Infant Behavior & Development 36 (2013) 642–649

Contents lists available at ScienceDirect

Infant Behavior and Development

The influence of babbling patterns on the processing of speech

Rory A. DePaolis[a,*], Marilyn M. Vihman[b], Satsuki Nakai[c]

[a] *Communication Sciences and Disorders, James Madison University, MSC #4304, Harrisonburg, VA 22807, USA*
[b] *University of York, England, UK*
[c] *Queen Margaret University, Scotland, UK*

ARTICLE INFO

Article history:
Received 27 June 2012
Received in revised form 10 June 2013
Accepted 27 June 2013
Available online 31 July 2013

Keywords:
Speech production
Speech perception
The headturn preference paradigm
The articulatory filter
Babble
Action–perception link

ABSTRACT

This study compared the preference of 27 British English- and 26 Welsh-learning infants for nonwords featuring consonants that occur with equal frequency in the input but that are produced either with equal frequency (Welsh) or with differing frequency (British English) in infant vocalizations. For the English infants a significant difference in looking times was related to the extent of production of the nonword consonants. The Welsh infants, who showed no production preference for either consonant, exhibited no such influence of production patterns on their response to the nonwords. The results are consistent with a previous study that suggested that pre-linguistic babbling helps shape the processing of input speech, serving as an articulatory filter that selectively makes production patterns more salient in the input.

1. Introduction

The developmental strands of vocal production and speech perception, which together constitute the critical underpinning for language acquisition, are necessarily deeply interconnected, yet they are traditionally studied in isolation. The primary reason may be methodological. Production studies are typically observational, involving small numbers of infants and a large investment in time per infant. The analysis often concentrates on individual differences. In contrast, perception studies examine groups, using techniques in which individual differences are often treated as noise that obscures the findings based upon group responses. Combining these two methods is difficult, since the attrition rate in perception experiments can be as high as 33% (Jusczyk & Aslin, 1995; Jusczyk, Cutler, & Redanz, 1993; Jusczyk, Houston, & Newsome, 1999), effectively making it risky to invest a good deal of time in each infant. This study combines the observational and experimental approaches to language acquisition by utilizing individual differences in a production study to guide the analysis of perception experiments that examine group differences.

There are compelling reasons to expect the link between perception and action to be strong in the infant. Motor activity is one of the earliest modes used to explore the world in activities such as kicking (Rovee-Collier, 1995), reaching (Witherington, 2005), crawling (Campos et al., 2000), touching (Striano & Bushnell, 2005), and mouthing (Gottfried & Rose, 1980). In the case of the perceptuomotor activity of speech, the action–perception link is supported by a wide range of empirical findings with adults (e.g., Fowler & Dekle, 1991; Kerzel & Bekkering, 2000; McGurk & MacDonald, 1976) and a smaller number of studies with infants.

* Corresponding author. Tel.: +1 540 568 3869.
E-mail address: depaolra@jmu.edu (R.A. DePaolis).

0163-6383/$ – see front matter
http://dx.doi.org/10.1016/j.infbeh.2013.06.007

For example, Kuhl and Meltzoff (1984) found that infants prefer to look at faces that are seen to be producing the vowels they are listening to than at faces seen to be producing incongruent vowels. Interestingly, this preference is affected by the type of lip movement the infant is making while looking and listening to these vowels (Yeung & Werker, 2013). In addition, the McGurk effect (McGurk & MacDonald, 1976) has been demonstrated in 2- and 5-month-old English learning infants (Burnham & Dodd, 2004; Rosenblum, Schmuckler, & Johnson, 1997). The recent finding that infants shift their gaze away from the eyes to the mouth at about the same time that they begin to babble constitutes further evidence for this link (Lewkowicz & Hansen-Tift, 2012).

Taking these studies as a starting point and considering speech in the context of other motor activities that inform cognitive processes, there is a distinct possibility of a link between an infant's babble and their perception of speech that incorporates elements of that babble. One way to think of such a link is to consider that the early production ability of an infant could provide enhanced accessibility to input that matches his or her production (Locke, 1993, p. 204). This process of matching production to perception has been suggested to act as an 'articulatory filter' (Vihman, 1996, p. 142; see also Vihman, 1991, 1993), making patterns that an individual infant regularly produces more salient in the input.

A recent study by DePaolis, Vihman, and Keren-Portnoy (2011) provided empirical evidence for this link. DePaolis et al. followed the babble of prelinguistic infants with regular recordings. Once there was evidence of a well established or 'favored' consonant, these investigators tested the infants' preference for passages with nonwords containing this favored consonant versus a passage with nonwords containing consonants that the infant was not yet frequently or consistently producing. The infants who were favoring two consonants in their babble exhibited a significant preference for the passage with nonwords comprised of the consonant that they were *not* producing. This effect was interpreted as a novelty response, based on the infants' over-familiarity with the consonants known to them from their own babbling practice. This interpretation was supported in a study by Majorano, Vihman, and DePaolis (in revision), which showed the complementary familiarity effect: Italian infants who show frequent use of only one consonant in babble showed a significant preference for words with that consonant over words with a consonant that they are not yet producing with any consistency.

Since it is possible that the frequency of consonants in the input speech could account for these effects, both DePaolis et al. (2011) and Majorano et al. (in revision) investigated the frequency of consonants used in the input and tested whether mothers' differing consonant use in infant directed speech (IDS) might be the source of their infants' babble patterns. Both studies found that while the infants' babbling patterns were marked by clear differences, the frequency of consonantal use in the mothers' IDS was very similar and not always related to the consonant use in their infant's babble. These findings were consistent with Vihman, Kay, Boysson-Bardies, Durand, and Sundberg (1994), who found, in each of three language groups, that infants' productions were highly variable while the mothers' sound patterns in any given language varied very little.

Nevertheless, since the input frequency of words affects lexical development (Goodman, Dale, & Li, 2008; Hart, 1991) and the production patterns of infants match characteristics of the IDS to which they are exposed (Lee, Davis, & MacNeilage, 2008), it is possible that the infant is affected by the input frequency of consonants. It could be that the frequency of consonants in the infant's environment affects perception in much the same way that the statistical or phonological properties of the language do (Saffran, Aslin, & Newport, 1996; Saffran & Thiessen, 2003). One way to examine this prospectively is to use experimental stimuli which highlight consonants that are either high or low frequency in the ambient language but potentially contrastive in the infants' production (frequently versus infrequently produced in babble).

Accordingly, the purpose of this study was to explore the effect of production on perception by matching the frequency of input-language occurrence of the speech stimuli presented in the experiment. We chose contrasts in Welsh and in British English that had a similar frequency of occurrence, but that we assumed would differ in their frequency of occurrence in the infants' vocal production. First infants' babbling patterns were documented; then the headturn paradigm was used to test for a correlation between an individual infant's production patterns and their responses to these patterns in the perception experiment. We hypothesized that the infants should respond to consonants that are prevalent in their own babble differently than to consonants that are not, as indexed by a difference in looking times to the two types of stimuli. Alternatively, if the frequency of occurrence of consonants in the IDS was a major factor but the infants' production patterns were not, then the infants should show no preference for either contrast.

2. Methods

The study consisted of two parts, a longitudinal observational study and a perception test. The first part was designed to provide a profile of each infant's production patterns so that the subsequent perception test could be analyzed with reference to the individual infant's production.

2.1. Participants

A total of 53 infants from the area around Bangor, Wales participated in the study. Each infant participated in both parts of the study. The English-learning sample included 27 infants (15 males) and the Welsh-learning sample consisted of 26 infants (13 males). Three additional Welsh infants were excluded from the study when it was observed that their parents used English with them during the recording sessions; no other participants had to be excluded. The lack of attrition in the perception experiments was unexpected, but consistent with low attrition in two of three experiments in a previous study (Vihman, Nakai, DePaolis, & Hallé, 2004) with 12 test trials; the current study had only eight test trials (see below).

Table 1
Nonword stimuli.

Spelling	IPA	Spelling	IPA
English			
tetty	/tɛti/	sessy	/sɛsi/
tooyer	/tujə/	sooyer	/sujə/
tawey	/tɑwi/	sawey	/sɑwi/
tutter	/tʌtə/	susser	/sʌsə/
teehy	/tihi/	seehy	/sihi/
tetter	/tɛtə/	sasser	/sæsə/
Welsh			
bewa	/bewa/	gewa	/gewa/
bawy	/bawɨ/	gowy	/gowɨ/
boia	/boja/	gwia	/guja/
bwba	/buba/	giga	/giga/
bibu	/bibɨ/	gagu	/gagɨ/
bybu	/bəbɨ/	gygu	/gəgɨ/

It is important to note that although the Welsh infants were learning Welsh only in the home, virtually all Welsh speakers are bilingual so it is likely that the infants were also exposed to English on a regular basis (Vihman, Thierry, Lum, Keren-Portnoy, & Martin, 2007). All infants began the study at 10.5 months of age. In addition, each infant passed a hearing screening with the National Health Service and none had any reported health problems at the time of the study.

2.2. Production study

Infants were recorded on audio and video in unstructured half-hour play sessions in the home with a caregiver, beginning at 10.5 months of age and continuing fortnightly until 12 months of age (four sessions). Recordings were made using a wireless microphone and transmitter (AKG PT60) placed in an inside pocket of a soft vest worn by the infants. The caregiver also wore a wireless microphone and transmitter. All audio recordings were digitized at 44.1 kHz using a Tascam DA-P1 DAT recorder.

Each session was transcribed from the audio signal by two independent transcribers, who agreed on 75% of the consonants for English and 82% of the consonants for Welsh. Only those consonants transcribed in the same way by the two transcribers were included in the consonant counts, to ensure reliable identification.

2.3. Perception study

2.3.1. Stimuli

Female native speakers of Southern British English and Northern Welsh recorded English or Welsh nonwords, respectively (see Table 1). The nonwords consisted of two CV syllables. Each nonword began with the target consonant, while the medial consonant varied between a repeat of the target consonant, a glottal (/h/), or a glide (/w,j/). The repeated consonant, glottal and glides were chosen to avoid 'competition' with the target consonant while ensuring a minimum of diversity in the stimuli. Both glottals and glides occur early in infant vocalizations and do not involve full supraglottal closure, while the target consonants are emergent in the age-range under investigation.

The target consonants, /t/ versus /s/ in English and /b/ versus /g/ in Welsh, were chosen to have comparable input frequency within each respective language, based on Mines, Hanson, and Shoup (1978) for English and on mothers' speech to 12-month-old infants for Welsh. The frequency counts for the latter were based upon 30-min observational sessions for five mothers of infants in an earlier study. The use of consonant counts from the mothers' speech is not ideal, since the corpus upon which the estimated exposure to consonants is so small and derives from IDS, unlike the English counts, but no independent consonant counts were available for Welsh.

In English /t/ is the most frequent early consonant while /s/ is infrequently used by infants at this age (McCune & Vihman, 2001), providing an excellent test of a production versus non-production contrast with similar frequency of occurrence in the input. In Welsh, the similar input frequency of /b/ and /g/ contrasts with our choice of test consonants for English, since these are both among the earliest consonants produced. However, we expected frequent production of the bilabial to precede that of the velar for most infants, with some infants possibly showing the reverse pattern (Locke, 1983; Vihman, 1996).

Each contrasting pair of nonwords in English and Welsh contained the same two vowels combined with contrasting consonants. An unintended result was that five of six English but only two of six Welsh nonwords differed with respect to consonants only, since the vowels used were identical (see Table 1). The contrast of these words across the English lists could create what amounts to minimal pairs (e.g., *tetty : sessy*). In order for an infant to detect such an effect, however, the infant would have to retain the words in short term memory long enough to compare and contrast words across lists. Although unlikely, the infant could then be expected to show interest in both word lists due to the novelty of experiencing

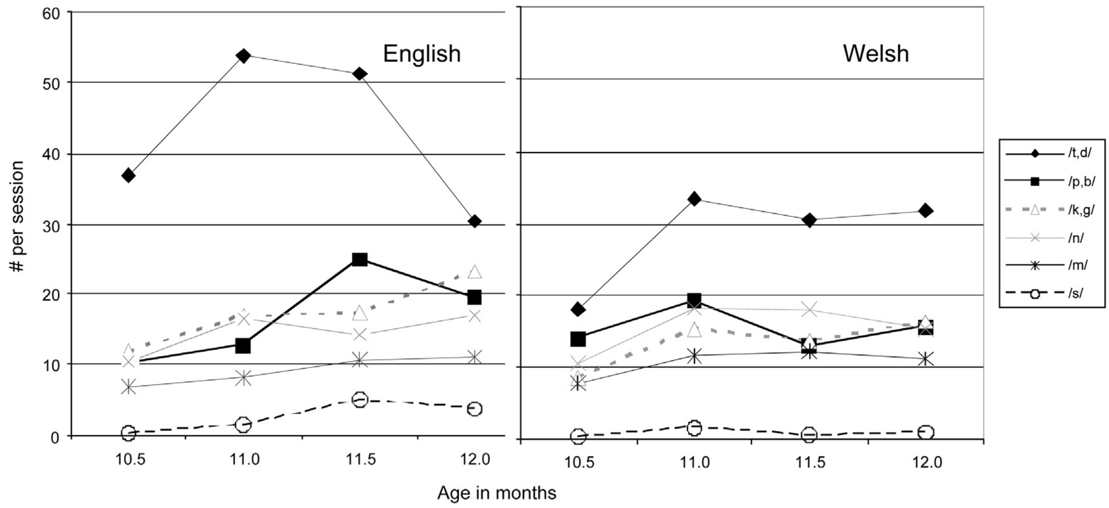

Fig. 1. Average number of consonants produced per child in English (left) and Welsh (right). Note the difference in English productions of /t,d/ versus /s,z/ and those of the Welsh productions of /p,b/ versus /k,g/.

the minimal pairs across the lists. Thus, any differences due to perceived minimal pairs should not affect the pattern of preferences reported below.

All items were recorded in a sound-treated room using a Sennheiser ME 66 microphone (with K6 power module) connected to a Tascam DA-P1 digital recorder sampling at 44 kHz. The stimuli were transferred digitally onto a PC hard drive for eventual output.

2.3.2. Procedure

The headturn preference paradigm (HPP) used was similar to that described in Kemler Nelson et al. (1995). Seated on the caregiver's lap in a quiet darkened room, the infants faced the central panel of a three-sided test booth where a camera and red light were mounted. A blue light and speaker were mounted on each side panel. A PC and video monitor were located in the adjoining room where the experimenter controlled stimulus presentation and recorded infant looking times by pressing the left and right mouse buttons. The computer initiated and terminated trials in response to signals from the experimenter. In each trial, the infant's gaze was centered by the blinking red light. The experimenter then initiated the computer run trial involving a blinking blue light on the left or right of the infant. When the infant was judged to orient to the blue light, a trial was presented from that speaker. If the infant looked away from the speaker for more than two seconds, the trial was terminated and another begun. Both the experimenter and the caregiver wore headphones playing music (constant level with no quiet passages) chosen to effectively mask the speech stimuli.

Each experimental session consisted of a familiarization and a test phase. In the familiarization phase the infant was presented with nonwords different from those used in the test trials. The test phase consisted of 8 trials, four each of the two test conditions (/t/ versus /s/ or /b/ versus /g/ for English and Welsh, respectively). Although many of our previous studies used 12 test trials with this paradigm (DePaolis, Vihman, & Keren-Portnoy, 2012; Vihman et al., 2004), the current study included only 8 test trials in order to avoid infant habituation to the (rather repetitive) stimuli. Each familiarization and test trial consisted of a randomized presentation of six nonwords. Each nonwords was presented twice such that each trial consisted of 12 nonwords. Four trials were presented in the familiarization phase (two of each contrast). For the test phase, the six lists of each stimulus type (/t/ versus /s/ and /b/ versus /g/ nonwords for English and Welsh, respectively) were pseudo-randomized such that each nonword was presented either first or second in one list to ensure that each infant heard each of the nonwords over the course of the test phase. Thus in the test phase eight different lists were presented (four pseudo-randomizations of each of the two stimulus types). In addition, during presentation no more than two lists in a row of the same stimulus type were allowed. Reliability of looking-time coding was assessed by an independent researcher coding eight of the infants' HPP video recordings offline ($r = 0.91$, $p < 0.05$).

3. Results

3.1. Production data

Fig. 1 presents the average production across infants of the most commonly produced phonemes (plus /s/) for Welsh and English. Voiced–voiceless cognates were counted as one phoneme class since there is little evidence that infants control voice onset time at this age (Macken, 1980). Thus, /t,d/, /k,g/, /p,b/, and similarly /s,z/ were considered one category each. Of particular interest are the phonemes /t,d/ and similarly /s,z/ in English and /k,g/ and /p,b/ in Welsh, since they are contrasted in the perception test. For English it is clear from Fig. 1 that /t,d/ are produced with the greatest overall frequency while /s/ and /z/ are rarely produced. For Welsh, the contrast used in the perception test is less clearly differentiated in production: On average, the infants

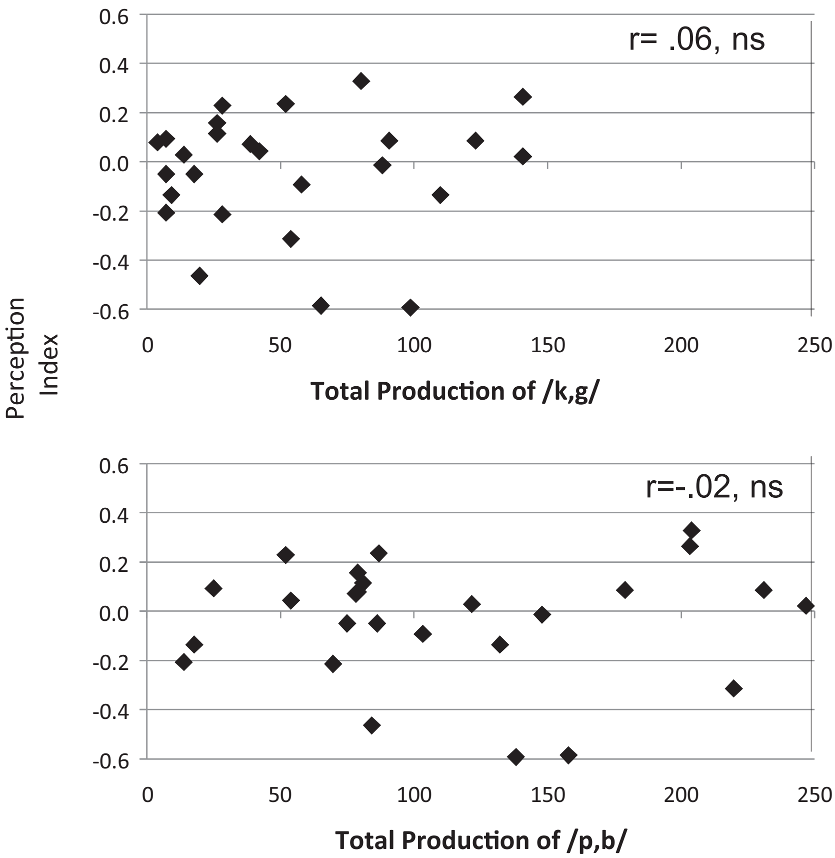

Fig. 2. The Welsh production of /k,g/ (top) and /p,b/ (bottom) versus preference (based on looking time) for /b/ non-words. The vertical axis is preference for /b/ non-words (values greater than zero). Neither correlation was significant.

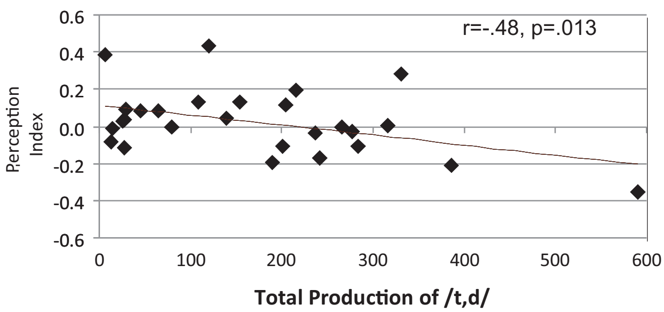

Fig. 3. The English production of /t,d/ versus the preference for looking at the /t/ non-words. Data points below zero indicate a preference for /s/ words. Note that the horizontal scale is different from that of Fig. 2.

produced both consonants with nearly equal frequency. Thus, counter to our expectations, the Welsh contrasts do not provide a contrast in consonant use in this sample.

3.2. Perception test

The English infants' mean looking time in seconds to the alveolar stop nonwords ($M = 5.43$, $SD = 1.74$) was not significantly different ($t[26] = 0.69$, $p = 0.50$) from alveolar fricative nonwords ($M = 5.20$, $SD = 1.67$). The Welsh infants' mean looking time in seconds to the bilabial stops ($M = 4.74$, $SD = 2.14$) was not significantly different ($t[25] = 0.24$, $p = 0.82$) from the velar stops ($M = 4.84$, $SD = 1.39$).

In order to test for an effect of production on perception we first computed indices of each. The perception index used for the English data was the difference between the looking time in response to /t/ versus /s/ divided by the total looking time to both phonemes ([t − s]/[t + s]). The production index was the total number of productions of [t,d] in the last session (two weeks before the headturn test), since the infants produced /s/ only minimally. The perception index used for the Welsh data was the difference between the looking time in response to /b/ versus /g/ divided by the total looking time to both phonemes ([b − g]/[b + g]). The production index used was the total number of productions of [k,g] or [b,g] in the last session (two weeks before the headturn test). Figs. 2 and 3 show the relationship between these indices plotted separately for each language.

Table 2
Summary of simple regression with the total number of alveolar stops produced predicting the results on the HPP test.

	Variable		
	B	SE *B*	*β*
Constant	0.115	0.049	
Alveolar stop production	−0.001	0.000	−0.438

$p = 0.022$.

A simple linear regression was used to examine the possibility of predicting the infants' preference for consonants based upon their production patterns (i.e., number of productions of a specific consonant was the predictor and consonant preference, using the perception index, was the dependent variable). For the Welsh infants neither the production of /p,b/ and preferential looking toward /p,b/ ($R = -0.025$, $p = 0.905$) nor the production of /k,g/ and preferential looking toward /k,g/ ($R = 0.057$, $p = 0.783$) were linearly related. However, there was a significant negative linear relationship between the production of and subsequent preference for alveolar stops ($R = -0.438$, $p = 0.022$) for the British English infants (see Table 2). Thus, infants who produced relatively fewer alveolar stops looked more in response to words including those stops, while the infants who produced the most alveolar stops looked less in response to the words including these stops and more in response to the words with alveolar fricatives.

4. Discussion

As expected, all of the infants produced considerably more alveolar stops than they did [s], but there were only minor differences in their production of [p,b], [k,g], or [m,n] across the two languages. Interestingly, [t,d] was the most common babbling pattern overall. This is consistent with previous studies in English (DePaolis et al., 2011; McCune & Vihman, 2001) and likely reflects the fact that of the supraglottal consonants, alveolar stops (including nasal stops) are both motorically accessible and abundantly modeled in the input, an interaction that is likely a factor in their early production in English. (Note that, in contrast, Italian infants produce labial stops the most frequently: Majorano et al., in press.)

The predominance of [t,d] does not carry over to perception, however, since the English infants show no group preference for either alveolar stop or alveolar fricative nonwords. Similarly, the Welsh infants show no preference for bilabial over velar stop nonwords. The lack of a difference suggests that the infants are not responding to any purely acoustic differences between the words. This is not surprising for the Welsh infants, given the acoustic similarity of the bilabial and velar stop, but the absence of a group effect for English infants in the face of the large acoustic difference between stops and fricatives suggests that signal-based differences between different manners of production do not in themselves elicit any particular interest or attention in infants of this age.

The results of the English perception experiment clearly support our hypothesis of a differential pattern of looking time based upon the infants' production patterns. The contrast of /t/ versus /s/ made it possible for us to investigate this link by testing the relation between the perception of a phoneme produced by nearly all the infants in the study and one which very few of them produced at all. The fact that the infants who produced the most alveolar stops looked less to alveolar stop nonwords than did those who produced fewer such stops provides strong evidence for a link between production and perception. The lack of any preference for either list of nonwords in the Welsh infants is not inconsistent with the hypothesis of a production–perception link since, counter to our expectations, the Welsh infants were producing both of the stops used in the nonword stimuli at about the same level.

These results parallel those of DePaolis et al. (2011), in which the consonantal production patterns of an infant's babble proved to be predictive of the infant's preference for passages of running speech that either included or did not include the infant's most produced consonants. In that study, as in this one, the more experience an infant had with producing one stop consonant, the more likely they were to show a preference for a different stop consonant in the stimuli presented.

Previous studies have documented the strong similarity in input frequencies for consonants from one mother to the next within a language group, in contrast to the strong individual production differences seen in infants. Thus, there is no reason to suspect individual effects on infant responses stemming from their parents' speech patterns. Instead, it must be the individual production pattern experiences of the infants that bias their attentional responses to particular stimuli. Nevertheless, in order to further explore the possibility that the infants' preferences might be influenced by their mothers' speech, we transcribed the infant directed speech (IDS) to four English infants. The infant–mother dyads included two infants who produced fewer than 30 and two that produced more than 200 alveolar stops over the course of three observational sessions. Inter-rater reliability was 98% for a 10-min segment from one of the infants. Fig. 4 demonstrates that the IDS to three of the infants is very similar: alveolar stops are the most frequently appearing consonants in the IDS, followed by nasals, then similar frequencies of both the velar and bilabial stops, and then the alveolar fricative. One parent shows a dissimilar trend, producing more nasals and more velar than alveolar stops. Interestingly, this infant was one of the high producers of alveolar stops, suggesting that IDS is not a good predictor of an infant's preferred production patterns. This data supports previous findings (DePaolis et al., 2011; Vihman et al., 1994) in which there is similarity among the caregivers' productions but a high degree of variability in the infants' productions.

These findings shed some light on the way that infants process speech in the pre-linguistic period. As infants begin to babble, their babbling patterns come to coalesce around the consonants of the ambient language (Kuhl, 2004) and they begin to favor consonants that are motorically accessible and supported by the input speech. These consonants are engaging for the infants, as highlighted by the fact that the English infants in our study who are just beginning to produce alveolar stops show

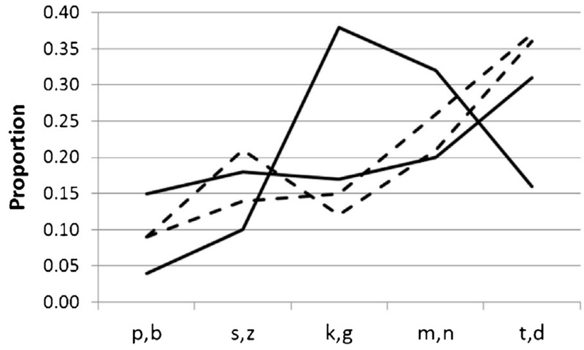

Fig. 4. The frequency of occurrence of consonants in the IDS directed to four infants. The dotted lines are for infants who produced less than 30 alveolar stops in three observational sessions. The solid lines are for the infants who produced more than 200 alveolar stops in three observational sessions.

a preference for alveolar stops. Continued production of that consonant, though, coupled with continued exposure to it in adult speech, leads to a waning of interest in the consonant. The alternate contrast, which has not yet become overly familiar through the infant's auditory experience of his or her own production, then begins to attract the infant's attention. The pattern of infant responses can be conceptualized in terms of the Hunter and Ames (1988) model of familiarity and novelty effects. Fig. 5 is redrawn from Hunter and Ames (1988) to characterize the English infants' responses in this experiment. The linear relationship is characterized as a familiarity effect for the infants who are producing small numbers of alveolar stops; this then shifts to a novelty effect as the infants begin to produce large numbers of alveolar stops.

This shift from familiarity to novelty is supported by recent work with Italian infants (Majorano et al., in press). In this study, a replication and extension of DePaolis et al. (2011) with a considerably larger sample, the infants who show repeated use of only one consonant in babble show a significant preference for words with that consonant over words with a consonant that they are not yet producing in large numbers. The infants who show repeated use of two or more consonants in babble show a significant preference for words with a consonant not yet being produced over words with a consonant that is repeatedly produced in babble. Thus, the infants' response changes from familiarity to novelty in the way predicted by Hunter and Ames (1988).

It may be that as babble coalesces around certain consonants in the ambient language, the processing of these consonants in the input becomes more efficient. A more fully instantiated memory for these consonants would result, leading to the observed switch from familiarity to novelty. The same effect has been observed in the visual modality. Roder, Bushnell, and Sasseville (2009) followed infants longitudinally and tracked their response to faces, objects, or kaleidoscope patterns. All but the latter stimulus type elicited a familiarity response followed by a strong novelty response as exposure to the visual stimuli increased. The current finding is similar but suggests that this shift from familiarity to novelty can occur in response to self-produced stimuli in a different modality.

Even more interesting is what this perception–production link might mean for development in general. Infant babble is initially a motor act supported, for example, by the environment, the inherent dynamics of the motor system, and the emerging sophistication of the auditory system. These factors, among others, interact to guide the infant's production patterns in babble toward first words, each relatively simple skill interacting to yield much more complex behaviors. The data in this experiment suggest that babble itself is one of these emerging skills. Thus, the emerging ability to produce motor patterns influences the way that infants perceive the speech patterns themselves. Consolidation of one consonantal pattern

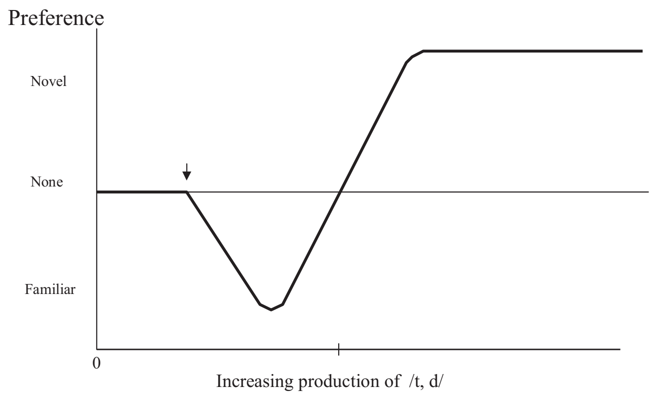

Fig. 5. Hypothesized time course of familiarity and novelty responses tied to the production of alveolar stops. The arrow indicates the onset of alveolar stop production. Figure redrawn from Fig. 2 in Hunter and Ames (1988).

enables the infant to focus on new patterns *both* motorically and perceptually. These ideas, and the data presented in this study, are consistent with a dynamic systems theory (DST) approach to development (Thelen, 1995; Thelen & Smith, 1994; see also Vihman, DePaolis, & Keren-Portnoy, 2009, for a description of DST as it applies to babble and first words).

In summary, the infants' apparent lack of interest in the consonants well established in their own production repertoire suggests a potentially important shift in attention. This mechanism could serve to guide the infant to "notice" not only what they are producing but also what they are not yet producing. Future studies could examine how this mechanism works to guide the infant to reach important milestones in language development.

Acknowledgements

This work was funded by a grant from the Economic and Social Research Council. In addition, a Marie Curie Fellowship afforded the first author time to analyze the data. Special thanks to Jarrad Lum for insights into the data analysis, Pam Martin and Eleanor Butcher for data collection, and Sarah Ilenko for help transcribing the IDS.

References

Burnham, D., & Dodd, B. (2004). Auditory–visual speech integration by prelinguistic infants: Perception of an emergent consonant in the McGurk effect. *Developmental Psychobiology: 45.*, 204–220.
Campos, J. J., Anderson, D. I., Barbu-Roth, M. A., Hubbard, E. M., Hertenstein, M. J., & Witherington, D. C. (2000). Travel broadens the mind. *Infancy: 1.*, (2), 149–219.
DePaolis, R. A., 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. A., Vihman, M. M., & Keren-Portnoy, T. (2012). When do infants begin recognizing familiar words in sentences? *Journal of Child Language*, http://dx.doi.org/10.1017/S0305000912000566
Fowler, C. A., & Dekle, D. (1991). Listening with eye and hand: Crossmodal contributions to speech perception. *Journal of Experimental Psychology: Human Perception and Performance: 1991.*, (17), 816–828.
Goodman, J. C., Dale, P. S., & Li, P. (2008). Does frequency count? Parental input and the acquisition of vocabulary. *Journal of Child Language: 35.*, 515–531.
Gottfried, A. W., & Rose, S. A. (1980). Tactile recognition memory in infants. *Child Development: 51.*, 69–74.
Hart, B. (1991). Input frequency and first words. *First Language: 11.*, 289–300.
Hunter, M. A., & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. *Advances in Infancy Research: 5.*, 69–95.
Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology: 29.*, 1–23.
Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development: 64.*, 675–687.
Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology: 39.*, 159–207.
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 Behavior and Development: 18.*, 111–116.
Kerzel, D., & Bekkering, H. (2000). Motor activation from visible speech: Evidence from stimulus response compatibility. *Journal of Experimental Psychology: Human Perception and Performance: 26.*, (2), 634–647.
Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience: 11.*, 831–841.
Kuhl, P. K., & Meltzoff, A. N. (1984). The intermodal representation of speech in infants. *Infant Behavior and Development: 7.*, 361–381.
Lee, S., Davis, B. L., & MacNeilage, P. F. (2008). Segmental properties of input to infants: A study of Korean. *Journal of Child Language: 35.*, 591–617.
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 of the United States of America: 109.*, 1431–1436.
Locke, J. L. (1983). *Phonological acquisition and change*. New York: Academic Press.
Locke, J. L. (1993). *The child's path to spoken language*. Cambridge, MA: Harvard University Press.
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*. New York: Academic Press.
Majorano, M., Vihman, M., & DePaolis, R.A. (in press). The relationship between infant production experience and their processing of speech.
McCune, L., & Vihman, M. (2001). Early phonetic and lexical development: A productivity approach. *Journal of Speech, Hearing and Language Research: 44.*, 670–684.
McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature: 264.*, 746–748.
Mines, M., Hanson, B., & Shoup, J. (1978). Frequency of occurrence of phonemes in conversational English. *Language and Speech: 21.*, 221–241.
Roder, B. J., Bushnell, E. W., & Sasseville, A. M. (2009). Infants' preferences for familiarity and novelty during the course of visual processing. *Infancy: 1.*, (4), 491–507.
Rosenblum, L. D., Schmuckler, M. A., & Johnson, J. A. (1997). The McGurk effect in infants. *Perception & Psychophysics: 59.*, (3), 347–357.
Rovee-Collier, C. (1995). Time windows in cognitive development. *Developmental Psychology: 31.*, 147–169.
Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month old infants. *Science: 274.*, 1926–1928.
Saffran, J. R., & Thiessen, E. D. (2003). Pattern induction by infant language learners. *Developmental Psychology: 39.*, 484–494.
Striano, T., & Bushnell, E. W. (2005). Haptic perception of material properties by 3-month-old infants. *Infant Behavior and Development: 28.*, 266–289.
Thelen, E. (1995). Motor development: A new synthesis. *American Psychologist: 50.*, 79–95.
Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.
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*. Hillsdale, NJ: Lawrence Erlbaum Associates.
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: Blackwell.
Vihman, M. M., DePaolis, R. A., & Keren-Portnoy, T. (2009). Babbling and words: A dynamic systems perspective on phonological development. In E. L. Bavin (Ed.), *The Cambridge handbook of child language* (pp. 163–182). Cambridge: Cambridge University Press.
Vihman, M. M., Kay, E., Boysson-Bardies, B. de, Durand, C., & Sundberg, U. (1994). External sources of individual differences? A cross-linguistic analysis of the phonetics of mothers' speech to one-year-old children. *Developmental Psychology: 30.*, 651–662.
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., Thierry, G., Lum, J., Keren-Portnoy, T., & Martin, P. (2007). Onset of word form recognition in English, Welsh and English-Welsh bilingual infants. *Applied Psycholinguistics: 28.*, 475–493.
Yeung, H. H., & Werker, J. F. (2013). Lip movements affect infants' audiovisual speech perception. *Psychological Science: 24.*, 603–612.
Witherington, D. C. (2005). The development of prospective grasping control between 5 and 7 months: A longitudinal study. *Infancy: 7.*, 143–161.