AN ANALYSIS OF PHONETIC DIFFERENCES BETWEEN GERMAN SINGING AND SPEAKING VOICES

Sigurd Rosenau
Institut für Phonetik und Sprachliche Kommunikation
Ludwig-Maximilians-Universität München
Schellingstr. 3, 80799 München

ABSTRACT
In this study an investigation of phonetic differences between singing and speaking voices is presented which focuses on vowel quality and the electroglottographic signal. It is based on the analysis of synchronous recordings of acoustic and electroglottographic signals of several singers that spoke and sung different German vowels at average speaking fundamental frequencies. Four different voice classes were considered ranging from bass to soprano. The analysis was done by determining qualitative differences of formant frequencies and the closed quotient between the spoken and sung vowels. It mainly showed that these differences are fairly independent of the voice class. The differences of the formant frequencies were clearly dependent on the vowel quality only for the formants F2 and F3. They were practically independent of the vowel quality for F4 (spoken and sung vowels) and F5 (sung vowels), and generally insignificant for F1. The differences of the values of the closed quotient were relatively small which suggests an influence of the singing voices on the speaking voices with regard to this phonatory feature.

1 INTRODUCTION
Differences between vowel qualities of speaking and singing voices may be caused by articulation or differences in the glottal wave form [Sundberg, 1970]. Articulatory differences have mainly been studied by analyzing formant frequencies (e.g. [Florig, 1995]) but X-ray pictures have also been used (e.g. [Sundberg, 1970][Sundberg et al., 1982]). Differences concerning the glottal waveform have also been examined whereby flow glottograms were often used (e.g. [Sundberg [1987]]). In most of these studies only one or few voice classes were considered.

The purpose of the present study is to make a first approach to the investigation of differences between speaking and singing voices for several different speaking voice classes, combining analyses of phonation and articulation of spoken and sung vowels. These phonetic differences are examined by analyzing synchronous recordings of acoustic and electroglottographic signals of German vowels spoken and sung by several singers at average speaking fundamental frequencies. The articulatory differences are characterized by qualitative determination of relative shifts of formant frequencies of sung vowels with respect to spoken vowels. The analysis of the electroglottographic signals for the analogous determination of phonatory differences is restricted to the closed quotient since this quantity turned out to be the most distinct feature of the corresponding signals of the speaking and singing voices.

2 MATERIAL AND METHOD

2.1 Subjects
This study was done with four professional singers of different voice classes (bass, baritone, mezzo-soprano and soprano). One of the singers is a soloist of the Bavarian State Opera (baritone), two of them are members of the choir of the Bavarian State Opera (bass and mezzo-soprano) and the fourth of them is an experienced soloist singer (soprano).

2.2 Material, recordings and measurements
Phonetic differences between spoken and sung vowels are characterized by analyzing formant frequencies and the closed quotient (C.Q.) of the eight monophthongal German vowels /ä, ö, ü, e, a, o, u, i/. These vowels were spoken and sung at the fundamental frequencies of the speaking voices of the singers. The values of the fundamental frequencies were 100 Hz (bass and baritone), 196 Hz (mezzo-soprano) and 220 Hz (soprano). The vowels were spoken in the symmetrical consonant context /p/ embedded in the carrier word gepCVCpe, and had a duration of 80 - 100 ms. The steady sung vowels were not embedded in a context and had a duration of at least 1 s. The singers were required to sing the vowels at a constant loudness level (mf). The recordings of the acoustic and electroglottographic signals of the spoken and sung vowels were done in an anechoic room using an electroglottograph and a microphone. The microphone was placed at a distance of about 30 cm from the mouth of the singer. The signals were recorded synchronously with a DAT-recorder and digitized with a frequency of 48 kHz. Four formants of the spoken vowels and five formants of the sung vowels were determined using an LPC analysis. The fifth formant of the spoken vowels was not determined because it could not always be seen clearly enough in the spectrum. The closed quotient of the electroglottographic signal was calculated by determining the quotient of the duration of the closed phase in the glottal cycle and the fundamental period, and averaging this value over the whole signal.

3 RESULTS

3.1 Bass
Formant frequencies The frequencies of the formants F4 and F5 of the sung vowels and F3 of the spoken vowel are quite independent of the vowel quality (see fig. 1). The formant frequencies of F4 of the spoken vowels are higher than the corresponding fourth formant frequencies of the sung vowels and
have on average more or less the same values as the corresponding fifth formant frequencies of the sung vowels.

The differences between the formant frequencies of spoken and sung vowels with respect to F$_2$ and F$_3$ depend on the vowel quality. The frequencies of F$_2$ and F$_3$ of the sung vowels /a:/, /ø:/ and /y:/ are higher than in speech or have practically the same value whereas the corresponding formant frequencies of all the other investigated vowels are lower than in speech.

The frequencies of F$_1$ of the spoken and sung vowels have more or less the same values, the formant frequencies of the sung vowels being generally somewhat lower.

Closed quotient For the description of the results the following quantities are used (see fig. 2 and 3): mean value and standard deviation of the C.Q. with respect to the vowel quality ($\bar{CQ}_{vq}$ and $\sigma_{vq}$), mean value of the standard deviations of the C.Q. of each vowel ($\bar{\sigma}_v$). The standard deviations are represented in the figures by the error bars.

The values of $\bar{\sigma}_v$, (0.021 for sung vowels and 0.018 for spoken vowels) are greater than corresponding values of $\sigma_{vq}$, (0.011 for sung vowels and 0.015 for spoken vowels). Moreover, the values of $\sigma_{vq}$ are relatively small and therefore the C.Q. is practically independent of the vowel quality.

The value of $\bar{CQ}_{vq}$ is slightly greater for the sung vowels (0.61) than for the spoken vowels (0.58).

3.2 Baritone
The differences between the spoken and sung vowels with respect to formant frequencies and the C.Q. are qualitatively nearly equal to the data of the bass singer (see fig. 4 - 6). One exception is the formant frequency of F$_3$ of the sung vowel /a:/, that is higher than the corresponding value of the spoken vowel. Furthermore, the frequencies of the fourth formants of the spoken and the sung vowels have, on average, similar values.

3.3 Mezzo-soprano and soprano
The results derived from the data of the singers of the female voice classes are similar to the results described above for the singers of the male voice classes (see fig. 7 - 12). The main qualitative difference is that the formant frequencies of F$_2$ and F$_3$ of the sung vowel /e:/ are higher than the corresponding values of the spoken vowels. The corresponding differences of F$_2$ and F$_3$ with respect to the vowel /æ:/ are dissimilar: the data of the mezzo-soprano and the soprano singer resemble qualitatively the equivalent data of the baritone and the bass singer, respectively.

4 DISCUSSION AND CONCLUSIONS
The presented results show that qualitative differences between spoken and sung vowels concerning formant frequencies and the C.Q. are fairly independent of the voice class. Since data of only one singer for each singing voice class was analyzed, more recordings need to be done in order to also evaluate the results quantitatively.

Formant frequencies The differences between the formant frequencies of the spoken and sung vowels depend on the order of the formants.

The formant frequencies of F$_4$ and F$_5$ of the sung vowels and the fourth formant frequencies of the spoken vowels are practically independent of the vowel quality. Moreover, their variations with respect to the vowel quality are relatively small. This can be explained for the sung vowels by the presence of the singer’s formant which is known to be produced by a clustering of the formants F$_5$ and F$_4$ (and F$_3$) [Sundberg, 1974]. Since the singer’s formant is a main feature of the sound quality of sung vowels its position in the spectrum should not depend on vowel quality. The small variations of the fourth formant frequencies of the spoken vowels suggest that the speaking voices of the singers are influenced by their singing voices in the corresponding spectral region. There is thus no influence to be seen concerning the fourth formant frequencies of the spoken vowels, since their relative differences to the formant frequencies of F$_4$ and F$_5$ of the sung vowels are dependent on the voice class.

The differences between the formant frequencies of the spoken and sung vowels of F$_2$ and F$_3$ depend clearly on the vowel quality. They are qualitatively identical for all the singers except for the vowels /æ:/ and /ə:/ of the data of the vowel /æ:/ depends on the voice class in an unsystematic manner. However a clear distinction between male and female voice classes can be made for the data of the vowel /ə:/ and the same result is not valid for clarifying the results with regard to these two vowels.

The differences between the first formant frequencies are generally insignificant and therefore no systematic relations for this data can be given.

Closed quotient The analysis of the C.Q. shows that the singers are able to control the C.Q. quite independently of the vowel quality. This result is remarkable since the forms of the electroglotographic signals can differ noticeably not only between spoken and sung vowels but also with respect to the vowel quality. Moreover, it was to be expected that the values of the closed quotient are greater for the sung vowels than for the spoken vowels, but the relatively small differences of the values are noticeable. Obviously the singing voices have highly influenced the C.Q. of the speaking voices of the singers.

REFERENCES
Figure 1: Formants of german vowels spoken and sung by a bass singer

Figure 2: Closed quotient (C.Q.) of german vowels sung by a bass singer

Figure 3: Closed quotient (C.Q.) of german vowels spoken by a bass singer

Figure 4: Formants of german vowels spoken and sung by a baritone singer

Figure 5: Closed quotient (C.Q.) of german vowels sung by a baritone singer

Figure 6: Closed quotient (C.Q.) of german vowels spoken by a baritone singer
Figure 7: Formants of German vowels spoken and sung by a mezzo-soprano singer

Figure 10: Formants of German vowels spoken and sung by a soprano singer

Figure 8: Closed quotient (C.Q.) of German vowels sung by a mezzo-soprano singer

Figure 11: Closed quotient (C.Q.) of German vowels sung by a soprano singer

Figure 9: Closed quotient (C.Q.) of German vowels spoken by a mezzo-soprano singer

Figure 12: Closed quotient (C.Q.) of German vowels spoken by a soprano singer