Infant Behavior & Development 36 (2013) 847–862

Contents lists available at ScienceDirect

Infant Behavior and Development

Mommy is only happy! Dutch mothers' realisation of speech sounds in infant-directed speech expresses emotion, not didactic intent

Titia Benders[a,b,*]

[a] Amsterdam Center of Language and Communication, University of Amsterdam, The Netherlands
[b] Center for Language Studies, Radboud University Nijmegen, P.O. Box 9103, 6500 HD Nijmegen, The Netherlands

ARTICLE INFO

Article history:
Received 29 October 2012
Received in revised form 29 August 2013
Accepted 11 September 2013
Available online 13 November 2013

Keywords:
Infant-directed speech
Mother–child interaction
Affective/emotional communication

ABSTRACT

Exaggeration of the vowel space in infant-directed speech (IDS) is well documented for English, but not consistently replicated in other languages or for other speech-sound contrasts. A second attested, but less discussed, pattern of change in IDS is an overall rise of the formant frequencies, which may reflect an affective speaking style. The present study investigates longitudinally how Dutch mothers change their corner vowels, voiceless fricatives, and pitch when speaking to their infant at 11 and 15 months of age. In comparison to adult-directed speech (ADS), Dutch IDS has a *smaller* vowel space, higher second and third formant frequencies in the vowels, and a higher spectral frequency in the fricatives. The formants of the vowels and spectral frequency of the fricatives are raised more strongly for infants at 11 than at 15 months, while the pitch is more extreme in IDS to 15-month olds. These results show that enhanced positive affect is the main factor influencing Dutch mothers' realisation of speech sounds in IDS, especially to younger infants. This study provides evidence that mothers' expression of emotion in IDS can influence the realisation of speech sounds, and that the loss or gain of speech clarity may be secondary effects of affect.

1. Introduction

Caregivers from most cultures use a different speech register for babies than for other adults (see for reviews Cruttenden, 1994; Ferguson, 1977; Soderstrom, 2007). This special way of speaking to an infant expresses positive emotions and maintains the infant's attention, but it also conveys the structure of the language (Ferguson, 1977; Fernald et al., 1989; Uther, Knoll, & Burnham, 2007). Caregivers' positive affect is mostly carried by the pitch characteristics of infant-directed speech (IDS, Trainor, Austin, & Desjardins, 2000; Uther et al., 2007). One linguistic aspect that caregivers from many languages seem to clarify in IDS as compared to adult-directed speech (ADS) is the auditory contrast between the corner vowels[1] (Andruski, Kuhl, & Hayashi, 1999; Bernstein Ratner, 1984; Burnham, Kitamura, & Vollmer-Conna, 2002; Kuhl et al., 1997; Liu, Kuhl, & Tsao, 2003; Uther et al., 2007). Uther et al. (2007) have claimed that the special realisation of speech sounds in IDS is a result of caregivers' teaching efforts, which are independent of any positive affect. The present paper challenges the proposed

* Correspondence to: P.O. Box 9103, 6500 HD Nijmegen, The Netherlands. Tel.: +31 24 3616133.
E-mail address: titia.benders@let.ru.nl

[1] The corner vowels are the vowels produced with the most extreme articulations physically possible. For most languages they are /i/ as in English *sheep*, /u/ as in English *shoe*, and one or two low vowels such as /æ/ as in English *sand* and/or /ɑ/ as in English *shark*.

0163-6383/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.infbeh.2013.09.001

dichotomy between didactic changes to the speech sounds and affective changes to the pitch, and shows that the expression of affect is the main determinant of the realisation of speech sounds in Dutch IDS.

1.1. Enhanced auditory contrasts in IDS

If the auditory contrast between the corner vowels is enhanced in IDS, infants hear clearer examples of their native language's phoneme categories and probably better intelligible and more precisely articulated speech (Bradlow, Torretta, & Pisoni, 1996). It has been hypothesised that mothers enhance speech-sound contrasts out of didactic consideration for their language-learning infant, because they similarly enhance their corner vowels in speech to adult second-language learners (Uther et al., 2007), but not in speech to pets (Burnham et al., 2002; but see Kim, Diehl, Panneton, & Moon, 2006). These clear pronunciations in IDS may indeed promote infant language acquisition, as mothers' degree of vowel-space enhancement in IDS is related to their infants' rate of phoneme perception development (Liu et al., 2003).

The occurrence of such vowel enhancement in English, Swedish, Russian, Japanese, and Mandarin IDS has led to the claim that it is a universal characteristic of this register (Kuhl et al., 1997; Uther et al., 2007). However, the expansion of the vowel space in IDS is not consistently replicated in studies of American English (Green, Nip, Wilson, Mefferd, & Yunusova, 2010) and not found in all languages (Dodane & Al-Tamimi, 2007; Englund & Behne, 2006; Van de Weijer, 2001). In Norwegian, the vowel space is crucially *smaller* in IDS than in ADS (Englund & Behne, 2006). Even if the contrast between the corner vowels is enhanced, this is not necessarily the case for closer vowel pairs or more central vowels (Cristia & Seidl, 2013; McMurray, Kovack-Lesh, Goodwin, & McEchron, 2013). Further evidence against the universality of clear speech in IDS is the dependence of the infant-directed vowel space on infant characteristics: Mothers speak more clearly to normally-hearing infants than to infants that cannot hear their mother as a result of actual or simulated deafness (Lam & Kitamura, 2010, 2012). Lastly, there is little evidence as to whether and how consonants are modified in speech to infants and children (see Cristiá, 2010; Soderstrom, 2007, for reviews). Cristiá (2010) has found that the spectral mean of /s/ is increased in American-English IDS to infants of 13, but not 5 months of age, so that the contrast between /s/ and /ʃ/ is enhanced for the older infants only. The voice-onset time (VOT) of voiced and voiceless plosives have been observed to become consistently shorter as well as consistently longer in IDS, without a change to the contrast (Sundberg & Lacerda, 1999; Englund, 2005). More recent results suggest that specifically the VOT of voiceless plosives gets longer in IDS, which is consistent with the hypothesis that the changes to the speech sounds are a side effect of the slower speaking rate in IDS (McMurray et al., 2013). To summarise, enhanced contrasts in IDS do not occur in all languages, to all children, or for all speech-sound contrasts. Consequently, enhanced auditory contrasts cannot be considered a universal characteristic of IDS.

1.2. Affective speech-sound changes in IDS

A second cross-linguistically attested pattern of change to infant-directed vowels is an overall increase of formant frequencies (Dodane & Al-Tamimi, 2007; Englund & Behne, 2005; Green et al., 2010). Formant frequencies roughly depend on the shape and size of the vocal tract and are independent of the fundamental frequency. As a rule of thumb, the first formant (F1) becomes higher as the mouth opens and the second and third formant (F2 and F3) increase as a results of a shortened vocal tract. During a smile, speakers open their mouth and shorten their vocal tract by retracting their lips (Shor, 1978). The expected acoustic consequence of speaking with a smile is a rise of formant frequencies and this has been observed repeatedly in smiled and happy speech (Tartter, 1980; Tartter & Braun, 1994; Waaramaa, Laukkanen, & Väyrynen, 2008; Zacher & Niemitz, 2003; cf. Fagel, 2010, showing that the acoustic effect of smiling is vowel-dependent, and Aubergé & Cathiard, 2003, suggesting that formants are lower in speech with amused smiles). The raised formant frequencies in happy speech can be regarded as biologically grounded acoustic carriers of positive affect: Animals express hostility with low-frequency sounds, which are associated with large bodies, and appease an opponent with high-frequency sounds, which are more likely to stem from a small body (Morton, 1977). Ohala (1980, 1984) proposes that humans similarly use the frequency-size relationship to signal their intentions and that the smile has become the facial expression of goodwill exactly because its acoustic consequence is a rise of the formant frequencies. IDS is a highly affective register and the predominant facial expression in interactions with infants is a very joyful smile (Chong, Werker, Russell, & Caroll, 2003; Stern, 1974). High formant frequencies of infant-directed vowels could well be a side effect of smiling (Englund & Behne, 2005), or, more generally, the result of caregivers' enhanced positive affect when they speak to their infant.

An investigation of fricatives in IDS can help to further determine whether speech sounds in IDS indeed reflect an affective speaking style. The spectral energy of fricatives is on higher frequencies in emotional speech than in neutral speech (Kienast & Sendlmeier, 2000). A single-parameter measure of the concentration of the energy distribution in a fricative is the centre of gravity (COG, also spectral mean or first spectral moment, Forrest, Weismer, Milenkovic, & Dougall, 1988). If COG is overall higher in IDS than in ADS, this would be further evidence for the influence of positive affect on the realisation of speech sounds in IDS.

Raised formant frequencies in IDS do not necessarily result from an affective speaking style, but could also show that caregivers attempt to imitate their infant (Dodane & Al-Tamimi, 2007): Infants have a smaller vocal tract than adults and therefore they produce their vowels with overall higher formants (Peterson & Barney, 1952). Conveniently, children realise fricatives with most spectral energy on lower frequencies than adults (Nissen & Fox, 2005). The affect hypothesis and the imitation hypothesis for raised formant frequencies in IDS thus provide competing predictions with respect to the realisation