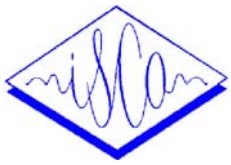


5th International Conference on Spoken Language Processing (ICSLP 98)
Sydney, Australia
November 30 - December 4, 1998

ARE YOU MY LITTLE PUSSY-CAT? ACOUSTIC, PHONETIC AND AFFECTIVE QUALITIES OF INFANT- AND PET-DIRECTED SPEECH

Denis Burnham, Elizabeth Francis, Ute Vollmer-Conna, Christine Kitamura, Vicky Averkiou, Amanda Olley, Mary Nguyen, and Cal Paterson

School of Psychology, University of NSW [1]

ABSTRACT

It has been suggested that infant-directed speech (IDS) is a special speech register conducive to gaining infants' attention, eliciting and expressing affect, and teaching infants about the phonology of the ambient language. Certain similarities have been observed between IDS and pet-directed speech (PDS), but until now these have not been systematically studied. In this study IDS, PDS, and adult directed speech (ADS) are compared on acoustic, phonetic, and affective measures. It has been suggested that IDS should be more tuned to a linguistic didactic function than PDS, but that the two should be similar acoustically and affectively. The results show that IDS and PDS are generally different from ADS, but that IDS and PDS do not differ substantially, either acoustically, or phonetically. As the expected difference between IDS and PDS was not obtained, it could be suggested that dependent variables thought to measure the functions of special speech registers may not necessarily do so unequivocally.

1. INTRODUCTION

Infant-directed speech (IDS) is a special speech register used by parents and others when talking to infants. IDS differs from adult-directed speech (ADS) in its higher overall pitch and pitch variation [2,3], longer duration of periodic components [4], greater affective salience [5], and hyperarticulation of vowels [4,6]. It has been suggested that IDS performs three functions: to attract and maintain attention; to convey and elicit positive affect; and to teach language [7].

Another special speech register thought to share many of the characteristics of IDS is pet-directed speech (PDS). Some studies have been conducted on PDS [8,9], but as yet, no detailed acoustic or phonetic analyses have been undertaken. In this study the speech of mothers speaking to their infant, their pet, and another adult was analysed acoustically and phonetically.

With regard to the three functions evident in IDS, it would be expected that PDS perhaps shares the attentional and affective functions of IDS, but that humans do not use PDS in a linguistic didactic manner, i.e., they do not try to teach their pets to speak [9,10]. It could be suggested that in IDS the higher pitch and greater pitch modulation might be used to attract attention, the greater affective salience to elicit positive affect in the infant, and the hyperarticulated vowels to teach infants about the phonology of the ambient language. Thus here it is expected that compared with ADS, both IDS and PDS should have similarly heightened pitch and pitch modulation, longer vowel duration, and greater affective salience. However, it would be expected that compared with ADS, vowels should be hyperarticulated in IDS, but not in PDS.

2. METHOD

2.1. Participants and Procedure

The speech from 16 mothers of 6-month-old infants, who also had a pet dog or cat were recorded in their home using a lapel microphone and a Professional Walkman. Due to recording errors and some mothers not completing all phases of the recording, only the data of 12 mothers were available for analysis. In the final sample there were 6 mothers of boys (4 with dogs and 2 with cats), and 6 mothers of girls (2 with dogs and 4 with cats).

On the first visit, the mothers were instructed in the use of the tape-recorder, and asked to record their speech to their child for about 10 to 15 minutes, and on a different occasion to their pet for the same period of time. In their speech they were asked to include nominal reference to three provided toys, a *sheep*, a *shoe*, and a *shark*, as often as possible. The tape-recorder was left with mothers for about 3 days to allow recording at a time convenient to them. On the second visit ADS was collected: the experimenter engaged the mother in conversation particularly about the three toys that had been used in the study.

2.2. Data Analysis

There were three groups of dependent variables: acoustic, phonetic, and affective. The first two were based on measurements taken from the vowels, /i/, /u/, and /a/, in the words *sheep*, *shoe*, and *shark*. These words were digitised at 10 kHz from the mothers' 10-15 minute tapes onto the Kay Elemetrics CSL hardware/software package, and then analysed. For the third variable, ratings of low pass filtered samples of the mothers' speech were made by introductory psychology students. Further deathless of each are given below.

Acoustic Analyses. All exemplars (free from background noise) of the three target words in the 10-15 minute samples of speech from each of the 12 mothers were digitised. The number of targets on which measurements were made is shown in Table 1, split for vowel type and speech register. Analysis of F_0 and duration were done via a series of macros in Kay CSL. In these, words were low-pass filtered at 1kHz and impulse (glottal pulse) markers were added, and words were trimmed so that only the voiced portion was analysed. Mean F_0, minimum and maximum F_0 (from which the pitch range measured in semitones was calculated), and duration were then measured. Thus the three acoustic variables were mean-F_0, pitch range, and duration.

Table 1: Exemplars in Acoustic Analysis

	IDS	PDS	ADS	Total
/i/	113	129	79	321
/u/	130	96	79	305
/a/	135	112	75	322
Total	378	337	223	948

Phonetic Analyses. The phonetic analyses centred on the values of the first and second formants in the /i/, /u/, and /a/ vowels in IDS, PDS, and ADS. This is a time-consuming process; to date only the data of 4 of the 12 mothers have been analysed. These mothers' data were chosen for having the clearest speech and the greatest number of target exemplars. There were 3 mothers of girls (2 with cats as pets and 1 with a dog), and 1 mother of a boy (with a pet dog). Analyses were conducted using Kay CSL macros, followed by MS-DOS batch programs based on the CSL library of routines.

Table 2: Exemplars in Phonetic Analysis

	IDS	PDS	ADS	Total
/i/	22	18	15	55
/u/	24	22	17	63
/a/	14	18	20	52
Total	60	58	52	170

From each target word, a portion of the vowel was marked and a narrow band spectrogram with formant plots generated. This spectrogram and the audio playback was used to isolate the largest possible portion of the vowel, which was saved for formant analysis via the batch program. Some vowels could not be used due to fragmented formants, and at the batch analysis stage some further vowels could not be analysed due to insufficient data. The number of exemplars in the final analysis, for vowel and speech register is given in Table 2.

Affective Analyses Affect in the three different speech styles was measured via the ratings of 20 undergraduate introductory psychology students. For these ratings the first 30 secs of continuous speech in IDS, PDS, and ADS from each of the 12 mothers (N = 3 x 12 = 36 samples) were digitised, and low pass filtered at 400 Hz. Twelve of the 36 samples were then recorded onto each of 3 audio-tapes, with 4 mothers x 3 speech types on each tape in quasi-random order. A practice tape with 4 samples (1 of IDS, 1 of PDS, & 2 of ADS) was also prepared. Raters, in small groups of 2 to 4, listened to the practice tape and then the three test tapes with order of presentation of the latter counterbalanced across groups of raters. These informants rated each speech sample using 5 scales of affect developed in our laboratory by Kitamura [11]. Of the 5 scales, the data from two, the intention of the speaker to 'express affection' and to 'direct behaviour' were analysed here, as it has been found that these are the two which contribute most significantly to the Affective and Attentional factors which have been derived from IDS data [11].

3. RESULTS

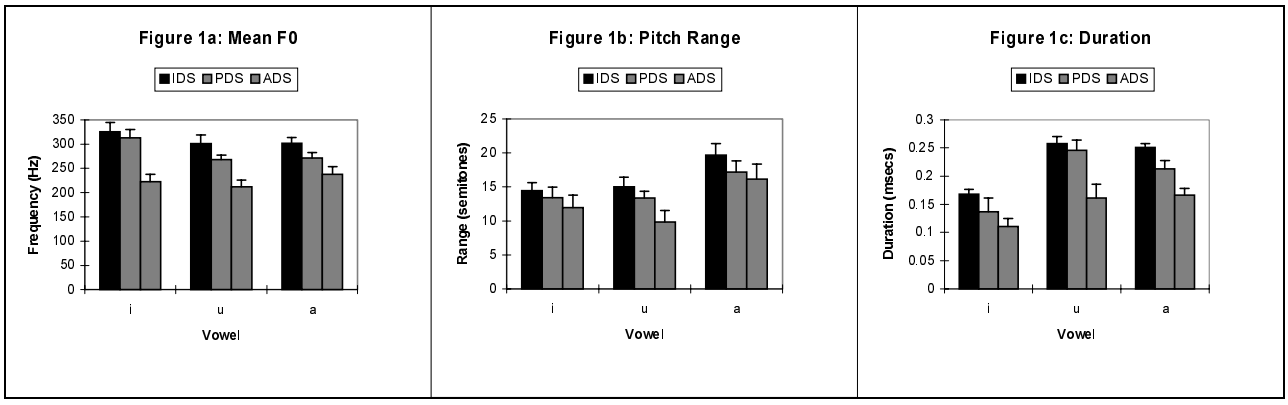
Results for acoustic, phonetic, and affective analyses are reported separately.

3.1 Acoustic Analyses

The data for the three acoustic variables, mean-F_0, pitch range, and duration, are graphically represented in Figures 1a, 1b, and 1c, respectively. These data were analysed in three separate gender of mother's child (boy, girl) x speech register (IDS, PDS, ADS) x vowel (/i/, /u/, /a/) analyses of variance (ANOVAs) with repeated measures on the last 2 factors. The critical F-value at α=.05 was 4.96. (Additional ANOVAs using species of pet (dog, cat) as the between-subject factor instead of sex of infant, revealed no effect of pet species or interactions of pet with the other factors).

Mean F_0 IDS and PDS had higher mean F_0 than ADS, F(1,10)=20.2, but were not significantly different from each other, F(1,10)=3.97. There were no interactions with vowels, so these results hold across all three vowels.

Pitch Range For pitch range, there was no overall difference between speech registers, $F(1,10)_{IDS+PDSvsADS}$=2.85, $F(1,10)_{IDSvsPDS}$=1.86, but there was a 3-way interaction of vowel, speech register and gender of infant, F(1,10)=20.81, showing that there was higher pitch range in IDS and PDS than ADS in the /u/ and /a/ vowels for mothers of girls and in just the /i/ vowel for mothers of boys. The reason for such a selective effect is unknown, but it is interesting to note that Kitamura has found significantly higher pitch range in IDS to 6-month-old girls than to 6-month-old boys.



Duration Overall IDS and PDS words were of longer duration than ADS words, F(1,10)=26.17, and interaction effects showed that this was especially so for the /u/ and /a/ vowels, F(1,10)=9.68, and then especially for mothers of girls, F(1,10)=10.72. Again it is interesting to note Kitamura's finding of heightened acoustic parameters of speech to girls compared with that to boys.

3.2. Phonetic Analyses

As only the data of four mothers were included in the phonetic analyses, statistical analyses of these data would be premature. The data collapsed over mothers are graphed in Figure 2 in terms of F1-F2 vowel space. As can be seen, the IDS triangle is expanded relative to the ADS triangle, as expected; but unexpectedly, the PDS triangle is also expanded relative to the ADS triangle and is even slightly larger than the IDS triangle. Nevertheless, the IDS and PDS triangles could be considered to be equivalent, and both larger than the ADS triangle. Part of this could be due to the relatively low value of F2 in the ADS /u/, and the relatively high value of F1 in the ADS /i/ compared with other measurements of female Australian vowels [12]. Definitive conclusions on the basis of the vowel triangle data will have to await analysis of the other 8 mothers' data, and subsequent statistical analysis. For now it appears that there is hyperarticulation of vowels in both IDS *and* PDS, compared to the values in ADS.

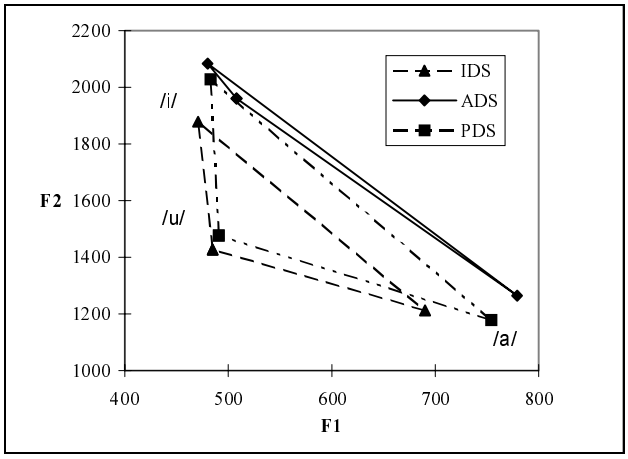


Figure 2: Vowel triangles for IDS, PDS, and ADS.

3.2. Affective Analyses

The variables 'express affection' and the 'direct behaviour were each rated on 1 (not at all) to 5 (extreme) scales. The mean data for these are graphically represented in Figure 3. These data were analysed in a Speech Register (IDS, PDS, ADS) x Express Affect / Direct Behaviour x 12 mothers ANOVA with repeated measures on all factors. Only contrasts on the first two factors are of interest in the analysis. The critical F at α=.05 is 4.38. There was no significant overall difference between Affect and Direct scores. Ratings, both affect and direction, were higher overall for IDS and PDS than for ADS, F(1,19)=86.01, and for IDS than PDS, F(1,19)=12.30. More importantly there were interactions of ADS vs (IDS + PDS) x Affect / Direct, F(1,19)=20.486, and of IDS vs PDS x Affect / Direct, F(1,19)=20.68. These indicate that in IDS the degree of Affect was relatively greater than that of Direction, that in PDS these two aspects of the speech register were statistically equivalent, and that in ADS there was more Direction than Affect.

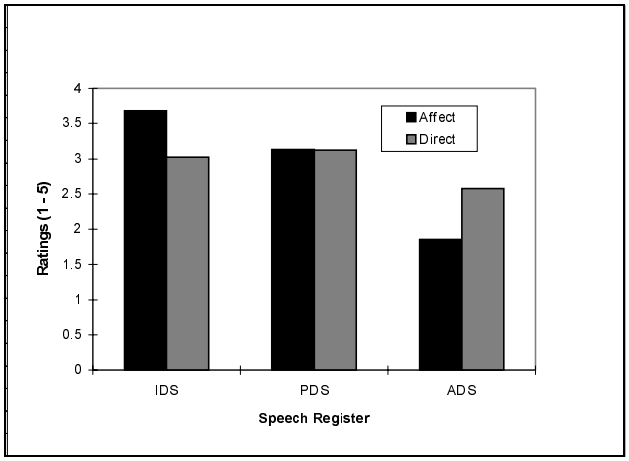


Figure 3: Ratings of 'Express Affect' and 'Direct Behaviour' in IDS, PDS, and ADS.

4. DISCUSSION

On the acoustic variables there is evidence that IDS and PDS are distinctly different from ADS: IDS and PDS contain words which have higher overall F_0, higher pitch range (though this is qualified by sex of infant and vowel), and longer duration. These characteristics are those usually found for IDS, and so it can be concluded here that PDS shares these characteristics. Over and above their joint difference from ADS, the further question of whether IDS and PDS differ from each other on these characteristics begs further investigation: here IDS consistently contains words with higher F_0, higher pitch range, and longer duration than those in PDS, but none of these effects are statistically reliable.

On the phonetic variable of vowel space, surprisingly both IDS and PDS have expanded vowel space relative to ADS. One possible reason for this is the similarity of adults' F1 and F2 values of /i/ and /u/ [see 12]. Thus further data needs to be analysed to determine whether the effect found here is a reliable one. If it *is* reliable then the argument that hyperarticulation of vowels in IDS is a didactic device would at least be weakened. What other possibilities are there? Hyperarticulation could be a didactic device in both IDS and PDS; it could be argued that in PDS we humans are trying to make our pets *understand* our speech to a limited extent. However, there may be a more basic phonetic reason for the hyperarticulation in both speech styles. Recall that the duration of the target words is longer in both IDS and PDS than in ADS. When vowels are shorter they tend to be more centralised, and when they are longer there is a greater possibility that the speaker can hit the ideal target. This could simply shift the focus, i.e, it might be that vowel lengthening is the didactic device. But this is by no means clear - longer vowel duration might be associated with gaining

infants' and pets' attention, or it may accompany the high degree of affect in IDS and PDS relative to ADS.

Expressed affect and direct behaviour were both expressed more in IDS and PDS than they were in ADS, as measured by adults' ratings of these messages in low-pass filtered speech. Thus the low frequency information in these two speech styles appears to carry more affective and directive information than ADS. This is understandable, as ADS can additionally use linguistic (syntactic and semantic) information. IDS and PDS differ in that there is a greater relative level of affect than direction in IDS and equivalent levels of each in PDS.

From this study it can be concluded that the linguistic didactic function of IDS is certainly not its sole or even its main function when the speech of mothers to their 6-month-old infants is considered. However, it is possible and quite likely that this function will increase with speech to older infants [11]. The results of this study also blur the edges of the dependent variables we use for looking at the different functions of special speech registers. Hyperarticulation, on the surface a transparently linguistic device for teaching the child about the vowel space of the ambient language, appears to be evident not only in IDS, but also in PDS. Further data analysis is required before we can confidently draw such a conclusion. For the present we can but question the origin and purpose of hyperarticulation of vowels.

5. REFERENCES

1. After February 1999, the first author will have a new affiliation: Macarthur Auditory Cognition Centre Sydney, Department of Psychology, University of Western Sydney Macarthur.

2. Fernald, A., T., Taeschner, T., Dunn, J., Papousek, M., De Boysson-Bardies, B., & Fukui, I. A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, *16*, 477-501, 1989.

3. Fernald, A., & Simon, T. Expanded intonation contours in mother's speech to newborns." *Developmental Psychology*, *20*, 104-113, 1984.

4. Kuhl, P.K., Andruski, J., Christovich, L., Kozhevnikova, E., Ryskina, V., Stolyarova,E., Sundberg, U., & Lacerda, F. Cross-language analysis of phonetic units in language addressed to infants. *Science*, *277*, 609-732, 1997.

5. Werker, J.F., & McLeod, P.J. 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, 1989.

6. Andruski, J.E. & Kuhl, P.K. The acoustic structure of vowels in mothers' speech to infants and adults. In H. Bunnell & W. Isardi (Eds) *Proceedings of the Fourth International Conference on Spoken Language Processing*, Philadelphia, *3*, 1541-1544, 1996.

7. Fernald, A. Meaningful melodies in mothers' speech to infants. In H. Papousek, U. Jurgens, & M. Papousek (Eds) *Nonverbal vocal behavior.* Cambridge University Press: Cambridge, 1992.

8. Levin, H., & Hunter, W. Children's use of a social register: Age and sex differences. *Journal of Language and Social Psychology*, *1*, 63-72, 1982.

9. Hirsh-Pasek, K., & Trieman, R. Doggerel: Motherese in a new context. *Journal of Child Language*, *9*, 229-237, 1982.

10. Brown, R. Introduction. In C.E.Snow & C.A. Ferguson (Eds) *Talking to children.* Cambridge University Press: Cambridge, 1977.

11. Kitamura, C & Burnham, D. Acoustic and affective qualities of IDS in English. In this Volume - *Proceedings of the Fifth International Conference on Spoken Language Processing*, Sydney, 1998.

12. Harrington, J., Cox, F., & Evans, Z. An acoustic phonetic study of broad, general, and cultivated Australian English vowels. *Australian Journal of Linguistics*, *17*, 155-184, 1997.