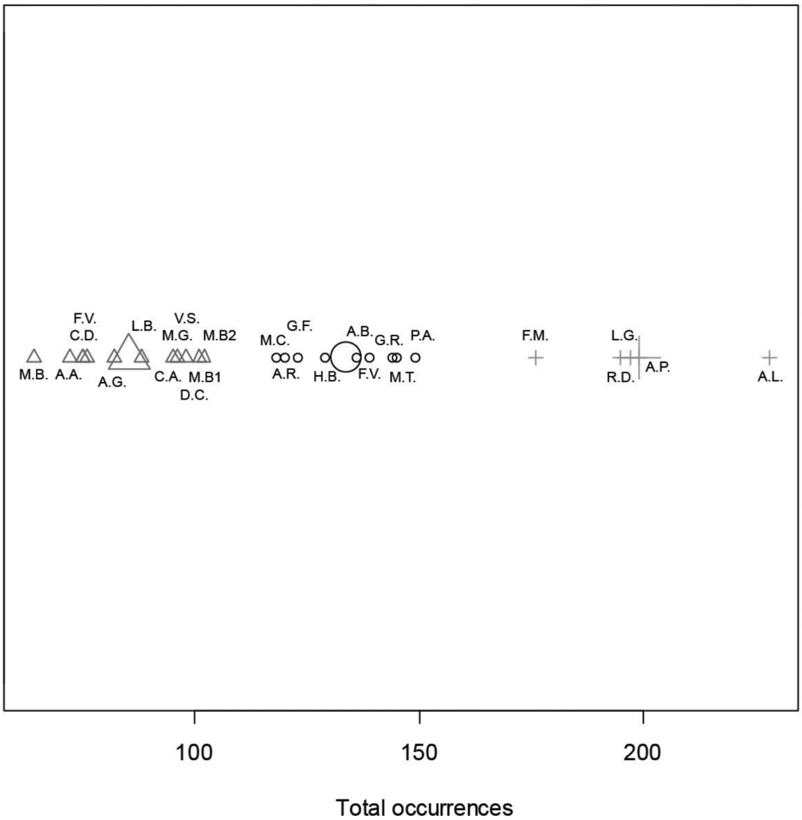

FIGURE 1 A plot of the three clusters identified using k-means analysis; 'total occurrences' is the sole variable analyzed.

indicated that single-VMS infants looked longer in response to Own-VMS than to Other-VMS: $t(12) = 3.20$; $p < .01$ in the second HPP. In the same HPP both groups of multiple-VMS infants (middle production and high production) looked longer to Other-VMS than to Own-VMS: $t(9) = 3.20$; $p < .01$ and $t(5) = 2.95$; $p < .05$. These effects are not seen in the first HPP, where neither the single- nor the multiple-VMS infants showed a preference for either list of nonwords ($t(12) = .05$; $p = .95$ for single-VMS infants, $t(9) = .05$; $p = .95$ and $t(5) = .58$; $p = .58$, for the other two groups). These results indicate a preference for nonwords based on the infants' production practice, as indexed by VMS, in the second HPP.

Since the /p,b/ list was preferred by most of the children, in order to test the alternative hypothesis that the preference was due to the specific consonant rather than to individual infant production, a similar mixed 3 × 2 × 2 (VMS categories [Own-VMS, Other-VMS, Non-VMS] × infant production [/p,b/ as VMS and /p,b/ as non-VMS] × session [6 months and second HPP]) was carried out. Analysis of variance (ANOVA) was performed on the test trials using looking time as the dependent variable. Analysis indicated no main effect for the single factors and no significant interactions [VMS category: $F(2, 48) = .52$; $p = .60$, $\eta^2 = .02$; infant production:

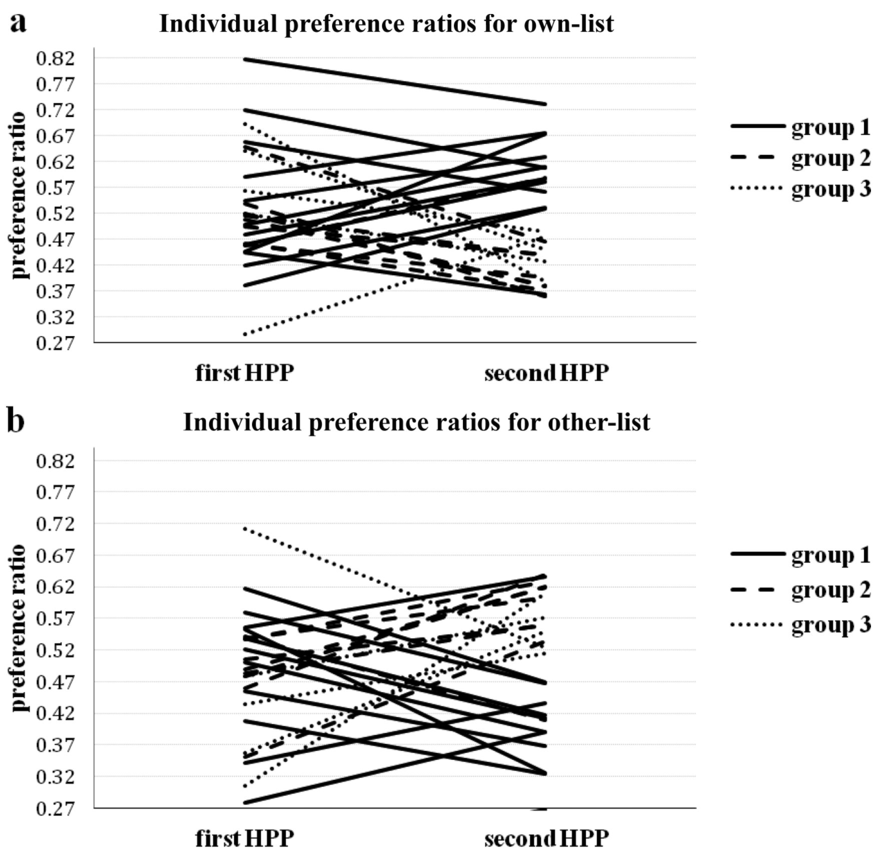

FIGURE 2 Individual plots for preference for stimuli containing the most often produced sounds, using preference ratios [looking time (own) / looking time (own + other)] (Figure 2a), and preference for stimuli containing less often produced sounds, using preference ratios [looking time (other) / looking time (own + other)] (Figure 2b).

$F(1, 24) = .02$; $p = .89$, $\eta^2 = .01$; experimental session: $F(1, 24) = .24$; $p = .62$, $\eta^2 = .01$; interaction between VMS categories, infant production, and experimental session ($F(2, 48) = 2.02$; $p = .13$, $\eta^2 = .08$)]. Thus, there is no relationship between the bilabial stop, the infants' favored production pattern, and their preference for bilabial stop consonants in the HPP.

In short, the lower production group (group 1, which corresponds to the single-VMS group) showed a difference in their stimulus preference in the second HPP compared with the other two groups; thus child experience with a single (group 1) vs. multiple VMS (groups 2 and 3)—a difference in experience which reflects different levels of articulatory competence—appears to be associated with different perceptual preferences toward the end of the first year but not at 6 months of age.

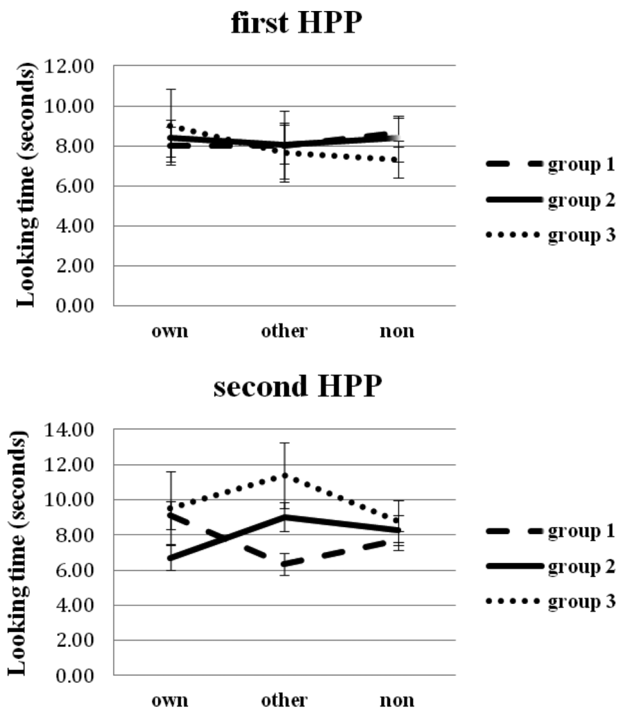

FIGURE 3 A plot of infant looking times to own-, other- and non-VMS in the two HPPs for the three groups. The bars represent ± 1 standard error.

Analysis of Infant Directed Speech

To determine whether individual infants' VMS were affected by their mother's speech the IDS of six mothers of infants (M.B.2; V.S; D.C.; A.B.; M.T.; G.R.) who produced different VMS (/p,b/; /t,d/; /k,g/; /p,b/ and /t,d/; /t,d/ and /k,g/; /p,b/ and /k,g/) were chosen for analysis. Both the 6-month session and the session in which VMS was identified were transcribed. The percentages of occurrences for stop VMS are reported in Table 4.

As shown in Table 4, the distribution of stops in maternal input speech in the pre-VMS and VMS sessions is quite similar. Figure 4 shows this relative uniformity (for onset consonants only, since the onset consonant, which generally falls in the accented syllable, is likely to be the most salient for children: Cf. Vihman, Nakai, DePaolis and Hallé, 2004). A series of Friedman's non-parametric tests for related samples was conducted to analyze differences in the proportions of the three pairs of stop consonants. Data analysis indicated no significant differences (χ^2 (2) ranging between .33 and 2.33) in any position analyzed. This means that although each of the infants first produced one or more of the stop pairs, there was no significant difference in their mothers' frequency of use of these stop pairs.

To more formally compare the distribution of stop-VMS for the different IDS measures with the children's production, however, a distance matrix was constructed, following the procedure

TABLE 4
Infant-directed Speech (IDS) of Six Infants' Mothers, Indicating Proportion of Occurrences for Each of the Stop-VMS Tested

			p/b				*t/d*				*k/g*			
Infant	*VMS*	*session*	*running speech*	*onset*	*conten words*	*geminates*	*running speech*	*onset*	*content words*	*geminates*	*running speech*	*onset*	*content words*	*Geminates*
M.B.2	p/b	pre-VMS	0.44	0.46	0.47	0.33	0.23	0.23	0.24	0.25	0.34	0.31	0.29	0.42
		VMS session	0.39	0.43	0.44	0.29	0.23	0.24	0.25	0.29	0.37	0.33	0.31	0.43
V.S.	t/d	pre-VMS	0.37	0.47	0.39	0.24	0.26	0.27	0.27	0.41	0.37	0.27	0.34	0.35
		VMS session	0.36	0.47	0.37	0.29	0.26	0.26	0.28	0.36	0.38	0.26	0.35	0.36
D.C.	k/g	pre-VMS	0.28	0.39	0.30	0.42	0.33	0.29	0.34	0.25	0.38	0.32	0.36	0.33
		VMS session	0.32	0.41	0.36	0.29	0.33	0.31	0.36	0.36	0.35	0.28	0.27	0.36
A.B.	p/b, t/d	pre-VMS	0.27	0.32	0.28	0.31	0.36	0.32	0.36	0.31	0.37	0.37	0.35	0.38
		VMS session	0.30	0.30	0.33	0.27	0.34	0.39	0.34	0.36	0.36	0.30	0.34	0.36
M.T.	t/d, k/g	pre-VMS	0.26	0.28	0.28	0.36	0.35	0.34	0.39	0.36	0.39	0.38	0.33	0.29
		VMS session	0.32	0.33	0.35	0.30	0.29	0.36	0.29	0.40	0.39	0.31	0.36	0.30
G.R.	p/b, k/g	pre-VMS	0.26	0.28	0.21	0.36	0.39	0.37	0.45	0.36	0.35	0.35	0.35	0.29
		VMS session	0.29	0.31	0.23	0.25	0.45	0.37	0.50	0.42	0.26	0.33	0.26	0.33

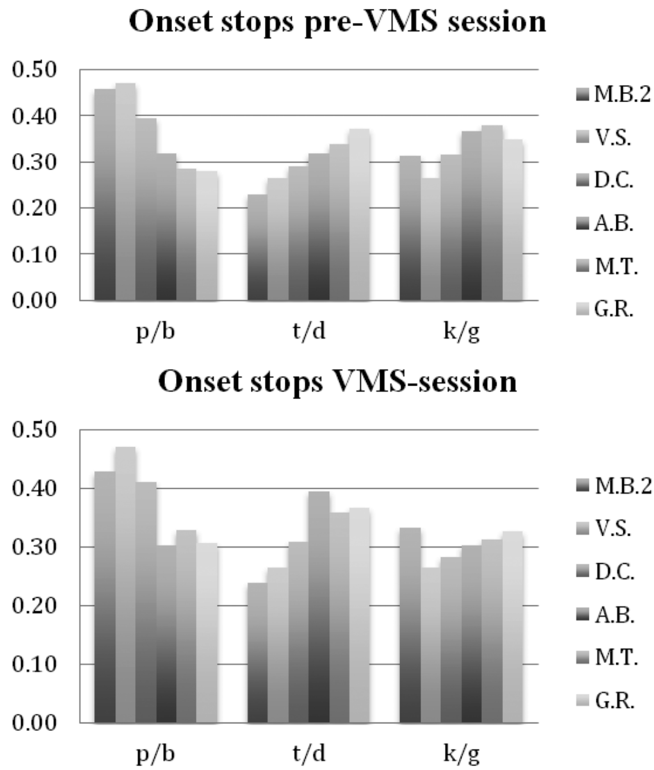

FIGURE 4 Stop-VMS occurrences (percentages) in IDS of six infants' mothers.

reported in Vihman et al. (1994). In this version of the test, the distance matrix tests the similarity (or distance), for pre-VMS and VMS sessions, between the proportion of stop consonants in mothers' running speech and content words, and in infant vocalizations. For running speech the mean ranks found for row (mother - child) and column (child - mother) were as follows: pre-VMS session 3.5, 3; VMS-session 3.8, 3.8; for content words: pre-VMS session 3.5, 3.7; VMS- session 3.7, 3.3. Non- parametric Mann Whitney's tests for independent samples were applied to these data to determine whether these means are different from chance (in this case the mean should be 3.5). None of the tests showed a significant difference (U ranging from 12 to 16; $p > 05$). This indicates that none of the within-dyad pairings (child A to mother A) had greater similarity than other-dyad pairings (e.g., child A to mother B).

Predicting Lexical Development

The Italian version of the MacArthur-Bates CDI (Caselli & Casadio, 1995) was administered to participants' mothers when the infants were 12 and 18 months of age. Table 5 shows group vocabulary sizes and normative data.

TABLE 5
Mean Lexicon Size of Participants in this Study and Normative CDI Data (Standard Deviation in Parentheses)

	normative group			*current research*	
	production	*comprehension*		*production*	*Comprehension*
12 months	8 (9)	109 (57)	12 months	7.07 (4.42)	116 (47.72)
18-19months	88 (88)	–	18 months	116.91 (84.54)	–

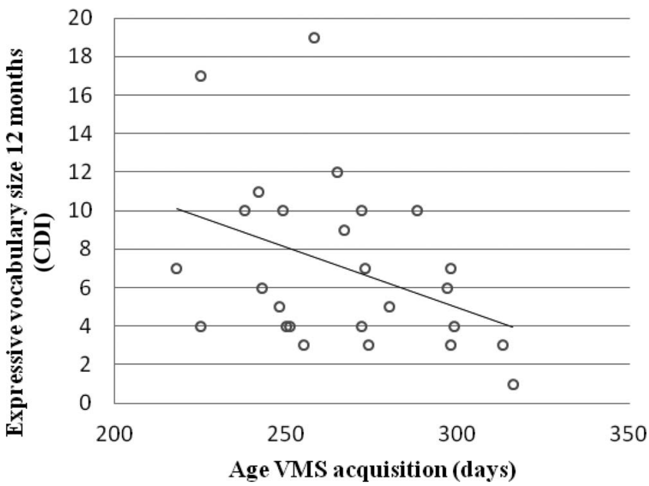

FIGURE 5 Vocabulary size in relation to age of VMS acquisition.

As shown in Table 5, the vocabulary sizes of the group fall within one SD of the normative data at both 12 and 18 months of age. Two linear regressions were performed to analyze the relation between early phonetic development (age of acquisition of VMS) and vocabulary size at 12 and 18 months of age. The results indicate an effect of the predictor on (expressive) lexical development at 12 months [$R^2 = .15$, $R^2_{adj} = .12$; $F(1, 24) = 4.29$, $p < .05$; ß $= -.39$ and $p < .05$] but not on receptive vocabulary [$F(1, 24) = 1.47$; $p > .05$] and not at 18 months [$F(1, 24) = .05$; $p > .05$]; see Figure 5.

The results show a relation between preverbal articulatory practice and the number of first words children produce. However, this association is limited to early words (12 months), and to expressive vocabulary, and is not detectable later (at 18 months), in contrast with the finding of effects on later vocabulary learning in other studies (e.g., McCune & Vihman, 2001; Stoel-Gammon, 1989). The difference in the findings here and in previous studies may be due to differences in the methods used to assess vocabulary size in the second year (direct observation plus CDI in the previous studies vs. CDI only in the present study). More careful child observation may be needed to test the effects of preverbal development on later child word acquisition.

Comparison of Italian and English Groups

To investigate a possible role for input characteristics in the association between perception and production, data from this study are compared with data of DePaolis et al. (2011), conducted with English infants. The English group comprised 18 infants. The procedure was similar to that used in this study: Infants were recorded regularly in their homes until at least one VMS was identified. Then the same infants were tested using HPP. The nonwords used as stimuli were similar to the stimuli of the present experiment (they contained /p,b/, /t,d/, /k,g/ and /v,f/), but were embedded in passages. In contrast with the Italian study, most of the infants in the English study produced /t,d/ as VMS. In order to compare the two studies the Italian participants were grouped by number of VMS identified (multiple- vs. single-VMS infants). The data from the English study and the second Italian HPP are plotted in Figure 6.

As indicated in Figure 6, a difference in looking times emerged between the two language groups. A mixed 3 × 2 × 2 (VMS categories [own-VMS, other-VMS, non-VMS] × infant production [single VMS and multiple VMS] × language group [English, Italian]) Analysis of Variance (ANOVA) was performed on the test trials, in order to analyze this difference between groups, using looking time as the dependent variable. Analysis indicated no main effect for the single factors but a significant interaction between VMS categories and infant production [$F(2, 80) = 10.30$; $p < .001$, $\eta^2 = .20$]. Posthoc analysis indicated that single-VMS infants looked longer to Own VMS: $t(20) = 3.49$; $p < .01$ and multiple-VMS infants looked longer to Other VMS: $t(22) = 4.79$; $p < .01$. In addition, an effect of the between-subjects variable "language group" emerged [$F(1, 40) = 15.48$; $p < .001$; $\eta^2 = .28$]. Post-hoc analysis indicated that the Italian group showed longer looking-times for all three lists of nonwords (t ranging between 2.81 and 3.29; $p < .05$). In short, the English and Italian groups display the same preference for auditory stimuli in relation to their articulatory practice, but the kind of stimuli presented (list of nonwords or passages) influenced looking times in the HPP.

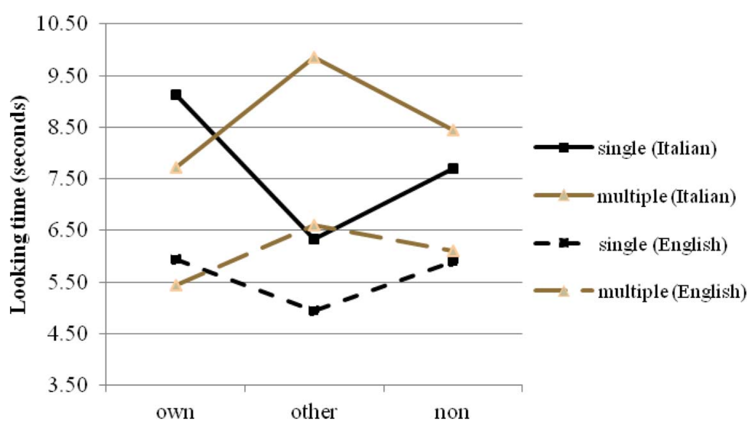

FIGURE 6 Looking times in the second HPP for the Italian infants compared with a previous study involving British English infants (DePaolis et al., 2011). (Color figure available online.)

DISCUSSION

The present study describes early production and perception in infants learning Italian. The production data indicate that the most frequent early VMS in Italian infants is /p,b/, in line with other studies of the phonetic characteristics of early child words in Italian (Keren-Portnoy, et al., 2009; Majorano & D'Odorico, 2011). This is in contrast with English and Welsh, where /t,d/ has been found to be the most frequently produced early consonant type (DePaolis et al., 2011; DePaolis et al., 2013b; McCune & Vihman, 2001). The difference between Italian and English early phonetic preferences can be explained by the complex interaction between individual experience and input-specific influence on early language development; it provides evidence against the existence of a universal bias in early consonant production.

Since Italian IDS has not yet been extensively documented, it is worth noting that Italian mothers used many words with initial bilabials (e.g., *bella* 'nice,' *brava* 'good,' *bimba* 'baby,' *bau* 'bowwow,' *buono* 'good,' *papa* 'daddy,' *pappa* 'food' (baby talk), *piccolo* 'little,' *prendere* 'to take'). The high frequency of words with labial stop onsets (as shown in Figure 4) could account for the frequent use of labial stops in the Italian infants' babbling and early words (also Zmarich, Dispaldro, Rinaldi, & Caselli, 2011, on the phonetic characteristics of Italian CDI, where early but not later acquired words have a high percentage of labials at word onset [44%]). Note as well that a higher percentage of velar sounds was found in IDS here as compared with adult Italian (Batinti, 1993), perhaps due to the observational context, in which mothers made high use of questions and deictics (*cosa* . . . ? 'what . . . ?', *questo* . . . 'this,' *quello* . . . 'that').

We now take up each of our hypotheses separately and then conclude with a general discussion that suggests that our results support the existence of an articulatory filter acting to mediate between production and perception in the prelinguistic infant.

As proposed in our first hypothesis, prior to the onset of babbling production (around 6 months of age) there is no evidence of systematic infant preferential attention to the particular nonword list preferred in the second HPP. Instead, as seen in the findings for the second HPP, there is an association between production practice and speech perception toward the end of the first year, which supports the second hypothesis. Infants producing fewer consonants to criterion (group 1, or the single-VMS group) displayed longer looking times in response to nonword lists that contained sounds they were consistently producing, while children with consistent and stable use of more consonants (groups 2 and 3, or the multiple-VMS infants) displayed the reverse pattern. This effect was not due to the acoustic characteristics of the stimuli, since both own- and other-VMS lists featured acoustically similar stop consonants, and the different groups of infants looked equally long in response to the non-VMS list, which featured the more acoustically salient fricatives.

Our findings also support the third hypothesis, consistent with previous studies that show that the distribution of consonants in the mothers' infant-directed speech is not the main determinant of infants' favored consonants in production (DePaolis et al., 2011, DePaolis et al., 2013b; Vihman et al., 1994). We see in our data no correspondence between the mothers' frequency of consonant use, including labials (the most common consonant produced early on by the infants in this study), and individual child production. Although the input patterns undoubtedly guide the infant's babble toward the adult model, the individual infant's favoured consonants in babble do not match the most frequently used consonants of their mothers.

Focusing on the cross-linguistic comparison, the results of the second HPP here closely parallel those of DePaolis et al. (2011) and thus support the fourth hypothesis. Both Italian and British English infants who had acquired a single VMS preferred the consonants that they were producing, while in both language groups those who showed evidence of two or more VMS preferred the production patterns that they were not yet producing. The difference between English and Italian infants' overall looking times can be explained by the difference in the mode of presentation of stimuli in the two experiments: While the Italian infants listened to isolated word lists, the English infants listened to words embedded in passages. As reported in DePaolis et al. (2013a), the ability to recognize familiar words without familiarization (or 'priming') in the laboratory emerges earlier for isolated words presented in lists than for words presented in passages. A more robust and stable representation of the words appears to be needed for the more challenging task of segmentation compared with simple word-form recognition. Here we see, similarly, that infants who are just beginning to produce VMS are sensitive to the occurrence of those consonants in words; their occurrence in isolated words, where they are less likely to be overwritten by succeeding words or syllables, seems to hold infant attention longer.

The fact that the overall pattern of infant responses is the same in British English and Italian suggests a possible underlying mechanism that is independent of the ambient language. Moreover, it may be independent of language itself. That is, the data are consistent with a more general cognitive account such as the intersensory redundancy hypothesis (Bahrick, Lickliter, & Flom, 2004), which suggests that an event that conveys coordinated time-locked information across multiple senses promotes selective attention to the underlying event. Thus, the proprioceptive feedback from the articulators, visual cues from the caregiver, and the auditory signal produced both internally, by the child's own vocalizations, and externally by an adult speaker would combine to bring the well-practiced consonant to the foreground, separating it from other consonants. As it relates to language, this idea amounts to the articulatory filter hypothesized for early infant production.

Finally, the expected relation between production and early lexical development (fifth hypothesis) was partially confirmed in this study by the regression analysis: Children who acquired VMS earlier showed more advanced (expressive) lexical development at 12 months. This is in agreement with other studies that have reported that an earlier start on consistent vocal practice with consonants correlates with earlier first word learning (e.g., McCune & Vihman, 2001; Stoel-Gammon, 1989). However, this early advantage does not necessarily lead to more rapid lexical development in the longer term, as suggested by the absence of a relation between VMS acquisition and expressive lexical development at 18 months. The regression analysis also failed to show a relationship between receptive lexicon and age of acquisition of VMS. Although unexpected, this result is not surprising since word comprehension, with situational support, begins well before first word use (Bergelson & Swingley, 2012).

General Conclusion

What do these results tell us about early language development? On the one hand, since we did not track the development of specific perceptual abilities reported as predictive of later advances in other studies, such as emergent sensitivity to input regularities in the middle of the first year, it is possible that skills of that type could make phonological forms more accessible independent of production advances. Thus the more perceptually advanced child might display more advanced phonological development as expressed by greater production of utterances tuned to more frequently occurring stimuli in the ambient language (Kuhl et al., 2008) or greater attention to novel phonetic forms (e.g., Werker & Curtin, 2005) and, accordingly, differential sensitivity to the consonants presented in the current experiment. Although this interpretation is possible, the growing evidence for connections between the development of production and perception (e.g., Lefkowicz & Hansen-Tift, 2012; Yeung & Werker, 2013), coupled with the lack of an effect in the HPP at six months in the current study, make it unlikely. While it is tempting to posit preëminence for perception over production (or vice versa), a more integrative approach exploring connections between the two seems more persuasive. Separate evidence for the activation of auditory (superior temporal) and motor (inferior frontal) areas of the brain in 6- and 12-months olds, but not neonates, in response to speech (Imada et al., 2006) provides compelling independent evidence for this alternative view.

In our interpretation, then, the findings can be taken to mean that children with greater articulatory experience are also better able to retain a sample of phonological forms from input speech, based on their familiarity with elements of those sequences from their own vocal practice. Following the articulatory filter hypothesis, babbling production can be considered an important condition for children to construct the phonological system of their ambient language: VMS production sensitizes children to a larger range of ambient language patterns and supports the development of more robust phonological representations of those patterns. The supraglottal consonants that are motorically the most accessible and that are abundantly modeled in the input are the most favored consonants in early production. Perception and production interact to guide infant's babble to coalesce around consonants in the ambient language.

What emerges from this study is a clearer picture of the likely path from babble to words and good evidence of interaction between production and perception at the end of the first year. From this period on, infants become increasingly sensitive to the phonetic categories of their native language and their production becomes more and more attuned to input characteristics (e.g., DePaolis, Vihman, & Kunnari, 2008; Kunnari, Nakai, & Vihman, 2001; Vihman, Nakai, & DePaolis, 2006; Vihman & Velleman, 2000), resulting in a more advanced level of phonological representation. According to this developmental model, while practice with a single VMS sensitizes infants to that pattern in the input (phonetic level), practice with more than one VMS leads children to begin to construct a phonological system based on differences between patterns.

This idea is in line with a broader non-linear developmental model (Thelen & Smith, 1994) that assumes that higher levels of complexity (phonology) can be explained through the integration of less complex elements (babble and attunement to native language sound patterns) to produce "cohesive patterns" (Smith & Thelen, 2003). Thus, the development of linguistic knowledge can be seen as the product of a complex interaction between factors that cannot be considered independent but must instead be understood as strictly associated; "no single element has causal priority" (Smith & Thelen, 2003, p. 344).

ACKNOWLEDGMENTS

The authors would like to thank Dr. Andrea Capra (PhD, Department of Industrial Engineering, University of Parma), who kindly gave his time to record the stimuli and help set up the experiment; Dr. Roberto Burro, who helped with the cluster analysis; and Tamar Keren-Portnoy, who provided insightful comments on an earlier draft. They are also grateful to the families who participated in this study.

REFERENCES

Anderson, J. L., Morgan, J. , L., & White, K. S. (2003). A statistical basis for speech sound discrimination. *Language and Speech*, *46*, 155–182.

Bahrick, L. E., Lickliter, R., & Flom, R. (2004). Intersensory redundancy guides the development of selective attention, perception, and cognition in infancy. *Current Directions in Psychological Science*, *13*(3), 99–102.

Batinti, A. (1993). *Il sistema fonologico dell'italiano. Proposte per ricerche statistiche nel lessico basico*. Perugia, Italy: Guerra Edizioni.

Benasich, A. A., & Tallal, P. (2002). Infant discrimination of rapid auditory cues predicts later language impairment. *Behavioral Brain Research*, *136* (1), 31–49.

Benasich, A. A., Thomas, J. J., Choudhury, N., & Leppänen, P. H. T. (2002). The importance of rapid auditory processing abilities to early language development: Evidence from converging methodologies. *Developmental Psychobiology*, *40*, 278–292.

Bergelson, E., & Swingley, D. (2012). At 6 to 9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences of the USA, 109*, 3253–3258. Retrieved from www.pnas.org/cgi/doi/10.1073/pnas.1113380109

Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation hypothesis. In J. C. Goodman & H. C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167–224). Cambridge, MA: MIT Press.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171–204). Timonium, MD: York Press.

Best, C. T., & McRoberts, G. W. (2003). Infant perception of nonnative consonant contrasts that adults assimilate in different ways. *Language & Speech*, *46*, 183–216.

Bloom, K. (1975). Social elicitation of infant vocal behavior. *Journal of Experimental Child Psychology*, *20*, 51–58.

Bloom, K., Russell, A., & Wassenberg, K. (1987). Turn taking affects the quality of infant vocalizations. *Journal of Child Language*, *14*, 211–227.

Caselli, M. C., & Casadio, P. (1995). *Il primo vocabolario del bambino: guida all'uso del questionario MacArthur per la valutazione della comunicazione e del linguaggio nei primi anni di vita*. Milan, Italy: Franco Angeli.

Chapman, R. (2000). Children's language learning: An interactionist perspective. *Journal of Child Psychology and Psychiatry*, *41*(1), 33–54.

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

Davis, B. L., & MacNeilage, P. F. (2000). An embodiment perspective on the acquisition of speech perception. *Phonetica*, *57*, 229–241.

Dehaene-Lambertz, G., Hertz-Pannier, L., Dubois, J., Mériaux, S., Roche, A., Sigman, M., & Dehaene, S. (2006). Functional organization of perisylvian activation during presentation of sentences in preverbal infants. *Proceedings of the National Academy of Science USA*, *103*, 14240–14245.

DePaolis, R., Vihman, M. M., & Keren-Portnoy, T. (2011). Do production patterns influence the processing of speech in prelinguistic infants? *Infant Behavior and Development*, *34*, 590–601.

DePaolis, R., Vihman, M. M., & Keren-Portnoy, T. (2013a). When do infants begin recognizing familiar words in sentences? *Journal of Child Language*. doi:10.1017/S0305000912000566

DePaolis, R. A., Vihman, M. M., & Kunnari, S. (2008). Prosody in production at the onset of word use: A cross-linguistic study. *Journal of Phonetics*, *36*, 406–422.

DePaolis, R. A., Vihman, M. M., & Nakai, S. (2013b). The influence of babbling patterns on the processing of speech. *Infant Behavior and Development*.

Eimas, P., Sequeland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, *171*, 303–306.

Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience*, *15*, 399–402.

Ferguson, C. A., & Farwell, C. B. (1975). Words and sounds in early language acquisition. *Language*, *51*, 419–439. (Reprinted in *The emergence of phonology: Whole-word approaches, cross-linguistic evidence*, by M. M. Vihman & T. Keren-Portnoy, Eds., 2013, Cambridge, England: Cambridge University Press.)

Fry, D. B. (1966). The development of the phonological system in the normal and the deaf child. In F. Smith & G. Miller (Eds.), *The genesis of language: A psycholinguistic approach* (pp. 187–206). Cambridge, MA: MIT Press.

Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin and Review*, *13*(3), 361–377.

Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science*, *19*, 515–522.

Gros-Louis, J., West, M. J., Goldstein, M. H., & King, A. P. (2006). Mothers provide differential feedback to infants' prelinguistic sounds. *International Journal of Behavioral Development*, *30*, 509–516.

Imada, T., Zhang, Y., Cheour, M., Taulu, S., Ahonen, A., & Kuhl. P. K. (2006). Infant speech perception activates Broca's area: A developmental magneto-encephalography study. *Neuroreport*, *17*, 957–962.

Jones, S. S. (2007). Imitation in infancy: The development of mimicry. *Psychological Science*, *18*, 593–599.

Jusczyk, P. W. (1992). Developing phonological categories from the speech signal. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 17–64). Timonium, MD: York Press.

Jusczyk, P. W. (1993). From general to language-specific capacities: The WRAPSA model of how speech perception develops. *Journal of Phonetics*, *21*, 3–28.

Jusczyk, P. W. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press.

Kemler-Nelson, D. G., Jusczyk, P. W., Mandel, D. R., Myers, J., Turk, A., & Gerken, L. (1995). The head turn preference procedure for testing auditory perception. *Infant Behaviour and Development*, *18*, 111–116.

Keren-Portnoy, T., Majorano, M., & Vihman, M. M. (2009). From phonetics to phonology: The emergence of first words in Italian. *Journal of Child Language*, *36*, 235–267.

Keren-Portnoy, T., Vihman, M. M., DePaolis, R. A., Whitaker, C., & Williams, N. A. (2010). The role of vocal practice in constructing phonological working memory. *Journal of Speech, Language, and Hearing Research*, *53*, 1280–1293.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, *5*, 831–843.

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Graxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B*, *363*, 979–1000.

Kuhl, P. K., & Meltzoff, A. N. (1996). Infant vocalization in response to speech: Vocal imitation and developmental change. *Journal of the Acoustical Society of America*, *100*, 2425–2438.

Kunnari, S., Nakai, S., & Vihman, M. M. (2001). Cross-linguistic evidence for the acquisition of geminates. *Psychology of Language and Communication*, *5*, 13–24.

Lewkowicz, D. J. (2011). The biological implausibility of the nature–nurture dichotomy and what it means for the study of infancy. *Infancy*, *16*, 331–367.

Lewkowicz, D. J., & Hansen-Tift, A. M. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *Proceedings of the National Academy of Sciences*, *109*, 1431–1436.

Liberman, A., Cooper, F., Shankweiler, D., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, *74*, 431–461.

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition 21*(1), 1–36.

Liberman, A. M., & Whalen, D. H. (2000). On the relation of speech to language. *Trends in Cognitive Sciences*, *4*, 187–196.

Lindblom, B. (2000). Developmental origins of adult phonology: The interplay between phonetic emergents and the evolutionary adaptations of sound patterns. *Phonetica*, *57*, 297–314.

Locke, J. L. (2001). First communion: the emergence of vocal relationships. *Social Development*, *10*, 294–308.

Macken, M. A. (1980). Aspects of the acquisition of stop systems: A cross-linguistic perspective. In G. Yeni-komshian, J. F. Kavanagh, & C. A. Ferguson (Eds.), *Child Phonology, I: Production* (pp. 143–168). New York, NY: Academic Press.

MacQueen, J. B. (1967). Some methods for classification and analysis of multivariate observations. In L. M. Le Cam & J. Neyman (Eds.), *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability* (pp. 281–297). Berkeley, CA: University of California Press.

Majorano, M., & D'Odorico, L. (2011). The transition into Italian language. *First Language*, *31*, 47–66.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*, B101–111.

McCune, L. (1992). First words: A dynamic systems view. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 313–336). Timonium, MD: York Press.

McCune, L., & Vihman, M. M. (2001). Early phonetic and lexical development: A productivity approach. *Journal of Speech, Language and Hearing Research*, *44*, 670–684.

Oller, D. K., Wieman, L. A., Doyle, W. J., & Ross, C. (1976). Infant babbling and speech. *Journal of Child Language*, *3*, 1–11.

Olswang, L. B., Stoel-Gammon, C., Coggins, T., & Carpenter, R. (1987). *Assessing prelinguistic and early linguistic behaviors in developmentally young children*. Seattle, WA: University of Washington Press.

Paul, R., & Jennings, P. (1992). Phonological behaviors in toddlers with slow expressive language development. *Journal of Speech and Hearing Research*, *35*, 99–107.

Piaget, J. (1936). *Origins of intelligence in the child*. London, England: Routledge & Kegan Paul.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, *27*, 169–192.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, *3*, 131–141.

Savinainen-Makkonen, T. (2007). Geminate template: A model for first Finnish words. *First Language, 27*, 347–359. (Reprinted in *The emergence of phonology: Whole-word approaches, cross-linguistic evidence*, by M. M. Vihman & T. Keren-Portnoy, Eds., 2013, Cambridge, England: Cambridge University Press).

Smith, L. B., & Thelen, E. (2003). Development as a dynamic system. *Trends in Cognitive Sciences 7*, 343–348.

Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Science*, *93*, 13515–13522.

Stella, V., & Job, R. (2001). Le sillabe PD/DPSS. Una base di dati sulla frequenza sillabica dell'italiano scritto. *Giornale Italiano di Psicologia*, *3*, 633–642.

Stoel-Gammon, C. (1985). Phonetic inventories 15-24 months — a longitudinal study. *Journal of Speech and Hearing Research*, *28*, 505–512.

Stoel-Gammon, C. (1989). Prespeech and early speech development of two late talkers. *First Language*, *9*, 207–224.

Thelen, E., Corbetta, D., & Spencer, P. (1996). The development of reaching during the first year: The role of movement speed. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 1059–1076.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.

Veneziano, E. (1981). Early language and nonverbal representation: A reassessment. *Journal of Child Language*, *8*, 541–563.

Vihman, M.M. (1991). Ontogeny of phonetic gestures: Speech production. In I. G. Mattingly & M. Studdert-Kennedy (Eds.), *Modularity and the motor theory of speech perception: Proceedings of a conference to honor Alvin M. Liberman* (pp. 69–84). Hillsdale, NJ: Erlbaum.

Vihman, M. M. (1993). Variable paths to early word production. *Journal of Phonetics*, *21*, 61–82.

Vihman, M. M. (1996). *Phonological Development: The origins of language in the child*. Oxford, England: Blackwell.

Vihman, M. M. (2002). The role of mirror neurons in the ontogeny of speech. In M. Stamenov & V. Gallese (Eds.), *Mirror neurons and the evolution of brain and language* (pp. 305–314). Amsterdam, The Netherlands: John Benjamins.

Vihman, M. M., Kay, E., Boysson-Bardies, B. de, Durand, C., & Sundberg, U. (1994). External sources of individual differences? *Developmental Psychology*, *30*, 651–662.

Vihman, M. M., Keren-Portnoy, T., Whitaker, C., Bidgood, A., & McGillion, M. (2013). Late talking toddlers: Relating early phonological development to later language advance. York Papers in Linguistics.

Vihman, M. M., Nakai, S., & DePaolis, R. A. (2006). Getting the rhythm right: A cross-linguistic study of segmental duration in babbling and first words. In L. Goldstein, D. Whalen, & C. Best (Eds.), *Laboratory phonology 8* (pp. 341–366). New York, NY: Mouton de Gruyter.

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

Vihman, M. M., & Velleman, S. L. (2000). The construction of a first phonology. *Phonetica*, *57*, 255–266.

Werker, J. F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Learning and Development*, *1*(2), 197–234.

Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology*, *24*, 672–683.

Werker, J. F., & Pegg, J. E. (1992). Infant speech perception and phonological acquisition. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 285–311). Timonium, MD: York Press.

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

Werker, J. F., & Yeung, H. H. (2005). Infant speech perception bootstraps word learning. *Trends in Cognitive Sciences*, *9* (11), 519–527.

Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*, *7*, 701–702.

Yeung, H. H., & Werker, J. F. (2009). Learning words' sounds before learning how words sound: 9-month-olds use distinct objects as cues to categorize speech information. *Cognition*, *113*, 234–243.

Yeung, H. H., & Werker, J. F. (2013). Lip movements affect infants' audiovisual speech perception. *Psychological Science*, *24*, 603–612.

Zmarich, C., & Miotti, R. (2003). The frequency of consonants and vowels and their co-occurrences in the babbling and early speech of Italian children. In M. J. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona* (pp. 1947–1950).

Zmarich, C., Dispaldro, M., Rinaldi, P., & Caselli, M.C. (2011). Caratteristiche fonetiche del primo vocabolario del bambino. *Psicologia Clinica dello Sviluppo*, *15*, 235–253.