Speech Clarity in Infant-directed Singing: an Analysis of German Vowels

Simone Falk

Ludwig-Maximilians-Universität München
simone.falk@campus.lmu.de

ABSTRACT

The findings discussed in this paper are part of a broader fieldwork study where the characteristics of infant-directed singing in natural interaction are analyzed. The aim of this part of the study was to determine whether vowels in infant-directed (ID) singing were of a clearer speech quality compared to adult-directed (AD) speech as it has been shown several times for ID speech. Six German speaking mothers sang for their children aged between 2 and 10 months. Stressed long vowels /aː/, /iː/, /uː/ of these songs were analyzed and compared to vowels in samples of AD speech. Results show that mothers use acoustically more extreme vowels when singing for their infants than in AD speech. This indicates that singing forms a subpart of the infant-directed register which is well-adapted to assist the child at an early stage in discovering and acquiring the linguistically relevant characteristics of his/her mother tongue.

Keywords: Singing, vowel space, infant directed speech.

1. INTRODUCTION

Singing is one of the earliest and most frequently used speech registers that parents use to address their infants [11]. However, while the properties of ID speech (also known as "Motherese", "Parentese" or "Ammensprache") have been extensively studied in the last decades, singing for children and its impact on language acquisition was nearly absent in the literature.

This paper therefore focuses on vowel space in ID singing. From the first month of life on, infants' attention seems to be especially attracted by vocalic elements of speech [6]. It has been shown cross-linguistically that the vowels /a/, /i/, /u/ in mothers' speech to infants during their first year of life differ significantly from the same vowels in adult-directed speech (for American English, Russian and Swedish see [4], Australian English [1], Norvegian [2], Mandarin-Chinese [5]).

This difference manifested itself in modifications of the first and second formants (F1, F2) resulting in an overall expansion of the vowel triangle in ID speech compared to AD speech (perhaps with exception of [2]). This was shown for speech in free natural interaction as well as for more controlled settings. Burnham et al. [1] and Kuhl et al. [4] interpreted this as a hyperarticulation of vowels in ID speech, which might make vowels more distinct from each other representing perceptually better categorical instances of the vowel type. This might help infants to acquire the vowel system more rapidly, aid them in their own productions as well as in their perception of vowel related prosodic processes. The expanded vowel space was also seen as a measure of mothers' speech clarity [5]. Infants might benefit from a clearer articulation by quickly discovering linguistic units like words and phrases. Liu et al. [5] actually found a correlation between the expanded vowel space in mothers' speech and infants' speech discrimination performance.

Until now sung vowels were mainly studied on material with professional singers [8], [9], [10]. These studies showed strong deviations from normal speech vowels revealing different formant patterns ("singer's formant") and overall poor intelligibility of vowel qualities in high frequency regions. Rosenau [8] assessed the qualities of isolated sung German vowels at a speech-like pitch range in singers with different voice types (e.g. soprano, bass). His findings indicate that vowel space in singing tends to be rather compressed.

The aim of this study is to determine whether vowels in ID singing show a pattern resembling the sung vowel space found by Rosenau [8] or whether they are more closely related to ID speech with a clearer vowel quality than in AD speech, thus probably enhancing speech perception by infants.

2. METHOD

2.1. Subjects

Six German speaking mothers (27-38 years, no specific dialect) living in or near Munich were recorded in a natural setting at their home while

singing and talking to their infants aged 2 to 10 months. An interview about child care and singing habits was also recorded. None of the informants was a professionally trained musician or singer, but all of them liked singing and sang regularly at home with their children or/and in choir.

2.2. Material and measurements

The material was recorded with a DAT recorder and digitized at a frequency of 48 kHz. The first and second formant of the long stressed German vowels /aː/, /iː/, /uː/, their duration and fundamental frequency were determined. F1 and F2 were taken at the steady-state temporal midpoint of each long stressed vowel in the songs and interviews. All words containing /aː/, /iː/, /uː/ were included in the study, except those vowels, where the formants were not clearly distinguishable in the spectrum, noises occurred simultaneously with the signal or the vowel was spoken with breathy voice. As the data originate from natural speech/singing recordings, there was a variation in phonetic context. Therefore the material was examined for left consonantal context, especially for place of articulation [3]. No bias of phonetic context could be found. All formant measurements were performed in PRAAT. We measured 600 vowels, (291 ID singing and 309 AD speech). There were at least 15 instances of /aː/ and /iː/ per person and Mode (AD speech, ID singing) and at least 12 /uː/ per person and Mode (words with long /uː/ are relatively rare in German). We additionally measured 115 vowels in ID speech of three of our subjects. All formant values were recorded in Hz and subsequently transformed in Bark in order to get a clearer picture of the perceptual quality of the measured vowels.

3. RESULTS

The duration of the vowels in ID singing was three (aː, iː) to four (uː) times longer than in AD speech. The median of F0 in ID singing (272 Hz) was slightly higher than in AD speech (226 Hz). The same is true for the overall pitch range (95 % of all values lying between 148-383 Hz in ID singing, 97-336 Hz in AD speech).

We considered the median of F1 and F2 for all vowels per subject and Mode. F2 was generally more affected by Mode differences than F1. We found clearly higher F2 values for /iː/ (up to 374 Hz) in all subjects in ID singing. Higher F2 values for ID vowel /aː/ (up to 201 Hz) and only slightly higher F1 values (up to 140 Hz) were found in all but one subject. /uː/ showed the least differences in ID singing compared to AD speech, but sometimes F2 was lower in ID singing than AD speech.

Figure 1: Vowel space of subject 6 in AD speech. F1 and F2 in Bark. 95 % confidence ellipses.

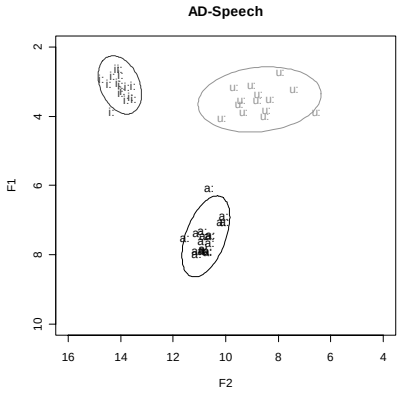
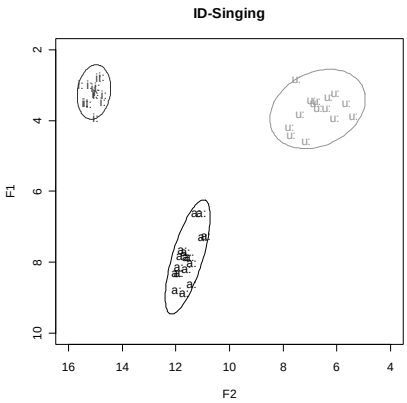


Figure 2: Vowel space of subject 6 in ID singing. F1 and F2 in Bark. 95 % confidence ellipses.

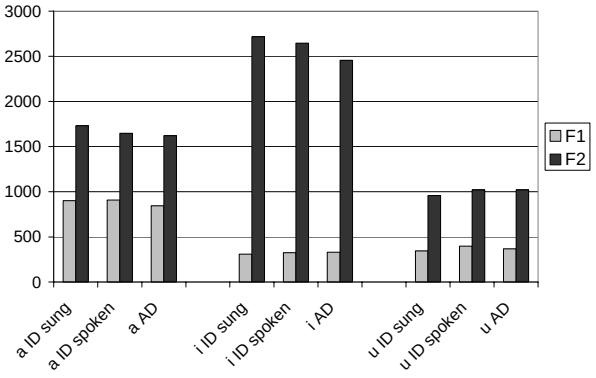


These changes in formant values in ID singing vs. AD speech suggest an expansion of the vowel space in singing as it can be seen exemplarily for subject 6 in Figs. 1 and 2. In this case the formant values (in Bark) are more peripheral in ID singing than AD speech, especially in the F2 dimension, showing that /uː/ is articulated more back, whereas /iː/ and /aː/ are fronted. /aː/ shows also raised F1 values that relate to a wider jaw opening in singing.

In addition the spoken ID vowels of three subjects were taken into account: approximately the same picture emerged as in the sung ID vowels (higher F2 in /iː/ and /aː/ and higher F1 in /aː/ than in AD speech). Nevertheless, it can be seen in

Figure 3 that sung ID vowels have even higher F2 values in /iː/ and /aː/ than spoken ID vowels and also lower F2 in /uː/.

Figure 3: Median of ID singing, ID speech and AD speech vowel formants in Hz, subjects 1, 5 and 7.



Next we calculated the Euclidean distance of the sung ID and spoken AD vowels from the vowel space center (the centroid). Euclidean distance is a measure of vowel centralization [12]. Fig. 4 indicates that all vowels in AD speech for all but subject 8 are more centralized.

These results were also confirmed by statistical analysis. We conducted a three-way repeated measures ANOVA with the main independent variables of Mode (AD speech / ID singing), Vowel (/aː/, /iː/, /uː/) and Formant Number (first or second). The effect of Mode for the overall data was highly significant ($p<0.01$). Then a one-way repeated measures ANOVA was performed for each formant and vowel to determine the effect of Mode in each condition separately. There were significant effects for F1 and F2 in /aː/ (F1: $p<0.01$; F2: $p<0.05$) and a significant effect for F2 in /iː/, ($p=0.001$). No significance was found for both formants of /uː/ and F1 of /iː/.

Table 2: Significantly higher formant values in ID singing than in AD speech.

Formant/ Vowel	Mode	Mean (Hz)	SD
F1 /aː/	ID	877.67	44.58
	AD	817.89	36.38
F2 /aː/	ID	1625.89	174.26
	AD	1492.94	144.86
F2 /iː/	ID	2665.71	82.51
	AD	2427.65	62.99

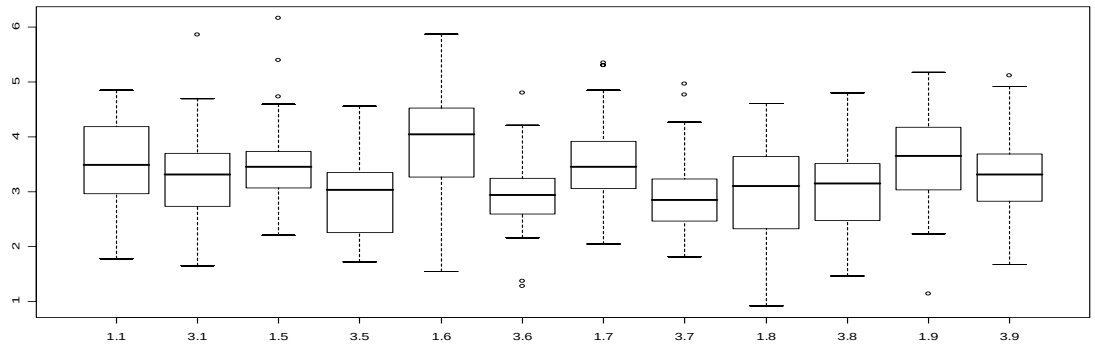
In conclusion, the ID vowel /aː/ was sung with significantly higher F1 and F2 than the AD vowel /aː/. The ID vowel /iː/ was sung with significantly higher F2 than spoken in AD speech (Table 1).

4. DISCUSSION AND CONCLUSIONS

The present data show that there is a general difference between formant frequencies in ID singing and AD speech which results in an expansion of the vowel space in ID singing. This means that the pattern of ID singing resembles the pattern of ID speech found in previous studies. In the following we compare some of the findings in more detail:

1. The raising of F2, especially in /iː/ and /aː/, corresponds to the results for ID speech vs. AD speech in [1], [4], [5]. In our own material this difference was even more pronounced in ID singing than in ID speech.

Figure 4: Boxplots of Euclidean distance (Bark) to the vowel centroid of the vowel plane F1xF2, calculated for all vowels. Two boxplots (ID singing:1, AD speech: 3) are shown per subject (on the x-axis: A.B=Mode.Subject).



2. The overall picture suggests that ID singing has a greater front-back vowel contrast than AD speech. Surprisingly we found only weak F1-raising in /aː/ which we had expected to be more pronounced, because studies of ID speech as well as studies of singing reported that the vocal tract producing /aː/ should be more open in either condition. This may be a difference between ID speech and ID singing: Englund and Behne [2], relying on the work of Lacerda, hypothesized that the beneficial qualities of ID speech may be primarily due to a more pronounced high-low vowel contrast where the front-back contrast is only a secondary, perceptually less relevant dimension. In ID singing it would be the opposite: F2 plays the crucial role in distinguishing the vowels. In our data this is most evident for /iː/ that is clearly fronted in ID singing in all subjects. /aː/ is also fronted whereas /uː/ is articulated more back in some subjects. As German has a complex 16-vowel system containing the front rounded vowel /yː/ the front-back contrast is especially important for the perception of the difference between peripheral /iː/ and /yː/. However, it will be necessary to further investigate this topic. At the moment we may conclude that both modes of infant-directed communication (ID singing, ID speech) provide more peripheral vowels working as a better perceptual anchor for categorical discrimination by infants than AD speech [7].

3. Finally, other factors like F0 and duration could influence the differences between ID and AD formants. The effect of higher F0 in singing might be negligible because in this study the pitch ranges in singing to infants and in speaking were not so different as it has been expected. However, in our data and in the studies cited above ID vowels tended to be longer than AD vowels. The more extreme nature of sung ID vowels may mostly be due to the fact that parents reach the vowel target more easily in slow articulation than in fast articulation. This raises the question if the vowels of ID speech or ID singing are really an expression of hyperarticulation [4] or if the vowels in AD productions simply undergo undershoot in faster speech. Whichever the answer to this question is, the effect that infants may perceive the vowel quality in either ID speech or ID singing more readily remains the same.

5. REFERENCES

[1] Burnham, D., Kitamura, C., Vollmer-Conna, U., 2002. What's new, pussycat? On talking to babies and animals. *Science* 296, 1435.

[2] Englund, K.T., Behne, D.M., 2005. Infant directed speech in natural interaction — Norwegian vowel quantity and quality. *Journal of Psycholinguistic Research* 34, 3, 259-280.

[3] Hillenbrand, J.M., Clark, M.J., Terrance, M.N 2001. Effects on consonantal context on vowel formant patterns. *J. Acoust. Soc. Am.* 109,2, 748-763.

[4] Kuhl, P.K. *et al.*, 1997. Cross-language analysis of phonetic units in language addressed to infants. *Science* 277, 684-686.

[5] Liu, H-M., Kuhl, P.K., Tsao, F-M., 2003. An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science* 6, F1-F10.

[6] Mehler, J. *et al*., 1996. Coping With Linguistic Diversity: The Infant's Viewpoint. In: Morgan, J.L., Demuth, K. (eds): *Signal to Syntax. Bootstrapping from Speech to Grammar in Early Acquisition*. Mahwah: Erlbaum, 101-116.

[7] Polka, L., Bohn, O-S., 2003. Asymmetries in vowel perception. *Speech communication* 41,1, 221-231.

[8] Rosenau, S., 1999. An analysis of phonetic differences between German singing and speaking voices. *Proc. 14th ICPhS, San Francisco,* .

[9] Scotto di Carlo, N., 2005. Contraintes de production et intelligibilité de la voix chantée. *Travaux Interdisciplinaires du Laboratoire Paroles et Langage* 24, 159-179.

[10] Sundberg, J., 1987. *The science of the singing voice*. Illinois: Northern Illinois University Press.

[11] Trehub, S. E., Trainor, Laurel, 1998. Singing to infants: Lullabies and play songs. In: Rovee-Collier, C., Lipsitt, L.P. (eds): *Advances in infancy research* 12, 43-78.

[12] Wright, R., 2003. Factors of lexical competition in vowel articulation. In: J. Local, R. Ogden & R. Temple (eds.): *Phonetic Interpretation: Papers in Laboratory Phonology VI*. Cambridge: Cambridge University Press, 75-87.

6. Acknowledgements

Thanks for their advice and comments on this paper to Melanie Moll, Jonathan Harrington, Elisabeth Leiss, Katalin Mady and Phil Hoole. This research project was financially supported by the Studienstiftung des deutschen Volkes and the International PhD Programme "Linguistics" (LIPP) of the LMU Munich and the DAAD.