Laughter and Other Non-Verbal Vocalisations Workshop 2020
5 October 2020, Bielefeld, Germany

Acoustic vowel quality of filler particles in German

Malte Belz
Humboldt-Universität zu Berlin
malte.belz@hu-berlin.de

Abstract

The vowel quality of filler particles (FP) is studied for 24 speakers of German who produced 666 instances of vocalic (*äh*) and vocalic-nasal forms (*ähm*) in spontaneous dialogues. The FP vowel quality is compared to reference vowels of a word list as well as to phonologically and graphematically similarly constructed lexical syllables. Filler particles show a complete overlap with the reference vowels [œ] and [ɐ], but overlap only partially with [ɛ] and [ə].

1 Introduction

In German, vocalic (V) and vocalic-nasal (VN) filler particles (FPs) are orthographically represented by ⟨äh⟩ and ⟨ähm⟩, which implies a phonological representation as /ɛ/ and /ɛm/, according to grapheme-phoneme correspondences (Fuhrhop and Peters, 2013). While FP vowel duration is often longer than those in lexical words (Shriberg, 2001), not much is known about their vowel quality. Some researchers assume a language-specific distribution, for example, the vowel /ə/ is used for FPs in English (Lickley, 2015) and /e/ in Spanish (Roggia, 2012), and both are used within the phonological system of their respective language.

This study investigates the vowel quality of FPs in German and is motivated by the difference between their auditorily assumed realizations and their acoustically measured vowel qualities. Although their quality has been assessed impressionistically in German and can be represented as [e eː ɛ ɛː æ ə əː ɐ ɐː œ œː ø] (Willkop, 1988; Rasoloson, 1994; Batliner et al., 1995), a systematic study of their acoustic quality with a sufficient amount of speakers is lacking.

The most recent acoustic study assesses the vowel qualities of ten male speakers of German as close to [ə ɐ œ øː ʏ ʊ] (Klug, 2013). Previously, Pätzold and Simpson (1995) hypothesize that FP vowels are phonetically different from vowels occurring in lexical items. They compare FP vowels of two German speakers to vowels in lexical items and report FP vowel qualities close to the reduced vowels /ə ɐ/ as well as significant differences as compared to the lexical vowels [ɛ ɐ ə ɑ], albeit some speaker-specific overlaps are observed. However, they fall short on "comparing [FP vowels] with vowels from prominent syllables in lexical items, i.e. those which are stressed" (Pätzold and Simpson, 1995, 514). Therefore, apart from reference vowels, the present study also uses phonologically and graphematically similarly constructed lexical syllables in order to better understand the location of FP vowels in the vowel space.

2 Method

2.1 Corpus

The corpus used for this study consists of 24 native speakers of German, 12 males and 12 females, speaking in 12 dialogues. They are recorded in a soundproof cabin via head-mounted microphones with a sample rate of 44 100 Hz. The interlocutors were seated facing each other. Each dialogue lasts approximately 15 min and was initiated by the experimenter, asking them about their experience with the university canteen. Parts of the data are published in the corpus BeDiaCo v.1 (Belz and Mooshammer, 2020).

At the beginning and end of the experiment, the participants read a list of disyllabic words embedded in the carrier sentence *Sage X bitte* ('Say X please'). The items permutated all 15 monophthongs of German in the first syllable. The second syllable contained either [ə] or [ɐ] (e.g., *bäte* ['bɛːtə] 'asked', *Täter* ['tɛːtɐ] 'offender').

The words *Äther* ['ʔɛːtɐ] ('ether') and *Ämter* ['ʔɛmtɐ] ('offices') have been included twice in

the list, as their first syllables are similar to the /ɛ/-vowel assumed for *äh* and *ähm*.

2.2 Annotation, correction, and query

FPs are annotated in Praat (Boersma and Weenink, 2019) following the guidelines of Belz (2019). The items in the word list are annotated for all full and schwa vowels on a separate tier. The corpus is transformed into an EMU database (Winkelmann et al., 2017) and queried with the R package *emuR* v. 1.1.1 (Winkelmann et al., 2018) in R, not considering glottal FPs (Belz, 2017). Formants are added to the database via the Praat formant trackers (Winkelmann, 2015), separately for each gender. F1 and F2 of all vowels in FPs and in the word list are manually corrected, if necessary. Formant values are obtained at the vowel midpoint.

2.3 Statistics

The degree of overlap of two vowel distributions is calculated by means of a multivariate analysis of variance (MANOVA) with F1 and F2 as combined dependent variable. Pairs of FP vowels and lexical vowels are taken as independent variable. The Pillai's trace is a test statistic from the MANOVA, ranges between 0 and 1, and represents the shared variation of F1 and F2 (Nycz and Hall-Lew, 2013, 5). The higher the Pillai's trace, the greater the distance of F1 and F2 between two vowels (or between two distributions). Conversely, for low values, two vowels overlap (Hay et al., 2006, 467).

15 cases for which participants pronounced [ˈeːtɐ] and one case of [ɛˈtɛɐ] were discarded; one participant skipped *Äther* once. A total of 96 cases of /ɛ/ in *Ämter* and 79 cases of /ɛː/ in *Äther* are included in the analysis, as well as 1 402 reduced and 1 439 full monophthongic reference vowels.

3 Results

3.1 Vowel quality of FPs

The corpus contains 37 263 word tokens on the transliteration tier, 306 vocalic and 360 vocalic-nasal FPs. Figure 1 shows the vowel space for female and male speakers containing reference vowels of the word list, FP vowels of V and VN forms, and the first vowels of the words *Äther* and *Ämter*.

The vowel quality of both vocalic and vocalic-nasal FPs in German completely overlaps with the ellipse of [œ]. For male speakers, the FP ellipses of V and VN forms overlap completely with [ɐ]. For female speakers, only VN forms show a complete overlap. Additionally, partial overlaps between the FP vowels and the reference vowels can be observed for [a aː ɛ ɛː ɔ ə ʏ ø ɪ ɐ ø ʏ].

The distributional differences between the word list vowels and the FP vowels are assessed by a MANOVA, comparing their F1/F2-distribution to the most adjacent vowels [a ɐ ə ɛ ɛː œ ø ʏ], respectively. Table 1 shows the results for each comparison, ordered by Pillai traces.

Table 1: Pillai traces of distribution overlap between vowels in vocalic (V) and vocalic-nasal (VN) FPs and lexical vowels in the wordlist (Vowel), with corrected p-values (Bonferroni), ordered by effect.

FP	Gender	Vowel	Pillai	Df	F	p
V	f	œ	0.026	109	1.430	= 1
V	f	ə	0.136	109	8.530	< 0.05
V	f	ɐ	0.140	109	8.810	< 0.01
V	f	ʏ	0.150	109	9.500	< 0.01
V	f	øː	0.171	109	11.150	< 0.01
V	f	ɛ	0.313	109	24.560	< 0.001
V	f	ɛː	0.316	109	24.930	< 0.001
V	f	a	0.423	109	39.620	< 0.001
V	m	œ	0.005	217	0.520	= 1
V	m	ɐ	0.044	217	4.960	= 0.3
V	m	ə	0.062	217	7.140	< 0.05
V	m	ʏ	0.112	217	13.650	< 0.001
V	m	øː	0.149	217	18.860	< 0.001
V	m	a	0.162	217	20.890	< 0.001
V	m	ɛ	0.277	217	41.310	< 0.001
V	m	ɛː	0.349	217	57.880	< 0.001
VN	f	œ	0.006	136	0.420	= 1
VN	f	ɐ	0.032	136	2.230	= 1
VN	f	ə	0.237	136	20.930	< 0.001
VN	f	ʏ	0.261	136	23.810	< 0.001
VN	f	a	0.291	136	27.700	< 0.001
VN	f	øː	0.299	136	28.780	< 0.001
VN	f	ɛ	0.439	136	52.840	< 0.001
VN	f	ɛː	0.445	136	54.020	< 0.001
VN	m	ɐ	0.001	244	0.150	= 1
VN	m	œ	0.032	244	4.060	= 0.6
VN	m	a	0.099	244	13.320	< 0.001
VN	m	ə	0.181	244	26.810	< 0.001
VN	m	ʏ	0.248	244	40.090	< 0.001
VN	m	ɛ	0.257	244	41.960	< 0.001
VN	m	øː	0.290	244	49.730	< 0.001
VN	m	ɛː	0.307	244	53.720	< 0.001

A complete overlap is found for [œ] for both genders and FP forms. [ɐ] overlaps completely for V and VN forms of male speakers, yet for only for VN forms of female speakers. Statistics repeated with bark-transformed values achieve similar results. Table 2 summarises the means and standard deviations of the vowels in FP forms and lexical items.

3.2 Comparison to similar lexical syllables

Table 3 shows the Pillai's traces for the differences between the read vowels of *Äther/Ämter* /ɛː ɛ/ and the vocalic part of FPs. Vocalic FPs are compared to /ɛː/, vocalic-nasal FPs are compared to /ɛ/. FP vowels are clearly different from the presumably

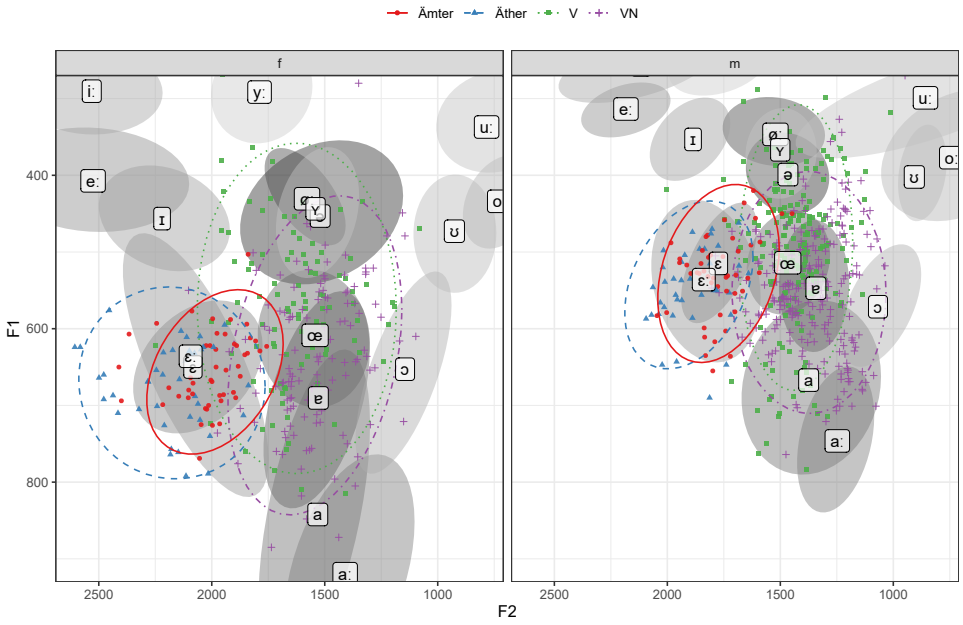

Figure 1: FP vowels in vocalic (V, green dotted) and vocalic-nasal (VN, violet dotdashed) forms as well as in *Ämter* [ɛ] (red solid) and *Äther* [ɛː] (blue dashed) within vowel spaces for women (f) and men (m). Data ellipses contain 95 % of all values.

Table 2: Means ($\overline{x}$) and standard deviations (s) of vowels in FP and lexical items (in Hz).

Origin		female		male	
		$\overline{x}_{F1}$	$\overline{x}_{F2}$	$\overline{x}_{F1}$	$\overline{x}_{F2}$
FP	V	572 (103)	1621 (235)	498 (94)	1435 (128)
FP	VN	634 (97)	1540 (181)	552 (76)	1365 (157)
Ämter	ɛ	653 (49)	2004 (149)	529 (54)	1770 (124)
Äther	ɛː	673 (54)	2186 (186)	548 (52)	1893 (122)
Lex.	ə	453 (38)	1520 (146)	399 (21)	1466 (73)
Lex.	øː	431 (27)	1579 (90)	347 (23)	1522 (90)
Lex.	ɐ	690 (56)	1528 (102)	547 (37)	1343 (63)
Lex.	œ	609 (41)	1543 (114)	514 (29)	1467 (71)
Lex.	a	842 (88)	1531 (125)	667 (54)	1375 (122)
Lex.	aː	919 (71)	1410 (121)	746 (38)	1248 (75)
Lex.	ɛ	650 (34)	2083 (133)	515 (31)	1774 (63)
Lex.	eː	407 (32)	2525 (184)	314 (15)	2183 (75)
Lex.	ɛː	636 (81)	2094 (147)	536 (49)	1839 (100)
Lex.	ɪ	456 (32)	2220 (132)	353 (22)	1886 (81)
Lex.	iː	290 (23)	2531 (142)	256 (20)	2116 (138)
Lex.	ɔ	653 (52)	1145 (91)	571 (37)	1068 (84)
Lex.	oː	434 (24)	726 (64)	377 (38)	741 (151)
Lex.	ʊ	473 (32)	926 (76)	402 (38)	907 (132)
Lex.	uː	337 (40)	784 (131)	299 (40)	858 (269)
Lex.	ʏ	444 (37)	1541 (87)	367 (26)	1500 (67)
Lex.	yː	292 (27)	1789 (98)	254 (17)	1772 (90)

similar vowels /ɛː/ and /ɛ/ in *Äther/Ämter*.

Thus, FP vowels in vocalic forms are produced significantly higher and further back in the vowel space than /ɛː ɛ/ in *Äther/Ämter* for both genders (female F1: Welch Two Sample t-test, $t = -7.6$, $df = 130$, $p < 0.001$, 95% CI = $[-128, -75]$; female F2: Wilcoxon rank sum test with continuity correction, $W = 61$, $p < 0.001$; male F1: $W = 2\,196$, $p < 0.001$; male F2: $W = 57$, $p < 0.001$). For FP vowels in vocalic-nasal forms only F2 differs significantly from /ɛ/ in *Ämter* (F2, female: $W = 114$, $p < 0.001$; F2, male: $t = -19.6$, $df = 81.2$, $p < 0.001$, 95% CI = $[-446, -364]$; cf. also Table 2).

Table 3: Pillai traces of distribution overlap between vowels in vocalic (V) and vocalic-nasal (VN) FPs and /ɛ ɛː/ in *Äther/Ämter*, with corrected p-values (Bonferroni).

Comparison	Gender	Pillai	Df	F	p
V:Äther	f	0.61	138	108.99	< 0.001
V:Äther	m	0.63	243	209.56	< 0.001
VN:Ämter	f	0.60	172	127.74	< 0.001
VN:Ämter	m	0.51	280	145.42	< 0.001

Further, /ɛ/ in /ˈɛm.tɐ/ is significantly lower in F2 (but not different in F1) than /ɛː/ in /ˈɛː.tɐ/ for both male ($t = 4.6$, $df = 79.9$, $p < 0.001$, 95% CI = $[70, 176]$) and female speakers ($W = 1\,578$, $p < 0.001$).

4 Discussion

FP vowels show a broader range of possible realizations in vowel space than lexical vowels. Although previous evidence led to believe that FP vowels cover the regions of [ə] and [ɛ] in German, this can only be partially confirmed. At least for this study, FP vowels completely overlay only the acoustic ranges of the two lexical vowels [œ] and [ɐ]. Never-

theless, a range of other vowel qualities is touched upon. Therefore, it seems that the dissimilarity between FP and lexical vowels is not as categorical as implied by the statement that FP vowels are "different from those employed in lexical items" (Pätzold and Simpson, 1995, 514).

In general, VN forms are produced with a higher F1 and a lower F2, thus being located more down and back than V forms. Although the similarity between FP vowels and the vocalic portions of the phonologically and graphematically similarly constructed syllables in lexical items could only be partially confirmed, the results are nonetheless interesting as they show that FP vowels behave just in the same way on the anterior-posterior dimension as lexical vowels do if there is a bilabial nasal following the first vowel – they are then located more to the back of the vowel chart. This kind of behaviour raises the question whether vowels in VN forms are not just randomly (or epiphenomenally) produced further back, but are already planned in this way, and "have targets of their own" (Gick et al., 2004, 231). F3, which has not been addressed here, could potentially explain more of the vocalic behaviour before bilabial nasals, as the vowels could be rounded to some extent.

In conclusion, from an acoustic point of view, FPs in German are most closely represented by the symbolic forms [œ œm ɐ ɐm], although FP vowels can in principle come from a relatively wide range in the central part of the vowel space. Future studies may investigate how these acoustic instances are linked to their auditory perception.

References

Anton Batliner, Andreas Kießling, Susanne Burger, and Elmar Nöth. 1995. Filled pauses in spontaneous speech. Verbmobil Verbundvorhaben (Report 88).

Malte Belz. 2017. Glottal filled pauses in German. In *Proceedings of DiSS 2017*, number 58(1) in TMH-QPSR, pages 5–8.

Malte Belz. 2019. Ein Schema zur phonetischen Annotation von Füllpartikeln und ihres Kontextes. Humboldt-Universität zu Berlin.

Malte Belz and Christine Mooshammer. 2020. Berlin Dialogue Corpus (BeDiaCo). Medien-Repositorium der Humboldt-Universität zu Berlin.

Paul Boersma and David Weenink. 2019. Praat. Doing phonetics by computer.

Nanna Fuhrhop and Jörg Peters. 2013. *Einführung in die Phonologie und Graphematik*. Metzler, Stuttgart.

Bryan Gick, Ian Wilson, Karsten Koch, and Clare Cook. 2004. Language-Specific Articulatory Settings: Evidence from Inter-Utterance Rest Position. *Phonetica*, 61:220–233.

Jennifer Hay, Paul Warren, and Katie Drager. 2006. Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics*, 34:458–484.

Katharina Klug. 2013. „Ähm" – Sind Häsitations-Partikeln sprecherspezifisch? Untersuchung der Parameter Grundfrequenz und Vokalqualität. In Lutz Christian Anders, Ines Bose, Ursula Hirschfeld, and Baldur Neuber, editors, *Aktuelle Forschungsthemen der Sprechwissenschaft 3*, number 43 in Hallesche Schriften zur Sprechwissenschaft und Phonetik (HSSP), pages 65–94. Peter Lang, Frankfurt am Main.

Robin J. Lickley. 2015. Fluency and Disfluency. In Melissa A. Redford, editor, *The Handbook of Speech Production*, pages 445–469. John Wiley & Sons, Inc, Hoboken, NJ.

Jennifer Nycz and Lauren Hall-Lew. 2013. Best practices in measuring vowel merger. In *Proceedings of Meetings on Acoustics 20, 060008*, pages 1–19. Acoustical Society of America.

Matthias Pätzold and Adrian P. Simpson. 1995. An acoustic analysis of hesitation particles in German. In *Proceedings of the XIIIth International Congress of Phonetic Science (ICPhS 95)*, volume 3, pages 512–515.

Janie Noëlle Rasoloson. 1994. *Interjektionen im Kontrast*, volume 22 of *Arbeiten zur Sprachanalyse*. Peter Lang, Frankfurt am Main.

Aaron B. Roggia. 2012. *Eh* as a polyfunctional discourse marker in Dominican Spanish. *Journal of Pragmatics*, 44:1783–1798.

Elizabeth E. Shriberg. 2001. To 'errrr' is human: ecology and acoustics of speech disfluencies. *Journal of the International Phonetic Association*, 31:153–169.

Eva-Maria Willkop. 1988. *Gliederungspartikeln im Dialog*, volume 5 of *Studien Deutsch*. Iudicium, München.

Raphael Winkelmann. 2015. PraatToFormants2AsspDataObj.R.

Raphael Winkelmann, Jonathan Harrington, and Klaus Jänsch. 2017. EMU-SDMS. *Computer Speech & Language*, pages 392–410.

Raphael Winkelmann, Klaus Jaensch, Steve Cassidy, and Jonathan Harrington. 2018. emuR. R package version 1.1.1.