

Individual Differences Between Vowel Systems Of German Speakers

Sebastian J. G. G. Heid
Ludwig-Maximilians Universität München
Institut für Phonetik und Sprachliche Kommunikation
heid@phonetik.uni-muenchen.de
http://www.phonetik.uni-muenchen.de/Mitarb/Heid.html

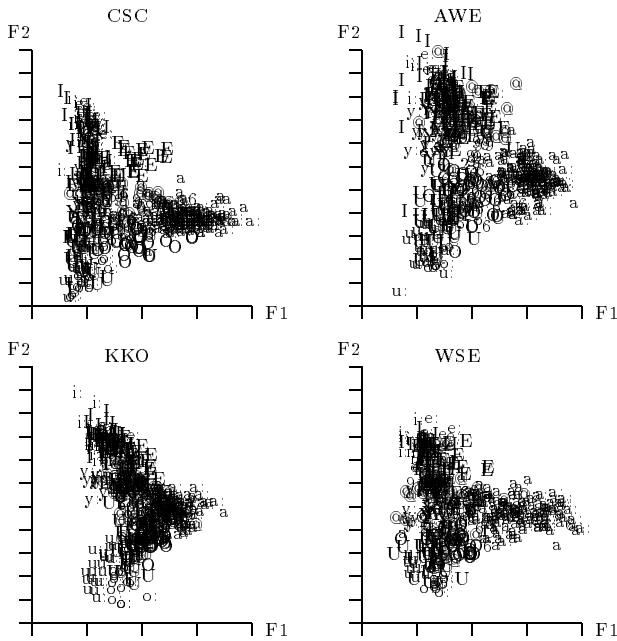


Figure 1: Vowel distributions of four speakers plotted in F1/F2, and F1/F3 formant plane

ABSTRACT

Based on formant measurements of more than 10000 vowels from 16 German speakers, vowel quality differences between the speakers have been analyzed. The main result of this investigation is that different speakers show not only different formant values (as one would expect due to individual differences), but exhibit different arrangements in their vowel systems. These differences are demonstrated by examples from the general distribution of vowels, the structure of vowel prototypes, and the differing number of degrees of tongue height for different speakers. The problem of vowel normalization for that data will also be demonstrated and discussed.

1 INTRODUCTION

In [HWD95] we presented a database with vowel formant data for more than 10000 German vowels from 16 Speakers. Although all speakers can be considered as speakers of standard German, they show perceptible regional differences, as the recordings took place in Munich (southern Germany), Bonn (middle Germany), Kiel (northern Germany) respectively. These regional influences on the speaking style lead to considerable differences in the structures of the vowel systems of the individual speakers, which will be displayed in this paper.

As figure 1 demonstrates, the general distribution patterns of the vowels in formant planes exhibit clear differ-

ences between the individual speakers. Speaker CSC and KKO represent most clearly the two main classes found for the distribution of the vowels in the F1/F2-plane: a kind of triangular shape on the one hand, and a more heart like shaped pattern, with left out edge between u- and i-region on the other. These overall distribution differences turned out to be closely connected to the distribution of the vowel prototypes of the individual speakers.

2 VOWEL PROTOTYPE DIFFERENCES

As vowel prototypes we calculated the average formant values over the realizations of one canonic vowel category. A canonic vowel is the vowel which is to be expected in careful realization of the words in question. For example, "Bären" (bears) is supposed to be produced with the vowel /E:/ as opposed to "Beeren" (berries) which is to be produced with the vowel /e:/. Many speakers of northern varieties of standard German do not produce that vowel opposition. For those speakers we find for the canonic vowel [E:] more or less the same formants as for [e:]. So although the canonic vowel categories roughly correspond to a phonemic representation, they still are phonetic categories, since they are defined as standard production patterns and not via an abstract phonological system of oppositions.

In his comparisons of "regional differences in the realizations of standard German vowels", Antti Iivonen ([Iiv87]) mainly reports differences in the degree of centralization for the lax vowels between speakers of East Middle German and speakers of Vienna German, finding that the tense-lax opposition for speakers from Vienna is mainly a question of length and not vowel quality. Although no speakers from Austria took part in our record-

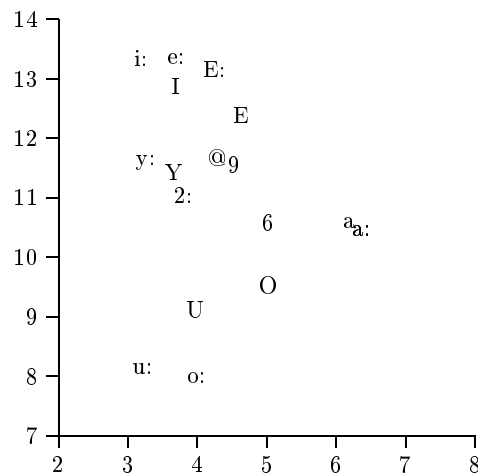


Figure 2: F1/F2-plane (bark scaled) with vowel prototypes averaged over all speakers

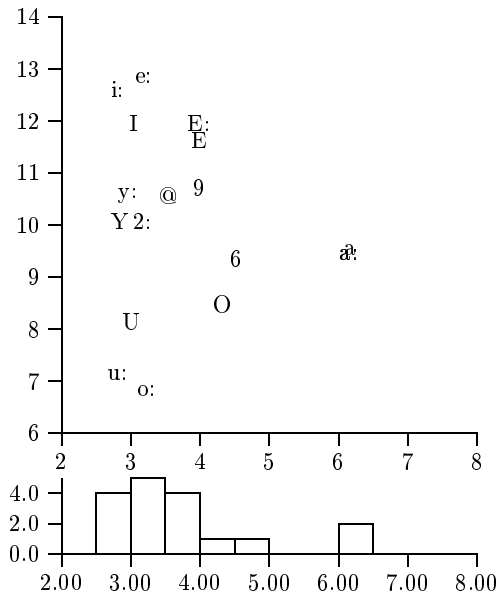


Figure 3: Vowel prototypes and F1-histogram for speaker CSC

ings, this difference is found in our data too, but only for one speaker (AWE) who comes from the south western area of Germany. This tendency of not centralizing lax vowels is also visible in the scatter plot in figure 1. The distribution of the vowels for AWE exhibits a kind of clearing in the central region. Speaker CSC (West Middle German), shows a similar tendency. In his case the lax vowels (/I/,/Y/,/U/) are produced with almost the same F1 as their tense counterparts.

Figure 2 shows a scatterplot of the vowel prototypes averaged over all speakers in a F1/F2-plane, figure 3 and figure 4 show the vowel prototypes for the speakers CSC and KKO, respectively. Comparing the prototypes for CSC and KKO one can see, that they display two different systems. There are differences not only in the specific formant values, but also a lot of clear differences in the relations between the vowels. For example, in the data of CSC one finds /y:/, /Y/, and /2:/ as one cluster in the central high region (low F1, medium F2), for KKO /Y/, /2:/, /@/, and /9/ cluster far away from /y:/ in the front mid region (medium F1, high F2). Differences like these are found for any pair of speakers in the database. For a systematization of the many individual differences it turned out to be useful to look at some key features of the individual vowel systems.

3 TYPICAL DIFFERENCES

As already mentioned the difference in the realisation of the “ä” as [E:] or [e:], is a long known marker for different types of vowel systems coexisting in standard German. As displayed in figure 3 and 4 speaker CSC does produce an /E:/ as an [E:], while speaker KKO produces /E:/ almost like [e:] and not like [E:]. Since both production types occur at nearly the same frequency, averaging over all speakers one consequently finds the /E:/ in the middle between /e:/ and /E:/, as shown in figure 2. This is also

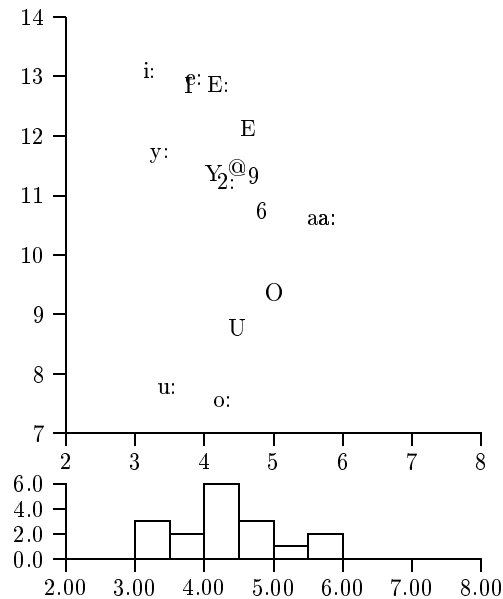


Figure 4: Vowel prototypes and F1-histogram for speaker KKO

true for the two speakers HGP and MKN who use both variants.

Another difference between speakers conveyed by the data is a kind of backing for the rounded high front vowels (y: and 2:). Normally one would expect them to be articulated in the front region, i.e. to show F2-values which lie nearer to /i:/ or /e:/ than to /u:/. Speaker KKO does exhibit this expected pattern, but for 6 speakers we found F2-values almost in the middle between /u:/ and /i:/, what could be a result of either a backing of the tongue, or more likely a different degree of lip rounding. Anyway, the realization of /y:/ and /2:/, too, can be regarded as a marker for a categorical difference between the vowel systems of the speakers.

4 TONGUE HEIGHT

One outstanding characteristic of the vowel systems of different speakers is the uneven distribution of the vowel prototypes along the F1-dimension, which is known to correspond quite closely to tongue height.

According to phonetic descriptions of the German vowel system, one could expect up to 7 possible degrees of tongue height [PM95]: high ([i:], [u:], [y:]), half high ([I:], [Y:], [U]), upper mid ([e:], [2:], [o:]), mid ([@]), lower mid ([E:], [E:], [9], [O]), half low ([6]), low ([a], [a:]). Although none of the speakers really exhibits 7 distinct classes of tongue height, each of the distinctions in the 7 degree system is realized by at least some speakers.

Most speakers exhibit 4 degrees of tongue height, with largely varying content, and some tend to exhibit only 3 degrees. Although it is hard to draw clear boundaries, the differences in the distribution of F1-values for the vowel prototypes of individual speakers seems to be a crucial feature, since it cooccurs with many of the other differences we found.

As a first way of distinguishing these distribution patterns we do suggest the shape of F1-histograms for the vowel prototypes (with a bin size of 0.5 bark) as plotted beneath the vowel scatterplots. It is easy to see that

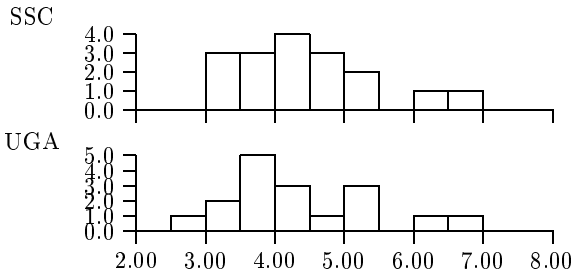


Figure 5: F1-distribution for speaker SSC and UGA, which are examples for distribution types G and TR.

Spk.	e:/E: dist.	y:/2: back	F1-distr.
AWE	+	-	TR
BMO	+	-	G
CHK	-	-	RS/TR
CSC	+	+	RS
DLM	-	++	G
HGP	0	0	G
HPT	-	-	RS
KKO	-	-	TL
KMA	+	-	G
MKN	0	+	TR
RTD	-	-	TL
SAT	+	-	G
SSC	+	-	G
TPO	+	+	TR
UGA	-	+	TR
WSE	-	+	G

Tab. 1: Vowel system features for the speakers in the database

the differences in the distribution of the vowel prototypes for speaker CSC and KKO lead to two distinct histogram shapes: a right skewed curve for CSC and a peak in the central regions with two side peaks for KKO. When inspecting the histograms for all speakers we could single out four class-building features for the shape of the F1-histograms:

G	≡	Gaussian shaped distribution
TR	≡	Trough right of maximum
TL	≡	Trough left of maximum
RS	≡	Right skewed distribution

There is one trough right of the maximum found for almost every speaker, since the two a's (/a:/, /a/) always form a kind of side peak, so TR means actually two troughs right of maximum, as shown for speaker UGA in figure 5. The Gaussian shaped distribution, again neglecting the side peak for the a's, is, for example, found for speaker SSC, as displayed in figure 5. Speaker CSC is an example for type RS and KKO for the type TL. All speakers with the feature TL, show like KKO, a large group of vowel prototypes in the central region, with distinct F1-values for the high tense vowels (/i:/, /y:/, /u:/).

5 VOWEL SYSTEM

The above mentioned features can be used to build up a feature matrix for the vowel systems of the individual speakers, which demonstrates the complete mixture of effects for the speakers. This suggests that the speakers of

standard German might reveal a mixture of different dialectal or sociolectal influences. A reason for the rather massive mixture of influences in this particular data might be that the speakers — being students or university personnel — came from all over Germany and not just from the three cities mentioned above, and have thus been exposed to different regional influences.

6 VOWEL NORMALIZATION

[PB52], [Tra81] and [SG86] suggest a method for vowel normalization which is based on bark scale transformation and on the observation that the formant values of different speakers covary with the F0-values of those speakers. The quality of this kind of transformation is typically shown with examples of vowels from male and female speakers, which exhibit large formant differences. Figure 6 displays the success of this method for our data, with formant values averaged over male and female speakers.

Nevertheless, the differences of the individual vowel systems as presented above are resistant against this kind of normalization, as demonstrated in figure 7.

This ineffectiveness of vowel normalization for reducing the observed differences supports the assumption that these differences were not due to general shifting or transformation, but result from inherent systematic structural differences. Normalization can only work successfully to reduce differences like those between between male and

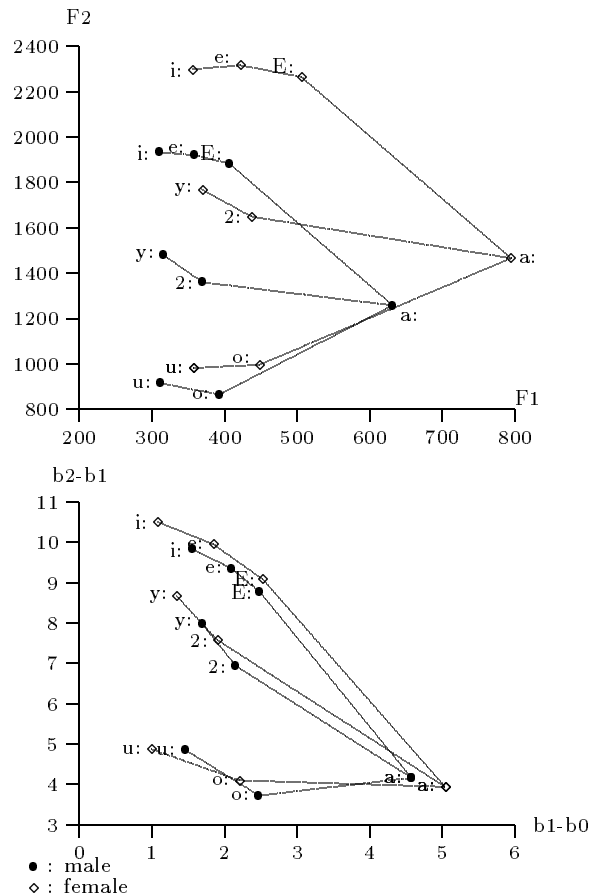


Figure 6: Tense vowel prototypes for male and female speakers, in Hz-scaled formant plane and in normalized formant plane

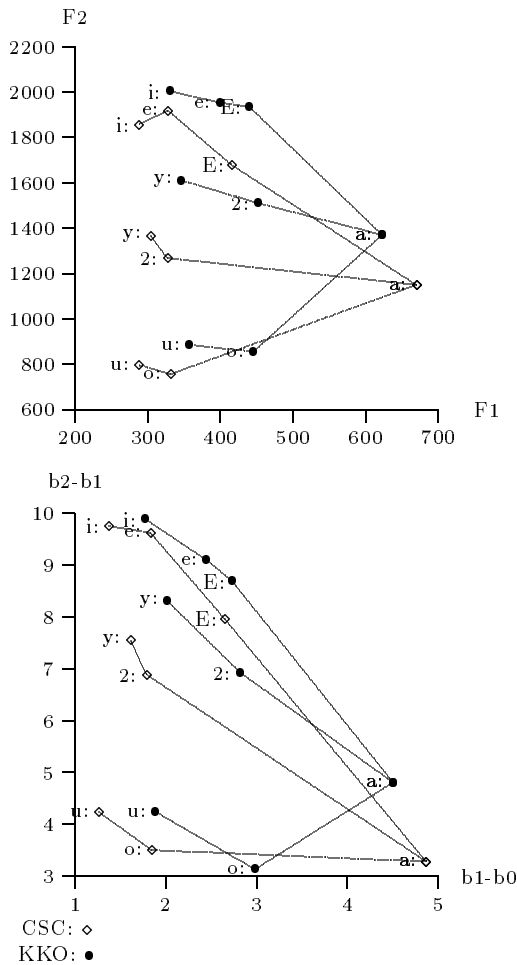


Figure 7: Tense vowel prototypes for the speakers CSC and KKO in hz-scaled and normalized formant plane

female speakers, since those can mainly be explained by the physiological differences.

7 CONCLUSION

Structural differences of individual vowel systems have been demonstrated and features for a first systematization were suggested. The idea of a vowel feature system seems to be useful for further research into regional differences between these individual vowel systems, although the impressionistic features proposed in this paper will of course need refinement. Future research might reveal new features, methods for quantification of the features, and better knowledge about the connection of the features to specific regional influences.

The failure of normalization shows that one has to differentiate between different formant patterns which nevertheless result in identical vowel quality and different formant patterns which result in perceivable different vowel qualities. It will need further explorations of the possibilities of modelling human perception of vowel quality, e.g [Pfi95], to gain a better understanding of this differences.

As a method for explaining or modelling the perception of linguistic categories normalization seems not to be useful, since as demonstrated, the speakers are quite free in their ways of producing the linguistic categories. This suggests that the connection (often proposed by phonolo-

gists) between linguistic categories and mental perception or production categories might need more extensive clarification.

REFERENCES

- [HWD95] Sebastian Heid, Barbara M. Wesenick, and Christoph Draxler. Phonetic analysis of vowel segments in the phondat database of spoken german. In Kjell Elenius and Peter Branderud, editors, *Proceedings of the XIIIth International Congress of Phonetic Sciences*, pages 4:416–19, Stockholm, 1995. KTH and Stockholm Univ.
- [Iiv87] Antti Iivonen. Regional differences in the realization of standard german vowels. In *Proceedings of the 11th International Congress of Phonetic Sciences*, volume 4, pages 161–4, Tallinn, 1987.
- [PB52] Gordon E. Peterson and Harold L. Barney. Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, 24:175–184, 1952.
- [Pfi95] Hartmut R. Pfitzinger. Dynamic vowel quality: a new determination formalism based on perceptual experiments. In *Proceedings of Eurospeech '95*, volume 1, pages 417–20, Madrid, 1995.
- [PM95] Bernd Pompino-Marschall. *Einführung in die Phonetik*. Walter de Gruyter, Berlin, 1995.
- [SG86] Ann K. Syrdal and H. S. Gopal. A perceptual model of vowel recognition based on the auditory representation of American English vowels. *Journal of the Acoustical Society of America*, 79:1086–100, 1986.
- [Tra81] Hartmut Traunmüller. Perceptual dimensions of openness in vowels. *Journal of the Acoustical Society of America*, 5:1465 – 75, 1981.