

Proseminar

Verarbeitung gesprochener Sprache

Timo Baumann

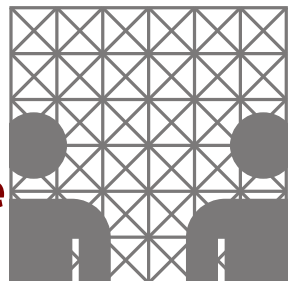
baumann@informatik.uni-hamburg.de



nats-www.informatik.uni-hamburg.de/VGS20/WebHome

Universität Hamburg, Department of

Informatics



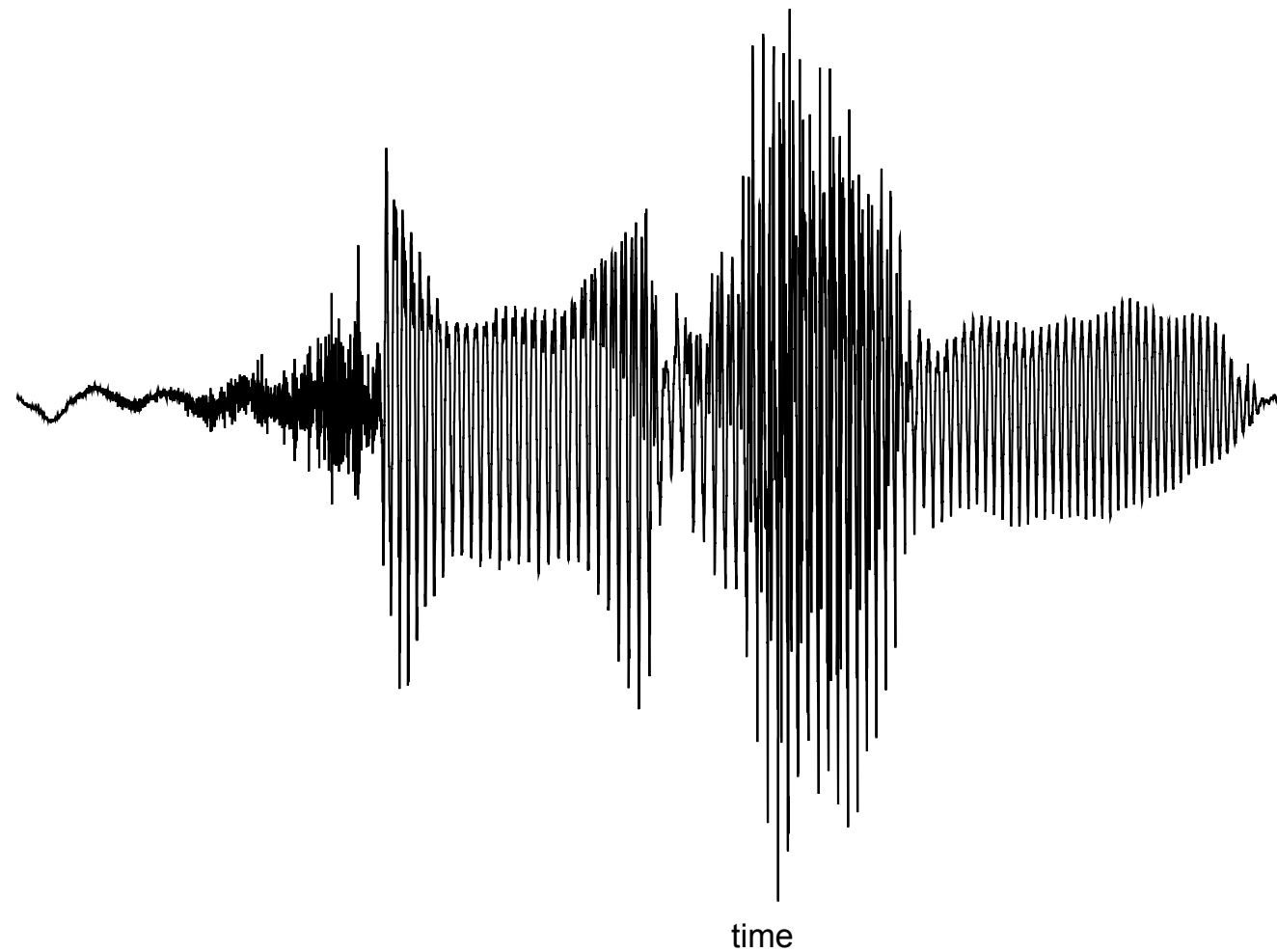
Heute

1. Themen – Klarheiten/Unklarheiten klären
2. Vortrag zum Aufbau der Sprache
3. Debriefing und Hausaufgaben

ein bisschen Phonetik

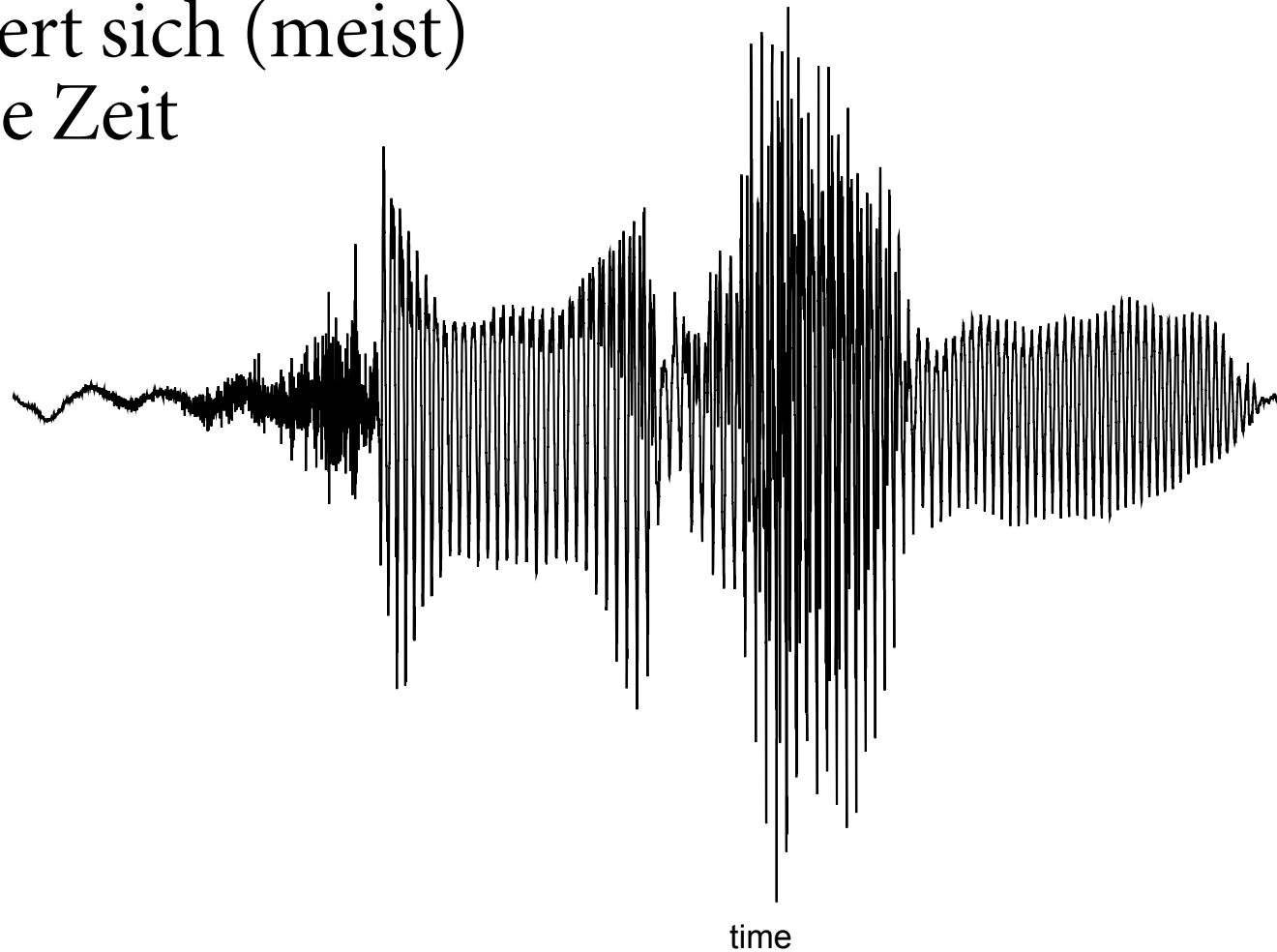
- Wissenschaft der gesprochenen Sprache
- unterschiedliche Blickwinkel:
 - akustische Phonetik (was ist Sprachschall, was zeichnet ihn aus)
 - artikulatorische Phonetik (wie entsteht Sprachschall)
 - auditorische Phonetik (wie wird Sprachschall wahrgenommen)
- interdisziplinär zwischen Linguistik, Akustik, Medizin, Neurologie (und Informatik?)
- Laute sind unterhalb der Wahrnehmungsschwelle, keine sinntragenden Einheiten → Abstraktion
- gleichzeitig sind allgegenwärtig → schwer zu handhaben

Akustik der gesprochenen Sprache (superkurz)



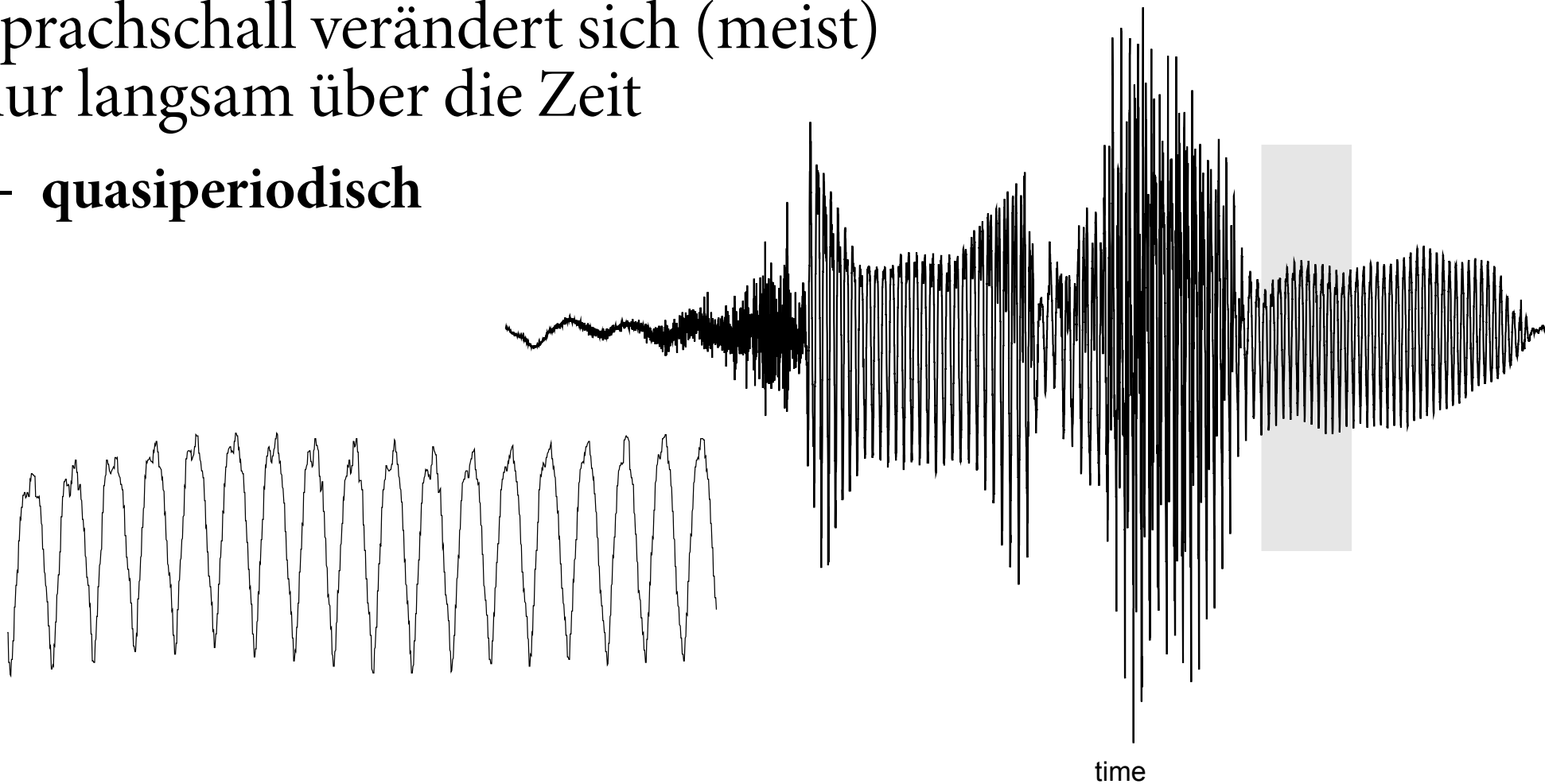
Akustik der gesprochenen Sprache (superkurz)

- enthält keine klaren Grenzen
- Laute nicht eindeutig unterscheidbar
- Sprachschall verändert sich (meist) nur langsam über die Zeit
 - **quasiperiodisch**



Akustik der gesprochenen Sprache (superkurz)

- enthält keine klaren Grenzen
- Laute nicht eindeutig unterscheidbar
- Sprachschall verändert sich (meist) nur langsam über die Zeit
 - **quasiperiodisch**

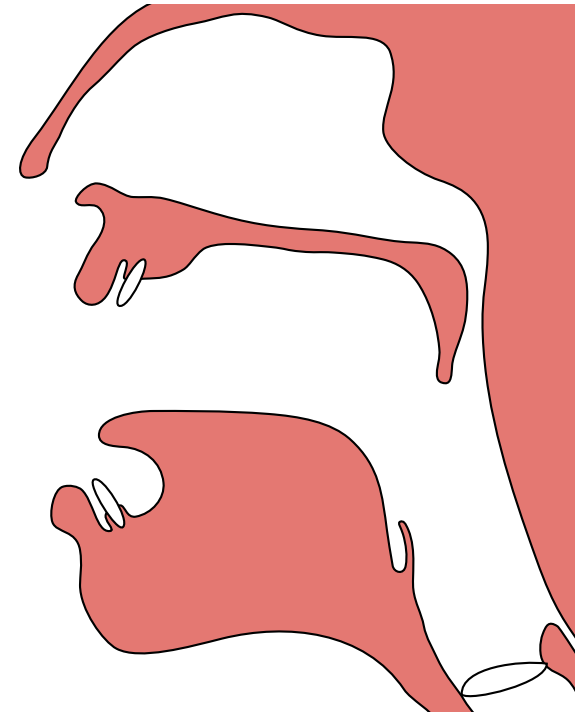


Artikulatorische Phonetik

- zwei Grundkomponenten
 - Generierung eines (stimmhaften) Primärschalls (**Phonation**) am Kehlkopf durch die **Stimmlippen**
 - Formung zum tatsächlichen Sprachschall (**Artikulation**) im **Ansatzrohr** (alles zwischen Kehlkopf und Lippen)

Quelle-Filter-Modell der Sprachproduktion

- Stimmlippen erzeugen einen (harmonischen) Primärschall
 - hoher Gehalt an Obertönen
- Das Ansatzrohr bildet Resonanzen (wie eine Orgelpfeife)
 - hauptsächlich negative Resonanz (Dämpfung)
- Das Modell stimmt primär für stimmhafte Laute
 - Annahme von weißem Rauschen für Frikative (obwohl dieses Rauschen tatsächlich größtenteils nicht an den Stimmlippen entsteht)
 - schlechte Modellierung von Plosiven



Phone (bzw. Laute)

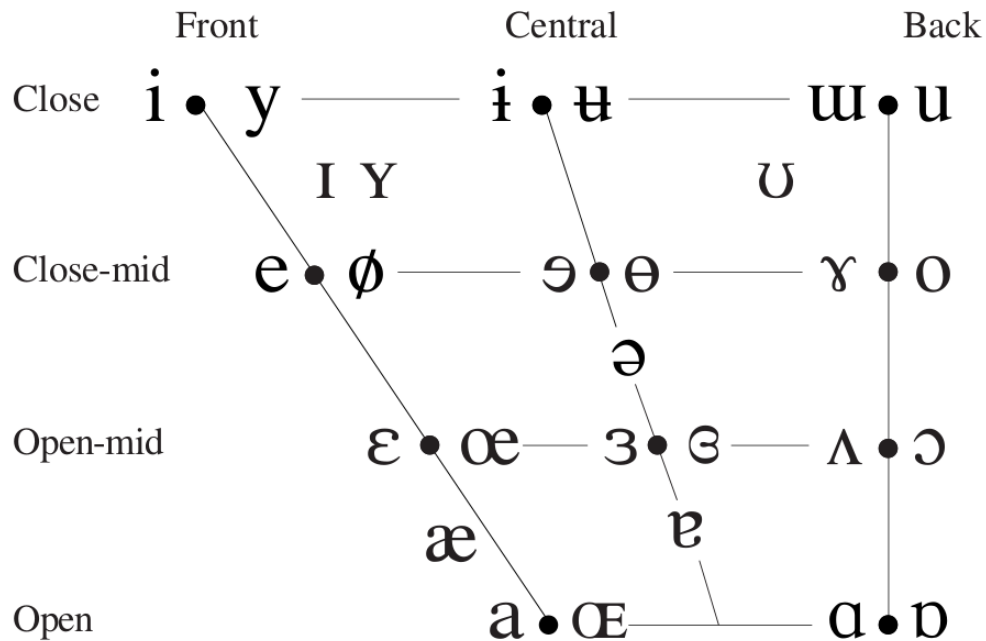
Sprachlaute werden durch die Möglichkeiten der Artikulation bestimmt

- zwei Grundarten:
 - Vokale (freier Luftfluss im Ansatzrohr)
 - Konsonanten (Obstruktion stört den Luftfluss)
- Distinktive Merkmale auf Basis der Artikulationsmöglichkeiten:
 - Stimmhaftigkeit, Dauer, Mundöffnung, Lippenöffnung, Zungenstellung, Nasalität, ...
- Klassifikation durch die International Phonetic Association

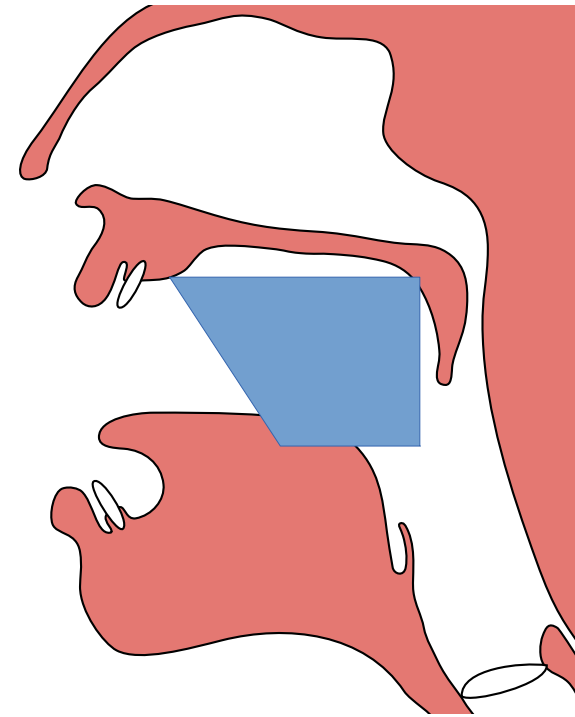


Vokalische Laute: Vokale

- Der Bewegungsraum der Zunge in der Mundhöhle wird durch ein Trapez (das Vokaltrapez) idealisiert
- Resonanz der Mundhöhle bestimmt Laut

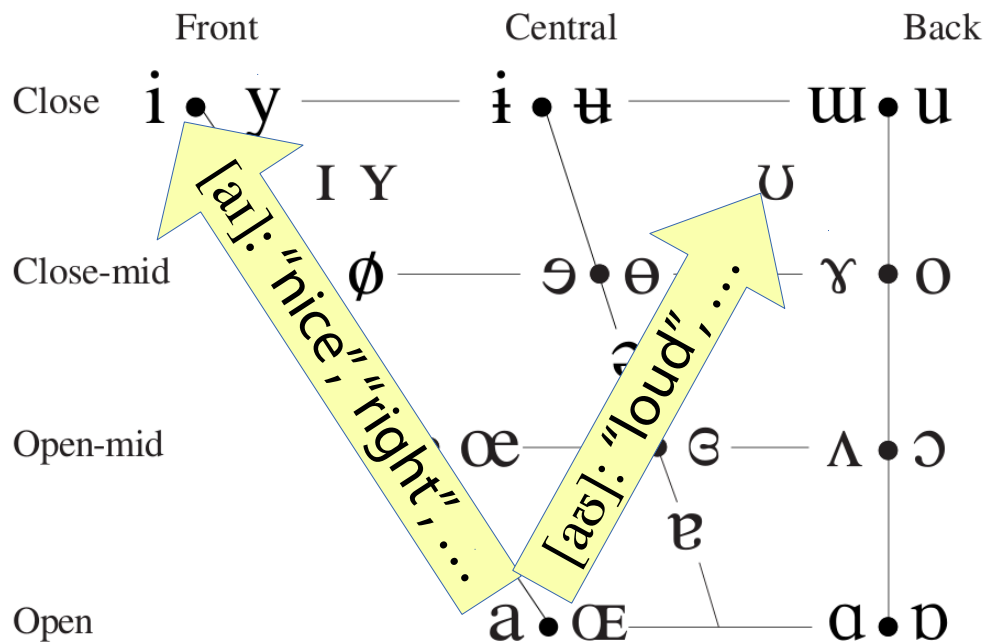


Where symbols appear in pairs, the one to the right represents a rounded vowel.

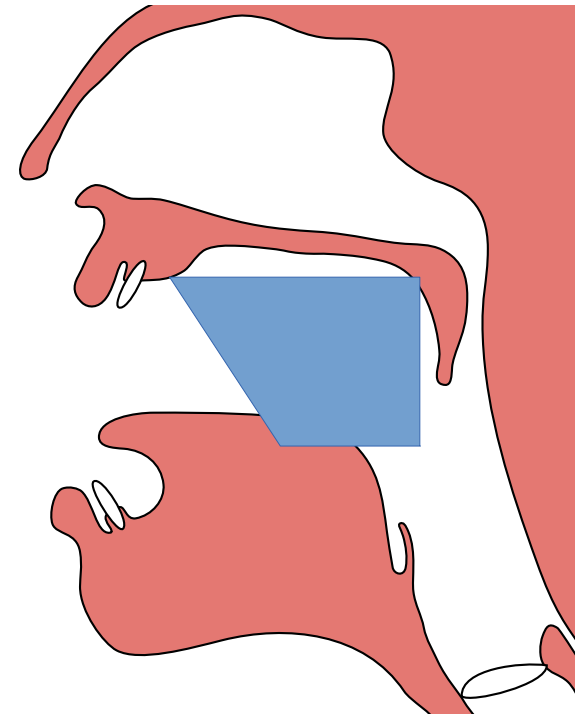


Vokalische Laute: Diphthonge

- die Zunge kann sich bewegen während der Phonation, sodass der Höreindruck sich verändert



Where symbols appear in pairs, the one to the right represents a rounded vowel.



Konsonanten

- Artikulationsart (Plosive, Nasale, Fricative, ...)
- Artikulationsort (Lippen, Zähne, ... Glottis)

	Bilabial	Labiodental	Dental	Alveolar	Post alveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k ɡ	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Trill	ʙ			r					ʀ		
Tap or Flap		ⱱ		ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateral fricative				ɬ ɮ							
Approximant		ʋ		ɹ		ɻ	j	ɰ			
Lateral approximant				l		ɭ	ʎ	ʟ			

Where symbols appear in pairs, the one to the right represents a voiced consonant. Shaded areas denote articulations judged impossible.

Übung (in Zweiergruppen):

1. transkribiere Deinen Namen im phonetischen Alphabet
2. transkribiere einige andere Wörter **ohne sie laut vorzusprechen**
3. tausche Deine Notizen mit dem Nachbarn aus, spricht Euch die Notizen vor, prüft auf Fehler/Ungenauigkeiten

Die Einheiten der Sprache: Phone vs. Phoneme

die Einheiten der Sprache: Phone vs. Phoneme

- Sprachlaute
(→ Phonetik)
 - unterscheid**bare** Einheiten
 - sprachunabhängig
 - **Signifiant**
- linguistische Symbole
(→ Phonologie)
 - unterscheid**ende** Einheiten
 - jede Sprache hat ihr Phonemsystem
 - **Signifié**
- Minimal Paare: “**b**at” – “**r**at” – “**c**at”
 - /**b**/, /**r**/, /**k**/ sind Phoneme im Englischen
 - sind Laute in irgendeiner Sprache unterscheidend, so sind sie jeweils Phone (also gibt es die Phone [**b**], [**r**], [**k**])
 - one's articulatory/perceptory capacities are shaped by the mother tongue(s)
 - different sounds may sound identical or be hard to pronounce

die Einheiten der Sprache: Phone vs. Phoneme

- Sprachlaute
(→ Phonetik)
 - unterscheidbare Einheiten
 - sprachunabhängig
 - **Signifiant**
- linguistische Symbole
(→ Phonologie)
 - unterscheidende Einheiten
 - jede Sprache hat ihr Phonemsystem
 - **Signifié**
- Notationskonvention:
"examples" in quotes
/phonemes/ in slashes
[phones] in brackets
- Minimal Paare: "bat" – "rat" – "cat"
 - /b/, /r/, /k/ sind Phoneme im Englischen
 - sind Laute in irgendeiner Sprache unterscheidend, so sind sie jeweils Phone (also gibt es die Phone [b], [r], [k])
 - one's articulatory/perceptory capacities are shaped by the mother tongue(s)
 - different sounds may sound identical or be hard to pronounce

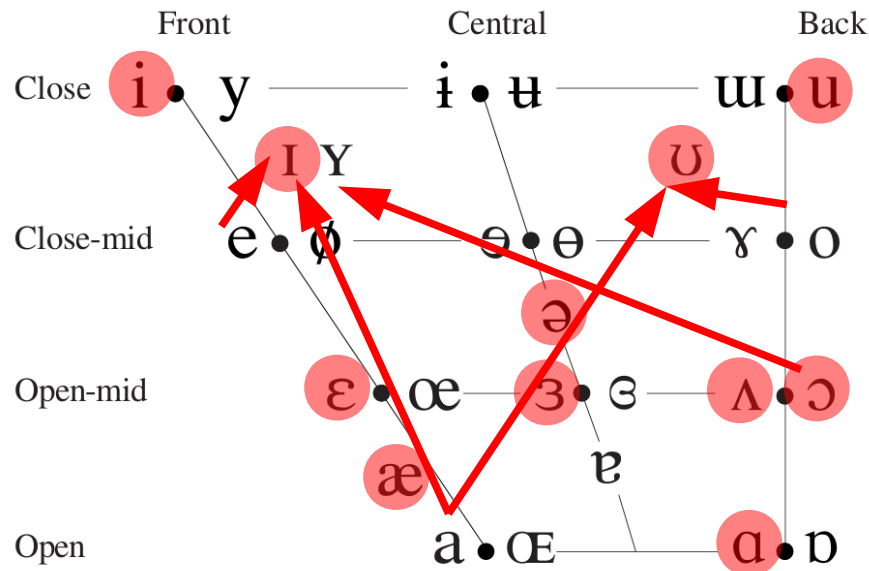
The Phonemic System of a Language

- only small subset of symbols in the IPA
- contextual rules determine phonetic realization
 - e.g. German [ç/x] (“ich”/“ach”) is a single phoneme /ç/
- context limitations (*Phonotactics*), often in combination with **syllabic structure**
 - syllable = onset + nucleus + coda
 - e.g. German nucleus must be a vowel and coda limited to 5 consonants; many rules for consonant sequences
 - e.g. English: no /ŋ/ in onset, no /h/ in coda, ...

N-American English Phoneme Set

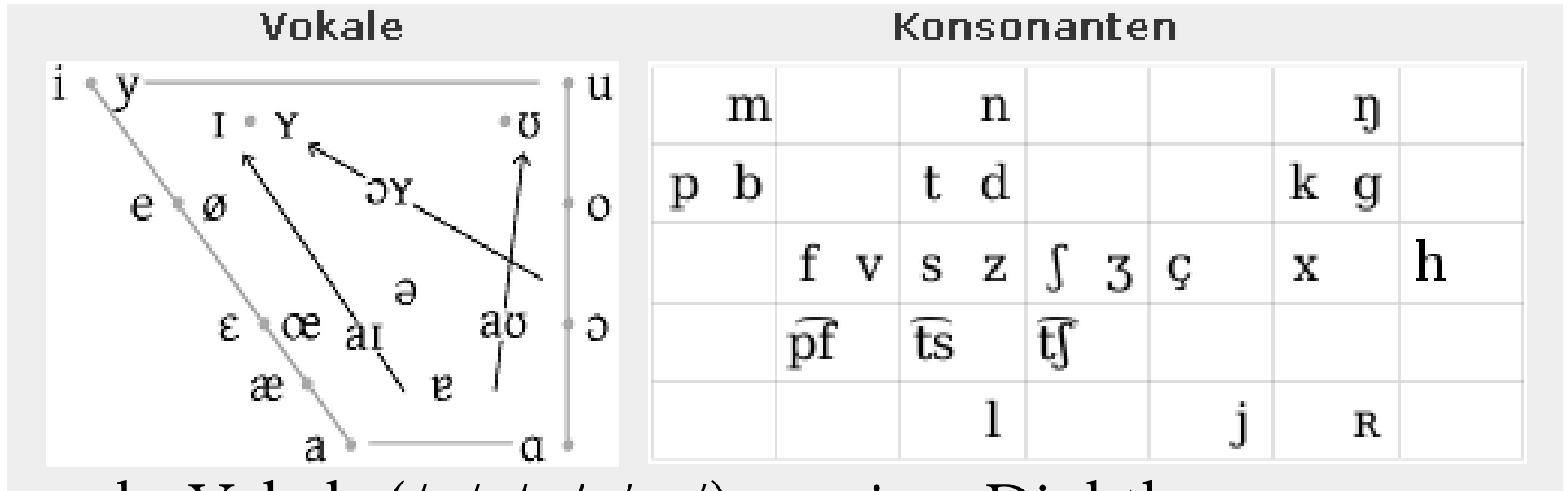
	Bilabial	Labiodental	Dental	Alveolar	Post alveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k ɡ	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Trill	ʙ			ʀ					ʀ		
Tap or Flap		ⱱ		ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateral fricative				ɬ ɮ							
Approximant		ʋ		ɹ		ɻ	j	ɰ			
Lateral approximant				l		ɭ	ʎ	ʟ			

Where symbols appear in pairs, the one to the right represents a voiced consonant. Shaded areas denote articulations judged impossible.



Where symbols appear in pairs, the one to the right represents a rounded vowel.

German Phoneme Set



- mehr Vokale (/y/, /ʏ/, /œ/), weniger Diphthonge
- ähnliche Konsonantenwahl
(allerdings im Detail unterschiedliche Realisierung, bspw. in immer aspirierte Plosive)

Phone vs. Phoneme in der Praxis

Mein sprachverarbeitendes System nutzt
A: Phone | B: Phoneme .

Wie wirkt sich das aus? Was sind Vor- und Nachteile?

Phone vs. Phoneme in der Praxis

Mein sprachverarbeitendes System nutzt
A: Phone | B: Phoneme .

Wie wirkt sich das aus? Was sind Vor- und Nachteile?

Fremdwörter?

Trainierbarkeit mit Beispielmateriale?

Anwendungskontext vs. Generalität?

Zusammenfassung

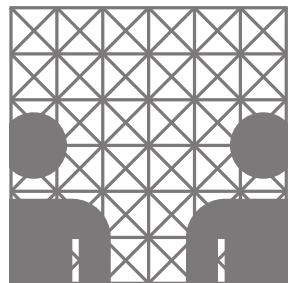
gesprochene Sprache: kontinuierliches Signal
eines symbolischen Systems

Aufgaben

- bis nächste Woche: Sie *haben* ein Thema
- bis 13.05.: Sie haben die angegebene Literatur gesichtet und nach weiterer nützlicher Literatur gesucht und die Ergebnisse an mich geschickt. Ihr Thema ist *eingegrenzt*.

Vielen Dank.

baumann@informatik.uni-hamburg.de



Further Reading

- Speech Signal Representation:
 - P. Taylor (2009): *Text-to-Speech Synthesis*. Cambridge Univ Press. ISBN: 978-0521899277. InfBib: A TAY 43070
 - D. Jurafsky & J. Martin (2009): *Speech and Language Processing*. Pearson International. InfBib: A JUR 4204x
- Phonetics:
 - M. Pétursson & J. Neppert (1996): *Elementarbuch der Phonetik*. Buske.
 - J. Neppert (1999): *Elemente einer akustischen Phonetik*. Buske.
- Phonology/Phonotactics/Phonological Systems:
 - E. Ternes (1999): *Einführung in die Phonologie*. Wiss. Buchgesellschaft. ISBN: 978-3534138708.

Notizen

Lernziele

- Speech is the continuous signal of a symbolic language
- Speech is ambiguous on multiple layers of analysis
 - elements of speech
 - quasistationary signal, Phone/Laute, Phoneme, Allophone
-