

Dokumentation der Forschungsdaten

GECO-FP v.1

Phonetische Füllpartikelannotation basierend auf GECO v.1.1

Malte Belz
Humboldt-Universität zu Berlin
Phonetik/Phonologie
`malte.belz@hu-berlin.de`

29. November 2019

Schlagwörter: Füllpartikeln, Gefüllte Pause, Mehr-Ebenen-Annotation, Phonetische Annotation, Dialogstruktur, Gesprochene Sprache, Phonetik, Korpuslinguistik

Key words: filler particle, filled pause, filler, spontaneous speech, dialogue structure, phonetics, multi-layer annotation, corpus linguistics

Inhaltsverzeichnis

1	Korpus	1
1.1	Zugrundeliegendes Korpus GECO v1.1	1
1.2	GECO-FP v.1	1
1.3	Struktur des Datensatzes	2
1.4	Annotation	3
	Literaturverzeichnis	5
	Anhang A Annotationsrichtlinien in GECO-FP	6
	Anhang B Zweifelsfälle	11
	Anhang C Vergleich IPA/SAMPA	12

1 Korpus

1.1 Zugrundeliegendes Korpus GECO v1.1

Das hier beschriebene Korpus GECO-FP v.1 (GECO-Füllpartikeln) basiert auf der GECO-Datenbank v.1.1 (GERman COversations, Schweitzer & Lewandowski 2013, PID: <http://hdl.handle.net/11022/1007-0000-0000-8E39-1>), die über eine persönliche Anfrage über das Institut für Maschinelle Sprachverarbeitung der Universität Stuttgart erhältlich ist (kontaktieren Sie bei Interesse geco@ims.uni-stuttgart.de).¹ GECO-FP v.1 enthält ausschließlich anonyme Annotationen und anonyme Aggregate. Die Erlaubnis für die Veröffentlichung der modifizierten Transliterationsebene *wordscor* wurde durch die Herausgeberinnen erteilt.

Die Nutzung von GECO v.1.1 und die vorliegende Veröffentlichung der annotierten Daten unter dem Namen GECO-FP v.1 erfolgte in Abstimmung mit Antje Schweitzer, der ich hierfür herzlich danken möchte.

1.2 GECO-FP v.1

GECO-FP (GECO-Füllpartikeln) baut auf einem Teil des GECO-Korpus auf. Es enthält Textdaten und Aggregate von acht Dialogen mit acht Sprecherinnen in der multidimensionalen Bedingung (die Sprecherinnen können sich sehen) mit 44 209 Token auf der *wordscor*-Ebene und basiert auf 197 Minuten gesprochener Sprache (ohne Pausen). Die Audiodateien aus GECO sind selbst *nicht* in den hier veröffentlichten Daten enthalten, da ihre Lizenzierung dies nicht erlaubt, sie sind aber über das Institut für Maschinelle Sprachverarbeitung der Universität Stuttgart erhältlich (s. o.). GECO-FP ist unter der Lizenz CC BY-NC-SA 3.0 DE veröffentlicht (<https://doi.org/10.18452/20794>).

¹<https://www.ims.uni-stuttgart.de/forschung/ressourcen/korpora/ims-geco/>, besucht am 27.11.2019.

Folgende Tonspuren aus GECO v1.1 wurden in GECO-FP ausgewertet.

1. multi_A-C_left
2. multi_A-C_right
3. multi_A-K_left
4. multi_A-K_right
5. multi_C-D_left
6. multi_C-D_right
7. multi_F-K_left
8. multi_F-K_right
9. multi_H-F_left
10. multi_H-F_right
11. multi_H-J_left
12. multi_H-J_right
13. multi_M-D_left
14. multi_M-D_right
15. multi_M-J_left
16. multi_M-J_right

1.3 Struktur des Datensatzes

Der Datensatz besteht aus TextGrids (Boersma 2001) und einer EMU-Datenbank (Winkelmann et al. 2017). Die TextGrid-Dateien enthalten alle sechs Annotati-
onsebenen und können in Praat (Boersma 2001) eingelesen werden. Die EMU-
Datenbank (Cassidy & Harrington 2001; Harrington 2010; Winkelmann et al. 2017)
enthält alle sechs Annotationsebenen und zusätzlich Grundfrequenz- und Formant-
trajektorien. Sie kann in R (R Core Team 2018) ausgelesen werden. Folgende Ver-
knüpfungen zwischen den verschiedenen Annotationsebenen sind in EMU hinzuge-
fügt worden:

```

1 add_linkDefinition(geco, type = "ONE_TO_MANY", superlevelName =
  "wordscor", sublevelName = "hes")
2 add_linkDefinition(geco, type = "ONE_TO_MANY", superlevelName = "hes",
  sublevelName = "segm")
3 add_linkDefinition(geco, type = "ONE_TO_MANY", superlevelName =
  "segm", sublevelName = "coart")
4 add_linkDefinition(geco, type = "ONE_TO_MANY", superlevelName =
  "func", sublevelName = "wordscor")
5 add_linkDefinition(geco, type = "ONE_TO_MANY", superlevelName =
  "func", sublevelName = "hes")

```

GECO-FP enthält keine weiteren Metadaten. Die Tools zum Auslesen der Daten (Praat und emuR) sind Open-Access-Tools.

1.4 Annotation

GECO-FP enthält Annotationen von Füllpartikeln (beispielsweise *äh* oder *ähm*) auf sechs Ebenen, deren Struktur und Annotationswerte in Anhang A beschrieben sind. Die Annotation erfolgt in Praat (Boersma 2001) mit TextGrids. Die TextGrids sind entsprechend ihren Tonspuren benannt, also beispielsweise *multi_A-C_left.TextGrid*.

1. Ebene *wordscor*. Beruht in größten Teilen auf der Transliteration in der Ebene *words* aus GECO, mit Korrekturen der Transliteration und Alignierung in Bereichen von Füllpartikeln (vgl. Tabelle A.1).
2. Ebene *hes*. Annotation der Füllpartikeln-Exponenten und Annotation ihrer sequenziellen Kontexte (Laute, Pausen, etc.) (vgl. Tabelle A.2).
3. Ebene *segm*. Annotation der enthaltenen Segmente (vgl. Tabelle A.3).
4. Ebene *coart*. Annotation der Phonationsart des Vokals (vgl. Tabelle A.4).
5. Ebene *vowel*. Referenzvokale der Sprecherinnen.
6. Ebene *func*. Annotation der Dialogstruktur (vgl. Tabelle A.5).

Anhang B gibt einen Überblick über Zweifelsfälle, Anhang C einen Vergleich zwischen IPA und SAMPA. Ein Beispiel für die Annotation findet sich in Abbildung 1.2.

Der Grundfrequenztrack in EMU wurde mit dem R-Paket *wrassp* (Bombien et al. 2018) und der Funktion *ksvF0()* mit geschlechtsspezifischen Einstellungen in R berechnet (mit einem Bereich von 80–640 Hz).

Die Formantentracks wurden mit dem Formantentracker von Praat (Boersma 2001) über das R-Paket *PraatR* (Albin 2014) mithilfe des Skripts *PraatToFormants2AsspDataObj* (Winkelmann 2015) mit geschlechtsspezifischen Einstellungen ermittelt und als Track in der EMU-Datenbank gespeichert (nach Burg, maximale Anzahl von fünf Formanten, obere Schwelle für den maximalen Formanten bei 5500 Hz für Frauen, Fensterlänge von 0,025 s, Voremphase von 50 Hz). Für alle Füllpartikelvokalsegmente wurden die ersten drei Formanten manuell in der EMU-Datenbank korrigiert (vgl. Abbildung 1.1). Dabei wurden klare Ausreißer an die umgebende Trajektorie angepasst.

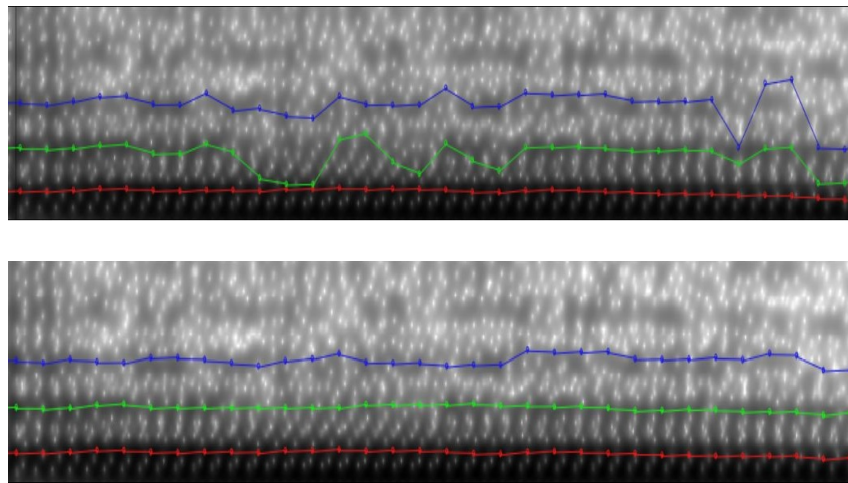


Abbildung 1.1: Beispiel für einen Vokal von 200 ms in *ähm* mit F₁ (rot), F₂ (grün) und F₃ (blau) nach der Formantenermittlung durch Praat (oberes Bild) und nach der manuellen Korrektur (unteres Bild), Sprecherin D, Dialog multi_C-D_right, Sek. 4,86–5,86.

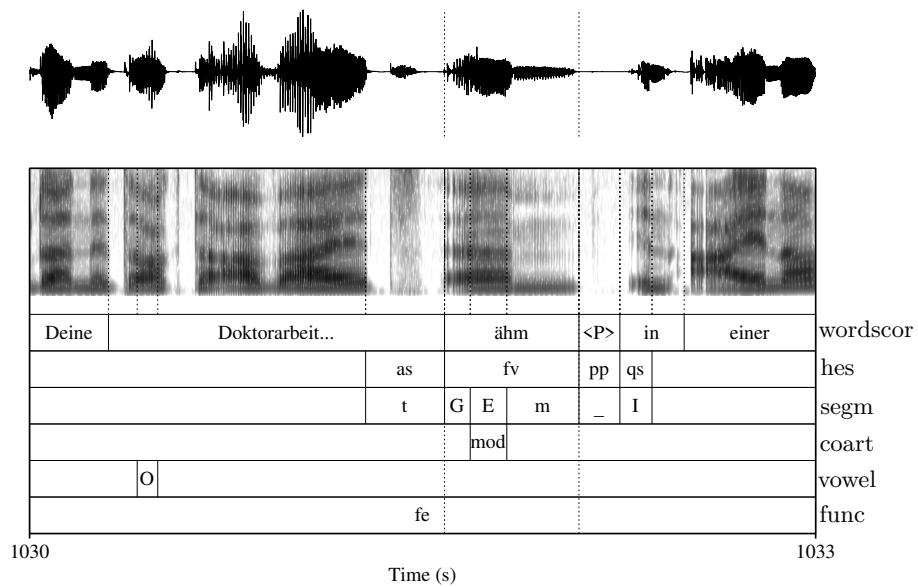


Abbildung 1.2: Ausschnitt aus GECO-FP v.1 *multi_A-C_left* mit den sechs Annotationssebenen.

Literaturverzeichnis

- Albin, Aaron (2014): PraatR. An architecture for controlling the phonetics software "Praat" with the R programming language. *Journal of the Acoustical Society of America* 135 (4). 2198. DOI: 10.1121/1.4877175.
- Boersma, Paul (2001): Praat, a system for doing phonetics by computer. *Glott International* 5 (9). 341–345.
- Bombien, Lasse; Raphael Winkelmann & Michel Scheffers (2018): *wrassp. An R wrapper to the ASSP Library*. R package version 0.1.8.
- Carletta, Jean; Amy Isard; Stephen Isard; Jacqueline Kowtko; Alison Newlands; Gwyneth Doherty-Sneddon & Anne H. Anderson (1996): *HCRC Dialogue Structure Coding Manual*. URL: www.hcrc.ed.ac.uk/publications/tr-82.ps.gz.
- Cassidy, Steve & Jonathan Harrington (2001): Multi-level annotation in the Emu speech database management system. *Speech Communication* 33 (1-2). 61–77. DOI: 10.1016/S0167-6393(00)00069-8.
- Harrington, Jonathan (2010): Acoustic Phonetics. In: Hardcastle, William J.; John Laver & Fiona E. Gibbon (Hrsg.): *The handbook of phonetic sciences*. 2nd ed. (= Blackwell handbooks in linguistics). Chichester, U.K. & Malden, MA: Wiley-Blackwell. 81–129.
- R Core Team (2018): *R: A language and environment for statistical computing*. Wien: R Foundation for Statistical Computing.
- Schweitzer, Antje & Natalie Lewandowski (2013): Convergence of Articulation Rate in Spontaneous Speech. In: *Proceedings of Interspeech*. 525–529.
- Winkelmann, Raphael (2015): *PraatToFormants2AsspDataObj.R*. URL: <https://gist.github.com/raphywink/2512752a1efa56951f04> (besucht am 07.03.2017).
- Winkelmann, Raphael; Jonathan Harrington & Klaus Jänsch (2017): EMU-SDMS. Advanced speech database management and analysis in R. *Computer Speech & Language*. DOI: 10.1016/j.cs1.2017.01.002.

A Annotationsrichtlinien in GECO-FP

Tabelle A.1: Transliteration/Signalannotation von Füllpartikeln.

Annotationsebene	wordscor.
Annotationsart	Spannenannotation auf akustischem Signal.
Beschreibung	Aussprachebasierte Transliteration von Füllpartikeln.
Annotationswert	Wertbeschreibung
Zeichenkette	Kann frei vergeben werden, beispielsweise <i>äh</i> , <i>öh</i> , <i>oahm</i> . Bei glottalen Ereignissen Beschreibung in spitzen Klammern, beispielsweise <i><knacken></i> , <i><knarren></i> .

Tabelle A.2: Annotation von Füllpartikeln und Mikrokontext.

Annotationsebene Annotationsart Beschreibung	hes, bezieht sich auf wordscor und das akustische Signal. Spannenannotation. Füllpartikeln und ihr Kontext. Die Annotationswerte sind immer zweistellig.	
Annotationswert	Wertbeschreibung	
f	Gehört zur Variablen, ist eigenständig abgrenzbar (keine Silbenprolongationen, keine stillen Pausen, keine Atmungspausen).	
a	Antezedens. Linkes Intervall von f. Stille Pausen oder das letzte phonetische Segment des vorangehenden Signals.	
b	Anteantezedens. Linkes Intervall von a. Wird nur dann annotiert, wenn in a eine stille oder Atmungspause steht.	
p	Postzedens. Rechtes Intervall von f. Stille Pausen oder das erste phonetische Segment des nachfolgenden Signals.	
q	Postpostzedens. Rechtes Intervall von q. Wird nur dann annotiert, wenn in p eine stille oder Atmungspause steht.	
f	v c g x	Ein vokalisiertes Phänomen ohne jene in fc und fg. Ein nicht näher spezifizierter Clicklaut. Eine glottalisierte Phase (kein Vokal, kein Click vorhanden). Nicht in fv, fc oder fg kategorisierbar.
a	s c p h e t	Im Antezedens ist ein Lautsegment enthalten. Im Antezedens ist ein Click enthalten. Im Antezedens ist eine stille Pause enthalten. Im Antezedens ist eine Atempause enthalten. Im Antezedens ist Extralinguistisches enthalten. Im Antezedens ist ein Turn des Dialogpartners enthalten.
b	s c p h e t	Im Anteantezedens ist ein Lautsegment enthalten. Im Anteantezedens ist ein Click enthalten. Im Anteantezedens ist eine stille Pause enthalten. Im Anteantezedens ist eine Atempause enthalten. Im Anteantezedens ist Extralinguistisches enthalten. Im Anteantezedens ist ein Turn des Dialogpartners enthalten.
p	s c p h e t	Im Postzedens ist ein Lautsegment enthalten. Im Postzedens ist ein Click enthalten. Im Postzedens ist eine stille Pause enthalten. Im Postzedens ist eine Atempause enthalten. Im Postzedens ist Extralinguistisches enthalten. Im Postzedens ist ein Turn des Dialogpartners enthalten.
q	s c p h e t	Im Postpostzedens ist ein Lautsegment enthalten. Im Postpostzedens ist ein Click enthalten. Im Postpostzedens ist eine stille Pause enthalten. Im Postpostzedens ist eine Atempause enthalten. Im Postpostzedens ist Extralinguistisches enthalten. Im Postpostzedens ist ein Turn des Dialogpartners enthalten.

Tabelle A.3: Segmentale Annotation von Füllpartikeln.

Annotationsebene	segm; bezieht sich auf hes und das akustische Signal.
Annotationsart	Spannenannotation.
Beschreibung	Annotation der Segmente, aus denen die Füllpartikeln bestehen sowie der Segmente, die in den Ante- und Postzedentia enthalten sind.
Annotationswert	Wertbeschreibung
SAMPA ^a	Jedes identifizierte Segment zu fp (<i>bs, as, fs, ps, qs</i>) wird mit seinem tatsächlichen symbolphonetischen Wert (weite Transkription) mithilfe von SAMPA annotiert. Die Segmentierung orientiert sich an der Ausdehnung der als unterscheidbar identifizierbaren Laute. Grenzen werden in Praat dort annotiert, wo der Laut gerade schon oder gerade nicht mehr als Laut einer bestimmten Klasse erkennbar ist. Grenzen werden am Nulldurchgang aufsteigender Flanken im Oszillogramm gesetzt. Die Segmentierung stimmt mit der Segmentierung der Mikrokontexte überein.
G	Einschränkung A: Für <i>fv</i> : Ein bis drei glottale Plosive werden mit <i>?</i> , mehr mit <i>G</i> annotiert. Die Abgrenzung zum Vokal in <i>fv</i> wird dort getroffen, wo die regelmäßige modale Phonation nach links hin endet. Unter <i>fg</i> wird <i>?</i> annotiert, wenn einzelne Plosive einen Abstand von mehr als 50 ms (entspricht 20 Hz) aufweisen, da diese als Einzellaute wahrgenommen werden. Mit <i>G</i> werden Phasen glottaler Plosive annotiert, die enger zusammenstehen und mehr als zwei sind. Falls Reste eines zugrundeliegenden Sonoranten erkennbar sind, kann dieser annotiert werden (bspw. [n]).
Q	Einschränkung B: Hohe gepresste Stimmlage vor dem Vokal in <i>fv</i> wird auf <i>segm</i> mit Q annotiert.
E	Einschränkung C: Vokale in <i>fv</i> werden immer als <i>E</i> annotiert. Ihre Segmentierung endet nach der letzten erkennbaren Vokalamplitude.
—	Stille Pausen (<i>_p</i>), Turn-Taking-Pausen (<i>_t</i> , der Gesprächspartner spricht) und Clicks (<i>_c</i>) ohne hörbare Atmung werden mit einem Unterstrich annotiert.
sil	Stille.
in	Atempausen (<i>bh, ah, ph, qh</i>) und Clicks, in denen eingeatmet wird.
ex	Atempausen (<i>bh, ah, ph, qh</i>) und Clicks, in denen ausgeatmet wird.

^a Zur Annotation mit SAMPA s. Tabelle C.1.

Tabelle A.4: Phonationsartannotation der Segmente.

Annotationsebene	coart; bezieht sich auf segm .
Annotationsart	Spannenannotation.
Beschreibung	Annotation der Phonation.
Annotationswert	Wertbeschreibung
glott	Bezogen auf segm für Segmente innerhalb des Intervalls <i>fv</i> von Ebene fp , wenn diese vollständig (ohne Bezugnahme auf das vorangehende Segment) glottalisiert sind.
mod	Bezogen auf segm für Segmente innerhalb des Intervalls <i>fv</i> von Ebene fp , wenn diese vollständig (ohne Bezugnahme auf das vorangehende Segment) mit modaler Phonation artikuliert sind.
lary	Bezogen auf segm für Segmente innerhalb des Intervalls <i>fv</i> von Ebene fp , wenn diese vollständig (ohne Bezugnahme auf das vorangehende Segment) laryngalisiert sind.

Tabelle A.5: Dialogzügeannotation.

Annotationsebene	func; bezieht sich auf wordscor.
Annotationsart	Spannenannotation.
Beschreibung	Annotation der Dialogzüge.
Annotationswert	Wertbeschreibung
	<u>Initiierende Züge</u>
	Prinzipiell werden syntaktisch oder pragmatisch eigenständige Fragen separat annotiert.
e	Erklärung/Erzählung/Statement. Information, die nicht (direkt) vom Gesprächspartner eliziert wurde.
e-t	Trunkierte Erklärung oder Erzählung. Sprecher A wird von Sprecher B unterbrochen. Bei Carletta et al. (1996) nicht vorhanden.
fv	Vergewisserungsfrage. Sprecher A möchte Information bestätigt haben. Bei Carletta et al. (1996) ‚CHECK‘: „requests the partner to confirm information that checker has some reason to believe, but is not entirely sure about“ (Carletta et al. 1996: 5)
fe	Entscheidungsfragen. Frage, die nur <i>ja</i> oder <i>nein</i> als Antwort zulässt, und syntaktisch Verberststellung oder syntaktisch Verbzweitstellung mit steigender Intonation aufweist. Auch Ein-Wort-Fragen können darunter fallen, wie <i>echt?</i> .
fe-t	Abgebrochene Entscheidungsfrage. Bei Carletta et al. (1996) nicht vorhanden.
fw	W-Fragen. Frage, die entweder mit einem W-Wort beginnt oder bei Ergänzung mit einem W-Wort beginnen könnte. Es wird ein Set aus möglichen Alternativen erfragt, dies kann auch mit einer Präpositionalphrase evoziert werden.
fw-t	Abgebrochene W-Fragen/Set-Fragen. Bei Carletta et al. (1996) nicht vorhanden.
r	Redebereitschaft.
i	Instruktion
	<u>Responsive Züge</u>
b	BACKCHANNELING. Rückversicherungssignal, zeigt, dass der Partner gehört oder verstanden wurde. Bei Carletta et al. (1996) ‚ACKNOWLEDGE‘.
ae-p	Positive Antwort ‚REPLY-Y‘.
ae-n	Negative Antwort ‚REPLY-Y‘.
ae-u	Unklare Antwort , die Antwort ist weder klar positiv noch negativ oder die Antwort wird nicht gewusst.
ae-t	Abgebrochene Antwort auf eine Entscheidungsfrage , wegen Turn-Unterbrechung des Gesprächspartners. Bei Carletta et al. (1996) nicht vorhanden.
aw	Komplexe Antwort. Antwort auf W-Frage, die nicht einfach <i>ja</i> oder <i>nein</i> bedeutet ‚REPLY-W‘.
aw-u	Unklare Antwort auf eine W-Frage. Sprecher ist sich über seine Absichten nicht im Klaren. Bei Carletta et al. (1996) nicht vorhanden.
av-p	Positive Antwort auf eine Vergewisserungsfrage. Bei Carletta et al. (1996) nicht vorhanden.
av-n	Negative Antwort auf eine Vergewisserungsfrage. Bei Carletta et al. (1996) nicht vorhanden.
av-u	Unklare, nicht gewusste Antwort auf eine Vergewisserungsfrage. Bei Carletta et al. (1996) nicht vorhanden.

B Zweifelsfälle

Bei adjazent vorkommenden Füllpartikeln kann sich unter Umständen das Post- oder Postpostzedens der vorangehenden Füllpartikel mit dem Ante- oder Anteantezedens der nachfolgenden Füllpartikel überlappen. Dies wird so gelöst, dass erst der Wert für die vorangehende Einheit vergeben wird, wonach dann der Wert für die folgende Einheit mit einem Unterstrich an den vorangehenden Wert angefügt wird (vgl. Beispiel 1).

(1)	wordscor		ähm	<P >	<C >	<P >	äh
	hes		fv	pp	qc_bc	ah	fv
							(GECO:multi_H-J_left 519 ms–521 ms)

Extrapositionen wie in Beispiel (2) werden trotz eigenständiger, steigender Intonationskontur nicht separat annotiert, sondern zählen zur vorangehenden Einheit.

(2)	wordscor		haben die dich dann berichtigt auch ↗ die Kinder ↗ <P>				
	func		fe				
	wordscor		so darfst du das nicht sagen oder so ↗				
	func		fe				
							(GECO:multi_H-J_left, 335 ms–341 ms)

Wenn ein Antezedens oder Postzedens sowohl in die Kategorie *Pause* (kleinere Ausdehnung) als auch in die Kategorie *Atmung* (größere Ausdehnung) fällt, wird die Atmungspause annotiert (größere Annotationsausdehnung).

Wenn ein Antezedens oder Postzedens sowohl in die Kategorie *Turn* (größere Ausdehnung) als auch in die Kategorie *Atmung* (kleinere Ausdehnung) fällt, wird die Atmungspause annotiert (kleinere Annotationsausdehnung).

C Vergleich IPA/SAMPA

Tabelle C.1: Vergleich zwischen IPA und SAMPA.

Konsonanten				Vokale			
IPA	SAMPA	IPA	SAMPA	IPA	SAMPA	IPA	SAMPA
p	p	x	x	i	i	o	o
b	b	v	v	y	y	ə	@
t	t	s	s	u	u	ɛ	E
d	d	z	z	ɪ	I	œ	9
k	k	ʃ	S	ɣ	Y	ɔ	O
g	g	x	x	ʊ	U	ɐ	6
ʔ	?	ɞ	r	e	e	a	a
m	m	h	h	ø	2		
n	n	j	j				
ŋ	N	l	l			fi	h\