AN ACOUSTIC ANALYSIS OF HESITATION PARTICLES IN GERMAN

M. Pätzold and A. Simpson
IPDS, Kiel, FRG

ABSTRACT

A spectral analysis of the vocalic portions of hesitation particles produced by three speakers of North German was undertaken. For two of these speakers vocalic portions of similar quality in lexical items were also analysed and found to be significantly different from the vocalic portions of hesitation particles.

INTRODUCTION

During the course of an interactional exchange movements in a speaker's vocal tract are fulfilling linguistic, interactional and primary biological functions. These functions can be carried out in temporal overlap, e.g. German *ja* produced on a pulmonic ingressive airstream is doing both linguistic and interactional work; counting out loud can be carried on while inhaling, allowing talk to continue while a primary biological function is performed.

In the majority of cases it is possible to assign the phonetics being produced in talk to one of these functional categories or, if need be, tease the various components apart if two or more functions are being accomplished simultaneously.

For a small number of items in German, however, the assignment of the phonetics to one or the other category is not always transparent. Hesitation particles in German are an example of this. By hesitation particles we mean a syllable comprising a vowel (plus bilabial nasal) employed by some speakers at trouble spots in talk, often represented in conversational transcripts with *uh(m)* (English) or *äh(m)* (German). While it is clear that hesitation particles only do interactional work, they do consist of a vowel (plus nasal for some speakers) and the question arises as to whether the phonetics which make up these particles are correlates of the same phonological systems and structures which make up lexical items.

Levelt [1,2] suggests that both possibilities must be entertained: while the vowel in hesitation particles may represent the neutral position of the oral cavity for many languages, the [ɛ] quality found in hesitation particles in Swedish may be one consequence of acquiring a form of derived lexical status [1:74].

In this paper we would like to provide tentative acoustic evidence from German that shows hesitation particles to be phonetically different from lexical items, i.e. although they make use of the facilities provided by the vocal tract hesitation particles do not take part in the phonology of the language. However, we will also distance ourselves from Levelt's claim that the vowels in these items represent the neutral position of the oral cavity.

We examine the acoustic details of the vocalic portions in the hesitation particles of three speakers of German and for two of these speakers we compare the vocalic portions in hesitation particles with those found in a selection of lexical items. We show that the vocalic portions in hesitation particles are significantly different from those of lexical items having similar quality.

The phonetic description of hesitation particles has received relatively little attention and has been largely restricted to Fo and duration [3,4]. The treatment of qualitative aspects is rare and restricted to brief impressionistic description [5:246].

DATA AND METHOD

The material analysed here was collected at the IPDS Kiel as part of the Verbmobil project [6]. Speakers were required to arrange a number of appointments within a two-month period displayed on a sheet they had before them. Appointments could not be made on certain days in the two-month period. These days were indicated by shaded areas which were different for each speaker. The speakers communicated via headsets and had to press a button if they wanted to talk, at the same time blocking the channel for the other speaker. This set-up elicited spontaneous data which excludes turn overlap, back channel responses, etc. For more details on the technical set-up used, elicitation materials, etc. see [7].

The hesitation particles produced by three speakers (henceforth TIS, OLV, GEP) were subjected to an LPC spectral analysis and the first two formant frequencies at around the mid-point of the vocalic portion were measured.

For two of the speakers (TIS and OLV) values for the first two formants were also obtained at the midpoint of vocalic portions in lexical items considered to lie in the vicinity of the vocalic portions in hesitation particles. The lexical vowels chosen for comparison are ɛ (e.g. *fest*), ɑ (e.g. *Tage*), ə (e.g. *bitte*) and ɐ (e.g. *wieder*). We will use ɜ to represent the vowel found in hesitation particles under investigation, although, as with the symbols being used for the other vowels, ɜ is not meant to directly represent the phonetic qualities found.

One of the problems of making a comparison of the vocalic portions in hesitation particles with those in lexical items are the considerable durational differences. The vocalic portions of hesitation particles are very long, having mean durations more than twice that of long open vowels in lexical items. So, whereas any contextual effects produced by neighbouring are likely to be negligible the same can not be said for vocalic portions in lexical items. In an attempt to minimize these contextual effects only vocalic portions in lexical items with a duration greater than 80ms were analysed.

RESULTS

Both GEP and OLV produced hesitation particles comprising only vocalic portions. These had central quality both in the height and front-back dimensions. The hesitation particles produced by TIS consist of a half-open central vocalic portion followed by a bilabial nasal. The vocalic portions in all cases were monophthongal.

Table 1: Mean and standard deviations of F1 and F2 of vocalic portions in hesitation particles for the three subjects. Measurements made at around the mid-point of the vocalic portion.

	F1		F2		
	$\bar{x}$	*s*	$\bar{x}$	*s*	n
OLV	520	38	1556	93	38
TIS	569	35	1273	46	40
GEP	377	28	1475	95	30

As we can see from Table 1 the impressionistic differences are supported by the inter-speaker differences in the formant values. TIS has a higher F1 and lower F2 indicating a vocalic portion more open in quality than those found for OLV and GEP.

Table 2: Mean and standard deviations of F1 and F2 of the vocalic portions in hesitation particles and a selection of vowels of similar quality. Measurements made at around the mid-point of the vocalic portion.

a) TIS

	F1		F2		
	$\bar{x}$	*s*	$\bar{x}$	*s*	n
ɛ	501n.s.	92	1797***	169	9
ɐ	556n.s.	77	1371***	120	21
ɔ	438***	62	1387**	183	18
ɑ	695***	46	1289n.s.	108	67
ɜ	569	35	1273	46	40

b) OLV

	F1		F2		
	$\bar{x}$	*s*	$\bar{x}$	*s*	n
ɛ	561***	44	1680***	155	26
ɐ	616***	81	1353***	133	34
ə	461***	53	1569n.s.	230	16
ɑ	659***	51	1236***	87	70
ɜ	520	38	1556	93	38

The comparison of the vocalic portions in hesitation particles with those in a selection of lexical items is shown in Table 2 for two speakers. For TIS the quality of the vocalic portions in hesitation particles is closest to that found for ɐ, for OLV it is the phonetic realisation of ə which is closest. Although in all cases either F1, F2 or both formant values are significantly different from those found in the vocalic portions of lexical items.

DISCUSSION

We set out to show that the phonetic quality of the vocalic portions in hesitation particles is significantly different from that of vocalic portions in lexical items. Our results show that for two speakers of German this would seem to be the case. The vocalic portions of hesitation particles have their own quality suggesting that they are phonetic correlates of a phonological system which is different from those employed in lexical items. This result is hardly surprising when one considers that these particles serve to indicate linguistic trouble, a function which can be successfully fulfilled by being different in form from surrounding linguistic material.

Our results in part lend support to Levelt's claim that hesitation particles are different from words. But Levelt's claim is stronger than this. He suggests that the vowel of a hesitation particle is a neutral sound which varies phonetically with "the neutral position of the oral cavity from different languages" [1:74], a vowel which he slightly later refers to as schwa. This claim is far harder to substantiate since it is not clear how one would go about ascertaining the "neutral position of the oral cavity" for a language, or even an individual speaker. Levelt is presumably referring to a cavity position which is dependent upon the articulatory setting [8] of a language and not to an independently motivated articulatory (e.g. [9:137]) or acoustic construct [10:49f]. Levelt's later use of schwa is equally problematic as it can only be referring to the auditory product of the neutral position and not to the phonetic vowel category [ə] or the phonetic correlate of a phonological item such as /ə/ often proposed for languages such as English and German.

On the basis of our data, we would like to make a claim which can be tested and refuted. The vowel quality found in hesitation particles is different from vowel qualities found in lexical items. A consistent difference is maintained, although the exact nature of this difference varies from speaker to speaker. The three speakers we investigated all produced vowels which were central, but with considerable inter-individual variation in height.

Indeed, it would be possible, to test our claim on a language, such as Swedish in which the non-central [ɛ] quality of the vowels in hesitation particles led Levelt to claim that such items were taking part in the phonology of the lexis. It would be interesting to see if the vowel quality found is different from that found in lexical items such as *lära, läkare*.

Problems of Comparability

One of the biggest problems we encountered in this study was the degree to which the items we are comparing are indeed comparable. Hesitation particles are in general prominent, brought about by factors such as length and loudness. We would therefore have been justified in comparing hesitation particles with vowels from prominent syllables in lexical items, i.e. those which are stressed. Although this comparison would have been simple, we wanted to go one step further and demonstrate that the vowels of hesitation particles are even different from those of central vowels in the language. This causes a serious problem since the central vowels in German (ɐ and ə) are always unstressed and the quality varies greatly, not least because of the often short duration in various consonantal and vocalic contexts. We therefore imposed a minimum duration of 80ms in an attempt to minimize these effects, while still being able to get a sufficient number of tokens. However, the setting of a lower duration is also not without problems as it almost exclusively returns central vowels in open syllables, which in the case of ə were also utterance final. One way of overcoming this problem may be to record the same speakers producing spoken prose. This would allow the structures and frequency of occurrence to be controlled.

Other aspects of hesitation particles

In the course of analysing the vowels of hesitation particles, we made a number of other observations which suggest that hesitation particles make use of a different set of phonetics from lexical items. So, for instance, the hesitation particles produced by TIS consisted of a vowel plus nasal sequence, however, the vowel was rarely nasalized and the soft palate was often lowered shortly after bilabial closure was made for the nasal, leading to nasal plosion. What is of interest is that this complex was different from similar complexes in lexical items, i.e. either vowel-stop-nasal sequences (e.g. *eben*) or vowel-nasal sequences (e.g. *gemeinsam*).

We suspect that as the amount of phonetic material gathered on hesitation particles in German and other languages grows, so too will the catalogue of differences between them and linguistic items.

ACKNOWLEDGEMENT

The data collection was partially funded by the German Federal Ministry for Education, Science, Research and Technology (BMBF) under grant 01IV101M7. The responsibility for the contents of this study lies with the authors.

REFERENCES

[1] Levelt, W.J.M. (1983), "Monitoring and self-repair in speech", *Cognition*, vol. 14, pp. 41-104.

[2] Levelt, W.J.M. (1989), *Speaking: from intention to articulation*, Cambridge: MIT Press.

[3] O'Shaughnessy, D. (1992), "Recognition of hesitations in spontaneous speech", *Proc. Int. Conf. on Acoustics, Speech and Signal Processing*, vol. 1, pp. 593-596.

[4] O'Shaughnessy, D. (1992), "Locating disfluencies in spontaneous speech: an acoustical analysis", in *Proc. European Conf. on Acoustics on Speech Communication and Technology*, vol. 3, pp. 2187-2190.

[5] Wittkop, E.-M. (1988), *Gliederungspartikeln im Dialog*, Munich: iudicium.

[6] Karger, R., Wahlster, W. (1994), *VERBMOBIL Handbuch*, Verbmobil Technisches Dokument 17, Saarbrücken: DFKI.

[7] Pätzold, M., Scheffers, M., Simpson, A., Thon, W. (1995), "Controlled elicitation and processing of spontaneous speech in Verbmobil", *Proc. XIIIth ICPhS.*

[8] Honikman, B. (1964), "Articulatory settings", in D. Abercrombie, D.B. Fry, P.A.D. MacCarthy, N.C. Scott, J.L.M. Trim (eds.) *In Honour of Daniel Jones*, London: Longman Green, pp. 73-84.

[9] Laver, J. (1994), *Principles of phonetics*, Cambridge: CUP.

[10] Fant, G. (1960), *Acoustic theory of speech production*. The Hague: Mouton.