Kjellrun T. Englund

Characteristics of and Preference for Infant Directed Speech

Norwegian University of Science and Technology

Doctoral thesis for the degree of philosophiae doctor 2005

Faculty of Social Sciences and Technology Management,

Department of Psychology

Elektronisk: ISBN 82-471-7168-6

Trykt: ISBN 82-471-7169-4

CONTENTS:

ACKNOWLEDGEMENTS

Many people have contributed to the work on completing this thesis. Starting chronologically, I owe my gratitude to the late Leif Stinessen, whose enthusiasm and interest in cognitive psychology and psychology of language inspired my interest in research.

There is one person without whom the current research would never have been conducted. I wish to express my deepest gratitude to my supervisor, Dawn Behne, who is secondary author on three of the four papers included in the current work. Her open and interested mind and broad knowledge in several fields as well as endless support has guided and inspired me through the whole process with completing this research. I also thank my secondary supervisor, Audrey van der Meer for valuable guidance on how to conduct perception experiments with infants.

Further, I wish to thank the infants and their parents who participated as subjects in the current research, who were kind enough to let me collect the data in their homes. In addition, The Midwife Service in Trondheim offered support throughout the recruitment process for the production study. I also wish to thank the students who have been involved in the current research. In addition to Tine Stokke and Linda Rydningen, Helga Valør Wist was central in collecting data, and together with her son Herman she was involved in making the stimulus material for the perception study. In addition, I

owe my gratitude to Rein Ove Sikveland for help with acoustic measurements for the production study.

I also wish to express gratitude to my colleagues at the Department of Psychology for supporting the project, especially to my fellow doctoral students. Special thanks to Nanna Kayed for kind support throughout the research process, as well as for providing pictures for the experimental settings used in the current studies.

Last, but not least, I wish to thank my family, not for enduring long hours without me, but for reminding me that the most important thing in life throughout any process is to treasure the people close to you. I therefore dedicate the current thesis to my family.

ABSTRACT

Infant directed speech (IDS) has different characteristics than adult directed speech (ADS). Studies of acoustic-phonetic aspects of IDS show, by measures of the first two formant frequencies, that vowels have more extreme articulations in IDS than ADS (Burnham, Kitamura & Vollmer-Conna, 2002; Kuhl, Andruski, Chistovich, Chistovich, Kozhevnikova, Ryskina, Stolyarova, Sundberg, Lacerda, 1997; Bernstein Ratner, 1984). Research has shown that stop consonants are articulated with a shorter VOT in IDS than in ADS (Sundberg & Lacerda, 1999; Malsheen, 1980). In addition, Sundberg (1998) propose that IDS contains adaptations to the infant's development. An experimental setting may not convey natural IDS. In addition, no studies have been made of the acoustic-phonetic content of Norwegian IDS. In a longitudinal production study using a natural interactive setting with 10 recordings per infant over a period of six months, recordings were made of mothers alone with their infants (IDS) and in adult conversations (ADS). Both to accommodate the families that participated in the study and to secure a natural setting for the productions, both IDS and ADS recordings were made in the participants' homes. An IDS recording was typically made in the bathroom when the mother changed infants' nappies and a typical ADS recording was made in the living room during a conversation between the mother and the experimenter. From words within these productions, segments were selected from both ADS and IDS recordings at each point in time, and acoustic analyses carried out for the vowels /a:/, /i:/ and /u:/, and their short counterparts /a/ /i/ and /u/, the voiced stop consonants /b/, /d/, /g/ and the voiceless /p/, /t/, /k/, as well as for the fricative /s/. Statistical analyses were based both on data for each segment that was merged across time points and on

analyses of the change in acoustic-phonetic attributes of segments between different points in time.

Repeated measures analyses show Norwegian IDS from a natural setting to be different than results from comparable previous research. In Norwegian IDS, formant frequencies in back vowels show that articulation of the back vowels /a:/, /a/, /u:/ and /u/ are articulated more front compared to ADS, while the front vowel quality /i:/-/i/ does not differ between speech types. For vowels, only the difference in duration between IDS and ADS changed over time, but duration was nonetheless longer in IDS than ADS throughout the six months. In addition, also contrary to previous research (Sundberg & Lacerda, 1999), voice onset time (VOT) in Norwegian IDS is longer than in ADS for alveolar and velar voiced and unvoiced stops as well as an unvoiced bilabial stop, but the voiced bilabial stop /p/ has a similar VOT across speech types.

The results from the production study show that Norwegian IDS is different from Norwegian ADS, possibly corresponding to an enhancement of speech to infants. Further, in addition to enhancing auditory cues to the infant, IDS could contain specifications that also enhance visual cues to the infant.

To address this as a possible explanation for the results from Norwegian IDS, a perception study was carried out with 23 infants from 2-8 months old. Infants' perception of audio-visually matching (congruent) or audio-visually mismatching (incongruent) syllables in ADS and IDS was tested to see whether the possible visual enhancement in IDS is important for perception of audio-visual processing. Enabling

testing of infants in their homes, an adapted method with a visual fixation task was used. To obtain visual fixation time, infants were video-recorded while they were watching a video where a woman was speaking to her infant and to an adult. She articulated the CV-syllables (/pi:/, /pa:/ and /pu:/). The syllables were analysed acoustically to ensure that the vowels /a:, i:, and u:/ had the same formant frequencies for F1, F2 and F3 as the mean formant frequencies resulting from the production study. An auditory syllable (/pa:/) was either paired with the corresponding visual syllable (/pa:/) or one of the other visual syllables (/pi:/ or /pu:/) produced in the same speech type. This audio-visual matching was done for both ADS and IDS. Infants watched the resulting eighteen trials, and their visual fixation time was logged. Results from a repeated measures analysis showed that infants from 2 to 8 months old looked longer at IDS than ADS but unexpectedly longer at incongruent than congruent syllables in the study. While it is uncertain why incongruent syllables are preferred to congruent syllables, infants possibly prefer audio-visual IDS to ADS due to both auditory and visual characteristics of IDS vowels. By investigating production and perception of IDS in natural settings, the current research may be a step forth in uncovering how the characteristics of IDS may be connected to infants' perception of speech.

AIMS AND OUTLINE

This thesis studies infant directed speech (IDS), and the aims of the current thesis are threefold. The first aim is to contribute to uncovering the acoustic-phonetic characteristics of vowels and consonants in IDS. The second aim is to investigate the possible changes in acoustic-phonetic characteristics of IDS over time, more precisely during the first six months of an infant's life. The third aim is to contribute to the research on perception of IDS, by investigating audio-visual perception of IDS compared to adult directed speech (ADS).

There is a need for studies of the acoustic-phonetic characteristics of Norwegian IDS, and to accomplish the above aims a longitudinal study of Norwegian mothers' IDS to their infants was conducted together with a perception experiment of 23 infants. A common goal for the components of the current research has been to ensure natural settings, both for the infants and parents in the mother-infant dyads in the production study as well as for the infants and their parents participating in the perception experiment.

Four papers are included in the thesis. The first reports acoustic-phonetic characteristics of Norwegian IDS vowels. The second reports acoustic-phonetic characteristics of Norwegian IDS stops and their development over an infant's first six months. The third reports how acoustic-phonetic characteristics in Norwegian IDS vowels, as well as in a Norwegian fricative, develop during a child's first six months. The first three papers lay

the foundation for the fourth paper, which investigates infants' perception of audio-visual syllables, reporting how congruent vs. incongruent audio-visual IDS and ADS syllables are perceived by infants.

The introductory chapter of this thesis will first provide an overview of speech perception by infants showing that infants discriminate speech segments well at an early age. Following this will be theoretical accounts of speech perceptual development and IDS. After this, research is presented showing differences in the characteristics of IDS and ADS. This will be followed by an introduction to some methodological issues. Based on the introduction and results from the four papers, rounding off the thesis will be a final discussion with concluding remarks about characteristics of speech directed at infants and perception of IDS by infants, as well as suggestions for future research.

INTRODUCTION

One of the main ways in which human beings communicate with each other is through speech. Speech may be generally defined as sound made by the human speech organs in order to convey a message. This message is conveyed differently by and to different people. In the analects of Confucius (900 BC) there is a principle for how to use speech, saying that "*not to speak with one who can be spoken with is to waste a person and to speak with one who cannot be spoken with is to waste one's words"* (quoted in Chang, 1997). Although this is probably intended to refer to communication between adults, applying this principle of Confucianism to the kind of speech we use with a small infant, speech would be considered a waste, because an infant cannot answer in words. This is close to one of the traditional views of infant directed speech (IDS), where the reason why it may be seen as a waste is that to an infant IDS is not of more use than is adult directed speech (ADS) (Pinker, 1994). From another angle IDS can be seen as useful to an infant's language development, and the description of what has been termed parentese, baby talk, motherese, fatherese, child-, or infant directed speech, has covered semantic, syntactic, phonological as well as phonetic studies (reviewed in Gallaway & Richards, 1994).

In order for communication to be successful humans depend upon both speech perception and speech production. When an infant acquires the skill of speech communication, he/she will need both. Regardless, for a small infant, speech communication may start with perceiving the world around (Vihman, 1996; Haynes and Shulman, 1998).

How does an infant come to know the world of language?

Potential answers to this question are related to theoretical accounts of development as based either in nativism or learning theory. A fundamental discussion between Skinner (1957) and Chomsky (1959) was centred on language as either being innate or learned. In Chomsky's nativist views (Chomsky, 1959; 1965), language unfolds itself as a complex system and each child comes into the world with a set of rules in which the ambient language should fit. Language acquisition is viewed as closely related to maturation and shows similar developmental milestones for children around the world (Shulman, 1998a). From this perspective, the input infants receive is seen as inadequate compared to the complex system it eventually results in, and children must therefore have innate linguistic abilities to acquire language. On the other hand, the behaviourist perspective, fronted by B.F. Skinner, look upon language as being learned through experience and by general learning principles (Skinner, 1957). In his book 'Verbal Behavior', he equals language use with other kinds of human behaviour, claiming that all learning is accomplished through selective reinforcement from people around the child. Other branches of learning theory (Bandura, 1989) have added imitation learning to learning theory, believing the child is imitating the language from parents and others in the immediate environment. Although the dichotomous views in their original form no longer form the sole basis for research on language development, the discussion of infant perceptual and cognitive capacities is strongly coloured by the original debate (Spelke, 2000).

It is not the intention of this thesis to go deeply into the debate about whether language is innate or learned. The dispute between the two theoretical points on language

development has nevertheless been basis for the detailed descriptions of the speech of mothers to their babies, because showing that the characteristics of IDS is adapted to the child's linguistic needs (Snow, 1972; 1977) may increase the likelihood that the ambient language is used when the child develop language.

From nativist and learning viewpoints, the interactionist approach has emerged. This theoretical stance does not deny an innate language system, but the role of learning in language development is recognised as important. An infant's urge to communicate might be what is innate, and language is acquired just like any other skill. One of the most recognised early interactionists was Jerome Bruner (1982). One of the processes he proposed as part of infants' language development was a system governing the language interaction between parent and infant. This system was called the language acquisition support system (LASS) (Bruner, 1982). In essence, this, and other interactionist views (Shulman, 1998b), holds that while humans are born with prerequisites to learn language, input is necessary in order for this to happen. Through the LASS, daily routines with children might help the language learning process by providing a framework for discovering the functionality of language. By saying this, Bruner (1982) presupposes that the LASS makes an adult adjust language to a child's cognitive level. A more general interactionist view on development is Gottlieb's epigenic theory (Gottlieb, 1991a, b), which describes how genes and environment interact for development to come about. The epigenic view builds on the idea that there are bidirectional interactions between brain and behavioural development. The epigenic view is related to systems theory which analyses behaviour on multiple levels, where no level has priority over another. In the vast range that Gottlieb's work covers, he set

forward the idea that imprinting in ducklings was not an all-visual phenomenon: a species-specific maternal call made imprinting in ducklings stronger. Interestingly, he also observed that ducklings would follow the maternal call of their species without having been previously exposed to it (Gottlieb, 1961, 1968). Although this can be attributed ducklings' instinct to follow the maternal call, one may ask if these ducklings had had any auditory perceptual experience before hatching. The same question could be asked for human foetuses.

When does speech perception start?

In speech perception, what we hear may be essential, but when does hearing first begin? After five and a half months gestation, the cochlea is sufficiently developed to enable auditory perception, and the neural structures necessary for processing speech are present (Reviewed in Kent, 1997). Although well developed, a neurologically and physiologically mature auditory system does not necessarily mean a foetus can hear what is happening inside or outside the womb. As recent research has shown (Printer, Freeman, Perez & Sohmer, 2003), since a foetus' middle ear is filled with fluid, although improving with the absorption of amniotic fluid by the tissue, hearing must be much reduced before birth (Printer et al., 2003).

Still, attempts to show foetal response to auditory stimuli have proven successful. In a classical study, Shahidullah and Hepper (1993) tested foetal motor activity at 20 and 25 weeks of gestation by ultrasound while a loudspeaker on the mother's abdomen presented pure-tone stimuli. Any observable movement by the foetus counted as a

response. Foetuses responded to acoustic stimuli at 20 weeks, albeit with few movements. At 25 weeks, the responses had become higher in number and foetuses responded immediately and clearly to the auditory stimulation. An additional study (Shahidullah & Hepper, 1994) showed that foetuses discriminated between 250 Hz and 500 Hz stimuli at 27 and 35 weeks of gestation. A study by Hepper and Shahidullah (1994) showed that foetuses at 27 weeks gestational age responded by movement to 250 and 500 Hz tones but not to 1000 and 3000 Hz tones. Moreover, all foetuses responded to these last two frequencies by 33 and 35 weeks of gestation. This study pointed out that the intensity level required for a response decreased as the foetus matured. Together, these studies support that infants are capable of perceiving auditory signals in the womb.

Although responded to, is the sound a foetus perceives similar to what it perceives outside the womb? While it may seem unlikely that much of another person's voice reaches the foetus' auditory system, the mother's voice is probably heard, and foetal responses to a mother's voice has been observed both in bird embryos (Lecanuet & Granier-Deferre, 1993), and in human foetuses (Spence & DeCasper, 1982). The sound from the mother's vocal apparatus must travel through tissue, bone mass and amniotic fluid to reach the foetus' ear. Therefore, it may be very different from what an infant will encounter after being born. Still, if a newborn infant responds differently to its mother's voice compared to other women's voices after birth, it is a good indication that the infant has perceived some aspect of sound before birth. Interestingly, infant preference for the maternal voice to other women's voices has been observed (DeCasper & Fifer, 1980). This implies that human foetuses can perceive their mothers' voices

despite the quality of this sound likely being very different from what they encounter after they are born.

What in the mother's voice is perceived by an infant's response? To test if the answer to this question could be fundamental frequency, Spence and Freeman (1996) tested infants' (14-73 hours old) responses to low pass filtered maternal voice and maternal whispered speech. Low pass filtering below 500 Hz excludes linguistic information that is conveyed mostly by high frequencies, but keeps prosodic characteristics of speech (Kent, 1997). Infants responded to their mothers' low pass filtered voices, but not their mothers' whispered speech. This suggests that low frequencies in the voice may travel through bone mass and tissue, reach the foetus who perceives it, and may retain some aspect of it after birth. A foetus in the womb may therefore perceive parts of voiced sounds. This is partly confirmed in studies showing that, in addition to the fundamental frequency of the voice, characteristic voice patterns may also be picked up. Foetuses seem capable of remembering rhymes repeatedly read aloud by their mothers during the last four weeks of pregnancy shown by a decrease in heart rate compared to rhymes never read by their mothers (DeCasper, Lecanuet, Busnel, Granier-Deferre & Maugeais, 1994). In addition, newborns respond differently to a speech passage recited by the mother during the last six weeks of pregnancy than a passage they have not been exposed to during pregnancy (DeCasper & Spence, 1986).

These studies indicate a likelihood of foetal audio perception, enabling foetuses to acquire parts of voiced sounds that can be recognised after birth. If the foetus hears and retains certain aspects of voiced sounds from in the womb, how may it come in handy

after birth? In addition to the capabilities mentioned above, infants are very good at discriminating speech sounds after birth.

Speech perception after birth

Vowels

As will be reported below, several studies demonstrate well-developed discrimination of vowel contrasts by infants between six and twelve months. In addition to vowel duration, which is perceived well and may even be a particularly salient acoustic cue to infants (e.g., Bohn & Polka, 2001), vowel quality is discriminated by one-to four-month-old infants (Trehub, 1973). If a sound is perceived categorically, listeners show insensitivity to within-category differences relative to between-category differences (Ryalls, 1996). Two-to-three-month-old infants can discriminate tense /i/ versus lax /I/ as was shown in a study by Swoboda, Morse and Leavitt (1976). By using the high-amplitude sucking procedure, sixty infants were tested on their ability to distinguish pairs of vowel stimuli differing in equal steps along the /i/ to /I/ continuum. Compared to a control condition, infants perceived within and between phonetic category conditions equally well. This indicates that infants at this age may perceive vowel quality as continuous rather than categorical.

If pitch is varied (either using monotone pitch or a rise-fall) when vowels from the same category are articulated and this is done for several vowel categories, five-to-six-month olds can overlook pitch and still discriminate vowel categories (Kuhl & Miller, 1982). Speech by different talkers has also been used to see if between-talker differences affect

infant vowel perception. In a study (Kuhl, 1983), five-to-six-month-olds were trained to discriminate /a/ and /ɔ/. Following this, vowels spoken by adult men and women as well as children were presented to the infants, who independent of speaker still categorised the vowel contrast (Kuhl, 1983). To do this, an infant must be able to overlook talker differences. As a vowel articulated by different talkers can be very different physically (Peterson & Barney, 1952), this shows how some attributes of the speech signal are overlooked while others are attended to by infants.

Infants do not, however, seem to perceive all members of a vowel category equally well. When rating category goodness of variants of a vowel, adults rate some instances as more (proto) typical than others (Kuhl, 1991). Moreover, compared to adults, infants have a greater tendency to generalise to non-prototypical members of a category (Kuhl, 1991). This finding led to the introduction of the notion of a 'magnet effect' in speech perception, namely that the category goodness of speech sounds strongly influences perception of these speech sounds (Kuhl, 1991; Iverson & Kuhl, 1995). More specifically, adult American English participants rated the perceived goodness (typicality) of different exemplars of the vowel /i/. The sounds that were rated as especially good exemplars were named prototypes of the /i/ while the sounds that were rated as especially poor exemplars were named non-prototypes of the /i/. It was shown that six-month-old infants found it more difficult to discriminate prototypes from near prototypical exemplars than non-prototypes from near non-prototypical exemplars. In this way, the prototypical exemplars functioned like magnets, drawing exemplars closer to them with the result that they became difficult for infants to discriminate from the actual prototypes (Grieser & Kuhl, 1989; Kuhl, 1991). Both Swedish and American

infants have shown a magnet effect that is stronger for their native language compared to a language they are not familiar with by six months of age. This implies that language experience alters vowel perception during the first six months (Kuhl, Williams, Lacerda, Stevens & Lindblom, 1992).

Consonants

Like for vowels, consonantal contrasts are found to be well discriminated by infants. In the classical study by Eimas, Siqueland, Jusczyk & Vigorito (1971), the ability of one- and four- month olds' to discriminate consonants was tested. The between category phonetic voicing contrast represented by /ba/ and /pa/ was compared to within category differences by different /ba/s. Results demonstrated that infants could discriminate the former but not the latter, showing that the ability to discriminate members of the same stop consonant category is poor compared to the ability to discriminate members of different voicing categories. Compared to the findings by Swoboda et al. (1976), consonants could be perceived in a more categorical manner than vowels.

Infants are also capable of discriminating differences in places of articulation. Morse (1972) found that two-month olds discriminated bilabial /ba/ from velar /ga/, and Moffitt (1971) found the same with five-month-olds. The difference between the labio-dental and interdental fricatives /fa/ and /θa/ was found to be discriminated by six-month-olds (Holmberg, Morgan & Kuhl, 1977). Infants as young as two months old have also been shown to discriminate place of articulation between the bilabial and

alveolar sonorant consonants /ma/ and /na/ (Eimas & Miller, 1980), as well as /wa/ and /ja/ (Jusczyk, Copan & Thompson, 1978).

Infants also discriminate differences in manner of articulation. Two-month-old American infants can discriminate /r/ and /l/ (Eimas, 1975) in a nearly categorical manner. Another study (Eimas & Miller, 1980) showed that two- to four-month-old infants could discriminate /ba/ from /ma/, but the discrimination was not categorical, meaning that different /ba/s were discriminated equally well. Another finding by Miller and Eimas (1983) demonstrated the discrimination between a stop and glide (/ba/ and /wa/) to be categorical by eighty three-and-four-month-old infants.

Some contexts and contrasts may be easier for infants to discriminate than other contrasts. A study by Goodsit, Morse, Ver Hoeve and Cowan (1984) showed that six-month old infants detected syllables within trisyllabic sequences better if the pattern of the other two syllables were repeated than in the case of a varied pattern. A secondary finding from this study was that the syllable /ba/ seemed easier for infants to detect than /du/.

Some studies have shown that infants may have difficulties with voicing discrimination among fricatives. In three experiments, Eilers and Minifie (1975) observed that infants had some difficulties to discriminate between /s/ and /z/, but no problems with /s/ versus /v/ or /s/ versus /f/. Levitt, Jusczyk, Murray and Carden (1988) have shown that two-month-old infants discriminated voiced and voiceless pairs of fricatives, while Eilers (1977) has shown that three-month-olds have some problems discriminating a voicing

contrast in /sa/ and /za/. This indicates that very young infants have some difficulty with discriminating the voicing contrast.

However, overall, infants seem to have few difficulties with discriminating most categorical differences and discriminate most contrasts at a remarkably early age. This capacity is affected by an ambient language, so that discrimination abilities of contrasts not present in that language gradually become poorer. By one year of age, infants have lost the ability to discriminate some non-native speech contrasts. The classical study to show this was done by Werker and Tees (first published in 1984, republished in 2002), where infants in three groups; 6-8 months, 8-10 months and 10-12 months were tested using a head-turn task with contrasts in Nthlakampx, a Native American language. Findings were that young infants could discriminate phonetic contrasts across languages, but that this ability declines during development to become restricted to the contrasts in an infant's ambient language. The paper was published again in 2002 (Werker & Tees, 2002), and additional research from the same lab has supported the findings, albeit with German vowel contrasts showing that by six months, English learning infants could not discriminate the German vowel contrasts as well as English vowel contrasts (Polka & Werker, 1994).

The studies above show that the ability to discriminate vowel and consonant contrasts decline if infants are not exposed to them. These studies were all testing infant auditory speech perception. Still, more often than not, when we speak with another person, we both see and hear our conversational partner. This is also the case for an infant.

Visual perception

Speech perception is usually thought of as unimodal, relying mainly on auditory cues. However, when an infant perceives speech what he/she sees may also be important, depending on how good an infants' visual perception is. Research has shown that an infant is about three to four months old before her/his visual ability is comparable to an adult's; visual acuity, colour vision and binocular vision gradually develops during the first year (Hainline, 1998). However, newborns perceive both size (Slater, Mattock & Brown, 1990), and shape (Slater & Morison, 1985) well, and like adults, three-to-four-month-old infants construct contours where there is none (Ghim, 1990). This is shown in an experiment where four circles with a quarter cut out of each presented together, one in each corner make out a pattern that is perceived as a square (Ghim, 1990). In addition, even newborn infants extract visual cues of a human face with a preference for their mothers' faces (Bushnell, Sai & Mullin, 1989), and an ability to imitate facial gestures (Meltzoff & More, 1977). There is a possibility that face-like patterns have a special perceptual role for infants, as is shown in preference for faces by infants nine minutes old, without any previous experience with faces (Johnson & Morton, 1991). Together these studies point to well developed visual perception in infants, and in addition a possible special interest in the human face.

A recent study has even speculated that a foetus may have some visual processing while still in the womb (Eswaran, Wilson, Preissl, Robinson, Vrba, Murphy, Rose & Lowery, 2002). The study of ten foetuses with gestational age of 28-36 weeks, used magnetoencephalography of the womb to record tiny fluctuations in foetal brain activity, and showed brain activity responses to lights flashed outside the abdomen in

four of them. Although this topic will not be taken any further here, it is worthwhile for the current purposes to open for the possibility for vision and hearing in utero. Evidence for this exists from birds, where quail embryos' ability to learn maternal calls is facilitated by redundant synchronous stimulation; the embryos receiving audio-visual information learn faster than those who received only auditory exposure (Lickliter, Bahrick & Honeycutt, 2002). This suggests that prenatal learning based on multimodal stimulation is possible and may lead to faster learning than learning from unimodal stimulation, at least in some species.

Whether foetal multimodal perception is possible or not, it seems that a newborn infant may possess the ability to perceive visual stimuli that are relevant to speech, namely the face of the speaker.

Multimodal perception

Although we perceive almost any event through different sensory modalities, research shows that adults we do not generally experience it as a collection of independent percepts, but rather as one. This is shown in the well-known ventriloquist effect, where the actual source of a sound deceives observers (Vroomen, Bertelson & deGelder, 2001; Bertelson, Vroomen, deGelder & Driver, 2000). Being so for adults how is it for infants? Research supports that infants have multimodal perception of auditory and visual cues of an event that may be of special interest for speech learning. In one study, three-week old infants were presented with white light followed by white noise, of different intensities. Infants responded to the auditory stimuli based on similarity in

intensity to the previously presented visual stimulus, matching relative intensity across perceptual modalities (Lewkowicz & Turkewitz, 1980). Infants also perceive temporal synchrony between auditory and visual attributes of multimodal events. Results from one study showed that when auditory and visual information was asynchronous, three-month-olds did not match the auditory with the visual material, while six and eight-month-olds did (Lewkowicz, 1986). In addition, a study showed that when 6- and 8-month-olds were habituated to two identical visual stimuli moving through each other together with a sound occurring at the time the visual stimuli coincided, the infants responded differently to the visual stimulus depending on whether the sound was presented before or after the point where the visual stimuli coincided. In other words, the introduction of a sound changed the infants' visual perception (Sheier, Lewkowicz and Shimojo, 2003).

The detection of temporal synchrony does not, however, directly demonstrate a match between auditory speech and articulatory movements. For adults, the 'McGurk-effect' showed that presenting a mismatch between a visual /ba/ and an auditory /ga/ syllable in most cases leads to the perception of /da/ or /ða/, which are intermediate in place of articulation to /ba/ and /ga/ (McGurk & MacDonald, 1976). Where there is correspondence between auditory and visually presented speech sounds, four-month-old infants will perceive this correspondence (Kuhl & Meltzoff, 1982). Evidence of the McGurk-effect has also been found with infants (Rosenblum, Schmuckler & Johnson, 1997). Five-month-old infants were habituated to an audiovisual presentation of /va/. Two different stimuli were then presented: audio /ba/-visual /va/ and audio /da/-visual /va/. Results showed that infants perceived the discrepant audio /ba/-visual /va/ as the

same as audio/va/-visual /va/, but they perceived audio /da/-visual /va/ as different. This indicates the labial and labio-dental articulated consonants were generalised based on the visual stimuli, implying that infants can integrate audio-visual speech much like the original McGurk-effect (Rosenblum et al., 1997). It is worth mentioning that one must be cautions about a direct comparison between the effects of audio-visual incongruence by adults and infants in these studies due to the differences in the methods employed.

In Rosenblum et al., (1997), infants were required to discriminate between stimuli, but other studies have used techniques that test infants' preference. This can be done by testing if infants choose a visual stimulus which matches an auditory presented stimulus. MacKain, Studdert-Kennedy, Spieker and Stern (1983) did a study where two displays presented to infants each presented visual information that either matched or did not match auditory presented CVCV syllables. Five-and-six month-olds looked longer at the display that matched the sound. It has also been shown that infants look longer at the face whose articulatory movement matches the sound they hear compared to when it does not match the sound, being observed both with two-month-olds (Patterson & Werker, 2003) and with four-month-olds (Patterson & Werker, 1999).
It has nevertheless also been demonstrated that infants do not always match audio and visual information. A recent study by Desjardins & Werker (2004) studied how strong infants' integration of auditory and visible speech is. They found that while both male and female four-month-old infants integrate auditory and visual speech, they do not do so all of the time. The authors concluded that the initial foundation for speech allows both integrated and separate audio and visual representations. Supporting the possibility of multimodal integration is research showing that some regions in the brain allow input

from two or more modalities to converge in multisensory neurons, which allows convergence of visual, auditory and somatosensory stimuli (e.g. Meredith, Wallace & Stein, 1992). Multisensory neurons, which are responsive to both visual and auditory stimuli, are quite common in this structure (Stein, Meredith & Wallace, 1994). The neurological studies of crossmodal processing have largely been carried out on animals, and suggest that networks of brain areas are involved in matching and integration of crossmodal input (reviewed in Calvert, 2001).

The above studies together support a general ability to perceive whether auditory and visual information match or not, but that although it may be integrated, this is not always the case.

Theoretical foundation

Traditionally the field of development has had opposing views on multimodal perception. In Piaget's theory perceptual experience is constructed from cognitive operations, with the sensory systems at first differentiated gradually becoming more unified into one system as the child's multisensory experiences increase (Piaget, 1952). According to this view infants have only a very simple repertoire of sensorimotor behaviours when they are born, but these behaviours are gradually integrated as infants develop. The other and quite opposite position is known as the differentiation view, of which the ecological theory of Eleanor Gibson's work from the 1950s and 60s has been central (Gibson, 1969). This position holds that different senses are part of one system at birth, but gradually become differentiated as the child develops. The function of the

perceptual system is closely attached to the concept of 'affordance'. Gibson's theory views perception as a direct response to stimulation. By this view, perception is unmediated by cognitive processes, being directly apprehended by an infant. The environment offers an affordance to a perceiving infant. In this view, the environment does not have different information corresponding to different sensory modalities.

The 'intersensory redundancy hypothesis' is based on the ecological view of development (Bahrick & Lickliter, 2000, 2002) and claims that redundant information has a special place in perception by infants. Redundant information is, for example, intensity and temporal attributes of an event across two or more sense modalities. Redundancy recruits infant attention, and facilitates perceptual differentiation of that information. In this view, redundant information which can be used by several sensory modalities at the same time is more effective than information, which can only be used by one sensory modality at a time.

Although these theoretical positions take different views on multimodal integration in infants, the research reviewed above show that information from different modalities is important for infants in speech processing. While it is possible that infants process some speech information unimodally, in the cases where they can see and hear the speaker, they have the opportunity to use both.

The research presented above show that at least some perceptual abilities may be developed before birth. After birth, infants have well developed auditory perception, and visual perception which enables them to perceive both auditory and visual aspects

of speech, shown through multimodal perception. In addition, infants' well developed capacities to discriminate both place and manner of articulation in vowels and consonants, and their ability to overlook talker differences, indicates well developed speech perceptual capacities. Infants' perceptual capabilities may be affected by the ambient language. If this is so, IDS may play an important role in the development of speech perception, and as will be reviewed above, several theoretical accounts have discussed how this may happen.

Theoretical accounts of speech perception and IDS

Among the many theoretical accounts of infants' perception of speech, a brief review of some of the more influential theories and models shows slightly different views on how speech perception generally works and how the language adults speak to infants might specifically fit perception by infants.

Historical perspective

In a paper by Aslin and Pisoni (1980) four possible mechanisms accounting for the role of experience in speech perception are described. The first is the assumption that there is a gradual unfolding of speech being largely uninfluenced by experience, called the maturation approach. In the second, the ability to discriminate speech sounds is thought of as unfolding and relatively uninfluenced by experience. This is called the induction approach. The third possibility, called the attunement approach, is that discriminative abilities are partially present before the onset of experience and will be shaped by experience. The fourth approach is that discriminative abilities are fully mature at the

onset of experience and that they are maintained or lost depending on presence or absence of experience, called the maintenance approach or universal theory. In a way, all later theories on speech perception relate to these perspectives on speech development. The capacity to discriminate VOT contrasts within plus/minus 20 milliseconds was proposed as a general sensory constraint on the mammalian auditory system. They argued that since adults are slightly worse at discriminating sounds in the voice lead region compared to the voice lag region, but infants have quite a great difficulty with discriminating VOT in the -20 msec region, that language experience plays a role in modifying the discriminability of speech stimuli. This finding, among other findings, motivated favouring the attunement approach (Aslin & Pisoni, 1980).

A level of representation is often included in some approaches to speech processing, however, recent approaches questions if such a level of representation is necessary (reviewed in Chemero, 1999). Anti-representationalism is present in the neural network approach by Damper and Harnad (2000) who stated that categorical perception could be explained by a general learning mechanism, where development of perceptual categories is described by learning, shown through network simulations. Elman (1993) also discusses the neural network approach in relation to learning and concludes that neural networks for speech learning may work best if they are most sensitive during the earliest period of learning for an infant. Being sensitive during early learning, networks will be able to learn a task (e.g., sentence processing) well if they are forced to begin learning with severe memory limitations. If Elman's (1993) findings that learning is most sensitive during the early period of learning transfers to humans, stimulation during infancy might be important for speech learning. Being a period when children

receive IDS, uncovering the characteristics of IDS can provide evidence for how speech learning and speech stimulation are related. We will now turn to more specific theoretical positions on how speech is developed in infants.

Word Recognition and Phonetic Structure Acquisition

Peter Jusczyk (1993, 1997) developed a model of language learning from an interactionist view of language development. Central to the Word Recognition and Phonetic Structure Acquisition model (WRAPSA) is the concept of innately guided learning (Gould & Marler, 1987; Marler, 1991). This concept connotes that learning is in some ways pre-programmed, for example, by drawing infants' attention towards particular aspects of speech. In addition, prenatal experience may enhance any basic bias that infants have to attend to sounds with human vocal properties. For an infant, auditory analysis may be particularly important when extracting characteristics of the speech signal. Supporting this is a new study where infants showed an auditory focus (Robinson & Sloutsky, 2004). According to Jusczyk (1997) the speech signal enters the auditory system, and acoustic properties are extracted (durations, bandwidths, intensities, and spectral changes). These analysers are seen as innate part of the human auditory system (Jusczyk, 1997). According to the model, the basic acoustic properties that are extracted are at first language general. WRAPSA states that as a language is gradually acquired, the auditory analysis is weighted to give prominence to those features of speech sounds that are most critical to making meaningful distinctions between words in the ambient language (Jusczyk, 1997). The model further states that during the first few months of life, the auditory analysis is expected to be quite active.

The information from auditory analysers decays quickly, and an infant has to attend to the information before it is lost. Infants' preference for IDS over ADS (Cooper & Aslin, 1990; Fernald & Kuhl, 1987; Werker & McLeod, 1989) may imply that infants more easily attend to IDS than to ADS and the information they get from this speech type may more easily be held in attentional focus than the information from ADS. If so, little information would be lost from IDS compared to ADS. The WRAPSA model is characterised by presupposing auditory analysis as a first step in speech learning.

Bootstrapping approach

Werker, Lloyd, Pegg and Polka (1996) proposed a theoretical view referred to as the 'bootstrapping approach'. The 'bootstrap' concept is borrowed from, among others, Pinker (1984). Concretely, a bootstrap is the part of a shoe used to pull the shoe on, and similarly parts of language content is used to pull further analysis of language (Morgan & Demuth, 1996). As part of the content in speech, acoustic-phonetic details of the ambient language (and consequently of IDS) forces further perceptual analysis. An infant tunes into the ambient language's phonology. Speech perception is seen as being both auditory and visible. Werker and Tees (1999) stress the bimodality of speech being present in infants but that this is influenced by the experience of producing speech. Both the quality and quantity of speech becomes important. Werker and McLeod (1989) showed that four-month-olds and seven-month-olds have more attentional and affective responsiveness to the prosody in IDS than ADS. In addition to the perceptual experience infants have, in this theoretical approach biological and cognitive factors are

decisive for how language develops (Werker et al., 1996). In particular, there is an interaction between genetically initiated neural substrates and normally occurring species-specific experience in the form of human speech. Werker et al. (1996) states that the initial genetic developing mechanism is formed by experience. Further, as the brain continues to mature during the first year of life the organisation of language is remodelled as the child hears and produces speech. Continuous transaction between the developing brain and experience is the mechanism that allows the emergence of perceptual processing. Vouloumanos and Werker (in press) showed that two-month-olds prefer speech to other sounds with comparable physical characteristics, and this supports this theoretical approach that speech has a special status for infants.

Native Language Magnet Theory

Kuhl has proposed The Native Language Magnet Theory (NLM), which is described in detail in Kuhl (1993 a and b). The NLM presupposes that the ability to categorise speech sounds is given by nature, implying that infants are born with the capacity to resolve the acoustic differences between sounds in different categories. In NLM, an infant's general auditory mechanism is innate. Adult Americans rated the perceived goodness of synthesised exemplars of /i/. The vowels that got a particularly high rating seemed to lie within a particular region of the vowel space (Kuhl, 1991). These prototypical exemplars of vowels, together with non-prototypical exemplars, were presented to six-month-old infants. Results showed that prototypes were not discriminated from other prototypes as well as non-prototypes were discriminated from other non-prototypes (Kuhl, 1991; Grieser & Kuhl., 1989). By six months old infants

have something more than the basic boundaries they were born with; they show evidence of a language-specific magnet effect. Results from a study of American and Swedish infants showed that six-month-olds showed a magnet effect for vowels in their own language, but treated foreign prototypical vowels like a non-prototype (Kuhl et al., 1992). The authors concluded that acquisition of native-language magnets alters the perception of differences in phonetic space. This is done by reducing the perceived distance between a magnet and its surrounding stimuli, and stretching the perceived distance in the region of the phonetic boundary. By this, the perceptual distinctions that are near the boundaries between two categories will be maximised, while those that are near the magnet attractors themselves will be minimised. Magnets cause boundaries for foreign language contrasts to decrease, but work on adults suggests that the boundaries do not disappear per se, since performance in discrimination of foreign-language contrasts can be increased with training (MacKain, Best & Strange, 1981). Kuhl et al. (1997) make the connection between vowel prototypes and the vowels observed in IDS, assuming that the vowels in IDS represent prototypical exemplars of vowel categories.

The NLM is not a modality specific theory, and holds that, although speech representations are initially auditory, they become polymodal as infants acquire information about the production of speech (Kuhl, 1993). Speech representations are initially shaped by auditory input, since early learning results in a perceptual representation that gradually serves as a guide to speech production, giving infants something to aim for when they start to produce speech themselves (Kuhl et al., 1992). Because of their experience listening to, watching and attempting to produce speech, infants' speech representations become polymodal in nature. Infants relate auditory

information to the visual information that accompanies speech and relate this to muscular movements required for producing speech. Although the polymodal nature of speech representations is discussed in the NLM theory, it maintains that speech representations are initially auditory. According to the NLM, IDS will probably enhance auditory cues to an infant, for example, by expanding the vowel space in IDS (Kuhl et al., 1997). A study by Liu, Kuhl and Tsao (2003) support this by the finding that Mandarin mothers with acoustically exaggerated vowel articulations in IDS was positively correlated with how well their infants discriminated speech sounds in a head-turn task.

An alternative explanation of the magnet effect

In an attempt to explain the magnet effect, Lacerda and Lindblom (1997) propose an exemplar-based learning perspective, where a representational space is created in infant memory from the stimuli presented. An important aspect of this perspective is that every stimulus input produces a trace in memory. Incoming speech stimuli lead to neuronal activity, and when presentation of a single stimulus, for example a vowel, is repeated, the final vowel space representation will be influenced by all repetitions. A point not mentioned by the authors, is that the repetitive nature of IDS (Snow, 1972) leads to the representational trace gained from IDS as being based on a larger number of exemplars than is the case with ADS. IDS is, however, mentioned as potentially facilitating the process of learning language. It catches and keeps an infant's attention, keeping an infant focused on the speech signal to a greater extent than is the case with

ADS. In addition, the authors stress that multi-modal information adds to the memory traces mad for a speech sound, so that each trace is by nature multimodal.

Mother-Infant Phonetic Interaction Model

The Mother-Infant Phonetic Interaction model (MIPhI) is a theoretical approach which specifies how IDS may be related to infant phonetic development (Sundberg, 1998; Sundberg & Lacerda, 1999) (implicit in its name is the fact that the mother is the main source of language influence at an early age, however, for the current purpose, the term parent is used). Taking a functionalistic starting point (Vihman, 1996), IDS adaptations are regarded as adjustments to an infant's needs in the development of speech perception, vocal production, attention, affection and communicative intentions. Adults adapt prosody, vowels and consonants differently as a function of the infant's age and development. Age is seen as the developmental variable of interest, and the model predicts that during the infant's first year, prosody and vowels in IDS are adapted, becoming more and more like prosody and vowels of ADS. In contrast, consonants are underspecified in IDS to a young infant, overspecified to the one-year old and gradually become more and more like the consonants of ADS. The MIPhI model is based on research showing that IDS vowels and consonants are different from those in ADS (e.g., Kuhl et al., 1997; Sundberg & Lacerda, 1999).

A source of influence upon the MIPhI model is Lindblom's Hyper/Hypo theory (Lindblom, 1990; 1992), which explains the phonetic trade-off between two speakers. Hyperarticulation or 'clear speech' corresponds to overspecification in IDS, while

hypoarticulation corresponds to an underspecification of speech. Lindblom (1990, 1992) describes a message may be more or less predictable to a listener, determining the strategy that the speaker employs. Predictability is defined according to the listener's perceptual, attentional and affectional needs and constraints. When predictability is low, the speaker will articulate more clearly than when predictability is higher. For a small infant, predictability will be low due to lack of experience with speech and IDS is therefore expected to be overspecified. When predictability is high, a speaker's articulation will be more relaxed. For an older infant, predictability will be higher due to some experience with speech and IDS would consequently be underspecified. As an infant's linguistic comprehension develops, the MIPhI model predicts that IDS will be phonetically modified, becoming gradually similar to ADS. This will be evident in longer vowel durations, expansion of the vowel space and a generally shorter VOT in stops for IDS compared to ADS.

The MIPhI model presupposes that both a parent and an infant possess a phonetic filter. A parent's filter evaluates the infant's speech capacity and articulation of speech passes through this filter. Speech to the infant is being modified according to the infant's needs, and affecting the phonetic content in IDS. The infant's filter is a similar mechanism, affecting both perception and production of speech, by the infant becoming sensitive to the language-specific properties of a mother's IDS. The infant's filter will in turn influence infant vocal production which again affects the ambient language, including the parent's IDS. Although the MIPhI model may seem to stress the auditory modality in speech learning, the visual modality is also recognised as being important for speech development (Sundberg, 1998) although it is not heavily stressed.

The current review of theoretical accounts for the development of speech perception shows how these have moved from generally acknowledging that the ambient language (and thereby IDS) is important for developing language, to more detailed accounts for how the characteristics of this ambient language affects development of speech in infants. Giving a comprehensive and detailed account of how IDS adaptations may be explained in the frame of the interactive setting between infant and parent, the theoretical framework of the MIPhI model is the foundation for the predictions in the current thesis.

How does an adult adapt language to an infant?

Interaction between infant and environment

Research on language addressed to children went from focusing on the structure of language in the 1960s, to focus more on the semantics of language in the 1970s (Shulman, 1998b). Catherine Snow has been a pioneer in studying caretaker-child interaction (Snow, 1972; 1977). Her work uncovered that IDS utterances are short, well formed and have few subordinate clauses. Tempo is usually slow (e.g., Garnica, 1977). Utterances have exaggerated intonation and are generally high-pitched (reviewed in Gallaway & Richards, 1994). Some researchers (e.g. Cross, 1977; Solokov, 1993) have proposed that the language children hear is 'tuned' to their linguistic needs, based on the finding that a parent's mean length of utterance (MLU) increases with a child's MLU, indicating a simplification of speech (Cross, 1977; Solokov, 1993). However, that IDS

is tuned to infant needs would have less significance if infants had scarce experience with IDS.

From birth, most infants hear a substantial amount of speech in everyday situations. Cameron-Faulkner, Lieven and Tomasello (2003) estimated in a study of three-year old children that they heard about 7000 utterances each day. Probably most of this speech is spoken by adults directed towards other adults; but infants also hear speech in situations where speech is directed towards them. Weijer (2002) reported that IDS constituted only 14 % of the total amount of speech that an infant heard during a day. If we follow up on Cameron-Faulkner et al. (2003), this would amount to children hearing 980 IDS utterances each day. This may seem a small amount, however, since IDS is spoken close to and directly to an infant, its repetitive nature has the potential to be relatively salient to an infant compared to ADS (Lacerda, Klintfors, Gustavsson, Lagerkvist, Marklund & Sundberg, 2004). As for the quality of IDS, Noam Chomsky has described the language infants receive from their parents as: "grossly defective - full of false starts, grammatical errors, and misleading pauses "(Chomsky, 1959; 1965). Therefore, it may represent a very poor sample of the language which a child must eventually learn. IDS has been studied in a variety of cultures, and shown that in almost all cultures studied, infants receive this type of speech (reviewed in Lieven, 1994). In addition, the research showing that some cultures do not use IDS to infants (Ochs, 1982; Pye, 1986) can be criticised. Since some studies are anthropological in nature, the literature may have relied on cultural norms or what caregivers say they do when interacting with infants instead of recording IDS. As there seems to be a discrepancy between what caregivers

say they do and what they really do, one must be careful about the conclusions drawn from these studies (Haggan, 2002).

Fathers' vs. mothers' IDS

Mothers are not the only caretakers that speak IDS; fathers also adapt their language when speaking with their infants. Early work by Berko Gleason (1975) proposed that because fathers may be less sensitive to infants' linguistic needs, their IDS may be more challenging to infants than mothers' IDS, leading to faster or better language learning than with mothers' IDS (Berko Gleason, 1975). This 'Bridge-hypothesis' was not supported by early work of syntactic differences between mothers' and fathers' child directed language (e.g. Golinkoff & Ames, 1979). Both similarities and differences between mothers' and fathers' IDS have been found. One study showed that both Japanese mothers and fathers spoke with higher fundamental frequency and greater F0 range to their three-to-seven-month old infants than they did when interacting with each other (Niwano & Sugai, 2003). Another study showed that fathers' IDS may be slightly different from mothers' IDS. Sixteen British fathers showed that compared to mothers' IDS, fathers' IDS had significantly raised fundamental frequency, but the large variability in fundamental frequency typical for mothers' IDS was not as evident for fathers (Shute & Wheldall, 1999).

A general description of IDS would include both mothers' and fathers' IDS. In addition, it has been shown that children as young as four years old also modify their speech to infants (Weppelman, Bostow, Shiffer, Elbert-Perez & Newman, 2003). Although it is

changing, in most Western cultures, the mother has traditionally been the primary caregiver during a child's first year. The mother is therefore a dominant language source for an infant during this period. Although infants can discriminate male voices, research indicates that infants do not prefer their fathers' voices to unfamiliar male voices at four months (Ward & Cooper, 1999). Infants do, however, prefer their mother's voice to that of a strange woman (DeCasper & Fifer, 1980; Spence & Freeman, 1996). In cases where the aim is to study the language infants are in contact with during the first few months, it would be natural to study the mother's IDS.

If the caregivers around an infant use IDS, infant preference may increase its significance as source of influence upon an infant's speech perceptual development.

Infant preference for IDS

It has repeatedly been shown that infants prefer IDS in the first months after birth. Newborns and one-month-olds (Cooper & Aslin, 1990), as well as four-month-olds (Fernald, 1985) prefer IDS to ADS. In addition to a general preference, Werker and McLeod (1989) found that this preference also holds for fathers. Four-month-old infants responded more eagerly to videotapes showing IDS to four-month-olds than videotapes showing IDS to seven-month-olds. Results from Werker and McCloud (1989) indicate that IDS may change somewhat over the first few months, and that the preference for IDS is independent of gender of the speaker. When isolating fundamental frequency, amplitude and duration in IDS in a preference study, it was shown that the fundamental

frequency patterns of IDS were decisive for four-month-old infants' preference for IDS (Fernald & Kuhl, 1987).

It may also be that infants at certain ages prefer IDS while others do not. A study of Japanese infants aged from four to fourteen months showed that from four to six months old, infants preferred IDS to ADS. Between seven to nine months old, the preference decreased, whereas infants from ten to fourteen months old again preferred IDS (Hayashi, Tamekawa & Kiritani, 2001). Others have also shown that the period around four months is a time when infants prefer IDS. Cooper, Abraham, Berman and Staska (1997) found that one-month-old infants did not prefer IDS over ADS, whereas four-month-old infants did. Infants around three and four moths old, engage in conversation-like interaction with their caregivers (Shulman, 1998a). This period (between ten and sixteen weeks) has traditionally been called 'turn-taking' (Boysson-Bardies, 1999), and is marked by an exchange of vocalizations, where the caregiver speaks to the child (or the child to the caregiver), who in turn vocalizes, upon which the other responds. This exchange appears as an early form of "conversation" between the two.

There are some indications that this period around the fourth month is one where certain prosodic features change in IDS, exemplified in a study by Stern, Spieker, Barnett and MacKain (1983). By studying IDS when infants were newborn, four, twelve and twenty-four months, this study showed that during the 4-month period, compared to the other ages studied, a particularly high F0 found in IDS compared to ADS.

Together these studies indicate that IDS is not preferred at all ages, and that one of the reasons for this could be that IDS is different at different ages.

Attention and affect in IDS

One way of studying preference for IDS is by testing how attentive infants are to IDS. Attention in a general sense is being alert and maintaining focus on one event. Werker and McLeod (1989) have shown that infants are more attentive to IDS than ADS both when the speaker is male and female. Attention allows focusing on what is important and to ignore what is not in an environment of sights, sounds, smells etc (Kellogg, 2003). As such, that which is attended to may be learned. When IDS is attended to, infants may learn speech communication cues from it. One type of cue that may be learned particularly well from IDS is emotional cues. Kaplan, Jung, Ryther, and Zarlengo-Strouse (1996) have discussed the possible associative learning of the emotional content of IDS. They observed that four-month-old infants more effectively learned to associate IDS with a smiling or a sad face, compared to a fearful or an angry face. If infants learn to associate IDS with emotional content in IDS, what in IDS reveals emotional cues?

Results from studies on acoustic correlates of the expression of emotion have found pitch to be important for expressing different affective states (e.g. Fairbanks & Pronovost, 1939; Williams & Stevens, 1972). Pioneering research on IDS, Fernald and her colleagues studied how affect is conveyed in IDS (Fernald & Simon, 1984; Fernald, 1989; Fernald, Taeschner, Dunn, Papoušek & De Boysson-Bardies, 1989), showing

fundamental frequency to be higher and pitch range to be larger in IDS compared to ADS. Papoušek and Papoušek (1995) related the emotional characteristics of the interaction pattern between infant and mother as part of a biological caretaking mechanism (intuitive parenting).

The change in pitch is an important cue for affective messages in IDS (e.g. Slaney & McRoberts, 2003), and not only do infants prefer to listen to IDS, they prefer the positive emotions compared to negative emotions that are conveyed through IDS. One study by Papoušek, Papoušek and Symmes (1991) showed that four-month-old infants preferred to listen to the prosodic patterns in IDS which expressed positive affect, such as approval, compared to when it expressed prohibition. Infants also attend longer to and smile more when listening to the exaggerated intonation of IDS compared to their reaction to ADS by the same speaker (Fernald, 1985). Further investigation of the preference for IDS supports that it may not be the mere intonation of IDS that is preferred by infants as proposed in Fernald (1989), but rather the positive affect in IDS. In line with the finding of Fernald (1989), it may be that the positive affect is what is special about IDS. Trainor, Austin and Desjardins (2000) found that the widespread expression of emotion in IDS makes it special to infants, and contrasts it with ADS where expression of emotion is more inhibited. In addition, a series of five experiments showed that if degree of affect in the speech message was held constant across ADS and IDS, six-month-old infants showed no preference for either speech type, and if ADS was more positive in affect than IDS, infants preferred ADS to IDS (Singh, Morgan &

Best, 2002). This research shows that infants prefer the positive affect in IDS maybe because it attracts their attention.

Facial expressions in IDS

One way to obtain information about the emotional state of a speaker is to look at facial expressions. One study (Chong, Werker, Russell & Carroll, 2003) video-recorded Canadian and Chinese mothers' IDS to their four to seven-month-old infants and compared it with their ADS. From these recordings, adults identified three distinct IDS expressions, which were not present in the ADS recordings. Using the Facial Action Coding System (Ekman & Friesen, 1978) muscle movements involved in the IDS expressions were identified as different from those in ADS. All were unlike any standard adult expressions. The first expression was characterised by puckered lips slightly apart, slight smile with lip corner pull and raised eyebrows. The second expression was characterised by inner and outer brow raise, open and stretched mouth and a hint of a smile with lip corner pull. The third expression was characterised by a smile with lip corner pull, cheek raise and slightly open mouth by lip parting or jaw dropping. Although this last expression did resemble an adult 'happy' expression, when this expression and an adult happy expression were compared, adult raters were able to say if this was taken from speech to an infant or to an adult. Infants were also able to discriminate these expressions taken from IDS and ADS.

Together these studies show that emotional information as well as prototypical facial expressions can be important for infants when perceiving IDS.

Segments in IDS

In addition to conveying emotional content, the study of IDS may also convey acoustic-phonetic characteristics of speech. Acoustic phonetics investigates various facets of production of speech, by describing and analysing physical aspects of the speech signal (Kent, 1997). One of the entities research on acoustic-phonetic characteristics of IDS studies is the segment. For the current purposes, a working definition of the segment is a short stretch of speech with relatively unchanging phonetic characteristics (for an elaborate description, see Laver, 1995). Describing the speech signal in terms of phonetic characteristics means to describe it in terms of, for example, voicing, frication, nasality etc. (Barry & Fourcin, 1992). Presented below is research on the acoustic-phonetic content in IDS segments.

Vowels

As primary carriers of prosody, vowels are more often studied in IDS than are consonants. IDS vowels are generally longer in duration than ADS vowels. Bernstein Ratner and Luberoff, (1984) observed nine mother-child dyads with infants with an age range of nine months to two years and three months old. The vowel durations of /i, I, ɛ, ʊ, ɔ, ʌ/ and /u/ were generally longer in IDS than ADS. Interestingly, the vowels /i/ and /æ/ were most similar between speech types. Studying six mothers, Bernstein Ratner (1984) found that vowels were produced with more first (F1) and second (F2) formant frequency separation in IDS, however, unlike the other front vowels studied in IDS to

preverbal listeners, /i/ did not have a higher F2 frequency than in ADS. In addition, the first and second formant frequencies were increased and decreased resulting in less overlap between the vowels in the space for IDS compared to ADS. These results were for preverbal infants. For infants who were at the one-word-stage in linguistic development, the same increase in F2 and clearer separation between vowels in the vowel space area were evident in IDS compared to ADS. In addition, acoustic-phonetic attributes of /ɛ/ characteristics were largely changed, so that the degree to which it overlapped with /I/ and /æ/ was reduced in IDS compared to ADS. The last group of infants had an MLU between 2.0 and 3.5. The IDS to this group deviated even more compared to ADS. The point vowels had least overlap in IDS compared to ADS of all the three groups studied. As there were only two dyads in each group, drawing conclusions about an age effect is difficult.

Indications of a clearer separation between point vowels in IDS compared to ADS have also been evident in other studies (Burnham et al., 2002). In a study of Australian English IDS and speech to pets (PDS), formant frequencies for the first two formants were selectively increased and decreased, resulting in a larger vowel space for IDS than for both ADS and PDS (Burnham et al., 2002). The authors suggest that special speech registers differing in vowel hyper articulation may be due to the speaker's intuitive perception of the linguistic and emotional needs of the audience. By this view, intuition about the audience is used to automatically adjust IDS accordingly. Further evidence for a larger vowel space in IDS compared to ADS come in a cross language study by Kuhl and her colleagues (Kuhl et al., 1997) who recorded mothers from three different

language backgrounds (American English, Russian and Swedish) speaking to their two- to five-month old infants. In each language, the vowels /a, i/ and /u/ F1 and F2 were more extreme in IDS than in ADS, producing a larger vowel space in IDS than in ADS. The authors draw the connection between a larger vowel space and an extreme articulation in IDS, which may make IDS perceptually salient to a child's language acquisition (Kuhl et al., 1997). The central studies of vowels in IDS presented above show that in several languages, vowels are produced with more extreme articulation in IDS than ADS. There are, however, no studies of the acoustic-phonetic content of Norwegian IDS.

Consonants

In addition to vowels, consonants in IDS have been studied, most frequently stop consonants. There are, however, discrepancies in the results. Baran, Zlatin Laufer and Daniloff (1977) recorded American English IDS to three children who were approximately 12 months old, finding no differences in the time between release of a stop and voicing of the following vowel (voice onset time (VOT)) between IDS and ADS. However, in 4.1 % of the data, a shorter VOT was found in IDS than in ADS, showing a discrepancy in results. Malsheen (1980), who investigated six mother-infant dyads engaged in conversational interaction in their homes, found different results. For infants aged 15-16 months, a greater difference was found between voiced and voiceless stops in IDS compared to ADS, mainly due to the longer VOT duration for voiceless stops. For infants aged 6-8 months and children aged two-and-a-half to five years, VOT

was the same between speech types. This suggests that children who are a little older than one year receive IDS where VOT may be more prominent than in ADS.

More recent findings on VOT in IDS have shown shorter VOTs in Swedish IDS to three-month-old infants than ADS (Sundberg & Lacerda, 1999). A laboratory setting with a play situation was used to elicit speech from mothers. In IDS, the difference in VOT between voiceless and voiced stops was smaller than it was in ADS. Instead, there was a gradual successive increase in VOT from /b/ to /d/ to /g/ to /p/ to /t/ to /k/. One of the explanations the authors offer is that consonantal properties follow a different developmental route compared to vocalic and prosodic characteristics of IDS.

Consonants are obstruent sounds, articulated with an obstruction of the vocal apparatus. From one point of view, they may fill a function of complementing the sonorant sounds, which are melodic, continuous and smooth. With shorter VOT, the onset of a vowel after a stop consonant will follow faster. This means that the consonants become less prominent in the speech stream, while vowels may be more salient to the infant. The MIPhI model proposed by Sundberg (1998) presupposes that VOT in IDS to 1-year olds is overspecified. Malsheen (1980) supports this. Support for this interpretation is also presented in a working paper by Sundberg, (2001). She reports longer VOT for stops in three mothers' IDS to infants 11-14 months old compared to ADS. Together, the studies presented above show that while some have shown a longer VOT in IDS compared to ADS, others have found the opposite.

The studies on vowels and consonants in IDS pictures a situation with differences between speech types for the languages studied. For consonants, the findings are discrepant, while for vowels the expanded vowel space seems to be evident in the most recent studies on IDS.

Does what parents do affect language development?

Attention and positive affect in IDS is mentioned above as possible indirect influences on infant speech perceptual development. However, Bortfeld (2004) states that the mere exposure to the prosodic, phonological and statistical consistencies of language would be sufficient to learn a language. Evolutionary, this superfluous speech evolutionary be the precursor for all language use (Falk, 2004). A review of research on vocal, visual, gestural and tactile interaction between infant and adult chimpanzees (pan troglodytes) and bonobos (pan paniscus) gives rise to speculation about how IDS has evolved. Bipedalism may have led to primate mothers putting their babies down to do other things, and this may have led to replacing cradling and carrying with prosodic vocalisations to reassure infants of the mothers presence, as well as securing that the infant kept close to the mother for protection. These vocalisations may in turn have developed to more elaborate forms of speech (Falk, 2004). Though the points made in Bortfeld (2004) question the value of IDS for an infant, studies have been conducted to try and show a connection between IDS and language learning.

Although uncovering a causal relationship between the ambient language and language development is difficult if not impossible, it has nevertheless been proposed that there is an association between IDS and infant language learning generally. Dominey and Dodane (2004) discuss the connection between child directed speech and language acquisition. They argue that prosodic cues are used for linguistic structural organisation, for example, boundaries between sentences, clauses and words (Dominey & Dodane, 2004). As mentioned earlier, infants prefer prosodic cues in speech (Fernald & Kuhl, 1987), which would draw an infant's attention toward the cues that will help to uncover the boundaries between structural entities in fluent speech. For example, the clause-final syllables are lengthened in English (Klatt, 1975). This cue is made salient in child directed speech by lengthening these syllables more than is done in ADS (Bernstein Ratner, 1986). Similarly, in the visual domain, infant attention is drawn towards the surrounding environment by uncovering the intended object referent of a word. This typically happens around an infant's first year, when dyadic interaction changes to a triadic exchange between an infant, a caregiver and a third object (Tomasello, 1992). It has been shown that mothers actively engage their infants' attention to a toy or other object during interaction with their one-year old infants, resulting in larger vocabularies for these infants (Tomasello & Todd, 1983). These studies point to the connection between what parents do and the outcome for infants' language acquisition.

Being focused on basic perceptual processes, DeBoer and Kuhl (2003) used a computer model to study how a system would better learn from IDS, compared to ADS. Recordings were made of ten mothers' natural speech to their infants and to an adult. From these, words with /a, i/ and /u/ in IDS and ADS were selected and used as input

for a computer model. The model was supposed to learn vowel qualities by determining the shape of the distribution of each vowel category in the vowel space. When means of the learned vowel representations based on IDS resulted in a representation of the vowel space closer to estimated standard formant values than vowel spaces based on ADS, this indicates that vowels might be more learnable from IDS than from ADS. A major assumption in the model was that the system knew how many vowels needed to be learned which would not be the case for a child. The results indicate that vowel quality is easier to learn from IDS than ADS, supporting the hypothesis that the characteristics of IDS serve an adaptive purpose. This was an artificial system, and not a direct model of how infants learn vowels, and we cannot directly apply these results to infant perceptual learning. However, the study shows that IDS works better than ADS for computer learning (DeBoer & Kuhl, 2003).

Further attempts have been made to show the possible connection between the characteristics of the IDS received and children's language development. Liu, Kuhl and Tsao (2003) found that Mandarin mothers increase the acoustic distance between vowels in IDS, showing acoustically exaggerated vowel articulations. The degree to which mothers had a larger vowel space in IDS compared to ADS was positively correlated with how well their infants discriminated speech sounds in a head-turn task. The results were shown for two groups of infants: 6-8-month-olds' and 10-12-month-olds' (Liu et al., 2003). This study establishes a relationship between a large vowel space in IDS and discrimination of speech sounds. However, being better at perceiving speech sounds need not mean that there is an effect of IDS on language development beyond infancy.

Another study by the same authors (Tsao, Liu & Kuhl, 2004), further showed a correlation between speech perceptual skills in infants at six months and language abilities at 13, 16, and 24 months. Language abilities were tested by using the MacArthur Communicative Development Inventories (CDI). Although it is important to acknowledge that a correlation is not evidence for a causal relationship, together, and by using different approaches, these studies may indicate that IDS may have an effect on speech learning and later language development.

Together, the research presented on speech perception in infants, show well developed abilities to perceive speech sounds. The main anticipation from the MIPhI model is that IDS affects development of speech communication through an interactive process between infant and parent. This is supported by the findings on the relationship between characteristics of IDS and later language development. A question still remains however, whether the acoustic-phonetic characteristics of segments in IDS can also be observed for Norwegian and whether there is a development in these characteristics over time observed for the same subjects. Ethical considerations as well as some central concepts relevant for the current research on speech production and perception are introduced below.

Methodological issues

Ethical considerations

The current research was conducted after being approved from the committee for medical research ethics (Regionaletisk Komite for Medisinsk Forskningsetikk (REK) – Midt-Norge) and from the Norwegian Social Scientific data register (Norsk Samfunnsvitenskapelig Datatjeneste (NSD)). Access to health care centres was gained through an approval from the leading MD for The Council of Trondheim. An agreement of silence was signed with the chairman for the social administration of the Council of Trondheim before presenting the project in maternity groups.

The two studies included in the current research involved infants and their parents. Formal adherence to ethical regulations (The Nuremberg Code and Helsinki-declaration), requires that the researcher has responsibility for maintaining the dignity and welfare of the participants, and to protect them from discomfort. Ethical considerations were particularly important for the production study, where parents (mothers) were informed about the study before the birth of their infants. This had to be done without inducing any pressure to participate.

A challenge for the production study was to collect the data without informing the participants about the projects detailed purposes to an extent that would affect the data. The participants were therefore informed that the current research studied how adults communicate with infants. Caution was, however, taken to conduct a proper debriefing

with them after the last recording was done, and in addition participants received two written reports after the study to inform them about the results.

Speech outcome depends on situation

A study is ecologically validity to the degree to which its findings have relevance to real-life situations for the participants involved (Corsini, 1994). It is important to assure the ecological relevance to different parts of an experimental investigation: the nature of the experimental setting, the stimuli under investigation and the measured response (Krøjgaard, 2001; Schmuckler, 2001).

As part of human interaction, IDS involves the study of speech as well as the intimate communicative context in which it occurs. This context potentially influences the content and manner by which information is conveyed, for example, via the parent's gaze, facial expression and tactile contact (Hertenstein, 2002; Muir & Hains, 1993), and therefore, the experimental setting is important in studies of the interaction between a parent and an infant.

The environmental circumstances of a setting affect how a person speaks (e.g., Tetnofski & Damico, 2001; Greasley, Sherrard & Waterman, 2000). The environmental circumstances are consequently important in studies of mother-infant interaction. Although some studies on non-linguistic interaction have shown no effect of difference in experimental setting (e.g., Borduin & Henggeler, 1981), a study by Christensen and Lingwall (1982) designed to explore the reduction in stuttering found a difference

between the laboratory and home setting. While two-year-old children seem to speak similarly in a laboratory and at home (Bornstein, Haynes, Painter & Genevro, 2000), another study of speech to infants in two settings found that mothers spoke more slowly to their infants in a home setting than they did in a laboratory setting (Stevenson, Leavitt, Roach, Chapman et al., 1986). This shows that speech to infants is adapted according to the situation, and consequently it is vital that comparable settings are used for studies of speech type differences.

Although the acoustic-phonetic variability in speech is limited by the necessity to preserve the phonological system of a language, there is a high degree of such variability, such as, for example, the reduction and assimilation of vowels and consonants due to, for example, stylistic factors in the communicative context (Duez, 2001). When different contexts affect the acoustic-phonetic content of speech, the goal of studies of speech to infants would be a high degree of external validity. The intention of studying IDS that would naturally occur should therefore be reached by studying IDS in natural interaction, by observing conversational partners where interaction naturally occurs.

IDS production by parents may be studied in a controlled environment offered by a laboratory setting. This is important for obtaining recordings with low noise levels, which can be analysed more easily, and thereby lead to more accurate measurements. In addition, the interaction between parent and infant can be directed in order to obtain repetitions of selected words. This will in turn provide consistent acoustic-phonetic

contexts for segments. Nevertheless, the natural interaction between, in this case, a mother and infant, is the situation where IDS occurs spontaneously.

The current research is carried out with an overall goal of using a more externally valid setting than a laboratory would provide. For the production study, mother-infant dyads as well as adult-adult dyads are studied at home where speech interaction occurred naturally. Presented below in Figure 1, a and b, are pictures of the two situations used in the current production study.

Figure 1a. Experimental setting for recording IDS in the production study.

Figure 1b. Experimental setting for recording ADS in the production study.

The perceptual demands put on a parent and infant in a laboratory setting may be quite different from those experienced in everyday life. Everyday settings usually include moving objects and sounds, such as those by parents, siblings or other persons, or by traffic, radio and TV, all of which require a perceiver's attention. Such attentional demands may not be present in a laboratory, where experiments are often designed to control the attentional demands made on an infant. Even in such a context, it may be difficult to hold an infant's attention over time. For instance, in the standard visual fixation procedure, an infant is seated in a caregivers lap facing a video monitor, surrounded by a uniform non-distracting background, in some cases with a fixation light mounted above or below the screen (Jusczyk, 1997). The caregiver typically wears headphones to avoid that he/she receives the same stimuli as the child (Jusczyk, 1997). This may seem like an artificial situation to both the adult and the infant. It can be

argued that if such a controlled context does not provide reliable and valid information about perception by infants, there is not much likelihood of obtaining results form a more natural noisy environment. Nevertheless, with the likelihood that the experimental outcome is different in the two situations due to the different attentional demands, the perception study included in the current research aimed to use a setting that resembles the everyday perceptual processing as an infant encounters it. Consequently, for the perception study, an aim was to explore using a context that would allow a relaxed setting for a parent and infant. Figure 2 below exemplifies the setting used for the perception study.

Figure 2. Experimental setting for the perception study.

Although the laboratory setting may in some ways be less distracting for an infant, it is relatively artificial since few communicative situations in everyday life resemble it. The

aim of studying speech perception by infants in a natural setting would motivate using a home setting, where an infant typically is exposed to speech during the first months.

Statistical analysis

Some studies of IDS have used a multivariate analysis of variance for similar designs as was employed for the production study (e.g., Burnham et al., 2002). The current research uses repeated measures analyses. With a multivariate approach, the variation across levels of the independent variables would statistically be interpreted as variance between subjects. This between subjects variance goes into, and increases, the error of the F-statistic. When the F-statistic becomes smaller as a result, the chance of committing a Type II error (deciding in favour of a null hypothesis that is false) is higher. A repeated measures analysis removes the between subject variability from the error, thereby reducing the chance of a Type II error (Stevens, 1992). In both current studies, the same participants provided data for several independent variables. In this vein, employing repeated measures as method of analysis was appropriate for the current quasi-experimental designs.

PAPERS I-IV

The general aims of the thesis are reflected in the four papers. The study reported in the first two papers meet the first general aim by investigating the acoustic-phonetic characteristics of IDS generally and by exploring such aspects in Norwegian IDS specifically by describing vowels and consonants in Norwegian IDS. A further investigation in the same study reported in the third (and in part the second) paper meets the second aim of investigating possible changes in IDS during the first six months of an infant's life. The study reported in the fourth paper meets the third aim of investigating perception of IDS.

Papers are not included due to copyright.

DISCUSSION

Presented below are summaries of the specific results in the four papers. After the summary, the main findings will be discussed in the context of theoretical and methodological topics relevant for the current research.

Summary of findings for papers I-IV

Paper I

Infant directed speech in natural interaction – Norwegian vowel quantity and quality.

Published in Journal of Psycholinguistic Research (2005) vol. 34(3).

The first paper recorded 6 mothers' IDS and ADS in a naturalistic home setting over six months. Acoustic-phonetic analyses of mothers' natural speech to their infants and to an adult showed that vowels in Norwegian IDS differ from ADS vowels. Both long and short vowel quantities were longer in duration in IDS than in ADS, but the relative durational difference between long and short vowels was equal between speech types. Vowel spectra were not equal between speech types, but showed unexpected results. While both long and short /i:/-/i/ were no different between speech types, formant frequencies for both /a:/-/a/ and /u:/-/u/ were different between speech types, producing a formant triangle that were shifted upwards in a two-dimensional representation of the vowel space, with F1 along the x-axis and F2 along the y-axis. The shifted vowel space

was probably a result of producing vowels more front, possibly resulting from both visual and auditory aspects of speech being enhanced in IDS.

Paper II

Voice onset time in infant directed speech over the first six months.

Published in First Language (2005) vol. 25(2).

In the second paper, the same corpus of data was used as for the first paper. Acoustic-phonetic analyses were made of voice onset time in stop consonants in IDS and ADS. Contrary to the expectation, it was observed that VOT in Norwegian stops was longer in IDS compared to ADS. However, one stop did not show the same pattern of results as the others. VOT for the unvoiced bilabial /p/ was equal between speech types. In addition, an analysis of the development of the difference between speech types over the six months showed no change over the first six months after infants were born. The selective overspecification observed for VOT in Norwegian IDS stops may result from adaptation of IDS to an infant's perceptual capacities by making stops easier to perceive in the speech stream, possibly enhancing auditory and visual cues simultaneously.

Paper III

Changes in infant directed speech in the first six months.

Submitted to: Infant and Child Development.

Further results for vowels from the same corpus as the first paper were reported in the third paper, with the addition of duration for a fricative /s/. The developmental trend in the differences between speech types were tested for vowel duration and fricative duration over six months. In addition, the sizes of the vowel spaces were calculated for the six months. Results for the size of the vowel space for Norwegian vowels showed an unexpected generally smaller vowel space in IDS compared to ADS. This was consistent over the six months. Only vowel duration showed a reliable change over time. The generally longer vowel duration in IDS than ADS was not so distinct during months three and four as it was before and after this period. Results from the third paper suggest in addition to the relatively stable over-, and underspecification of segments in IDS, the language in question may guide which adaptations are made in IDS and that IDS is possibly adapted to perceptual, neurological and social communicative needs during an infant's first six months.

Paper IV

Audio-visual congruence: Infant perception of adult and infant directed speech – a pilot study.

Unpublished manuscript. Abstract published in Child Language Seminar Abstracts, 12-14 July 2004, University of the West of England, Bristol, p. 29.

Paper IV emerged from the aim to study perception of audio-visual IDS that had the characteristics that are found in Norwegian IDS. Using a novel method that would result in low attrition rates and employ a naturalistic setting for studying speech perception by infants, this preliminary investigation explored whether infants more readily would notice whether the auditory and visual information in an audio-visual syllable match or not when presented in IDS than would be the case for ADS. 2 to 8-month old infants were videotaped while they watched a screen showing a woman pronouncing audio-visual syllables in IDS and ADS. Contrary to the expectation, audio-visual syllables where auditory and visual information did not match were attentionally preferred compared to congruent syllables. As infants treated congruency equal in IDS as in ADS, the current study could not verify whether the proposed visual cues of Norwegian IDS affects the perception of congruence. However, the preference infants from two to eight months showed for audio-visual IDS that is comparable to ADS with respect to smiling, fundamental frequency and pitch fluctuations, justifies continued speculation that Norwegian IDS may contain visual articulatory cues that are perceptible to infants throughout this period.

Limitations of the present findings

The current research aimed to explore the production and perception of IDS in relatively natural situations for mothers and infants. However, testing participants at home limits the degree of experimental control. It is essential to limit the degree of noise, by turning off radio and TV, as well as closing windows to limit traffic noise. The mothers in the production study (paper I-III) typically recorded their IDS in the bathroom, where water

was occasionally running. In such cases, that portion of the recording with noise was unsuitable for further analyses. Therefore, the experimental setting led to excluding a few parts of the recordings, but when these were only a few, this can be seen as a minor problem for the current research.

In addition to objects and events, noise may also have distracted the participants in the perception study (paper IV). The infants may have become engaged in other things than the computer screen on which they were to fixate their looks. A consequence could be that they did not attend to the stimulus material as much as they might have in a laboratory. However, familiar sounds, objects and events may be easier to overlook and consequently more attention may be directed towards the experimental material presented on the screen. In addition, this would be a problem for both speech types tested, and could not systematically have affected the pattern of results. In this way, the current familiar setting may still lead be an experimental situation suitable for the current purpose of studying infant perceptual processing.

Another problem arises from the many phonetic contexts that segments were selected from when doing acoustical measurements in the production study. The natural setting proved not to be an optimal situation for obtaining the same words for IDS and ADS across the time points when the mothers were recorded. Segments were selected from a variety of phonetic contexts. Selecting a vowel surrounded, for example, by stop consonants, would result in a spectrographic representation where the start and end of the vowel would be more evident than if that vowel were surrounded, for example, by glides. The reason is that the transition from a stop to a vowel is more abrupt and

evidently visible than the transition from a glide to a vowel, which is more gradual (see Kent & Read, 1992). As this is a general problem for both speech types it is not likely to have affected the pattern of results.

Throughout the four papers, the visual articulatory cues in speech are repeatedly mentioned. It is important to note that the direct study of visual cues in IDS has not been done in the current production study, since it would have involved audio-visual recordings of the mothers' faces while speaking to their infants. As is discussed in paper I, maintaining the intimacy of mother-infant interaction may be important for obtaining valid data. Introducing a video camera into the experimental setting could prove disturbing for some participants, resulting in less natural IDS. In addition, as will be evident from the discussion below, it is not an isolated finding that suggests enhancement of visual articulatory cues of IDS compared to ADS in the current research. The general pattern of results from the production study together leads to the tentative conclusion that visual articulatory cues are enhanced in Norwegian IDS compared to ADS.

Characterising Norwegian IDS

Papers I through III show that Norwegian IDS is substantially different from ADS. Compared to ADS, IDS has longer vowels and fricative /s/. VOTs in stops are longer in IDS than ADS, except from in a bilabial /p/. As seen from the results in papers I and III, vowel spectra are different, but contrary to previous research (e.g. Kuhl et al., 1997), the vowel space is smaller in Norwegian IDS than ADS. Instead, the vowel space is shifted,

possibly corresponding to a fronting of vowels. As this was not observed for all vowels, but seemed to include mainly back vowels, added to the pattern of results. Likewise, velar and alveolar stops were enhanced in IDS with longer VOTs than in ADS. For vowels, stops and the fricative /s/, these characteristics of IDS were mostly stable throughout the six months after birth.

The lack of studies of acoustic-phonetic characteristics of Norwegian IDS leaves no foundation for direct comparison. Studies of IDS have included several languages (Burnham et al., 2002; Kuhl et al., 1997), although the vowel systems in the languages that represents the foundation for comparison in the current production study, Swedish and English, are not the same as in Norwegian (Kristoffersen, 2000), characteristics of different speech types means in essence within language comparisons. Therefore, on the one hand, comparing speech types will not be affected by between language differences. However, on the other hand, the differences between languages may be important when comparisons are made. As mentioned in paper III, the characteristics of one language may found the characteristics of IDS in this language (for example, Kitamura, Thanavishuth, Burnham and Luksaneeyanawin, 2002).

The possible fronting of back vowels and the enhancement of VOT in velar and alveolar stops make the enhancement of back segments a common denominator for Norwegian IDS. In papers I-III it is discussed that this pattern may be related to visual articulatory cues in IDS. If mothers articulate back segments further front in IDS, to what may this be related?

Front segments

While back vowels were fronted (paper I), front vowels were no different between speech types. In addition, a bilabial stop was the exception from the longer VOT observed for stops in paper II. In the current research, one of the explanations for this is that back vowels and stops are overspecified in IDS. Front phonemes has been studied by Miall (2001, 2003), who measured vowels and consonants based on degree of frontness in articulation. Although very different from the acoustic-phonetic approach taken in the current research, in a poetic analysis of a dyadic interactive conversation between a mother and two-month-old son, he concluded that front articulation in speech was used to connote intimacy between parent and infant. Miall (2001, 2003) also pointed to the finding that front phonemes appear first during the child's speech acquisition, and proposed that the reason for this may be because infants are presented with fronted speech sounds early.

Adults may overspecify back segments to front them and the reason could be that infants have a preference for front phonemes. In Aldridge, Stillman and Bower (2001), the second of three experiments showed an order of newborn infants' preference for segments, /i/>/u/>/y/>/w/, with /i/ being most preferred and /w/ last preferred. As the more preferred phonemes are front phonemes adults could be fronting segments in IDS because of this preference.

As there may be a relation between the degree of frontness in phonemes and their visibility, visibility could be an underlying factor for the fronting observed in studies I through III. The idea that infants are sensitive to some changes in vowel spectra more

than others has also been proposed in Lacerda and Sundberg (2001). They found that a change in F1 is easier to perceive by infants than a change in F2, suggesting a perceptual bias for the high-low dimension of the vowel space. When this result could be an indication that infants use the degree of openness in perception, this supports the notion that the visibility of speech may be important in speech perception for an infant.

Together, the fronting of speech sounds may serve several purposes, stressing the intimacy of an interactive situation, and in addition, overspecifying segments, making them easier to perceive for an infant, possibly by enhancing visual articulatory cues. One way of stressing emotional cues and at the same time possibly enhancing visual articulatory cues in speech is through smiling.

Smiling

In face-to-face interaction between infants and their mothers, it has been shown that if the mother is smiling, an infant will have more speech-like vocalisations (Hsu, Fogel & Messinger, 2001). A side effect of smiling is alternation of the vocal tract from its neutral position such that the mouth widens and its opening enlarges (Tartter, 1980; Shor, 1978). In paper I it was mentioned that subjective evaluation of the emotional state of the mothers indicated that the mothers in the production study could be smiling while IDS recordings were made. The likelihood that the mothers are smiling in face-to-face interaction is strengthened by the findings in Chong et al. (2003). As described in the introduction, three facial expressions were identified from the IDS recordings in Chong et al. (2003). From the description of the muscle movements, all three

expressions involved lip corner pull. In two expressions, this pull was only slight. In the third expression, this pull was extensive. If these expressions are used in interaction between mothers and infants in the current production study, the mothers may regularly have a slight smile while speaking to her infant.

A smile would be emotionally preferable to an infant, but may also have particular effects on both the auditory and visual information revealed in speech. However, while the smiling hypothesis could explain parts of the findings for vowels, it is not obvious how it would account for why the fricative /s/ was longer in IDS compared to ADS and why VOT was longer for velar and alveolar stops, but not for a bilabial /p/. Therefore, smiling may be considered a secondary explanation for the findings from the production study.

Duration and syllables

While the focus in the current research is on segments on mothers' IDS, segments are usually parts of larger speech contexts (Laver, 1995). It has been shown that infants are sensitive to syllables (Jusczyk & Derrah, 1987; Bijeljac-Babic, Bertoncini & Mehler, 1993). In the production study, segments were the focus of study, and findings for duration included longer vowels and fricative /s/ as well as longer VOTs in stops. This is interpreted as IDS being enhanced by making segments longer and thereby easier for infants to perceive. In paper II, a slower speaking rate (Miller, 1981) was discussed as possible explanation for the longer VOTs. However, the difference in syllable duration between speech types was tested post hoc in paper II, and as syllable duration was no

longer in IDS than in ADS, speaking rate could not be sole explanation for the longer VOT in IDS. Therefore, results reported in papers I and II present a puzzle. Paper I revealed longer vowel duration in IDS than ADS and paper II revealed a longer VOT for stops in IDS than ADS, but syllable duration as such did not differ. When vowels, VOT and the fricative /s/ duration are longer, how can syllable duration be similar for the two speech types? VOT is only part of a stop consonant; the other part being the closure, or stop gap. This corresponds to the occlusion of the vocal tract and its duration typically ranges from 50 to 150 milliseconds (Kent & Read, 1992). Auditorily it is characterised by silence. If the vowel is longer in IDS compared to ADS and the voice onset time is longer in IDS compared to ADS, then the stop gap must accordingly be extensively shorter in IDS than in ADS in order for the whole syllable to have the same duration between speech types. To verify this, a closer analysis of the syllable-internal timing and a direct study of the stop gap in stop consonants would be necessary.

Above it was mentioned that infants may be sensitive to the syllabic entity in speech and syllables were also used as stimuli in the current perception study.

Perception of audio-visual IDS

In paper IV it was reported that infants look longer at IDS than ADS audio-visual syllables, most probably do so because they prefer to look at and listen to them. Degree of smiling, average fundamental frequency, as well as pitch fluctuations were comparable for the two speech types in the perception study. Therefore, the infants' preference must be based on other characteristics of IDS. Formant frequencies of the

vowels in the stimuli matched the ones reported for the vowels /a:, i:/ and /u:/ in paper I. Since vowel duration and VOT for stops in the stimuli were equal in the IDS and ADS stimuli, the reason for IDS preference may rather have resulted from the differences in formant frequencies for the vowels in the two speech types. As it also was anticipated in paper I and II, there is a possibility that Norwegian IDS enhances visual cues to infants, and this may also have attracted infants' attention in IDS.

However, as some of the infants in the perception experiment were only two months old, and as mentioned earlier, visual acuity develops over the first few months (Hainline, 1998), infants may not have been able to see the stimuli presented on the screen. However, the fact that the current stimuli were of the human face, may have facilitated their attention and vision. Goren, Sarty and Wu (1975) showed that newborn infants turned their heads more to follow the picture of a face-like pattern than facial features in wrong arrangements. A very recent study (Turati, Valenza, Leo & Simion, 2005), has also shown that 3-month-olds also prefer natural face stimuli compared to schematic ones. Although the youngest infants in the perception study (paper IV) had an overall longer fixation time than older infants, their pattern of results for both speech types and congruence resembled the pattern found for the older infants. Together, these results indicate that in the current research, infants' visual perceptual capacity is sufficient to process the speech stimuli presented in the perception study (paper IV), and that this may be related to infants' preference for the human face.

The meaning of enhancement

Throughout this thesis, the term 'enhancement' of speech is used. The term 'enhancement' generally refers to improvement or advancement (Webster's Dictionary, 1989). In connection with the characteristics of IDS, in papers I-III the term is used to explain the characteristics of vowels and consonants in Norwegian IDS. However, its connotations related to the characteristics of IDS can be slightly different depending on which result is being discussed. For example, longer vowel duration in IDS than in ADS possibly represents an enhancement of speech to infants. It has been shown that duration may be a preferred cue to children (Bohn & Polka, 2001; Burnham, 1986), and it is indicated that the durational cue may be preferred by infants (papers I-III). Sundberg (1998) also refers to highlighting cues in IDS that already have a preferential advantage over other cues. In her MIPhI model, prosodic cues are highlighted in IDS to small infants, taking advantage of their preference for prosodic cues in speech perception. In this vein IDS enhances speech by taking advantage of an infant's perceptual preferences. On the other hand, as seen in the current production study, VOT duration is longer in IDS than ADS (paper II). As, compared to vowels, stops are short and less prominent in the speech stream, this suggests that IDS may highlight that which is more of a perceptual challenge for an infant. In this case, enhancement is a result of highlighting that which is perceptually difficult for an infant.

In paper III, the size of the vowel spaces for vowels were smaller in IDS than ADS. This finding can be interpreted as the opposite from enhancement, and goes against predictions in Kuhl (1991; 1993). Hyperarticulated vowels resemble prototypical exemplars of vowels, which have been associated with the magnet effect in infants, it is

suggested that since the ambient language contains a larger vowel space, cues from IDS help infants learn speech and is associated to better speech perception (Liu et al., 2003). In this vein, vowel space size reduction could make IDS vowels more difficult for an infant to perceive. Unlikely as it may seem, this is a possible explanation of the result of a smaller vowel space in Norwegian IDS as reported in paper III. However, the smaller vowel space is also shifted, possibly being a result of fronting back vowels. In this way visual cues may be enhanced, and while some auditory speech cues may not be enhanced in Norwegian IDS, durational cues may be enhanced as may visual articulatory cues.

Evaluating results for the vowel space together with the results for vowel duration, fricative /s/ duration and VOT in stops, another way of describing the picture of results for papers I-III is that Norwegian IDS tones down spectral cues, while it enhances durational cues. In this way, both visual and auditory cues may be enhanced in IDS to infants, visual cues because they may be more difficult for infants to make use of (Hainline, 1998) and auditory cues because these are the preferred cues in perception by infants (Robinson & Sloutsky, 2004). Together, findings from the current research supports that IDS may enhance both that which is perceptually easy and that which is perceptually difficult to an infant.

Theoretical points related to current findings

The theoretical positions presented in the introduction (Jusczyk, 1997; Werker, Lloyd, Pegg & Polka, 1996; Kuhl, 1993; Lacerda and Lindblom, 1997; Sundberg, 1998) as

background accounting for the development of speech perception are not seriously opposing views. They acknowledge input as important for development of speech perception, and can be seen as complimentary. Some recognise IDS as particularly successful in the process of learning language (Liu et al., 2003). However, for the current research the multimodality of speech has proven an important topic, and while multimodality is mentioned (e.g. Lacerda & Lindblom, 1997) few discuss how it may be exploited. Gogate, Bahrick and Watson (2000), who use the concept 'multimodal motherese', have done this. They studied how mothers taught three new words to infants in three age groups: 5-8 months, 9-17 months and 21-30 months. They found that mothers most likely highlighted the relations between the words and the referents for the infants, by object motion and touch together with the word that they said. This concept could be further developed to contain visual articulatory cues in speech and seen in connection to the current data, where both auditory and possibly visual articulatory cues are highlighted in IDS.

Speech learning and Ecological theory

If visual cues are enhanced in IDS as shown in the current production study (paper I-III), why may this be the case? If perception is global or amodal at birth, infants would probably benefit from multimodal cues being enhanced. Norwegian IDS, where several cues may be enhanced to an infant (paper I-III) could exploit the perceptual capacity of an infant. The possible incorporation of multimodality in a model of language acquisition is made in an emerging theoretical approach (Lacerda et al., 2004). Lacerda (2004) presents the Ecological Theory of Language Acquisition, where early language

acquisition is described as a natural consequence of the multi-sensory events in adult-infant interaction settings. Although described as involving multimodal word-object coupling, this might also be applied to coupling between sound and articulatory movement in face-to-face speech interaction. The equivalent process in speech will then incorporate associations between sounds and articulatory movements of the speech organs during a face-to-face interactive learning process. This theoretical position opens for thinking about multimodality in speech development in a wider sense. Tactile stimulation is involved in mother-infant communication (Arnold, 2002), and if coupled with auditory and visual cues for speech, touch may be part of IDS multimodal stimulation.

Time dimension and MIPhI model

The current production study generally shows that the difference between IDS and ADS is relatively consistent over the first six months of an infant's life (paper Ii and III). This means there is little fluctuation in the acoustic-phonetic characteristics of segments in the IDS an infant receives during this period. The only significant decrease in difference that occurred in vowel duration shown in paper III appeared between months three and four. Since the MIPhI model (Sundberg, 1998) specifically predicts a developmental change in IDS that did not occur in the current production study (paper II and III), the model's developmental predictions were not supported from the current research. Anyhow, other qualities of Norwegian IDS, such as overspecification of vowel duration support the MIPhI model, and it serves a solid explanatory basis for IDS research. In paper III, adjustments are suggested (paper III) to complement the foundation of the

MIPhI model (Sundberg, 1998). While some adjustments in IDS are language-general, some may be language-specific. Findings from the current production study of Norwegian IDS (paper I-III), support the inclusion of the visual modality of speech as a language-general means of overspecification of IDS.

IDS in different settings

The method used to study acoustic-phonetic attributes of IDS, as well as perception by infants, may be important for the research outcome. The findings from the current production study (papers I-III), together with the results for vowels reported in Burnham et al., (2002), both included a shifted vowel space. While not discussed in Burnham et al., (2002), both these studies recorded mothers at home, in natural interaction. This may be the cause of the proposed enhancement of visual cues in IDS being elicited by the interactive setting.

The idea that the quality of IDS is different in different settings leads us to rethink the idea of possible better learnability from IDS stimuli (DeBoer & Kuhl, 2003). Different settings may inherently make available different aspects of speech to an infant, and the face-to-face contact of the early months would provide an opportunity for an infant to be familiarised with the visual cues of articulatory movement in speech through continuous speech. At the same time, when an infant is sitting up or even walking, interaction between an infant and adult may be directed toward different kinds of play with toys, singing or reading. These situations may provide the opportunity for an infant to focus on auditory cues as well as audio-visual cues, for example, hearing a toy's

name while it is being played with involving the sound of the word and the sight of the object. In these terms, an age related effect might be shown where the underlying effect may be the interactional setting. Whereas a four-month-old infant may receive IDS which promotes visual cues, a twelve-month-old may receive IDS which promotes auditory cues, because the IDS was encountered in different settings.

CONCLUSIONS AND FURTHER RESEARCH

Norwegian IDS differs from ADS

One of the aims of the current research was to uncover if segments of Norwegian IDS are different from those of ADS. The conclusions from paper I, II and III included that Norwegian follows the results from other languages by showing that IDS differs from ADS. Acoustic-phonetic characteristics of both vowels and consonants differed between speech types. Vowel spaces are smaller in IDS than ADS, and instead of an expansion, they are shifted. Vowel duration, fricative /s/ duration and VOT are longer in IDS than ADS.

Perceptual cues in IDS

The durational differences between speech types observed in studies I-III for vowels and consonants indicate an auditory enhancement of speech in IDS. An additional interpretation of the current findings is that IDS enhances visual cues to infants. Direct investigation of this would come in an extension of the current research by simultaneously recording and videotaping mothers' speech to small infants. Caution

needs to be taken, however, to still keep the situational context as natural as possible. Furthermore, videotaping the infant's face at the same time would give additional information on how the infant responds to stimuli that are adapted to enhance visual cues from its own mother. In addition, one could extend the current research to cover other vowels and consonants than those studied in the current thesis. For example, studying differences between speech types for other front vowels in Norwegian could verify if the exception of the /i:/-/i/ is carried over to other front vowels, for example, /y/. This would bring us further in verifying the visual cues explanation.

Perception of IDS

Although different from ADS, IDS was preferred by infants when audio-visually presented (paper IV). However, the perception study did not verify whether the possible visual enhancements in Norwegian IDS affect infants' perception of audio-visual congruence in syllables. However, the finding that infants preferred IDS to ADS when IDS had auditory and emotional comparable characteristics, justifies continued speculations that visual articulatory cues may be enhanced in Norwegian IDS.

Experimental situation

The findings in the production study (paper I-III), together with previous research on IDS (e.g., Burnham et al., 2002; Kuhl et al., 1997) indicate that the setting in which one studies IDS is important for the outcome. In the current production study, a natural setting was used for doing recordings. With the privacy of their own homes, being alone when recorded, the possible effects on the quality of the mothers IDS due to being

under study were considered minimal. This may have resulted in a high degree of naturalness in the study. The same can be said for the perception study (paper IV), where infants were video-recorded in their homes while doing the visual fixation task. However, to verify the anticipation of differences between a natural setting and a laboratory, a study need to be conducted where the same participants in both situations, where mothers' IDS are recorded/infants' perception are studied both at home and in a laboratory setting.

Changes in IDS over time

In the current study, findings based on the longitudinal development of segmental aspects of IDS from the same mothers have added to what we know about development in IDS over the first six months. For the acoustic-phonetic characteristics of segments studied here, small changes were evident over time. There are, however, important periods of development after these first months, for example, the period in which infants start babbling and speaking their first words (Vihman, 1996). A longitudinal study that extends beyond the six months studied here would give valuable information about what happens with acoustic-phonetic characteristics of segments in IDS through these important periods as language continues to develop.

Conclusive summary

With a well developed perceptual system, infants come into the world already having some experience with speech. Both attention to and preference for IDS in infancy may be related to its positive affective function, partly revealed through facial expressions in

face-to-face interaction. Explaining some parts of the results from Norwegian IDS, the MIPhI model presupposes a close adaptation between a mother's speech and an infant's perception. In addition, results from the current research show a malleability in IDS that are potentially tied, among other things, to the setting in which it is studied. The acoustic-phonetic characteristics of IDS speech may affect how well infants perceive speech sounds and consequently make speech easy to learn. Throughout the first six months, the IDS spoken in face-to-face interaction in a natural setting, has acoustic-phonetic characteristics that are probably adapted to an infant's emotional, social and neurological prerequisites and closely related to characteristics of the language at hand. Related to these prerequisites, infants' speech perception may benefit from both auditory and visual enhancement of speech.

Round off

The current findings for acoustic-phonetic characteristics of Norwegian IDS contributes to the study of the possible benefit IDS may have for an infant, by proposing joint use of enhancement of auditory and visual cues in IDS. The ecologically valid methods have in the current thesis proved beneficial to uncover fine-grained aspects of speech in mother-infant interaction, and infant processing of speech. Selectively distributing acoustic-phonetic characteristics of IDS points in the direction of parents having a sensitive ability to adapt the speech to what may be considered to be infant needs.

In the introduction, Confucianism was interpreted as possibly labelling IDS as being 'blameworthy' (Chang, 1997). From a different philosophical point of view, IDS can be

seen as very useful. Originating from the teachings of Siddhartha Gautama, around 500 BC, Zen Buddhism (reviewed in DeMartino, 1983) provides a very different philosophical point of view on speech communication. Applying Zen terms on IDS would characterise it as transmission from mind to mind, containing intuitivism, circularity and repetitiveness. IDS is repetitive in nature (Snow, 1972) and turn-taking in IDS exemplifies circularity (Boysson-Bardies, 1999). Therefore, in Zen terms, IDS will by its intuitivism make words, feelings and thoughts develop in the course of communication (De Martino, 1983).

REFERENCES

Aldridge, M.A., Stillman, R.D., Bower, T.G.R. (2001). Newborn categorization of vowel-like sounds. *Developmental Science*, 4(2), 220-232.

Arnold, S.L. (2002). Maternal tactile-gestural stimulation and infants' nonverbal behaviors during early mother-infant face-to-face interactions: Contextual, age, and birth status effects. *PhD thesis*, Concordia University, Montreal, Qubec, Canada.

Aslin, R.N. & Pisoni, D.B. (1980). Some developmental processes in speech perception. In: G.H. Yeni-Komshian, J.F. Kavanagh, & C.A. Ferguson (Eds.) Child Phonology, Vol. 2, Perception, New York, Academic Press.

Bahrick, L.E., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental Psychology,* 36, 190-201.

Bahrick, L.E., & Lickliter, R. (2002). Intersensory redundancy guides early perceptual and cognitive development. In: R.Kail & H.Reese (Eds.) *Advances in child development and behavior,* 30, 153-87. New York, Academic Press.

Bandura, A. (1989). Social cognitive theory. In: R. Vasta (Ed.), Annals of child development: Six theories of child development, Vol. 6, Greenwich, CT, JAI Press.

Baran, J.A., Zlatin Laufer, M., & Daniloff, R. (1977). Phonological contrastivity in conversation: A comparative study of voice onset time, *Journal of phonetics, 5*, 339-350.

Barry, W.J. & Fourcin, A.J. (1992). Levels of labelling. *Computer Speech and Language*, 6, 1-14.

Berko Gleason, J. (1975). Fathers and other strangers: men's speech to young children. *Georgetown University Roundtable on Language and Linguistics.* Washington, Dc, Georgetown University Press.

Bernstein Ratner, N. (1986). Durational cues which mark clause boundaries in mother-child speech. *Journal of Phonetics,* 14, 303-309.

Bernstein Ratner, N. (1984). Patterns of vowel modification in mother-child speech. *Journal of Child Language,* 11, 557-578.

Bernstein Ratner, N. & Luberoff, A.(1984). Cues to post-vocalic voicing in mother-child speech. *Journal of Phonetics*, 12, 285-289.

Bertelson, P., Vroomen, J., deGelder, B. & Driver, J. (2000). The ventriloquist effect does not depend on the direction of deliberate visual attention. *Perception and Psychophysics,* 62(2), 321-332.

Bijeljac-Babic, R., Bertoncini, J. & Mehler, J. (1993). How do four-day-old infants categorise multisyllabic utterances? *Developmental Psychology,* 29, 711-721.

Boer, B de, Kuhl, P.K. (2003). Investigating the role of infant-directed speech with a computer model. *Acoustics Research Letters On-line* 4(4) 129–134.

Bohn, O.S. & Polka, L. (2001). Target spectral, dynamic spectral, and duration cues in infant perception of German vowels. *Journal of the Acoustical Society of America,* 110, 1, 504-515.

Borduin, C.M. & Henggeler, S.W. (1981). Social class, experimental setting and task characteristics as determinants of mother-infant interaction. *Developmental Psychology,* 17(2), 209-214.

Bornstein, M.H., Haynes, O.M., Painter, K.M. & Genevro, J.L. (2000). Child language with mother and with stranger at home and in the laboratory: a methodological study. *Journal of Child Language,* 27, 407-420.
Bortfeld, H. (2004). Which came first: Infants learning language or Motherese? *Behavioural and Brain Sciences,* 27, 505-506. 491-541.
Boysson-Bardies, B. (1999). How language comes to children – from birth to two years. The MIT Press.
Bruner, J.S. (1982). Formats of language acquisition. *American journal of semiotics,* 1, 1-16.
Burnham, D. (1986). Development of speech perception: Exposure to and experience with a first language. *Applied Psycholinguistics, 7*(3), 207-39.
Burnham, D., Kitamura, C. & Vollmer-Conna, U. (2002). What's new pussycat? On talking to babies and animals. *Science,* 296, May.
Bushnell, I.W.R., Sai, F., & Mullin, J.T. (1989). Neonatal recognition of the mother's face. *British Journal of Developmental Psychology,* 7, 3-15.
Calvert, G.A. (2001). Crossmodal processing in the human brain: Insights from functional neuroimaging studies. *Cerebral Cortex,* 11, 12, 1110-1123.
Cameron-Faulkner, T., Lieven, E. & Tomasello, M. (2003). A construction based analysis of child directed speech. *Cognitive Science,* 27, 6, 843-873.
Chang, Hui-ching (1997). Language and words: Communication in the analects of Confucius. *Journal of Language and Social Psychology.* 16, (2), 107-131.
Chemero, A.P. (1999). How to be an anti-representationalist. PhD thesis, Indiana University, USA.
Chomsky, N. (1965). Aspects of the theory of syntax. Cambridge, MA, MIT Press.
Chomsky, N. (1959). A review of B.F. Skinner's Verbal Behavior. *Language,* 35, 26-58.
Chong, S.C.F., Werker, J.F., Russell, J.A. & Carroll, J.M. (2003). Three facial expressions mothers direct to their infants. *Infant and Child Development,* 12, 211-232.
Christensen, J.E. & Lingwall, J.B. (1982). Verbal contingent stimulation of stuttering in laboratory and home settings. *Journal of fluency disorders,* 7, 359-368.
Cooper, R.P.; Abraham, J.; Berman, S.; Staska, M. (1997). The development of infants' preference for motherese. Infant behavior and development, 20 (4), 477-488.
Cooper, R.P.., & Aslin, R.N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development,* 61, 1584-1595.
Corsini, R.J. (Ed.) (1994). Encyclopaedia of psychology. New York, Wiley & Sons.
Cross, T.G. (1977). «Mothers' speech adjustments: the contributions of selected child listenervariables. In: C.E. Snow and C.A. Ferguson (Eds.) Talking to children: language input and acquisition. Cambridge University Press.
Damper, R. I., & Harnad, S. R. (2000). Neural network models of categorical perception. *Perception and Psychophysics,* 62(4), 843-67.
DeCasper, A.J., & Fifer, W.P. (1980). Of human bonding: newborns prefer their mother's voice. *Science,* 208, 1174-1176.
DeCasper, A.J., Lecanuet, J.P., Busnel, M.C., Granier-Deferre, C., & Maugeais, R. (1994).Fetal reaction to recurrent maternal speech. *Infant Behaviour and Development,* 17, 159-164.
DeCasper, A.J. & Spence, M. (1986). Prenatal maternal speech influences newborns'

perception of speech sounds. *Infant Behaviour and Development,* 9, 133-150.
DeMartino, R. (1983). On Zen communication. *Communication,* 8, 13-28.
Desjardins R.N. & Werker, J.F. (2004). Is the integration of heard and seen speech mandatory for infants? Developmental Pscyhobiology, 45(4).
Dominey, P.F. & Dodane, C. (2004). Indeterminacy in language acquisition: the role of child directed speech and joint attention. *Journal of Neurolinguistics,* 17, 2-3, 121-145.
Duez, D. (2001). The acoustic and phonetic manifestations of the reduction and contextual assimilation of conversational speech segments. Revue Parole, 17-19, 89-111.
Eilers, R.E. (1977). Context-sensitive perception of naturally produced fricative and stop consonants by infants. Journal of the Acoustical Society of America, 61, 1321-1336.
Eilers, R. E., & Minifie, F. D. (1975). Fricative discrimination in early infancy. Journal of Speech and Hearing Research, 18, 158-167.
Eimas, P.D. (1975). Auditory and phonetic coding of the cues for speech: Discrimination of the [r-l] distinction by young infants. Perception and Psychophysics, 18, 341-347.
Eimas, P.D., and Miller, J.L. (1980). Discrimination of the information for manner of articulation. Infant Behaviour and Development, 3, 367-375.
Eimas, P.D., Siqueland, E.R., Jusczyk, P.W., Vigorito, J. (1971). Speech perception in infants. *Science*, 171, 303-306.
Elman, J. (1993). Learning and development in neural networks: The importance of starting small. *Cognition,* 48(1), 71-99.
Englund, K. (2005). Voice onset time in infant directed speech over the first six months. *First Language,* 25, 2.
Englund, K. & Behne, D. (2005). Infant directed speech in natural interaction – Norwegian vowel quantity and quality. *Journal of Psycholinguistic Research,* 34, 3.
Englund, K. & Behne, D. (submitted). Changes in infant directed speech in the first six months.
Eswaran, H., Wilson, J.D., Preissl, H., Robinson, S.E., Vrba, J., Murphy, P., Rose, D.F. & Lowery, C.L. (2002). Magnetoencephalographic recordings of visual evoked brain activity in the human foetus. *The Lancet,* 360, 9335, 779-780.
Fairbanks, G. & Pronovost, W. (1939). An experimental study of the pitch characteristics of the voice during the expression of emotions. *Speech Monographs,* 6, 87-104.
Falk, D. (2004). Prelinguistic evolution in early hominins: Whence Motherese? *Behavioural and Brain Sciences,* 27, 491-541.
Fernald, A. (1989). Intonation and communicative intent in mother's speech to infants: Is the melody the message? *Child Development, 60,* 1497-1510.
Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development,* 8, 181-195.
Fernald, A. & Kuhl, P.K. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development,* 10(3), 279-293.
Fernald, A., Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, 20 (1), 104-113.

Fernald, A., Taeschner, T., Dunn, J., Papoušek, M. & DeBoysson-Bardies, B (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language,* 16(3), 477-501.
Fintoft, K. (1961). The duration of some Norwegian speech sounds. *Phonetica,* 7, 19-39.
Gallaway, C., Richards, B. (Eds.) (1994). Input and interaction in language acquisition. Cambridge: Cambridge University Press.
Garnica, O. (1977). Some prosodic and paralinguistic features of speech to young children. In: C.E. Snow & C. A. Ferguson (Eds.) *Talking to children: Language input and acquisition.* New York, Cambridge University Press.
Ghim, H.-R. (1990). Evidence for perceptual organization in infants: Perception of subjective contours by young infants. *Infant Behaviour and Development,* 13, 221-248.
Gibson, E.J. (1969). Principles of perceptual learning and development. New York, Appleton.
Gogate, L.J., Bahrick, L.E. &Watson, J.D. (2000). A study of multimodal Motherese: the role of temporal synchrony between verbal labels and gestures. *Child Development,* 71(4), 878-894.
Golinkoff, R. & Ames, G. (1979). A comparison of mothers' and fathers' speech to young children. *Child development,* 50, 28-32.
Goodsit, J.V., Morse, P.A., Ver Hoeve, J.N., & Cowan, N. (1984). Infant speech recognition in multisyllabic contexts. *Child Development,* 55, 903-910.
Goren, C.C., Sarty, M., & Wu, P.Y.K. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics,* 56, 544-549.
Gottlieb, G. (1991a). Experiential canalization of behavioural development: Theory. *Developmental Psychology,* 27(1), 4-13.
Gottlieb, G. (1991b). Experiential canalization of behavioural development: Results. *Developmental Psychology,* 27(1), 35-39.
Gottlieb, G. (1968). Species recognition in ground-nesting and hole-nesting ducklings. *Ecology,* 49, 87-95.
Gottlieb, G. (1961). Developmental age as a baseline for determination of the critical period in imprinting. *Journal of Comparative and Physiological Psychology,* 54, 422-427.
Gould, J.L., & Marler, P. (1987). Learning by instinct. *Scientific American,* 256, 62-73.
Grieser, D. & Kuhl, P.K. (1989). Categorization of speech by infants: Support for speech-sound prototypes. *Developmental Psychology,* 25, 577-588.
Haggan, M. (2002). Self-reports and self-delusion regarding the use of Motherese: implications from Kuwaiti adults. *Language Sciences,* 24(1), 17-28.
Haynes, W.O. & Shulman, B.B. (Eds.) (1998). *Communication Development*, *foundations, processes and clinical applications.* Pennsylvania, Prentice Hall.
Hainline, L. (1998). The development of basic visual abilities. In: A. Slater (Ed.). Perceptual development. Visual, auditory , and speech perception in infancy. Psychology Press, East Sussex, UK.
Hayashi, A., Tamekawa, Y. & Kiritani, S. (2001). Developmental change in auditory preferences for speech stimuli in Japanese infants. *Journal of Speech, Language and Hearing Research,* 44(6), 1189-1200.

Hepper, P.G. , Shahidullah, B.S. (1994). Development of foetal hearing. *Archives of Disease in Childhood* - Fetal and Neonatal Edition, 71, F81-F87.
Hertenstein, M.J. (2002). Touch: Its communicative functions in infancy. *Human Development,* 45(2), 70-94.
Holmberg, T.L., Morgan, K.A., & Kuhl, P.K. (1977). Speech perception in early infancy: Discrimination of fricative consonants. Paper presented at the meeting of the Acoustical Society of America, Miami Beach, Fla., December.
Hsu, H.C., Fogel, A. & Messinger, D.S. (2001). Infant non-distress vocalization during mother-infant face-to-face interaction: Factors associated with quantitative and qualitative differences. *Infant Behavior & Development,* 24, 107-128.
Iverson, P. & Kuhl, P.K. (1995). Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. *Journal of The Acoustical Society of America,* 97(1), 553-562.
Johnson, M.H. & Morton, J. (1991). Biology and cognitive development. Oxford, Blackwell.
Jusczyk, P.W. (1993). How word recognition may evolve from infant speech perception capacities. In: G. Altman & R. Shillcock (Eds.) *Cognitive models of speech processing: The second Sperlonga meeting,* 27-55, Mawah, NJ, Lawrence Erlbaum Associates.
Jusczyk, P.W. (1997). *The discovery of spoken language.* Cambridge, Massachusetts, MIT Press.
Jusczyk, P.W., Copan, H. & Thompson, E. (1978). Perception by two-mont-olds of glide contrasts in multisyllabic utterances. *Perception and Psychophysics,* 24, 515-520.
Jusczyk, P.W. & Derrah, C. (1987). Representation of speech sounds by young infants. *Developmental Psychology,* 23, 648-654.
Kaplan, P.S., Jung, P.C., Ryther, P.C. & Zarlengo-Strouse, P (1996). Infant-directed versus adult-directed speech as signals for faces. *Developmental Psychology,* 32, 1-12.
Kellogg, R.T. (2003). Cognitive psychology, 2nd ed. London, U.K. Sage Publications.
Kent, R., (1997). The Speech Sciences. Singular Publishing Group, London, G.B.
Kent, R. & Read, C. (1992). *The acoustic analysis of speech*. San Diego, California, Singular Publishing Group.
Kitamura, C., Thanavishuth, C., Burnham, D., Luksaneeyanawin, S. (2002). Universality and specificity in infant-directed speech: Pitch modifications as a function of infant age and sex in a tonal and non-tonal language. *Infant Behavior and Development,* 24(4), 372-392.
Krøjgaard, P. (2001). Økologisk validitet og ekserimentell spædbarnsforskning/Ecological validity and experimental infancy research. *Psyche and Logos*, 22, 635-661.
Klatt, D. (1975). Vowel lengthening is syntactically determined in connected discourse. *Journal of Phonetics,* 3, 129-140.
Kristoffersen, G. (2000). *The phonology of Norwegian*. Oxford, Oxford University Press.
Kuhl, P.K. (1993a). Infant speech perception: a window on psycholinguistic development. *International Journal of Psycholinguistics,* 9(1), 33-56.

Kuhl, P.K. (1993b). Early linguistic experience and phonetic perception: Implications for theories of developmental speech perception. *Journal of Phonetics,* 21(1-2), 125-139.
Kuhl, P.K. (1991). Human adults and human infants show a "perceptual magnet effect", for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics,* 50, 93-107.
Kuhl, P.K. (1983). Perception of auditory equivalence classes for speech in early infancy. *Infant Behavior and Development,* 6, 263-285.
Kuhl, P.K., Andruski, J.E.. Chistovich, I.A., Chistovich, L.A., Kozhevnikova, E.V., Ryskina, V.L., Stolyarova, E.I., Sundberg, U. & Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science,* 277, 684-686.
Kuhl, P.K., & Meltzoff, A.K. (1982). The bimodal perception of speech in infancy. Science, 218, 1138-1141.
Kuhl, P.K., & Miller, D. (1982). Discrimination of auditory target dimensions in the presenceor absence of variation in a second dimension by infants. *Perception and Psychophysics,* 31(3), 279-292.
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., Klintfors, E., Gustavsson, L., Lagerkvist, L., Marklund, E. & Sundberg, U. (2004). Department of Linguistics, Stockholm University, Sweden.
Lacerda, F. & Lindblom, B. (1997). Modelling the early stages of language acquisition. In: Olofsson, Å & Strömqvist (Eds.) COST-A8, Cross-linguistic studies of dyslexia and early language development, European Commission, 14-33.
Lacerda, F. & Sundberg, U. (2001). Biases in early language acquisition. In: F.Lacerda, C von Hofsten, & M. Heimann (Eds.) Emerging cognitive abilities in early infancy. Mahwah, New Jersey, Lawrence Erlbaum Associates.
Laver, John (1995). Principles of phonetics. *Cambridge University Press Great Britain.*
Lecanuet, J.-P., & Granier-Deferre, C. (1993). Speech stimuli in the foetal environment. In: B de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage, & J. Morton (Eds.) Developmental neurocognition: Speech and face processing in the first year of life. Dordrecht, The Netherlands, Kluwer Academic Publishers.
Levitt, A., Jusczyk, P.W., Murray, J. & Carden, G. (1988). The perception of place of articulation contrasts in voiced and voiceless fricatives by two-month-old infants. *Journal of Experimental Psychology: Human Perception and Performance,* 14, 361-368.
Lewkowicz, D.J. (1986). Developmental changes in infants' bisensory response to synchronous durations. *Infant Behavior and Development,* 9(3), 335-353.
Lewkowicz, D.J. & Turkewitz, G. (1980). Cross-modal equivalence in early infancy: Auditory-visual intensity matching. *Developmental Psychology,* 16, 597-607.
Lickliter, R., Bahrick, L., & Honeycutt, H. (2002). Intersensory Redundancy Facilitates Prenatal Perceptual Learning in Bobwhite Quail (Colinus virginianus) Embryos. *Developmental Psychology,* 38(1), 15-23.
Lieven, E.V.M. (1994). Crosslinguistic and crosscultural aspects of language addressed to children. In: Gallaway, C. & Richards, B.J. (Eds.) Input and interaction in language acquisition. Cambridge university press, Cambridge, Great Britain.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In: W.Hardcastle & A. Marchal (Eds.) *Speech Production and Speech Modelling*. Kluwer Dordrecht, 403-439.
Lindblom, B. (1992). Phonological units as adaptive emergents of lexical development. In: C.A. Ferguson, L. Menn & C. Stoel-Gammon (Eds.) Phonological development. Models, research and implications. Timonium, MD, York Press.
Liu, H.M., Kuhl, P.K. & Tsao, F.M. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science,* 6(3), F1-F10.
Malsheen, B.J. (1980). Two hypotheses for phonetic clarification i the speech of mothers to children. In: G.H. Yeni-Komshian, J.F. Kavanagh, C.A. Ferguson, *Child phonology*, Vol. 2, Perception, 173-84, San Diego, Academic Press.
MacKain, K., Best, C. & Strange, W. (1981). Categorical perception of English /r/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics,* 2, 369-390.
MacKain, K., Studdert-Kennedy, M., Spieker, S., & Stern, D. (1983). Infant intermodal speech perception is a left hemisphere function. *Science,* 28, 396-405.
Marler, P. (1991) (ed.). *The instinct to learn.* Hillsdale, N.J. Erlbaum.
McGurk, H. & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature,* 264, 746-748.
Meltzoff, A.N., & Moore, M.K. (1977). Imitation of facial and manual gestures by human neonates. *Science,* 198, 75-78.
Meredith, M.A., Wallace, M.T. & Stein, B.E. (1992). Visual, auditory and somatosensory convergence in output neurons of the cat superior colliculus: multisensory properties of the tecto-reticulospinal projection. *Experimental Brain Research,* 88, 181-186.
Miall, D. (2003). The poetics of babytalk. *Human nature,* 14(4), 337-364.
Miall, D. (2001). Sound of contrasts: An empirical approach to phonemic iconicity. *Poetics,* 29, 55-70.
Miller, J. (1981). Effects of speaking rate on segmental distinctions. In: P. D. Eimas & J. L. Miller, *Perspectives on the study of speech*. Hillsdale, NJ, Erlbaum.
Miller, J.L. & Eimas, P.D. (1983). Studies on the categorization of speech by infants. *Cognition,* 13, 135-165.
Moffit, A. (1971). Consonant cue perception by twenty- to twenty-four-week-old infants. *Child Development,* 42(3), 717-731.
Morgan, J.L. & Demuth, K. (1996). Signal to syntax: An overview. In: J.L. Morgan & K. Demuth (Eds.) Signal to syntax. Bootstrapping from speech to grammar in early acquisition. Mahwah, NJ, Lawrence Erlbaum Associates.
Morse, P.A. (1972). The discrimination of speech and nonspeech stimuli in early infancy. *Journal of Experimental Child Psychology,* 13, 477-492.
Muir, D.W., Hains, S.M.J. (1993). Infant sensitivity to perturbations in adult facial, vocal, tactile, and contingent stimulation during fact-to-face interaction. In: B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage, J. Morton (Eds.) Developmental neurocognition: Speech and face processing in the first year of life. New York, US, Kluwer Academic/Plenum Publishers.
Niwano, K., & Sugai, K. (2003). Pitch characteristics of speech during mother-infant and father-infant vocal interactions. *Japanese Journal of Special Education,* 40(6), 663-674.

Ochs, E. (1982). Talking to children in Western Samoa. *Language in Society,* 11, 77-104.
Papoušek H. & Papoušek, M. (1995). Intuitive parenting . In: M.H. Bornstein (Ed.) Handbook of Parenting, Vol. 2, Biology and Ecology of Parenting. 148-163. Mahwah, NJ, Lawrence Erlbaum Associates.
Papoušek, M., Papoušek, H. & Symmes, D. (1991). The meanings of melodies in motherese in tone and stress languages. *Infant Behavior and Development,* 14, 415-440.
Patterson, M.L. & Werker, J.F. (2003). Two-month-old infants match phonetic information in lips and voice. *Developmental Science,* 6 (2), 191-196.
Patterson, M.L. & Werker, J.F. (1999). Matching phonetic information in lips and voice is robust in 4.5-month-old infants. *Infant Behavior and Development,* 22, 237-247.
Peterson, G.E. & Barney, H.L. (1952). Control methods used in a study of the vowels. *The Journal of The Acoustical Society of America,* 24 (2), 175-184.
Piaget. J. (1952). *The construction of reality in the child.* New York. Basic Books.
Polka, L. & Werker, J.F. (1994). Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance,* 20, 421-436.
Printer, R., Freeman, S., Perez, R., Sohmer, H. (2003). The neonate has a temporary conductive hearing loss due to fluid in the middle ear. *Audiology & neuro-Otology,* 8(2), 100-110.
Pye, C. (1986). Quiché Mayan speech to children. *Journal of Child Language,* 13, 85-100.
Rosenblum, L., Schmuckler, M.A., Johnson, J.A. (1997). The McGurk effect in infants. *Perception and Psychophysics,* 59 (3), 347-357.
Ryalls, J. (1996). A basic introduction to speech perception – speech science series. San Diego, California. Singular Publishing group.
Schmuckler, M.A. (2001). What is ecological validity? A dimensional analysis. *Infancy*, 2(4), 419-436.
Shahidullah, S. & Hepper, P.G. (1994). Frequency discrimination by the fetus. *Early Human Development,* 36(1), 13-26.
Shahidullah, S. & Hepper, P.G. (1993).The developmental origins of fetal responsiveness to an acoustic stimulus. *Journal of reproductive and infant psychology,* 11(3), 135-142.
Sheier, C., Lewkowicz, D. & Shimojo, S. (2003). Sound induces perceptual reorganization of an ambiguous motion display in human infants. *Developmental Science,* 6(3), 233-241.
Shor, R.E. (1978). The production and judgement of smile magnitude. *Journal of General Psychology,* 98(1), 79-96.
Shulman, B.B. (1998a). Child development. In: W.O. Haynes & B.B. Shulman (Eds.) Communication development. Foundations, processes and clinical applications. Pennsylvania, Prentice Hall.
Shulman, B.B. (1998b). Theories of communication development. In: W.O. Haynes & B.B. Shulman (Eds.) Communication development. Foundations, processes and clinical applications. Pennsylvania, Prentice Hall.

Shute, B. & Wheldall, K. (1999). Fundamental frequency and temporal modifications in the speech of British fathers to their children. *Educational psychology,* 19 (2), 221-233.
Singh, I., Morgan, J.L. & Best, C.T. (2002). Infants' listening preferences: Baby talk or happy talk? *Infancy,* 3(3), 365-394.

Skinner, B.F. (1957). Verbal behaviour. New York, Appleton-Century-Crofts.
Slaney, M. & McRoberts, G. (2003). Baby Ears: A recognition system for affective vocalisations. *Speech Communication,* 39, 3-4, 367-384.
Slater, A., Mattock, A., & Brown, E. (1990). Size constancy at birth: Newborn infants' responses to retinal and real size. *Journal of Experimental Child Psychology,* 49, 314-322.
Slater, A. & Morison, V. (1985). Shape constancy and slant perception at birth. *Perception,* 14, 337-344.
Snow, C.E. (1977). Mother's speech research: from input to interaction. In: C.E. Snow and C.A. Ferguson (Eds.) *Talking to children. Language input and acquisition*, Cambridge, Cambridge University Press.
Snow, C.E. (1972). Mothers' speech to children learning language. *Child development*, 43, 549-65.
Spelke, E.S. (2000). Nativism, Empiricism, and the origins of knowledge. In: D.Muir & A.Slater (Eds.) Infant development, the essential readings. Oxford, UK, Blackwell Publishers.
Spence, M.J. & DeCasper, A.W. (1982). Human foetuses perceive maternal speech. *Paper delivered at the International Conference on Infant Studies,* March, Austin, Texas.
Spence, M.J., Freeman, M.S. (1996). Newborn infants prefer the maternal low-pass filtered voice, but not the maternal whispered voice. *Infant Behavior and Development,* 19, 199-212.
Stein, B.E., Meredith, M.A. & Wallace, M.T. (1994). Development and neural basis of multisensory integration. In: D.J. Lewkowicz & R. Lickliter (Eds.) The development of intersensory perception: comparative perspectives. Lawrence Erlbaum Associates, Hillsdale, New Jersey.
Stern, D.N.; Spieker, R.K.; Barnett, R.K.; MacKain, K. (1983). The prosody of maternal speech: infant age and context related changes. *Child Language,* 1-15.
Stevenson, M.B., Leavitt, L.A., Roach, M.A., Chapman, R.S, et al. (1986). Mothers speech to their 1-year-old infants in home an laboratory settings. *Journal of Psycholinguistic Research,* 15(5), 451-461.
Sundberg, U. (2001). Consonant specification in infant-directed speech. Some preliminary results from a study of voice onset time in speech to one-year-olds. Lund University, Department of Linguistics, *Working Papers,* 49, 148-151.
Sundberg, U. (1998). Mother tongue - phonetic aspects of infant-directed speech. Dissertation, PERILUS, Stockholm.
Sundberg, U., Lacerda, F. (1999). Voice Onset Time in speech to infants and adults. *Phonetica,* 56, 3-4, 186-99.
Swoboda, p., Morse, P.A., & Leavitt, L.A. (1976). Continous vowel discrimination in normal and at-risk infants. *Child Development,* 47, 459-465.
Tartter, V.C. (1980). Happy talk: Perceptual and acoustic effects of smiling on speech.

Perception and Psychophysics, 27, 24-27.
The Helsinki Declaration. The 18th World Medical Assembly, Helsinki, Finland, June 1964. Amended by The 41st WMA, Hong Kong, September 1989. http://onlineethics.org/reseth/helsinki.html
The Nuremberg Code. Official trial record: *Trials of War Criminals before the Nuremberg Military Tribunals under Control Council Law No. 10.* Nuremberg, October 1946 - April 1949. Washington D.C.: U.S. G.P.O, 1949-1953. http://www.ushmm.org/research/doctors/index.html
Tomasello, M. (1992). The social bases of langauge acquisiton. *Social Development,* 1(1), 67-87.
Tomasello, M. & Todd, J. (1983). Joint attention and lexical acquisition style. *First Language,* 4, 3(12), 197-211.
Trainor, L.J., Austin, C.M. & Desjardins, R.N. (2000). Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychological Science,* 11, 3.
Trehub, S.E. (1973). Infants' sensitivity to vowel and tonal contrasts. *Developmental Psychology,* 9, 91-96.
Tsao, F.M., Liu, H.M. & Kuhl, P.K. (2004) Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child Development,* 75(4), 1067-1084.
Turati, C., Valenza, E, Leo, I. & Simion, F. (2005). Tree-month-olds' visual preference for faces and its underlying visual processing mechanisms. *Journal of Experimental Child Psychology,* 90, 3, 255-273.
Vihman, M.M. (1996). Phonological development. The origins of language in the child. Cambridge, Massachusetts, USA, Blackwell Publishers.
Vouloumanos, A. & Werker, J.F. (in press). Tuned to the signal: the privileged status of speech for young infants. *Developmental Science.*
Vroomen, J., Bertelson, P., DeGelder, B. (2001). The ventriloquist effect does not depend on the direction of automatic visual attention. *Perception & Psychophysics,* 63(4), 651-659.
Webster's Encyclopedic Unabridged Dictionary of the English Language. New revised addition (1989). Gramercy Books, New York.
Weijer, J. van de (2002). How much does an infant hear in a day? In: J. Costa and M. João Freitas (Eds.) *Proceedings of the GALA2001 Conference on Language Acquisition*, 279-282. Lisboa, Associação Portuguesa de Linguistíca.
Werker, J.F., Lloyd, V.L., Pegg, J.E. & Polka. L. (1996). Putting the baby in the bootstraps: Toward a more complete understanding of the role of the input in infant speech processing. Morgan, James L. (Ed); Demuth, Katherine (Ed). Signal to syntax: Bootstrapping from speech to grammar in early acquisition, 427-447. Hillsdale, NJ, England, Lawrence Erlbaum Associates.
Werker, J.F. & McLeod, P.J. (1992). Preference for infant-directed over adult-directed speech: Evidence from 7-week-old infants. *Infant Behavior and Development,* 15(3), 325-345.
Werker, J.F. & McLeod, P.J. (1989). Infant preference for both male and female infant-directed talk: A developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology,* 43, 230-246.

Werker, J.F. & Tees, R.C. (2002). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development,* 25(1), 121-133.
Werker, J.F. & Tees, R.C. (1999). Influences on infant speech processing: Toward a new synthesis. *Annual Review of Psychology,* 50, 509-535
Weppelman, T.L., Bostow, A., Shiffer, R., Elbert-Perez, E. & Newman, R.S. (2003). Children's use of the prosodic characteristics of infant-directed speech. *Language & Communication,* 23(1), 63-80.
Williams, C.E. & Stevens, K.N. (1972). Emotions and speech: some acoustical correlates. *Journal of Acoustical Society of America,* 52(4-2), 1238-1250.