Institute of Phonetic Sciences,
University of Amsterdam,
Proceedings 17 (1993), 65-78.

CYCLIC EFFECTS OF INFANT SPEECH PERCEPTION, EARLY SOUND PRODUCTION, AND MATERNAL SPEECH

Florien J. Koopmans-van Beinum

Abstract

During the last few decades more and more phoneticians have shown interest in infant speech development, and many stimulating results have been obtained in this field. These results concern early infant perception as well as infant sound production and mother-child interaction. However, achievements from these separate areas have only incidentally been considered in relation to each other. In this paper we have tried to combine knowledge from these areas into a broader picture of the development of infant speech communication.

1. INTRODUCTION

One of the most challenging problems for a phonetician is, in my opinion, to understand how young children master the basic elements of speech communication within the span of only one year. Around the time of their first birthday children usually understand most simple information and instructions, and they often use a few words with clear referential meaning, understandable by the people in their direct environment. What are the sources of speech input that an infant can make use of on its way of becoming a full-fledged partner in speech communication? What kind of speech sound properties and inter- as well as intra-speaker variabilities does the child have to cope with in its attempts to detect the structures of speech and language?

Before birth, in its mother's womb, the senses of a child start to function gradually. Neuronal impulses are transmitted to its brain, and little by little some impulses passing and passing again along the same way are connected into patterns, determining the neuronal basis of what will be the child's conceptual memory (Penfield & Roberts, 1952). Especially the auditory and tactile sensory systems register and transmit quite a number of impulses, some of them in a continuous stream of patterns and with great regularity: a never-ending rhythmic beat. Thus, there is no question about it, indeed "all God's children got rhythm" (Bickley, Lindblom & Roug, 1986), in fact they have it from the very beginning of life. The heart beat of the mother is probably the most basic sensation for an unborn child.

But the auditory system registers more. Apart from all sorts of bowel-noises of the mother, stretches of melodic sound alternated with silence are registered as well, in combination with the sensation of the mother's breathing movements. In spite of the damping by womb tissue and amniotic fluid all speech penetrates the infant's auditory system, internally as well as externally. There are indeed more stretches of sound and silence from outside the mother-child system, but these are not combined with those special breathing movements which occur when the mother is speaking.

Then, as soon as the child is born, the whole auditory input situation changes. Many of the familiar sounds have disappeared and totally different sounds are heard instead. Only at times can the safe and regular beat be heard vaguely. The long stretches of sound and silence differ from those heard before and are now combined with visible shapes and sometimes with tactile movements. In other words: a different world altogether. Thus, it is not surprising at all that, immediately after birth, the infant prefers those sounds that belonged to the system it was part of until birth (Moon & Fifer, 1986; Fifer & Moon, 1989), this actually being the infant's own system.

Without going into the problems of separating congenital and environmental effects on early speech and language acquisition, we agree with Locke (1990) that *"innate predispositions (activated or maintained by environmental stimulation) bias the organism to pay special attention to certain sensory patterns and therefore to learn particularly about those things."* (p. 624). This will give children the good start on their way to socialization and communication which they need. For, at birth, the 'human cub' is of all mammals probably the most dependent. Therefore, it is of vital importance for the infant, from the very first moments onwards, to succeed in communicating its needs to its direct environment, which is most often one of the parents. Fortunately, producing sounds turns out to be a very effective means for that purpose: almost every human adult is inclined to react immediately in one way or another when he or she hears an infant crying.

Thus, parent and child (initially, most of the time: mother and child) tune in to each other in a process which is natural and universal. However, when comparing individual mother-infant pairs, there may be large differences in the way in which this mutual attuning actually takes place. Each pair has to develop its own tuning system towards a communication system that is acceptable to both partners, as well as to the rest of the world (Fernald, 1992; Van der Stelt, 1993). It is clear that this mother-infant interaction system does not always develop without problems. Van der Stelt in her study focuses on three, closely connected, basic aspects of developing communication: *intersubjectivity*, *intentionality*, and *turntaking*. Transmission of intentions assumes intersubjective tuning, whereas turntaking may follow upon transmitted intentions. One of the most important conclusions of her study is the claim that it is the mother-infant interaction in the first six months of life which determines whether the development of speech communication will be uncomplicated or not. In those first months, the foundation must be laid for an optimal development of language and speech, as these are the most important human means of communication.

It may not be accidental that after this period, at about six or seven months of age, infants start to produce canonical babbling (Van der Stelt & Koopmans-van Beinum, 1986). Unlike the types of sound productions in the preceding periods, these types of utterances consisting of long series of repetitive articulations, contain all basic elements of adult speech production and are thus basic to further language acquisition (Oller, 1986; Koopmans-van Beinum & Van der Stelt, 1986; Koopmans-van Beinum, 1990). Two pre-conditions ought to be fulfilled at that stage. Firstly, there has to be a well-developed affect system between mother and infant (or with any other direct caregiver), as this is essential for learning the basic rules of speech communication. Secondly, the child should have a (healthy) well-developed and well-trained speech production instrument, capable of making all speech movements essential for accurately producing the necessary sound distinctions of the mother tongue. At least from this moment onwards (and possibly, even earlier) we may be able to trace some influences of the environmental speech input in the speech output of the child.

However, what is this environmental speech input which the child will use to construct a stable picture of the sound patterns of the mother tongue? And how will these sound patterns be used in speech communication? It will be very instructive to consider the changing character of the speech input which a child deals with before the

age of one year. It is challenging to try to relate the type of this input to the infant's changing perceptual capacities in the course of the first year, as well as relating the type of input to its developing sound producing system (Bloom, 1988). None of these three aspects of the child's speech communication development can be seen as a constant and unchangeable building block. On the contrary, it is most probable that within the developmental process a mutual influence or cyclic effect is permanently in action, as is indicated schematically (and incompletely) in Fig. 1: the infant produces sounds, the mother reacts with (among other things) speech, the child perceives the mother's speech, analyses this speech, builds some global representation of communication, of speech, and of the mother tongue, the child experiences its own sound production system, produces sounds, the mother reacts, ... etc., etc.. Although we are not able to survey this entire complex research area in detail, we will try to put together some of the pieces of this intriguing jig-saw puzzle.

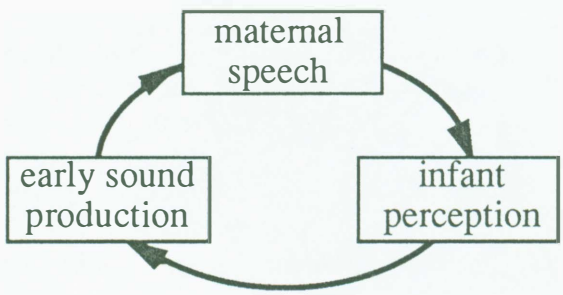

Fig. 1. Schematic representation of the assumed cyclic effect of infant speech perception, early sound production, and maternal speech, within the infant's speech developmental process.

To account for the way in which "the component processes that underlie word recognition in fluent speech evolve during the course of language acquisition", Jusczyk (1993) introduced the WRAPSA-model (Word Recognition and Phonetic Structure Acquisition). The main assumptions of the WRAPSA-model can be useful in our approach as well:

" ..1. Infants are born with a set of general auditory analyzers that underlie the perception of both speech and non-speech signals.
2. In acquiring a native language, the infant develops a scheme for weighting the information available through the auditory analyzers.
3. The weighting scheme is particular to the language and reflects the typical sound structure of words in the language.
4. Pattern extraction, in which words are segmented from the speech stream, is performed on the weighted output from the analyzers.
5. The representations provided by the pattern extraction process are structured into syllable-sized units containing prosodic marking and loosely organized featural information.
6. The representations form the probes to secondary memory where they are matched against traces of previous analyzed utterances, rather than against abstract prototype representations of the sound structure of known words." (p. 24).

In our attempt to describe the cyclic effects in the infant's speech developmental process, these main assumptions of the WRAPSA-model can be a good starting point at least as far as the infant's perception and the sound input are concerned, although the model does not show a cyclic structure.

Table 1. Overview of co-occurrences of speech characteristics in infant speech perception, maternal speech, and infant sound production, with global indication of periods of occurrence.

GLOBAL PERIOD INDICATION	INFANT SPEECH PERCEPTION	MATERNAL SPEECH	INFANT SOUND PRODUCTION
before birth	•speech and non-speech •intensity variations •intonation	•adult-to-adult conversational speech	•none
just after birth	•preference of the mother speaking •preference of low-pass filtered speech •preference of mother tongue	•soothing speech •grammatically normal, conversational speech with low intensity •many self-repetitions	•crying •vegetative sounds
first eight weeks	•all kinds of phonetic feature discrimination	•grammatically normal, conversational speech with low intensity •imitations of interrupted phonations •many self-repetitions	•uninterrupted (stage 1) and later on interrupted phonation with simple intonation and low intensity (stage 2)
nine to eighteen weeks	•all kinds of phonetic feature discrimination	•grammatically normal, conversational speech with high pitch •many self-repetitions	•one articulation movement with interrupted or uninterrupted phonation (stage 3)
eighteen to twenty-six weeks	•active reaction to language-specific intonation patterns at 5 months	•grammatically normal conversational speech with imitations of pitch movements and strong intensity variations •many self-repetitions	•all kinds of variations in phonation without almost any articulatory movement (stage 4)
after twenty-six weeks	•discrimination of language-specific vowel prototypes at 6 months •loss of all kinds of phonetic feature discrimination after 10 months •sensitivity for language-specific consonantal features after 10 months	•imitations of babble utterances and expansion towards existing words •low speech rate •much variation in intonation and intensity •many self-repetitions •clear pronunciation	•canonical babbling, reduplicated and variegated (stage 5)

In the next sections we will consider step by step, and throughout the first year of life, the role of the cyclic effects of the three main components (as indicated in Fig. 1) of the infant's speech developmental process. The steps coincide with the so-called 'jumps' in the infant's total development in the first year of life (Plooij & Van de Rijt-Plooij, 1989).

For each period we shall try to describe the three components in relation to each other, and to account for the findings reported in the literature. After the section concerning the last period, when the infant has started canonical babbling, we will discuss the possibility of infant speech representations being structured into syllable-sized units, and illustrate this section with some data from our own research. In the final section we shall evaluate the cyclic approach of the infant's speech developmental process, and conclude by indicating possible directions of future research.

Table 1 gives an overview of co-occurrences of speech characteristics in infant speech perception, maternal speech, and infant sound productions, together with a global indication of the periods of occurrence, as described in the following sections.

2. BEFORE AND JUST AFTER BIRTH

In the schematical display of Fig. 1 we assume a clear relationship between the three aspects: maternal speech, infant perception, and early sound production. Already before birth, as we described in the introduction, the child's auditory system registers all sorts of sound input, and, according to assumption 1 of Jusczyk (1993) quoted above, analyses speech and non-speech. This discrimination between the mother's speech and non-speech is facilitated by complementary information: the mother's speech is always combined with movements of the child's whole environment, viz. the breathing movements of the mother. The child can discriminate the mother's speech from other people's speech, apart from this breathing information, by the attenuation and filtering of this speech. Also, all registered mother sounds are accompanied by the rhythm of the mother's heart beat. This may account for the results which were found by Fifer & Moon (1989), mentioned in the introduction, who showed that newborn infants preferred to listen to their mother speaking over other female or male speakers.

The character of the mother's speech at this early stage, however, will not provide the child's auditory analyzers with very precise and clear information about prototypical aspects of the individual speech sounds of the mother tongue. The speaking style of the mother which an unborn child hears, is the normal speaking style as used in everyday adult-to-adult communication. Only exceptionally will the mother speak to her unborn child. Of course, speech to other children and toddlers may be heard, and sometimes perhaps some read aloud speech or professional, clear speech, but the main input for the child at that stage is normal, adult-to-adult conversational speech.

From our studies on speaking styles and vowel reduction (Koopmans-van Beinum, 1980) we know that in conversational speech the articulatory sloppiness of a speaker is maximal, and that contrasts between individual speech sounds are physically minimal, although sufficient in practice, because of the supplementary linguistic and situational information. In an infant that still has to learn everything about linguistic and situational functions, it is very unlikely that at birth any mental representation of segmental aspects of the mother's speech exists already. The only representations of her speech may concern global and overall statistical aspects like distributions of speech and pauses, long term average spectra (low-pass filtered) information, fundamental frequency range, and suprasegmental features like intonation contours and stress patterns. Even before birth the child may develop *"a scheme for weighting the information available through the auditory analyzers"*, as suggested by Jusczyk (1993, p. 24) in his assumption 2, although in a very global way. These are the schemes to be used by the infant just after birth, when it is involved in experiments as described by Mehler et al. (1988). It was proven then that newborn infants preferred to listen to their mother's language even when it was spoken by some other person. Results of experiments in which 3-day old infants turn out to be able to discriminate between two-syllable segments within or over word boundaries (Christophe, Dupoux & Mehler, 1993), may

also be accounted for by these overall statistical analysis and suprasegmental features. It can be concluded that maternal speech has influenced early perception in a global way.

We may assume that at birth the sound production system of the infant has not undergone much influence of the sound input. Producing sound at this stage of the infant's life is simply of life-saving interest.

3. THE FIRST EIGHT WEEKS AFTER BIRTH

As soon as the child is born, a totally new situation for mother and child begins. The mother has to accustom herself to a totally unknown new little being, that for its basic needs is completely dependent on her. And the newborn child has to learn that it is a person in itself now, separated from its protecting environment, and for its needs totally dependent on the persons in its direct environment. It has to learn that making sounds may be a good way of attracting attention of the care-givers. But making sounds will be rather counterproductive if no mutual affect system between mother and child exists and no clear intersubjective relationship develops.

Fortunately, most mother-infant pairs know by instinct how to create a basis of intersubjectivity, which is a precondition for an uncomplicated development of (speech) communication (Van der Stelt, 1993). And although many studies on infant sound perception have shown that children in the first months of life are very well equipped with capacities to discriminate between all kinds of small acoustic differences (cf. Aslin, Pisoni & Jusczyk, 1983; Jusczyk, 1986), the speaking style of the mother when speaking to the child during the first weeks after birth is almost totally aimed at soothing and comforting the infant.

Thus, the speech input provided by the mother will not contribute much to any detailed sound representation or weighting scheme at the segmental speech sound level in the infant. Since the distance between mother and infant at times is very small during these first weeks, the child learns that the safe rhythmic beat, and the voice combined with breathing movements from before birth, belong to the person that is its care-giver most of the time. The so-called social smile can be considered to be an affirmation of the growing intersubjectivity.

In order to study the relationship between the speech input of the mother and early sound productions by the infant, we scanned the speech material that has been video-recorded biweekly during the first year for six normal mother-infant pairs, from birth onwards, intended for another speech-developmental study (Koopmans-van Beinum, Jansonius-Schultheiss & Van der Stelt, 1990). The speaking style of all six mothers during these first weeks is the grammatically normal, conversational style, but with low intensity and few variations of suprasegmentals, intensity and intonation.

If the sound productions of the infant in the first year of life are described on the basis of the source-filter model (Koopmans-van Beinum & Van der Stelt, 1986; Koopmans-van Beinum, 1990), we find that with respect to the sound productions during these first weeks, apart from crying, the comfort sounds are almost exclusively laryngeal, on the basis of phonation only, without any supralaryngeal, articulatory movement (stage 1; for stages see also Table 1).

Although the first sounds may be reflexive in character, the infant will be able to stop and start phonation within one breath unit (stage 2), after about six weeks. It then succeeds in producing small rhythmic series of glottal sounds, initially without much intonational or intensity variation. If rhythm (or variation in sound intensity) is a primary sensation to an unborn child, and if after birth the child perceives the auditory feedback of its own sound productions, we may meet here a first relationship between maternal input and early infant sound productions.

Nevertheless, most mothers react only incidentally to this stop-and-start activity during phonation and the resulting sound series, by imitating them. Normally they keep on speaking fluently, in a low voice, in grammatically normal, conversational speaking style, although with many repetitions. The function of her speaking to the infant still seems to be mainly one of creating and preserving intersubjectivity.

4. FROM NINE TO EIGHTEEN WEEKS

Within the developing speech communication system of mother and child, the child seems to have the first move. This is quite evident at the age when the child starts making articulatory movements in combination with the phonation movements acquired earlier (stage 3), and since more complex intonation patterns are added, the resulting sound productions present slightly speech-like features and evoke a difference in the reactions of most of the mothers. Apart from a few segmental imitations, the most striking reactions consist in a strong raising of the pitch of the mother's voice. Only one of the six mothers mentioned above, did not raise the pitch of her voice when speaking to her child in this stage, but she did not often react by speaking either, before her child started canonical babbling. Without going into the problem of any cause or effect here, it is notable that this child had pronunciation problems at the age of two years (and even at six, after three years of speech therapy).

The mothers seem to be well aware of the fact that the content of what they say is not important at all. They do not try to speak clearly and to articulate carefully. On the contrary, the exaggerated way in which they raise their pitch is more likely to make their speech less intelligible. Moreover, mothers and infants at this age are still inclined to vocalize together, in unison (Van der Stelt, 1993), again with many self-repetitions. Although our systematic analyses of characteristics of the mothers' speech in relation to the type of infant sound productions are still in progress, we think that, at this stage, mothers are still directed towards attracting their infant's attention and towards preserving intersubjectivity. However, the means they use are adapted to the sound production capacities of the infants, much more than to the infant's speech perception capacities, which have been shown to be considerable with respect to discriminating all kinds of phonetic features.

Thusfar, no clear evidence can be given for any specific effect of the maternal speech on the infant's perception at this age.

5. FROM EIGHTEEN TO TWENTY-SIX WEEKS

After a period of much articulatory activity (often indicated as 'cooing', cf. Stark, 1980), when the infants extend their repertory of speech movements from back to front, a new stage sets in, this one with far less articulations. Although previous types of sound productions do not disappear completely, the infant now is strongly focused on all kinds of phonatory, laryngeal variations (often indicated as 'vocal play', cf. Stark, 1980), like squealing, yelling, growling, utterances with very long durations and many variations in intonation and intensity (stage 4).

So, after experiencing all 'filter' aspects, various 'source' aspects have to be trained, before both can be combined in the next stage. With respect to speech development this phonatory training is often left out of consideration, but it is quite plausible that it is one of the building blocks of speech, since long stretches of sound productions with much suprasegmental variation are basic in adult speech. Mothers, in this period, react to these phonatory variations by excessively applying comparable variations to their grammatically normal fluent speech when speaking to their infants.

With respect to the speech perception of the infants, their sound productions, and the maternal input, a relationship could now be demonstrated for the first time between intonational features of the mothers' utterances and subsequent infant vocalizations (Masataka, 1992). In Masataka's study, vocalizations of ten Japanese infants during the first five months are collected together with the immediately preceding mother utterances. Discriminant analyses showed a significant relationship between the intonation patterns of the mothers' and the infants' utterances, this becoming more clear as the infants grew older. As far as we know these are the first indications that infants try to imitate or adopt (suprasegmental) aspects of the mother tongue.

6. AFTER ABOUT TWENTY-SIX WEEKS

In a number of recent studies concerning the perception of vowels by young infants, it has been demonstrated that infants at the age of six months, when listening to vowel sounds, already display sensitivity towards the specific influence of their mother tongue (Kuhl et al., 1992). It is suggested that at this age infants possess an internal, mental representation of the vowel categories of their mother tongue. Just like in adults, the discrimination of synthetic vowel sounds around vowel 'prototypes' is less than around 'non-prototypes', indicating a 'perceptual magnet effect' of the language-specific prototypes. Differences in the reactions of English- and Swedish-learning infants on identical /i/- and /y/-like stimuli are accounted for on the basis of the language-specific vowel categories that the infants would have internalized.

It is not surprising that, after suprasegmental aspects, vowels are first in displaying language-specific influences in speech development, because of their longer duration, their periodic character, and their increased acoustic intensity. Nevertheless, quite a number of questions arise with respect to these results. Especially when we consider the maternal input thusfar, as described in the preceding sections, we may wonder how it may ever be possible that infants at that age have destilled a representation of vowel prototypes of the language environment into which they were born. The acoustic parameters of the prototypes in the experiments of Kuhl (1992) are derived from the best identified *male* vowels of the study of Peterson and Barney (1952). Why should infants use the acoustic parameters of clearly pronounced male vowels to build their internal representation of the vowel categories of their mother tongue? Most of the speech the infants could have registered and weighted so far, consisted of fluently spoken, grammatically normal conversational speech by a female voice, often with a very high pitch. Instead of clear vowel information they were presented with a large degree of vowel reduction. It is only after about six months that the infants' sound productions change dramatically. This is when canonical babbling sets in, which involves more turntaking behaviour and more imitations by the mothers (or parents), who expand the babble utterances of their children into the direction of existing words in the mother tongue.

Davis and Lindblom (1992) compared /i/ and /I/ vowel realizations of one mother in three conditions: when speaking to her six months old child, when speaking to an adult, and when producing the vowels in a /h/-/t/ context as has been done in the Peterson & Barney (1952) study. Although it is not clear from the Davis and Lindblom study in which speech-developmental stage the child was at that time, they found a number of 'prototypical' (female) vowel realizations in the speech of the mother when she was speaking to the child. But apart from that, a large variability in vowel realizations was found as well. This variability was slightly larger in the adult-directed speech than in the child-directed speech, which may point to the mother's application of more contrasts when speaking to her child. However, the authors also mention the high

pitch of the mother, causing a disturbing factor with respect to reliable formant analyses.

Other questions concerning the results found by Kuhl (1992) can be raised with respect to the type of stimuli, as they are synthetic vowel sounds, only with low F1. And what is the impact of the results found by Lacerda (1992) who demonstrated that infants were able to discriminate along the high/low (F1) dimension, but not along the front/back (F2) dimension? The results of Kuhl (1992), however, seem to be symmetric around the prototypical vowel.

With respect to the infant's sound production it is plausible that first a period is needed in which articulatory movements are trained, and subsequently a period of training all phonatory movements, before both can be combined in canonical babbling (stage 5), involving all basic aspects of adult speech. Mean utterance duration of canonical babbling (an utterance being defined as the sound production during one breath unit), is comparable with the mean duration of utterances that are typical of the previous, phonatory stage, although the range is much larger there: even sound productions of more than 6 seconds within one breath unit may occur in that stage!

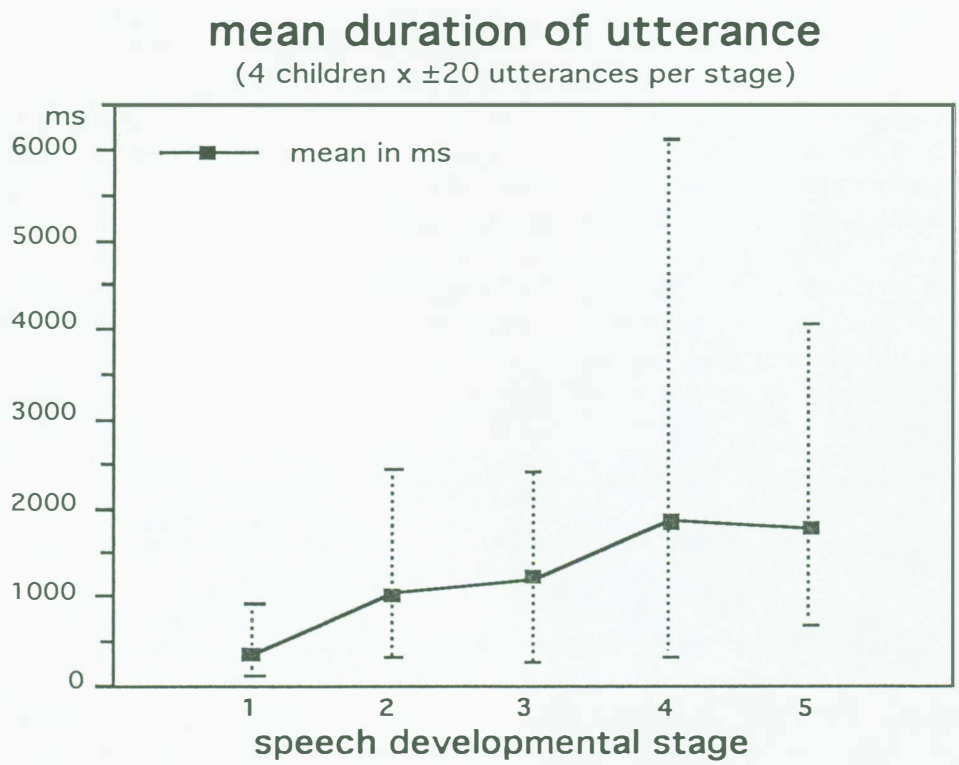

Fig. 2. Mean utterance duration and range (in ms) for utterances typical for five stages of speech sound development. The first twenty typical utterances (if present) of each stage of two male and two female infants are used.

Fig. 2 displays mean utterance duration of utterances being representative for the first five stages of speech sound development. These data are based on the first twenty typical utterances (if present) of each stage of two male and two female infants. The boys are the same as described by Koopmans-van Beinum & Van der Stelt, 1986; one of the girls has been mentioned in that study as well; the second girl has been added to this early infant sound data-base later on.

Most infants start canonical babbling, the production of long series of repetitive articulatory movements, at an age of about seven months, but the range in age in

normal hearing infants is found to be from 18 to 48 weeks (Van der Stelt & Koopmans-van Beinum, 1986). Other alternating movements like hammering, chewing, and thumping are mastered by the infant at about the same time. However, since the start of this babbling is so evident, and the sound productions are so speech-like, with clear syllabic structure, parents and other care-givers immediately feel invited to react as if the child utters meaningful speech now. As said above parents (fathers as well) start to imitate the babbling utterances, expanding them into the direction of existing words in their language like *papa*, *dada*, although initially the infant makes nothing else than series of alternating opening-closing movements, resulting in mainly accidental constrictions somewhere in the vocal tract.

Alternating sound production becomes more and more part of the game, and the speaking style of the parents includes many isolated words and short, clear sentences now, often repeated, providing the child with information about segmental aspects of speech, and with the acoustic characteristics of inter- and intra-speaker variability. Thus, from this time onwards, the child will learn to generalize or to normalize for speaker, for speaking style, and for contextual situation. This is a basic prerequisite for the acquisition of a specific language. At this stage the assumptions 3 to 6 mentioned above, as proposed by Jusczyk (1993), come into the picture. Now it becomes possible for the child to learn about the typical sound structure of the words in its native language. It is provided with more and more information to extract patterns which can be used for segmenting words from the speech stream.

Since the articulatory movements of the child gradually lose their arbitrariness, and since variegated babbling also occurs apart from repetitive babbling, the child gradually obtains the opportunity to tune in on the specific features of its mother tongue. At about ten months, the vocalic sound productions of the child become language specific (De Boysson-Bardies et al., 1989) and at about 10 to 12 months children will have lost the ability to discriminate all phonetic contrasts, and only preserve the discrimination ability necessary to perceive the contrasts in their mother tongue (Werker & Lalonde, 1988). Slight evidence of language-specific consonant-like sound productions in the period from babbling to one-word utterances (from 11 to 18 months) is provided by De Boysson-Bardies & Vihman (1991) and by Levitt and Utman (1992).

7. REPRESENTATIONS STRUCTURED INTO SYLLABLE-SIZED UNITS

One of the assumptions in Jusczyk's WRAPSA-model concerns the representations provided by the pattern extraction process to be structured into syllable-sized units (Jusczyk, 1993). Without going into a discussion about phoneme-sized or syllable-sized representations, we think that in our cyclic model of early speech development, syllable-sized units may be more useful in clarifying the relationship between early speech perception, infant sound production, and maternal speech input. Although Jusczyk's assumption concerns representation of input speech, we want to focus here on the syllabic structure of infant speech production.

In our study (in progress) concerning the acoustic characteristics of the sound productions typical for each of the five prelexical speech developmental stages mentioned above, we concentrated on syllable-like components within the utterances of two boys and two girls, audio-recorded weekly from birth onwards until about nine months of age (see above). For each of the five stages for each of the four children the first twenty representative utterances were selected and acoustically analysed (see also Fig. 2). Subsequently, syllable-like segments were determined on the basis of visual inspection of the wave-form, as well as by listening to the utterances. It should be kept in mind that when a new stage sets in, previous types of utterances are often present as well.

Table 2. Overview of listeners' agreement in indicating the number of syllables in utterances of five prelexical stages of one boy (number of utterances per stage = 20, number of listeners = 10).

	listeners' agreement	mean number of syllables per utterance
stage 1	89 %	1.1
stage 2	96 %	2.4
stage 3	84 %	2.2
stage 4	78 %	1.4
stage 5	79 %	3.8
overall	85 %	

Thus, the given data are not representative for all utterances produced during a specific stage, but only for the highest level utterances representative for the new stage.

Since it was not clear in advance, whether listeners would agree in their perception of a syllable in early infant sound productions, we presented 100 utterances (20 per stage) made by one of the boys, in random order to 10 listeners, who were asked to simply indicate for each utterance the number of syllables they believed to hear. Total agreement over the 10 listeners was 85%, but agreement per stage differed as represented in Table 2.

We may conclude that in spite of the absence of a linguistic content, listeners are well aware of the syllabic-like structure of early infant sound productions. The disagreements, apart from utterances in stage 4, almost always concerned a difference of one syllable, caused by what could be called an upbeat, or simply concerned a counting problem in the case of long utterances with many articulation movements.

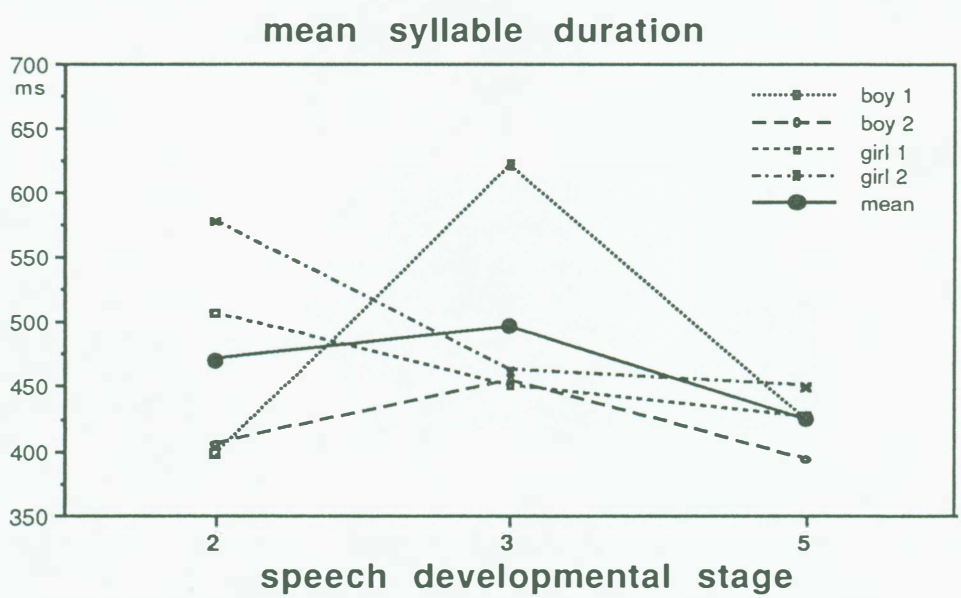

Fig. 3. Mean syllable durations (without silence) for utterances of each of the four children (see text), and average values calculated over the four children, for the three more-syllabic stages (see Table 1).

Utterances of stage 4 caused problems for all listeners because of an almost total absence of articulatory movements, combined with intonational and intensity variations, which made an unambiguous response almost impossible.

Mean syllable durations (without silence) for utterances of each of the four children, and average values over the four children, for the three more-syllabic stages, are given in Fig. 3. Initially, large differences exist between the mean values of syllable duration, but as soon as the stage of canonical babbling is reached, the mean durations are more or less equal for the four children, with a mean value of 424 ms, or an articulation rate of 2.36 syll./sec. However, the relative slowness of the production system at this age may be illustrated by the fact that this value is still about twice as long as the mean syllable duration given by De Boysson-Bardies et al. (1981), for a French child at the age of one and a half year. Den Os (1990) studying temporal properties in the speech of one Dutch child between one and three years of age, reports a mean articulation rate of about 3 syll./sec. For a Dutch professional adult speaker we obtained articulation rate values of 6.44 syll./sec. for read aloud speech and 6.71 syll./sec. for conversational speech (Koopmans-van Beinum, 1992). Indeed, mothers tend to slow down their speech in combination with speaking more clearly after about six months. However, since infants hear grammatically normal conversational speech during the first six months after birth, they will have to learn to normalize for speech rate as well. Parents' imitation of their baby's canonical babbling and their expansion of the utterances into the direction of existing words, will once again be of great help in this essential generalization or normalization process.

8. CONCLUDING REMARKS

We are well aware that our description of the cyclic effects in the infant speech developmental process remains very much incomplete. Nevertheless it provides us with the possibility to consider the main components in relation to each other and to formulate our research questions within a broader scope than the sub-area of either infant speech perception, early sound production, or 'motherese'.

It is striking that the earliest evidence of influence of the mother tongue in the infants' sound productions is found with respect to intonation, (between 5 and 10 months: Matasaka, 1992; Whalen, Levitt & Wang, 1991). Next, this evidence is found with respect to vocalic productions (between 5 and 10 months) and finally in consonantal productions (between 11 and 18 months). This yields the same order that was found for perceptual evidence of the influence of the mother tongue in infants. Although new research data should support the suggested cyclic model, sofar the results seem to affirm our theory concerning a relationship between infant speech perception, early sound production, and maternal speech.

However, many questions still remain. The role of the infant's auditory feedback is not very clear yet, since from the literature as well as from research in our own institute it turns out that severely hearing-impaired infants do not produce less vocalizations than normal hearing infants until the stage where normal hearing children start babbling (Clement & Den Os, 1993). Are mothers of severely hearing-impaired infants reacting to their infant's sound productions in another way than mothers of normal hearing infants? Is there indeed a systematic and permanent adaptation or development in the speech of mothers to their infants from birth onwards, concurrent with the development in the sound productions of the child and its perceptual capacities? Our ongoing research focuses on these problems, as well as on questions concerning the acoustic structure of the subsequent stages in early sound production. The problem of how children learn to master the basic elements of speech communication remains a phonetician's challenge in the children's corner.

9. REFERENCES

Ãslin, R.N., Pisoni, D.B. & Jusczyk, P.W. (1983): "Auditory development and speech perception in infancy". In: M.M. Haith & J.J. Campos (Eds.), *Infancy and developmental psychobiology*, Vol 2. Carmichael's handbook of child psychology (4th edition). New York: Wiley.

Bickley, C., Lindblom, B. & Roug, L. (1986): "Acoustic measurements of rhythm in infants' babbling or "All God's children got rhythm"", *Proceedings of the 12th International Congress on Acoustics*, Toronto, July 1986.

Bloom, K. (1988): "Quality of adult vocalizations affects the quality of infant vocalizations", *Journal of Child Lang*uage **15**: 469-480.

Christophe, A., Dupoux, E. & Mehler, J. (1993): "How do infants extract words from the speech stream? A discussion of the boothstrapping problem for lexical acquisition". In: E.V. Clark (Ed.), *The proceedings of the twenty-fourth annual Child Language Research Forum*, Stanford, 209-224.

Clement, C.J. & Den Os, E.A. (1993): "Development of vocalizations of severely hearing-impaired infants", Presentation at the *Third Congress of the International Clinical Phonetics and Linguistics Association, 9-11 August 1993.*

Davis, B. & Lindblom, B. (1992): "Prototypical vowel information in baby talk", *PERILUS* **15**: 119-124.

De Boysson-Bardies, B., Bacri, N., Sagart, L., Poizat, M. (1981): "Timing in late babbling". *Journal of Child Language* **8**: 525-539.

De Boysson-Bardies, B., Hallé, P., Sagart, L. & Durand, C. (1989): "A cross-linguistic investigation of vowel formants in babbling", *Journal of Child Lang*uage **16**: 1-17.

De Boysson-Bardies, B. & Vihman, M.M. (1991): "Adaptation to language: evidence from babbling and first words in four languages", *Language* **67**: 297-319.

Den Os, E.A. (1990): "Development of temporal properties in the speech of one child between one and three years of age", *Proceedings of the Institute of Phonetic Sciences Amsterdam* **14**: 39-52.

Fernald, A. (1992): "Human maternal vocalizations to infants as biologically relevant signals. An evolutionary perspective". In: J.H. Barkow, L. Cosmides & J. Tooby (Eds.), *The adapted mind*. New York: Oxford University Press, 391-428.

Fifer, W. & Moon, C. (1989): "Early voice discrimination". In: C. von Euler, H. Frossberg & H. Lagencrants (Eds.), *Neurobiology of early infant behaviour*. New York: Stockton Press, 277-286.

Jusczyk, P.W. (1986): "Toward a model of the development of speech perception". In: J.S. Perkell & D.H. Klatt (Eds.), *Invariance and variability in speech processes.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1-19.

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

Koopmans-van Beinum, F.J. (1980): *Vowel contrast reduction. An acoustic and perceptual study of Dutch vowels in various speech conditions.* Ph. D.-thesis, University of Amsterdam.

Koopmans-van Beinum, F.J. (1990): "The source-filter model of speech production applied to early speech development", *Proceedings of the International Conference of Spoken Language Processing 1990*, Kobe, Japan, Vol. 2, 1321-1324.

Koopmans-van Beinum, F.J. (1992): "The role of focus words in natural and in synthetic continuous speech: Acoustic aspects", *Speech Communication* **11**: 439-452.

Koopmans-van Beinum, F.J., Jansonius-Schultheiss, K. & Van der Stelt, J.M. (1990): "De invloed van een gehemelteplaatje op de spraakontwikkeling en -interactie in de eerste levensjaren van baby's met een gehemeltespleet", Final Report to The Netherlands Prevention Fund. *Report 110 of the Institute of Phonetic Sciences*, Amsterdam.

Koopmans-van Beinum, F.J. & Stelt, J.M. van der (1986): "Early stages in the development of speech movements". In: B. Lindblom & R. Zetterström (Eds.), *Precursors of Early Speech*. Wenner Gren. Int. Symp. Series 44, New York: Stockton Press, 37-50.

Kuhl, P. K. (1992): "Speech prototypes: studies on the nature, function, ontogeny and phylogeny of the 'centers' of speech categories". In: Y. Tohkura, E. Vatikiotis-Bateson & Y. Sagisaka (Eds.), *Speech perception, production and linguistic structure*. Tokyo: Ohmsha, 239-264.

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

Lacerda, F. (1992): "Young infants prefer high/low contrasts", *PERILUS* **15**: 85-90.

Levitt, A.G. & Utman, J.G.A. (1992): "From babbling towards the sound system of English and French: a longitudinal two-case study", *Journal of Child Language* **19**: 19-40.

Locke, J.L. (1990): "Structure and stimulation in the ontogeny of spoken language", *Developmental Psychobiology* **23**: 621-643.

Matasaka, N. (1992): "Early ontogeny of vocal behavior of Japanese infants in response to maternal speech", *Child Development* **63**: 1177-1185.

Mehler, J., Jusczyk, P., Lambertz, G., Halstead, N., Bertoncini, J. & Amiel-Tison, C. (1988): "A precursor of language in young children", *Cognition* **29**: 143-178.

Moon, C. & Fifer, W.P. (1986): "Newborns prefer the sound of mother's voice as experienced in the womb". Paper presented at the *International Society for Developmental Psychobiology, Annual Meeting*, Annapolis, November.

Oller, D.K. (1986): "Metaphonology and infant vocalizations". In: B. Lindblom & R. Zetterström (Eds.), *Precursors of Early Speech.* Wenner Gren. Int. Symp. Series 44, New York: Stockton Press, 21-35.

Penfield, W. & Roberts, L. (1959): *Speech and brain-mechanisms*. Princeton, N.J.: Princeton University Press.

Peterson, G.E. & Barney, H.L. (1952): "Control methods used in a study of the vowels", J. Acoust. Soc. Am. **24**: 175-184.

Plooij, F.X. & Van der Rijt-Plooij, H.H.C. (1989): "Vulnerable periods during infancy", *Ethology and Sociobiology* **10**: 279-296.

Stark, R.E. (1980): "Stages of speech development in the first year of life". In: G.H. Yeni-Komshian, J. Kavanagh & C.A. Ferguson (Eds.), *Child Phonology: Vol. I: Production*. New York: Academic Press, 73-92.

Van der Stelt, J.M. (1993): *Finally a word: a sensori-motor approach of the mother-infant system in its development towards speech.* Ph.D.-thesis, University of Amsterdam.

Van der Stelt, J.M. & Koopmans-van Beinum, F.J. (1986): "The onset of babbling related to gross motor development". In: B. Lindblom & R. Zetterström (Eds.), *Precursors of Early Speech.* Wenner Gren. Int. Symp. Series 44, New York: Stockton Press, 163-173.

Werker, J.F. & Lalonde, C.E. (1988): "Cross-language speech perception: Initial capabilities and developmental change", *Developmental Psychology* **24** (5): 672-683.

Whalen, D.H., Levitt, A.G. & Wang, Q. (1991): "Intonational differences between the reduplicative babbling of French- and English-learning infants", *Journal of Child Language* **18**: 501-516.