Forschungsberichte des
Instituts für Phonetik und
Sprachliche Kommunikation
der Universität München
(FIPKM) 35 (1997) 35-38

# FORMANT TRAJECTORY DYNAMICS IN SWABIAN DIPHTHONGS*

*Anja Geumann*
Institut für Phonetik und Sprachliche Kommunikation,
Universität München, Germany
E-mail: geumann@phonetik.uni-muenchen.de

## ABSTRACT

The vowel quality in some diphthongs of Swabian (an upper German dialect) was determined by measurement of first and second formant values. A minimal contrast could be shown between two different diphthong qualities /ɑɐ/ and /əi/, where for Standard German only one is assumed, viz. /ai/. The two diphthong qualities differ only slightly in onset and offset vowel quality, so a better understanding of their relationship was expected from an examination of their dynamic aspects. Our preliminary results suggest that there is indeed a difference in the temporal structure of the two diphthongs.

## 1. INTRODUCTION

In Swabian, a major German dialect, we differentiate on a phonological level between two diphthongs /ɑɐ/ and /əi/ which belong to the same phonological class in Standard German, i.e. are judged to be homophones. In linguistic descriptions the phonetic realisation of the phonological contrast between the two diphthongs is often given as a contrast of vowel quality [1]. The very preliminary results of Geumann & Hiller [2] support the idea of a minimal contrast, but indicate the durational pattern as relevant perceptual cue. In this paper we look more closely at parameters that might account for this distinction.

## 2. EXPERIMENT

5 adult male speakers of the Northern Central Swabian dialect aged between 20 and 40 were asked to read two lists of isolated words in Standard German orthography. These lists consisted of six minimal pairs. First they were given in completely randomized order. A second list gave the contrasting pairs in randomized order, i.e., n=12 items per diphthong and speaker, so n=60 in total. The acoustic signal was recorded in a quiet environment.

The data were judged informally by a native listener to ensure that there was in fact a minimal contrast produced.

Start and end of the diphthongs were determined in the acoustic signal at onset and offset of vowel periodicity. The vowel duration is given in table 1.

For the next step, formant values were calculated at 8 points: Start and end of the vowel, interval then divided by 6 equidistant points (deviation max. ± 2%). Values of first and second formants were determined. See the data in Figure 1.

The trajectories of the diphthongs measured were plotted with F1 against F2-F1 (mel scaling). (Figure 2).

Next, the distance between acoustic formant onset and velocity peak of F2 was determined for some of the speakers (see table 2).

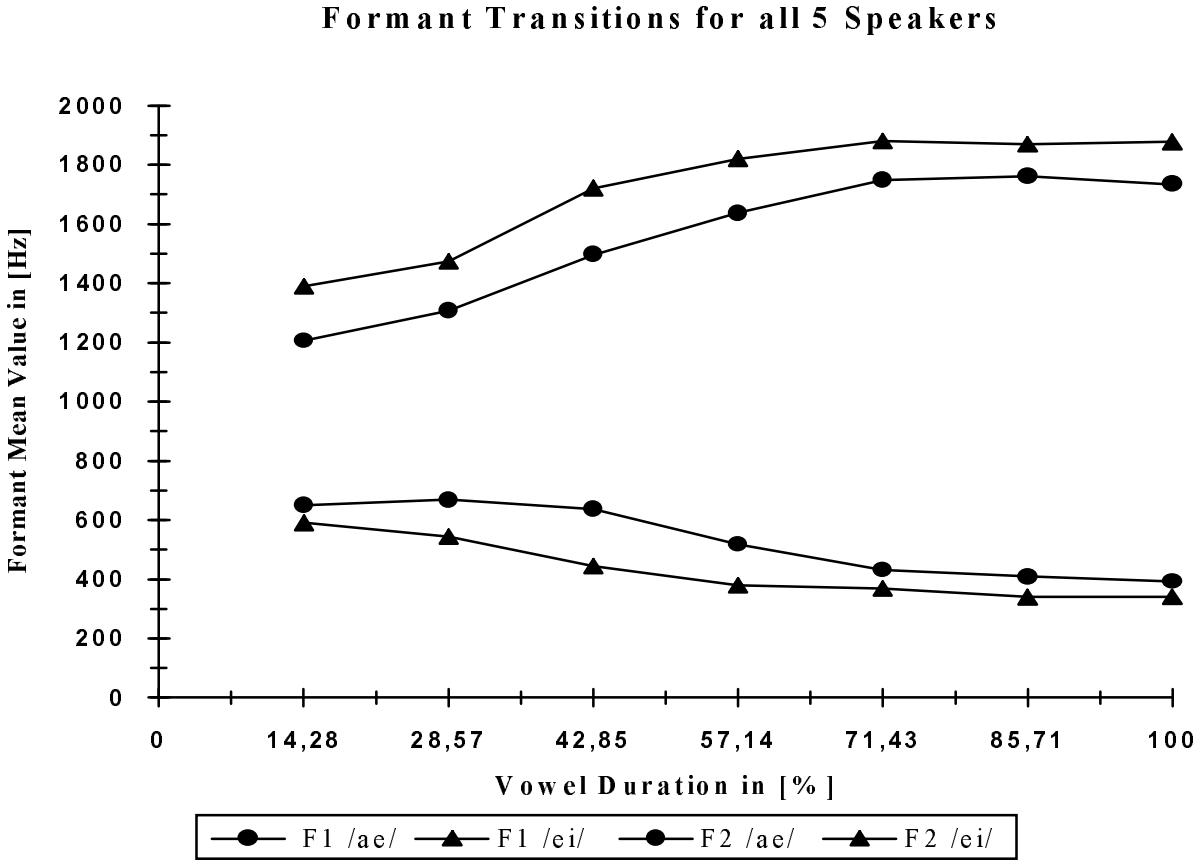
## 3. RESULTS AND DISCUSSION

Due to the small amount of data we confine ourselves to a qualitative discussion of the results.

Our data show close proximity of start and end points for the two contrasting diphthongs (see figure 1). A more vivid representation of the formant trajectories over time is given in figure 2. The same 7 equidistant points are plotted in this formant chart to shed light on the temporal behaviour of the trajectory. For both diphthongs there is clearly a movement towards the front high vowel space. The dots representing the /ɑɐ/ quality remain through points 2, 3, and 4 at an open-mid height with a slight front movement, and it is only at point 5 that a greater velocity is attained (which is represented by a larger distance between 4 and 5). The

little squares that represent the /əi/ quality start with point 2 at a slightly higher level (lower F1) and show at point 4 a somewhat faster upward movement. The trajectory already reaches a target region at point 5. This target region is a little bit higher and more fronted than the /ɑe/ target region represented with the dots 6, 7, 8.
It thus appears that the dynamical patterns require closer attention.

**Figure 1.** The formant values were determined at 8 equidistant points, with normalized time intervals. The first measurement point is left out because of excessive perturbation by coarticulatory effects of the preceding consonant.




It could be observed that /ae/ has a slightly longer overall duration than /əi/ for 4 speakers (speaker 3 is not very clear).
In fact the original formant patterns that are not normalized in time could at first sight be described in the foolowing way: The /əi/ type as a version truncated in the front part of the /ɑe/ type pattern. Whereas the /ɑe/ seems to consist of a steady-state-onset glide steady-state-offset pattern, the /əi/ type immediately starts with a glide phase resulting in a steady state offset. Peeters [3, 4] uses this 3 phase description to classify diphthongs across different languages.
So it might be argued that there is mainly a durational difference between the two diphthongs which is realized in the first part of the /ae/ diphthong. However, the durational effect is not too large, and might be thus attributable to the well known differences in inherent duration in vowels of different height (Lehiste [5], pp. 18-27). On the other hand the vowel height differences here measured are not that large.
A shortcoming of the Peeters' system of classifying diphthong dynamics is that it cannot easily be judged quantitatively what constitutes a steady state portion.
An interesting and maybe more promising measure seems to us to be the distance between acoustic vowel onset and the velocity peak in the formant transition. The data for some of the speakers for F2 peak velocity time are given in table 2. The absolute peak velocity timing might be a relevant cue, although it could be important to set this value in correspondence to the vowel duration too. Those values are given in the right column of table 2.
So one can evade the shortcomings of time normalization, that may mask relevant dynamic characteristics.

Figure 2. Trajectories of the diphthongs measured. The estimated values for cardinal vowels have been indicated.

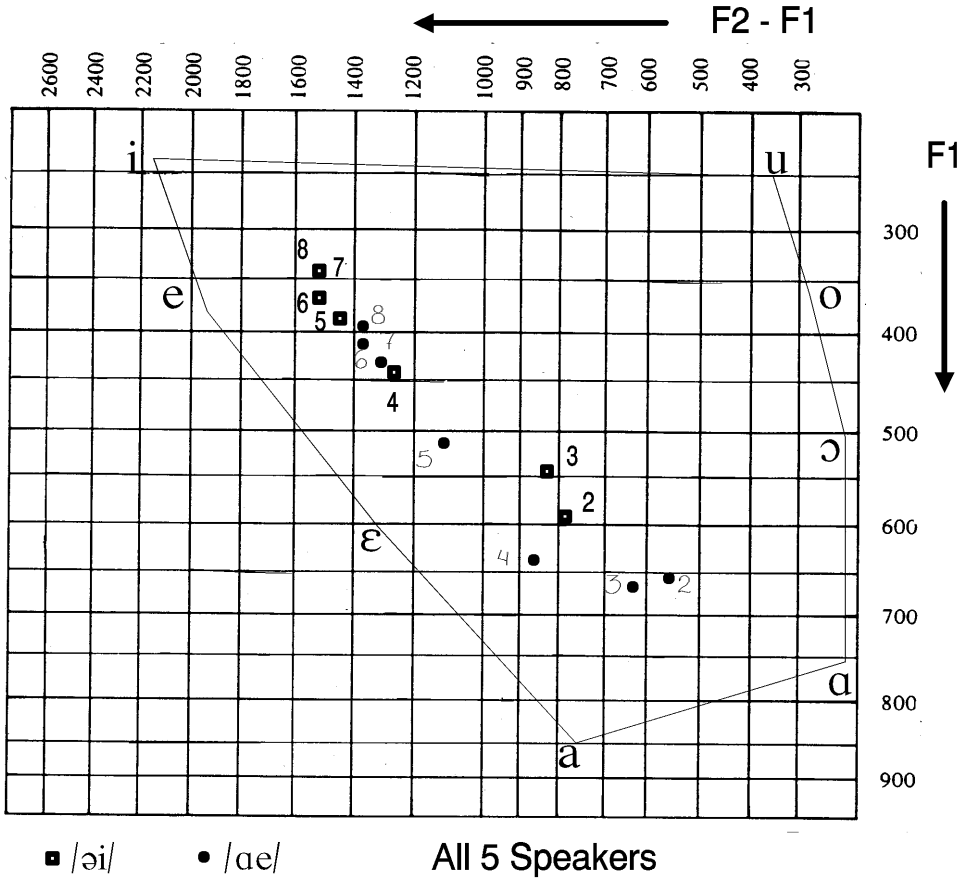



**Table 1.** Durations in [ms]

| Speaker | /ɑe/ | | /əi/ | |
|---|---|---|---|---|
| | mean value | σn | mean value | σn |
| **Speaker 1** | **226** | **19** | **192** | **17** |
| **Speaker 2** | **219** | **28** | **179** | **33** |
| **Speaker 3** | **217** | **28** | **211** | **31** |
| **Speaker 4** | **230** | **39** | **214** | **42** |
| **Speaker 5** | **253** | **46** | **227** | **44** |
| **All Speakers** | **229** | | **205** | |

**Table 2.** Interval between vowel onset and F2 peak velocity time in [ms]; this duration in [%] related to corresponding vowel duration

| Speaker | /ɑe/ | | /əi/ | |
|---|---|---|---|---|
| | peak velocity F2 | after percent duration | peak velocity F2 | after percent duration |
| **Speaker 4** | **94ms** | **41%** | **65ms** | **30%** |
| **Speaker 5** | **98ms** | **39%** | **73ms** | **32%** |

## 4. Conclusion

There is certainly not enough data here to give a clear judgement about whether the slightly differing vowel onset and offset qualities in the two diphthongs are already a sufficient cue for adequate differentiation. There are - as far as we know - no detailed acoustic analyses of this dialect. But linguistic descriptions as given by Russ [1] show that this is a very complex vowel system, even compared to Standard German. So for the speakers and listeners of the Swabian dialect even slight differences in quality might be a sufficient cue. On the other hand the shown dynamical distinctions should not be neglected. Furthermore it seems possible that interactions between vowel quality perception and the dynamic pattern are underlying, which are still not well understood.

Peeters ([3], pp. 306) observed that some subjects of his perceptual experiment remarked that some of the stimuli sounded like different diphthongs, or gave the impression of varying articulatory openness, although onset and offset frequencies were fixed and only the durational relationship among the pattern components had been varied.

More extensive analyses and perceptual experiments are planned.


## ACKNOWLEDGEMENTS

Thanks to Markus Hiller for his native listener judgements and support.

The author of this work was funded by the German federal ministry for research and technology (BMBF) in the framework of the VERBMOBIL project.


* This work was also published in the *Proceedings of Eurospeech* 1997, Rhodes (Greece).

## REFERENCES


[1] Ch.V.J. Russ, "Swabian" In: Ch.V.J. Russ, ed. *The Dialects of Modern German. A Linguistic Survey.* Routledge, London, pp. 337-363, 1990.

[2] A. Geumann and M. Hiller, "Diphthong Dynamics in Swabian" Third Joint Meeting of the Acoustical Society of America and the Acoustical Society of Japan, Honolulu, Hawaii. *JASA* vol.100, no.4, pt.2, p.2687, 1996.

[3] W.J.M. Peeters, *"Diphthong Dynamics. A cross-linguistic perceptual analysis of temporal patterns in Dutch, English, and German".* Diss. Utrecht. Mondiss, Kampen, 1991.

[4] W.J.M. Peeters and W.J. Barry *"Diphthong Dynamics: Production and Perception in Southern British English."* Proc. EUROSPEECH'89, Paris, vol. 1, pp. 55-58, 1989.

[5] I. Lehiste, *"Suprasegmentals"* MIT Press, Cambridge, MA, 1970.