

M.Sc. LST

Speech Science

Prosody

Dec 5 and 12, 2024

Bernd Möbius & Valentin Kany

Language Science and Technology
Saarland University



Prosody

- Prosody comprises properties of spoken language beyond single sounds
 - features extending over longer units/stretchers of speech (pitch, sentence melody, loudness, tempo, rhythm, ...)
- "prosodic", "suprasegmental", "nonsegmental" (often synonymous)
 - realized simultaneous with segments
 - longer than segments
 - manifestation of units higher in phonological hierarchy: syllables, words, phrases, sentences, discourses

Prosody: terminology

- Prosody: a subdiscipline of linguistics and phonetics
- Terminology debate at IPA conference at Kiel (1989): diversity of opinions on topic "Prosody", its symbolic representation, notation, transcription
- 3 terms: prosody, suprasegmentals, intonation
 - often used synonymously
 - unified definitions helpful (even if only in class)
- Literature:
 - Clark et al. (2007), ch. 9
 - Botinis et al. (2001)
 - Lehiste (1970), ch. 1
 - Möbius (1993), ch. 2

Remark on orthography

- Most writing systems:
 - only rudimentary markers of prosodic features: punctuation, typographic emphasis (underline, bold face, italic, spaced out)

Proposed systematicity

- Prosody:
 - linguistic subdiscipline, generic term
 - comprising linguistic and paralinguistic function
- Feature systems with linguistic function
 - intonation (tonal): linguistically relevant functions of fundamental frequency on levels of syllables, words, utterances, discourses
 - duration (temporal): linguistically relevant functions of absolute and relative duration of units
 - intensity (stress): linguistically relevant functions of energy-related features
- ("suprasegmentals"?)

More terminology

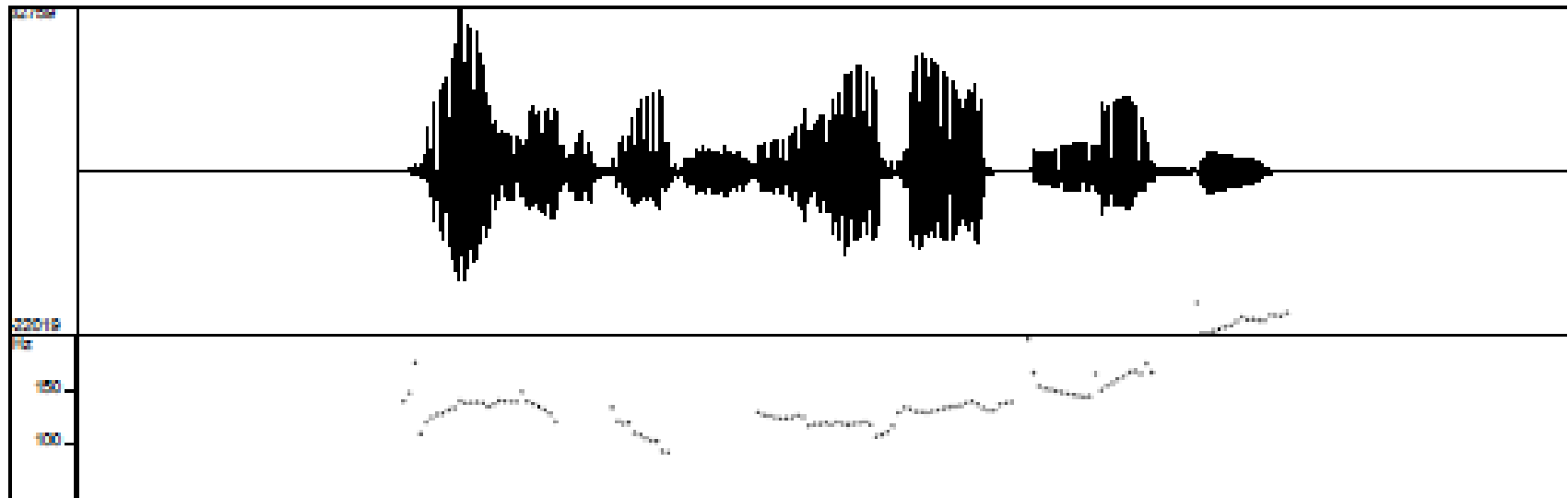
- Prominence: acoustic-phonetic and perceptual emphasis of parts of utterances; relative feature, binary, or gradual?
- Stress (syllabic stress, word stress): acoustic realization of (lexical) stress
- Accent: tonally marked (by F0) stressed syllable
 - special cases: contrast accent, emphasis
- Further prosodic features (somewhat under-researched): speaking rate (tempo), rhythm, pauses, hesitations, voice quality, phonation type, . . .
 - subject to conventions of linguistic community, may convey linguistic functions (e.g. laryngealization as boundary signal)
 - cue for speaking style, situative context, attitude, emotional state of speaker

Acoustic parameters

- Classical acoustic parameters
 - fundamental frequency (F0)
 - segmental duration
 - intensity
- Problems: each of these parameters
 - is also part of specification of speech sounds
 - supports several linguistic functions
 - also supports paralinguistic functions

Fundamental frequency contour

File: k01be041.wav Page: 1 of 1 Printed: Mon Dec 15 10:52:49



Grundfrequenzverlauf (F_0 -Kontur) der Äußerung
"Wer muss noch Schularbeiten machen?"

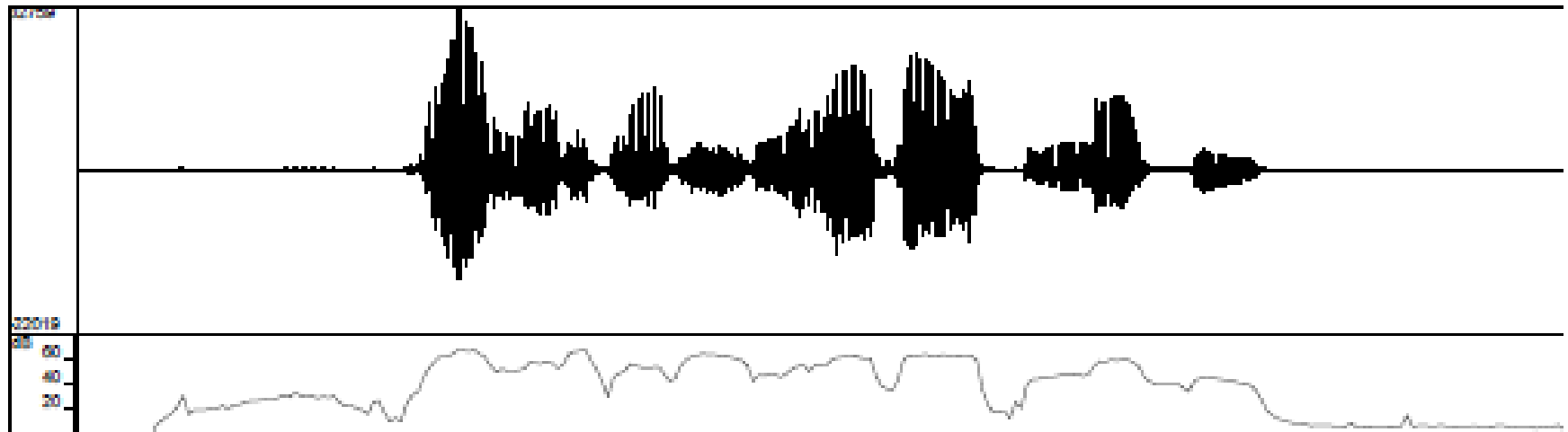


Perception of fundamental frequency

- Auditory-perceptual correlate of F0 variation is pitch variation - but this is not a bi-unique relationship!
 - pitch can be perceived even when fundamental is not present in signal
 - e.g. speech signal transmission: telephone bandwidth 300-3300 Hz
 - fundamental frequency is reconstructed from spectral harmonics, via integration of virtual and spectral pitch into combined sensory entity (Terhardt, 1974) or inferred from signal periodicity
- Fundamental frequency is only one of several acoustic correlates of perceived pitch!

Intensity contour

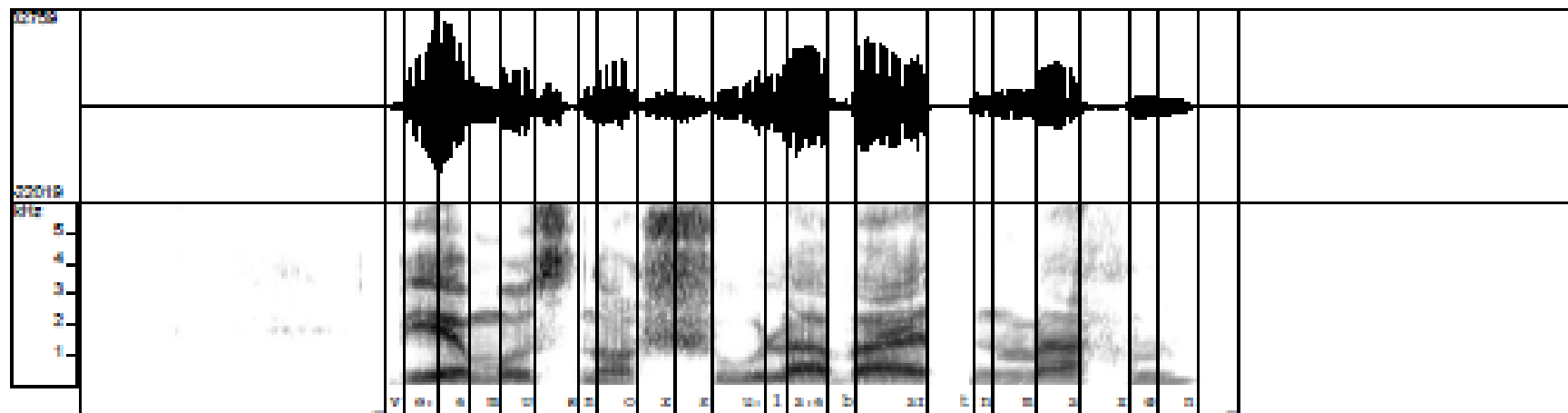
File: k01be041.wav Page: 1 of 1 Printed: Mon Dec 15 11:13:40



Intensitätsverlauf (RMS-Intensität) der Äußerung
"Wer muss noch Schularbeiten machen?"

Segmental durations

File: k015e041.wav Page: 1 of 1 Printed: Mon Dec 15 11:17:13



Dauer sprachlicher Einheiten (hier: Lautdauern) in der
Äußerung *“Wer muss noch Schularbeiten machen?”*

Word prosody

- **Fixed** word accent (predictable): WA always on specific syllable position in word, e.g. on first syllable in **all** words (Finn., Hung., Czech) or penultimate (Pol.) or last (French)
- **Flexible** word accent (not predictable): WA lexically associated with specific syllable, per word (Ger., Engl., Russ.); segmentally identical words may form prosodic minimal pairs:
 - DE umfahren – umfahren, übersetzen – übersetzen, Tenor, August, Konstanz
 - EN import – import, abstract – abstract
 - RU мыка ['mukə] "pain" – мыка [mʊ'ka] "flour",
замок ['zamək] "castle" – замоk [zə'mok] "lock"
 - ES término – termino – terminó
 - PS (Pashto) لارم laram "gone" – لرم larm "have"

Accent groups

- Accent group, sometimes called stress group
 - unit comprising an accented syllable and all subsequent unaccented syllables (if any)
 - independent of word boundaries
 - sensitive to phrase boundaries
 - e.g. | *unit com* | *prising an* | *accented* ...
- Stress-timed vs. syllable-timed languages? Isochrony?

Sentence/phrase prosody

- Phrasing
 - segmentation of utterances in sections characterized by their own intonation
 - intonation phrases do not have to be co-extensive with syntactic phrases
 - boundary signals: final fall, final rise, continuation rise, declination reset(ting), phrase-final lengthening, utterance-final lengthening
 - phrasing as cue for syntactic disambiguation?
 - "Ja zur Not geht's auch am Samstag"
 - "Flying planes can be dangerous."

Sentence/phrase prosody

- Phrasing and accenting affect meaning!

Ja zur Not geht's auch am Samstag

Ja! Zur Not geht's auch am Samstag.

Ja? Zur Not geht's, auch am Samstag.

Ja, zur Not. Geht's auch am Samstag?

Ja, zur Not geht's. Auch am Samstag!

Ja, zur Not geht's. Auch am Samstag?

...

- But it does not always disambiguate meaning!

Flying planes can be dangerous.

(To fly planes vs planes that are flying)

Sentence/phrase prosody

- Sentence mode
 - semantic/syntactic/functional definitions: relation between formal-grammatical sentence types and pragmatic function types, e.g.: declarative, interrogative, imperative, exclamative sentences
 - prosodic/intonational definitions: interrogative, progredient; terminal vs. non-terminal

Discourse and information structure

- Discourse and information structure
 - Carter called Nixon a Republican, and then he offended him.
 - Carter_i called Nixon_j a Republican, and then he_i offended him_j.
 - Carter_i called Nixon_j a Republican, and then he_j offended him_i.
- Given/new information
- Topic/Focus structure

Intonation

- Intonation: linguistically relevant functions of fundamental frequency on syllable, word, utterance, discourse levels
 - primarily lexical: on mora, syllable
(tone, tone languages, pitch accent languages)
 - primarily intonational: on words
(tone, pitch accent, accenting/accenuation)
 - primarily syntactic: on phrases and sentences
(phrase boundaries, phrasing)
 - primarily semantic: on phrases and sentences
(information structure, sentence mode)

Tone languages

- Some tone languages: e.g. Yoruba, Igbo, Thai, Vietnamese, Mandarin-Chinese (example below)



Word	Tone contour	Meaning
ma1 mā	high-level	mother
ma2 má	high-rising	hemp
ma3 mǎ	low or low-rising	horse
ma4 mà	falling	to scold
ma ma	neutral	(particle)

- Chinese tones (UCLA Sounds of the World's Languages)
- In tone languages, the tone contour is part of the lexical specification of *a//* syllables.

Pitch accent languages

- Some pitch accent languages: Japanese, Serbian, Croatian, Lithuanian, Latvian, Norwegian, Swedish (example below)

	Word	Tone contour	Meaning
	anden1	late rise, 1 st syllable	the duck
	anden2	2 rise-falls, 1 st and 2 nd syllable	the ghost

- Swedish pitch accents (Wikipedia)
- In pitch accent languages, the tone contour is part of the specification of syllables and words only in parts of the lexicon.

Intonational phonology

- Some intonation languages: English, German, French, Spanish, ...
- The domain of distinctive tone contours is the prosodic word, the phonological phrase, the intermediate phrase, and the intonation phrase.
- Autosegmental-metrical theory of intonation (Pierrehumbert 1980)
 - intonation is represented by sequence of high (H) and low (L) tones
 - H and L are members of a primary phonological contrast
 - hierarchy of intonational domains
 - **IP** – Intonation Phrase; boundary tones: H%, L%
 - **ip** – intermediary phrase; phrase tones: H-, L-
 - **pw** – prosodic word; pitch accents: H*, H*L, L*H, ...

Intonational phonology

- Finite-state grammar of well-formed tone sequences

$$\left(\left\{ \begin{array}{c} \%H \\ \%L \end{array} \right\} \left(\left(\left\{ \begin{array}{c} H* \\ L* \\ H* + L \\ H + L* \\ L* + H \\ L + H* \end{array} \right\} \right)^{+} \left\{ \begin{array}{c} H- \\ L- \end{array} \right\} \right)^{+} \left\{ \begin{array}{c} H\% \\ L\% \end{array} \right\} \right)^{+}$$

pw
ip
IP

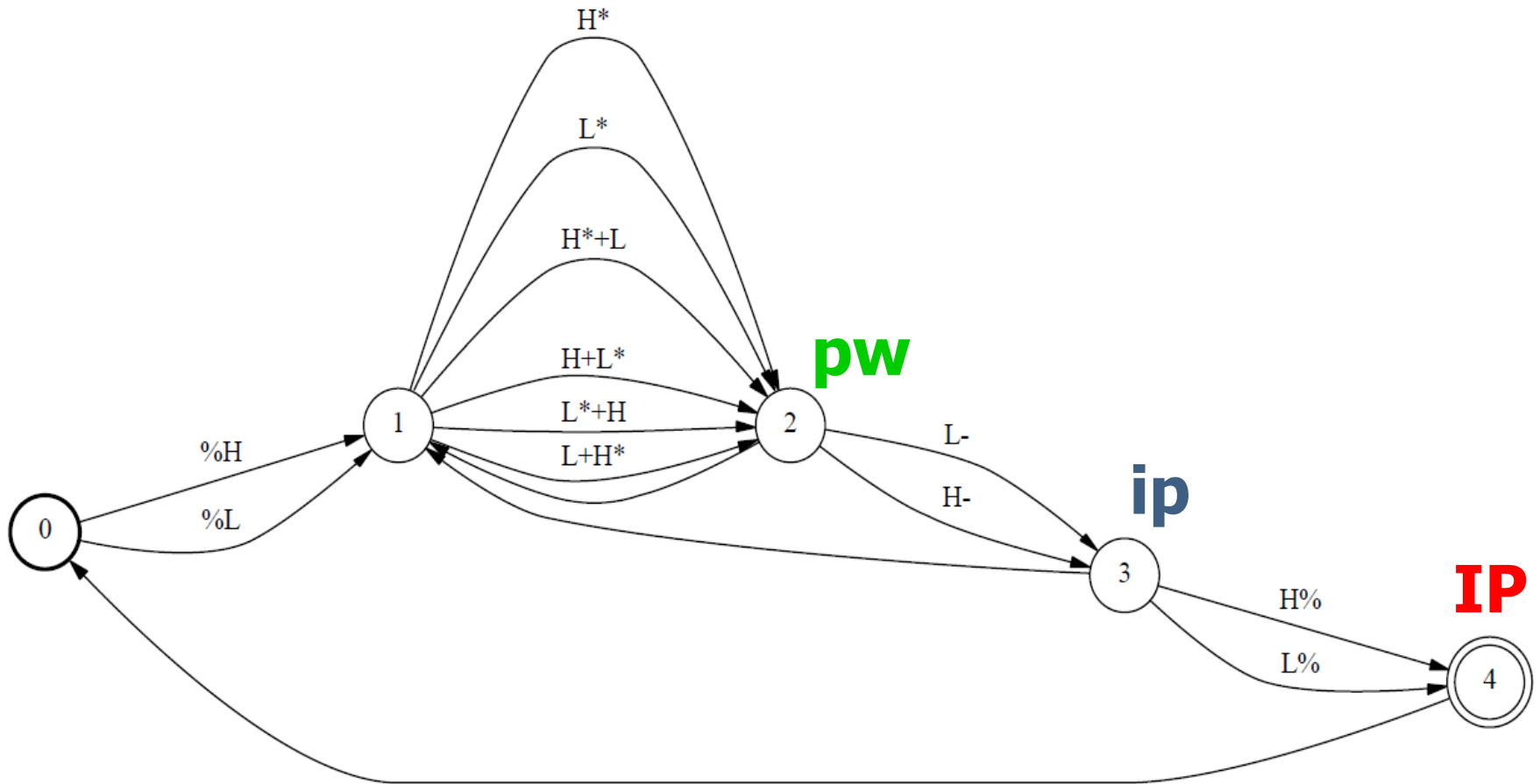
- Example [adapted from Pierrehumbert 1980, p. 276]

That's a remarkably clever suggestion.

%H
|
H*
|
H*L
L-
L%

Intonation phonology

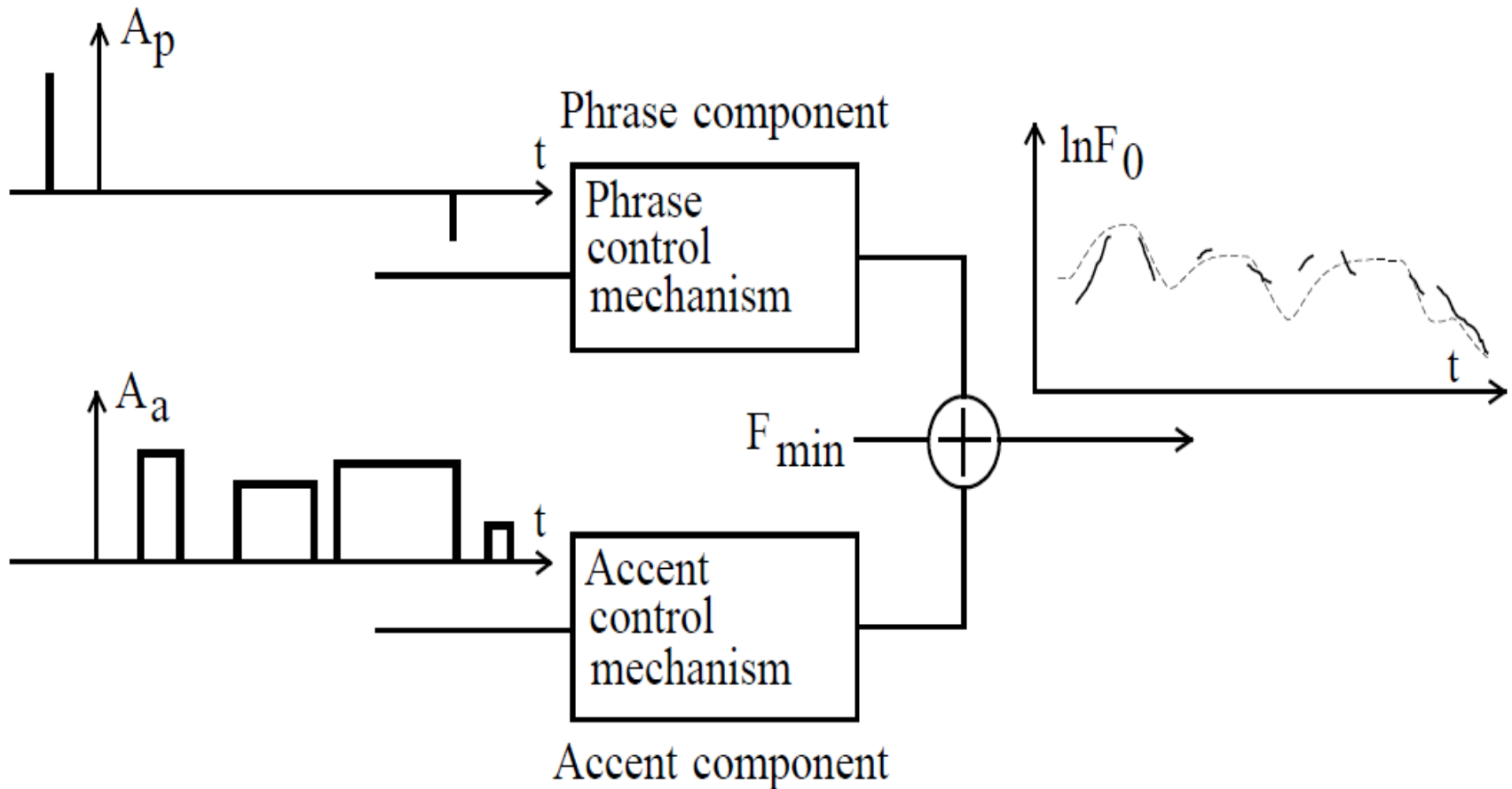
- Finite-state graph



ToBI: Tones and Break Indices

- Formalization of intonation model as transcription system [Silverman et al. 1992]
 - phonemic (=broad phonetic) transcription
 - originally designed for American English
 - limited applicability to other varieties/languages
 - language-specific inventory of phonological units
 - language-specific details of F0 contours
 - adapted to many languages (e.g. GToBI, JToBI, KToBI)
 - implemented in many TTS systems

Fujisaki's model

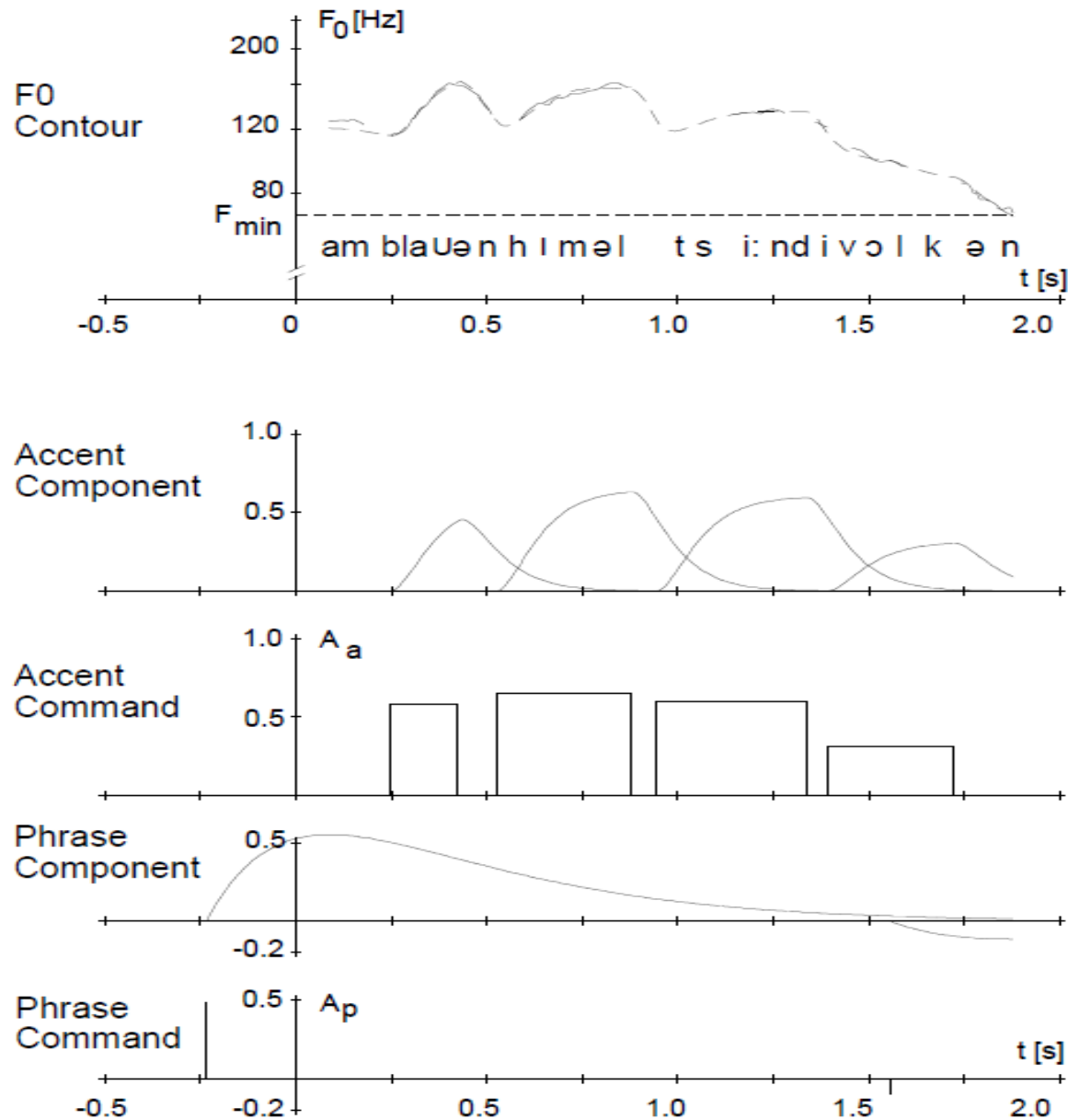


[Fujisaki 1983, 1988; Möbius 1993]

Fujisaki's model

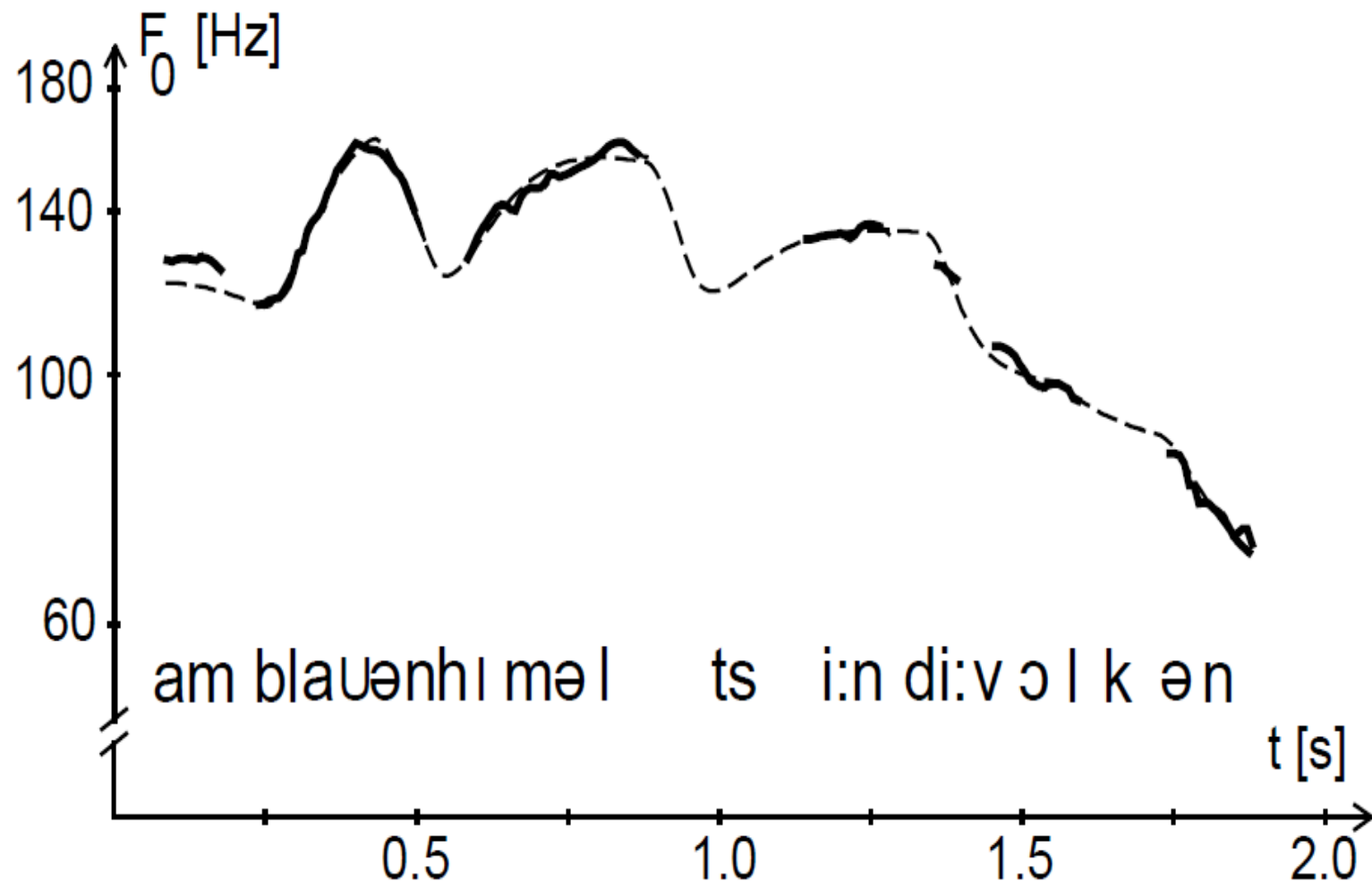
- Properties
 - superpositional
 - physiological basis and interpretation of components and control parameters
 - linguistic interpretation of components
 - applied to many (typologically diverse) languages
- Origins
 - Öhman and Lindqvist (1966), Öhman (1967)
 - Fujisaki et al. (1979), Fujisaki (1983, 1988)

Fujisaki's model



[Möbius 1993]

Fujisaki's model



Approximation of natural F_0 by optimal parameter values within linguistic constraints (accents, phrase structure) [Möbius 1993]

Many functions of F_0

- F_0 serves as a cue for many linguistic, paralinguistic, and extralinguistic factors, including:
 - lexical tones
 - word accent, syllabic stress
 - stress or accent groups
 - prosodic phrasing
 - sentence mode
 - information structure
 - discourse structure
 - pitch range, register
 - phonation type, voice quality
 - microprosody: intrinsic and coarticulatory F_0

Rhythm

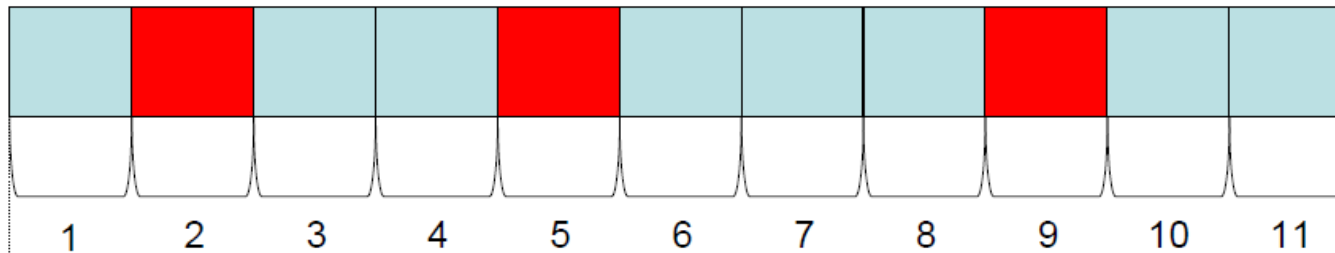
- Rhythm (of spoken language)
 - systematic organization of prominent and less prominent units of speech over time
 - units: syllables, vocalic intervals
 - prominence conveyed by duration, F0, intensity
- duration: same duration, then one long and two short beats
- F0: same F0, then one higher/lower and two lower/higher beats
- intensity: one louder/softer and two softer/louder beats



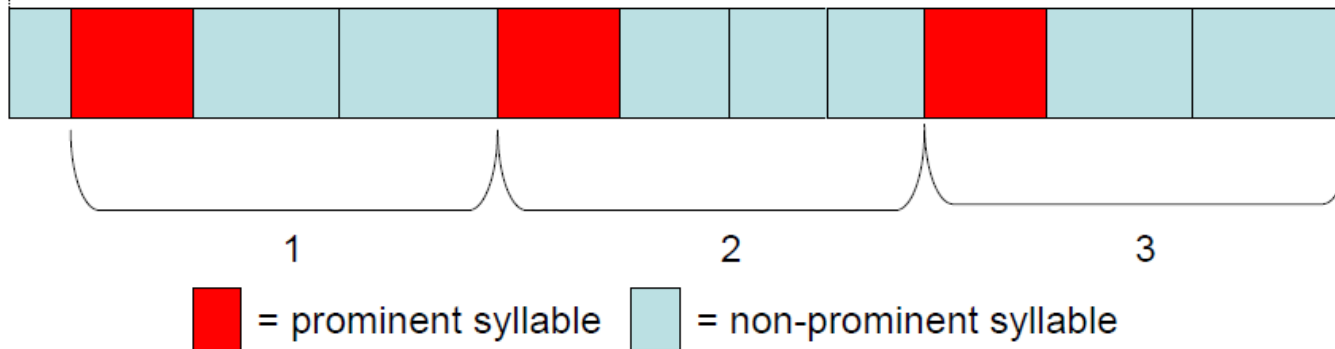
- isochrony hypothesis (since Pike 1945, Abercrombie 1967):
rhythm-based classification of languages
 - based on stress intervals: temporal distance between two stressed syllables approximately constant (morse code rhythm)
 - e.g. German, English, Dutch (West-Germanic languages)
 - based on syllable intervals: temporal distance between two syllable onsets approximately constant (machine gun rhythm)
 - e.g. French, Spanish, Italian (Romance languages)

The nature of syllable & stress timing

syllable timing: (syllable isochrony = here: 11 equally timed syllables)



stress timing: (foot or interstress isochrony = here: 3 equally timed feet)



adapted from Volker Dellwo (UCL, now U Zürich)

Rhythm

- Evidence for isochrony
 - experimental (acoustic, perceptual): rather weak tendency
 - many rhythm measures proposed and tested: no consensus
 - continuum rather than binary classification (cf. Polish)
- Speech rhythm not sufficiently well defined
- Question for **discussion**: why is phonetic research on speech rhythm interesting and important?

References

- IPA (ed.) (1999): Handbook of the International Phonetic Association. Cambridge University Press.
<https://www.internationalphoneticassociation.org/content/handbook-ipa>
- Hadumod Bußmann (²1990): Lexikon der Sprachwissenschaft. Kröner.
- Antonis Botinis, Björn Granström, Bernd Möbius (2001): "Developments and paradigms in intonation research". Speech Communication 33 (4), 263-296.
- John Clark, Colin Yallop, and Janet Fletcher (2007): An Introduction to Phonetics and Phonology. Blackwell.
- Ilse Lehiste (1970): Suprasegmentals. MIT Press.
- Bernd Möbius (1993): Ein quantitatives Modell der deutschen Intonation - Analyse und Synthese von Grundfrequenzverläufen. Niemeyer.
- Ernst Terhardt (1974): "Pitch, consonance, and harmony". Journal of the Acoustical Society of America 55, 1061-1069.

Thanks!

