# Basic notions of prosody

**Amalia Arvaniti**

**University of Kent**

Aix Summer School on Prosody

5-9 September 2016

University of Kent

# What is prosody?

- Umbrella term used to group together a number of phonetic parameters that typically extend over an utterance
  - Pitch (F0)
  - Duration (timing)
  - Amplitude
  - Phonation

**Phonetics**

- These parameters give rise to a set of phonological phenomena
  - Stress
  - Rhythm and timing
  - Phrasing
  - Intonation

**Phonology**

# Prosody or suprasegmentals?

- The term *suprasegmentals* has often been used to refer to prosody (e.g. Lehiste 1970)
- It is used as recently as Ladd (2008)
- It suggests that we have a two-layer organization of speech
- Segmentals: vowels and consonants
- Suprasegmentals: stuff that goes above

# So what?



# Icing on the cake?

- The view of prosody as suprasegmentals implies that the integrity of segmentals is not compromised by suprasegmentals
- This is how prosody if often taught
- This implies that one can study many facets of speech and language structure without reference to prosody
- Doing so is standard practice in research
- It all lumps together all aspects of *suprasegmentals*

# A better analogy

- Speech is more like a baklava
- It consists of many layers
- Prosody is the syrup that binds it together and turns it into a baklava
- Prosody is essential and **unavoidable**

# Components of prosody

- *Stress*
- *Rhythm & timing*
- *Phrasing*
- *Intonation*
- These phenomena are largely independent of one another but interact
- The same phonetic parameters, like F0, have several functions at the same time
- They also affect the realization of segments

# Some provisos

- Components of prosody share the same basic phonetic parameters; e.g.
  - Stress is manifested in many languages as additional duration and higher amplitude
  - Duration is also used to indicate phrasing
- Phonetic parameters are not used and organized in the same way across all languages; e.g.
  - The difficulty in defining *prominence* cross-linguistically
  - The role of F0 in Mandarin and English

# Provisos continued

- Components of prosody use the same basic **phonetic** parameters that are also used to encode **paralinguistic** information in speech
- **Linguistic**: systematic & structured; language-specific; categorial; meaningful
- **Paralinguistic**: more variable; less language-specific; continuous; e.g.
  - The higher the pitch range, the more excited the speaker is (but the essential meaning does not change as the pitch range changes)

# STRESS

# What is stress?

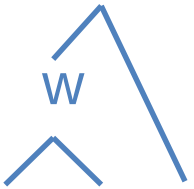
- We can define stress as relative prominence between constituents
- Stress is what makes some syllables in a word stand out relative to others
- Stress operates both lexically (at the word level) and post-lexically (at the utterance level); e.g.

banana vs. ripe banana

# Metrical phonology (Liberman & Prince 1977)

- Stress is relational = syntagmatic/ hierarchical (comes in layers)
- It does not make sense to say a syllable is stressed unless it can be compared to another syllable (though in English some syllables are inherently stronger than others)
- The relationship of strength among the syllables of a word does not change because of the context in which the word is found
- Syllables are designated as strong (s) or weak (w) and grouped in hierarchically structured trees


open

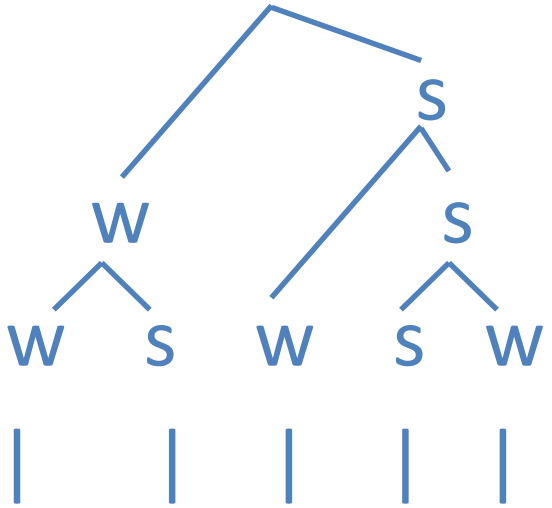



open door

# Some simple representations of stress


The sun is shining

GRID




The sun is shining

TREE

# The phonetics of stress

- How stressed syllable stand out is the result of
  - a combination of phonetic parameters; **there is no one phonetic property that gives rise to the percept of stress**; that is, stress is not like, say, nasality which is clearly associated with a lowered velum and nasal resonances.

  - These phonetic parameters are to a large extent language specific; therefore, percepts of stress differ by language

  - English [bəˈnænə]/[bəˈnɑnə]  Greek [baˈnana]

# Stress and the role of F0

- Early perceptual experiments on English by Dennis Fry (1955, 1958) have show that percepts of stress in ENGLISH change when F0 changes, in word pairs like

  *prótest* (N) ~ *protést* (V)

  [ˈproʊtəst] ~ [prəˈtɛst]

- This result has been interpreted as showing that stress is indicated by high or rising pitch
- This is simply a misunderstanding, confounding stress and intonation: stressed syllables often have high or rising pitch but this is not a necessary feature

# Stress as localized hyperarticulation

- Ken de Jong (1995): stress as **localized hyperarticulation**
- In English, stressed syllables are *hyperarticulated*:
  1. they have more carefully pronounced consonants
  2. they have more peripheral (non-centralized) vowels
  3. 1&2 make stressed syllables longer and often louder
  4. stressed syllables are also docking sites for F0 (pitch) movements (though not all stressed syllables show a pitch movement, and not all pitch movements occur on stressed syllables)

From Lindblom's H&H theory (Lindblom 1990)

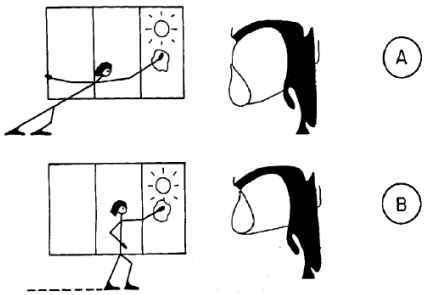



Figure 2. Movement economy and plasticity. A: Compensatory behaviors. B: Normal behaviors.

# Linguistic functions of stress

- Culminative function: refers to making elements in speech stand out; stress, by definition has culminative function
- Contrastive function: refers to changes in meaning as in *prótest* N vs. *protést* V
- Delimitative function: refers to the creation of groupings; e.g. in English stressed syllables are typically word-initial; in Finnish this always applies
- Stress can have any of these functions, depending on the language; e.g. Hungarian and Czech vs. English, vs. Greek, Spanish

# Stress in other languages

- Stress is manifested in a variety of ways
- In Spanish, Greek and Italian stressed syllables are longer than unstressed syllables, but the difference in duration is not as great as in English
- In Spanish and Greek all five vowels [i e a o u] can appear in both stressed and unstressed syllables; there are only small differences in quality between vowels in stressed and unstressed syllables

# An example

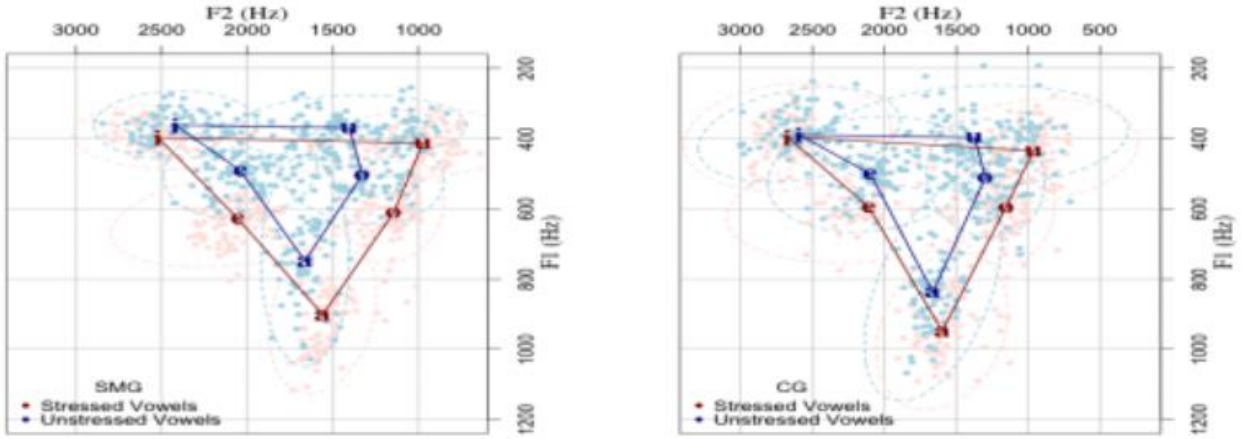



***Figure 2:*** *Stressed and unstressed SMG (Left Panel) and CG (Right Panel) vowels with 95% confidence ellipses.*

Charalambos Themistocleous & Angeliki Logotheti. 2016.
Standard Modern Greek and Cypriot Greek vowels: A sociophonetic study. Proceedings of MGDLT6.

# Stress in other languages

- Stress is manifested in a variety of ways
- In Spanish, Greek and Italian stressed syllables are longer than unstressed syllables, but the difference in duration is not as great as in English
- In Spanish and Greek all five vowels [i e a o u] can appear in both stressed and unstressed syllables; there are only small differences in quality between vowels in stressed and unstressed syllables
- In Italian some vowels do not appear in unstressed syllables:

  stressed vowels = [i e ɛ a ɔ o u], unstressed syllables = [i e a o u]
- Other languages, such as French, Korean, Bangla, Japanese, many varieties of Basque, do not reliably distinguish syllables on the basis of stress
- **It is not necessary for a language to have stress**

# Repercussions for analysis

- Necessary to question whether a language has stress when analyzing a new language
- Examining both phonetic (acoustic, perceptual, articulatory) and **phonological evidence** is essential
- The use of phonological evidence is often neglected
- Objectively measuring the acoustics of stressed syllables does not reflect native speaker perceptions; e.g.,
  - stress has a higher functional load in Spanish than in English, but Spanish cues to stress are more subtle than those used in English
- Consequences: for field work (language contact) / typology

# Some criteria for stress

Phonetic criteria

- In what ways do putatively stressed syllables stand out?
  - Longer duration
  - Higher amplitude
  - Greater **amplitude integral**
  - Hyperarticulation: e.g.
    - full vowel structure
    - spectral tilt
    - consonant quality
- Native intuitions are crucial

Phonological criteria

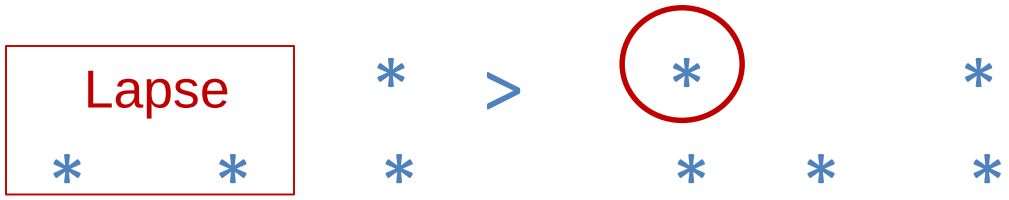
- Can stress distinguish word meaning?
- Lack of lenition and/or presence of fortition
- Different sets of vowels in stressed and unstressed positions
- Do differences consistently correlate with position within a phrase?

# STRESS, RHYTHM, AND TIMING

# Rhythm and stress

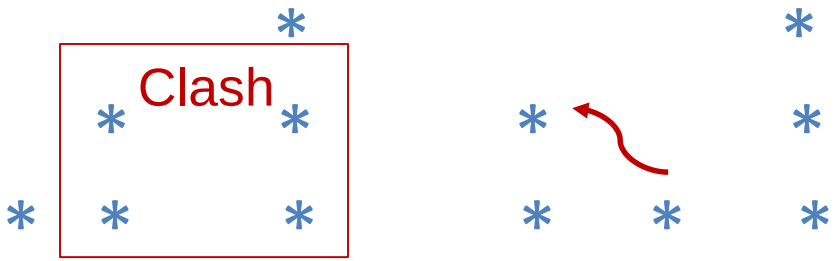
- Phonology: stress is used to create patterns of alternation between strong and weak constituents (syllables, feet, words)
- This alternation is thought to create rhythm
- Adjustments are made to stress patterns to retain eurhythmy, i.e. (relatively) regular alternation between strong and weak constituents

# How to achieve eurhythmy in English




**Beat Addition** to remedy the lapse

three red shirts > three red shirts




**Rhythm Rule** (Iambic Reversal) to remedy the clash

sixteen books > sixteen books

# Some problems with this view

- How does rhythm work in languages like Greek, Spanish or Italian in which such adjustments are not allowed?
- Possible answer relates to tempo: languages have a stressed syllable every 0.5 second or so (Dauer 1983, Tilsen & Arvaniti 2013; Arvaniti 2013)
- How is rhythm created in languages like Korean, Bangla, or French that do not have elements with culminative function?
- It appears that rhythm in those languages works in a similar fashion (cf. Chung & Arvaniti 2013 and Tilsen & Arvaniti 2013 on Korean) except that prominent syllables become so only postlexically (primarily through phrasing and initial strengthening)

# Speech rhythm in phonetics

- In phonetics **rhythm** has often been equated to **timing**
- **Timing is one element of rhythm, but rhythm and timing are not the same thing**
- Rhythm classes:
  - stress-timing (English)
  - syllable-timing (Spanish, French)
  - mora-timing (Japanese)
- A lot of research has been devoted (in vain) to finding **isochrony**: durational adjustments in feet, syllables, moras
- Maybe isochrony is perceptual? (Lehiste 1973, 1977)
- Maybe give up? (Roach 1982; Bertinetto 1989; Arvaniti 2009; 2012a; 2012b; Arvaniti & Rodriquez 2013; Chung & Arvaniti 2013; Horton & Arvaniti 2013; Tilsen & Arvaniti 2013)

# Dauer 1983, 1985

- No syllable-timing because it is not viable
- Stress salience creates a continuum from least- to most-stress based (NOT from stress- to syllable-timing)
- Eight criteria for stress salience
  - Phonetic: e.g. durational differences between stressed and unstressed syllables
  - Phonological: e.g. is the set of vowels used in stressed and unstressed syllables the same?
- Problem: assumes that there exist such 'objective' criteria

# Rhythm metrics

- In Ramus, Nespor & Mehler (1999), Dauer's eight criteria were reduced to two:
  - Syllable complexity
  - Vowel reduction (phonologized only)
- It was hypothesized that these two factors can be quantified and give us the distinction between stress-, syllable-, and mora-timing
- The typology is said to be the basis of acquisition and to be used for adult speech processing

# Popular rhythm metrics

Ramus et al. (1999):

ΔC = the standard deviation of consonantal intervals

%V = the percentage of vocalic stretches

Grabe & Low (2002): Pairwise Variability Indices

$$rPVI = \sum_{k=1}^{m-1} |d_k - d_{k+1}|/(m-1)$$

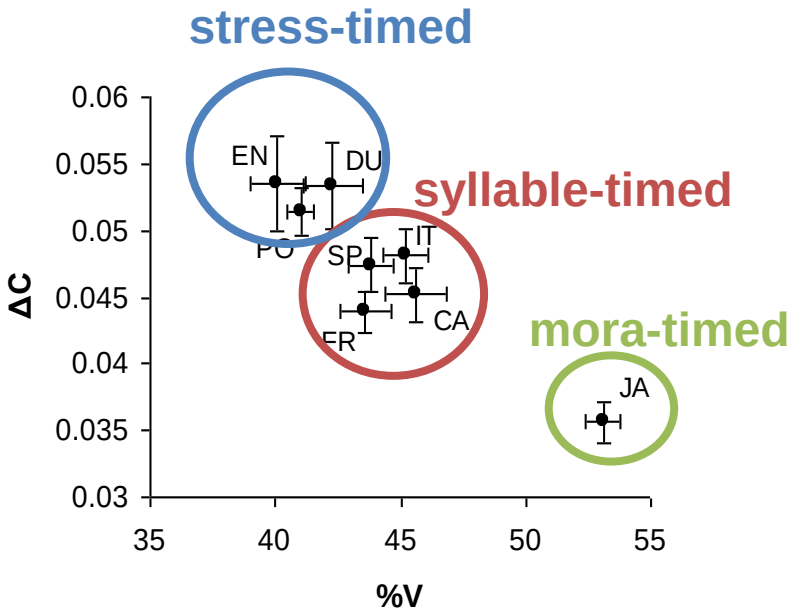
*Metric of consonantal variability*

$$nPVI = 100 \times \left[\sum_{k=1}^{m-1} \left|\frac{d_k - d_{k+1}}{(d_k + d_{k+1})/2}\right|/(m-1)\right]$$

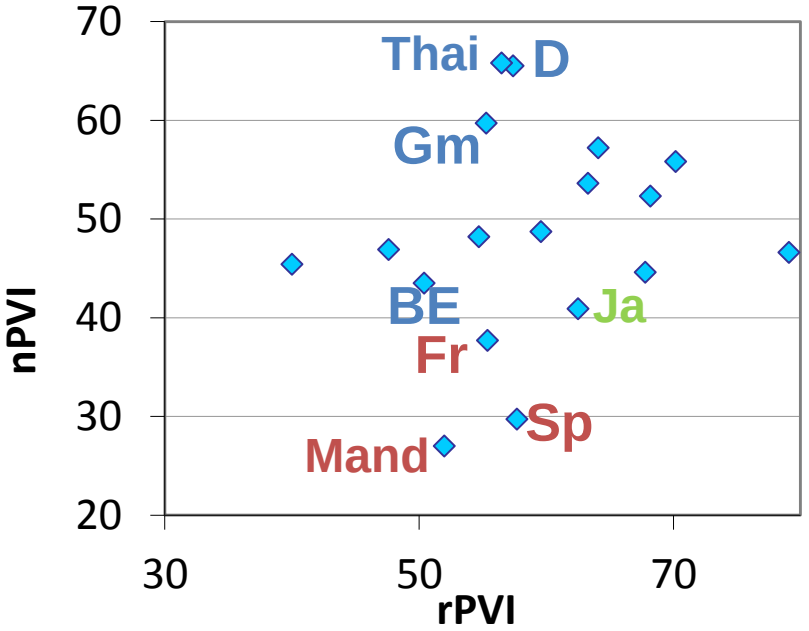
*Metric of vocalic variability*

Dellwo (2006): VarcoC and VarcoV (SD/Mean)

# Rhythm classification with metrics?




Ramus et al. (1999)



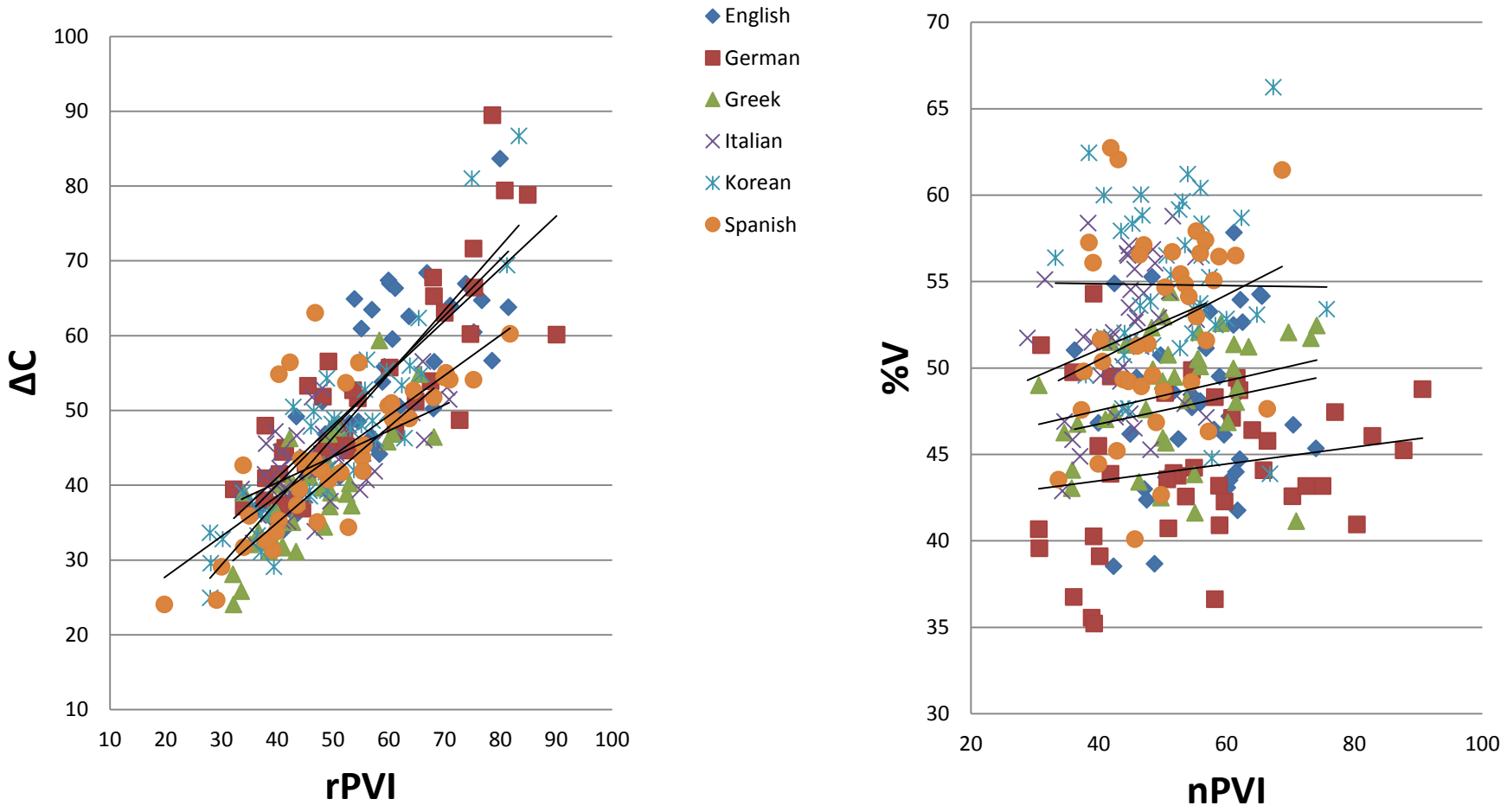

Grabe & Low (2002)

If metrics work,

- most languages should be classifiable (G&L classified just 5/18 languages)
- most studies beyond “the usual suspects” have yielded very poor results

# Cross-metric comparisons

Based on Arvaniti, 2012, *JPhon*



# Rhythm classes, production and perception

- Metrics don’t work in production
- They mostly capture syllable structure (Renwick 2012; Horton & Arvaniti 2013)
- Recent studies also show that languages are not discriminated in perception by rhythm class
  - Listeners cannot reliably perceive differences in rhythm (Arvaniti 2012b)
  - Discrimination is based primarily on speaking rate (Arvaniti & Rodriquez 2013)
  - Impoverished signals (like [sasasa]) are not ecologically valid (Arvaniti 2012b; Arvaniti & Rodriquez 2013)

# Rhythm and timing

- Neither the production or the perception of **timing** helped with rhythm classification
- No convincing evidence for rhythmic classes either in production or in perception
- Keep timing as a term that refers to the durational structure of utterances; this structure depends on stress patterns, phrasing and language-specific patterns to do with segment duration and coordination, realization of changes in speaking rate and so on (Arvaniti 2009)
- Define rhythm as an abstract structure based on prominence and grouping

# Phrasing

# Phrasing

- When we produce running speech, we do not just put words one after another
- Rather, we group them into hierarchical constituents: prosodic words, phrases, and utterances (simplifying a great deal)
- These constituents are not the same as syntactic constituents,

speech → [this is the cat] [that caught the rat] [that stole the cheese]…

syntax → this is [the cat that caught [the rat that stole] the cheese]]]…

- Differences the two are not always systematic and easy to account for; cf.

[My brother's wife's sister] [is a heavy smoker]

[John's a heavy smoker]

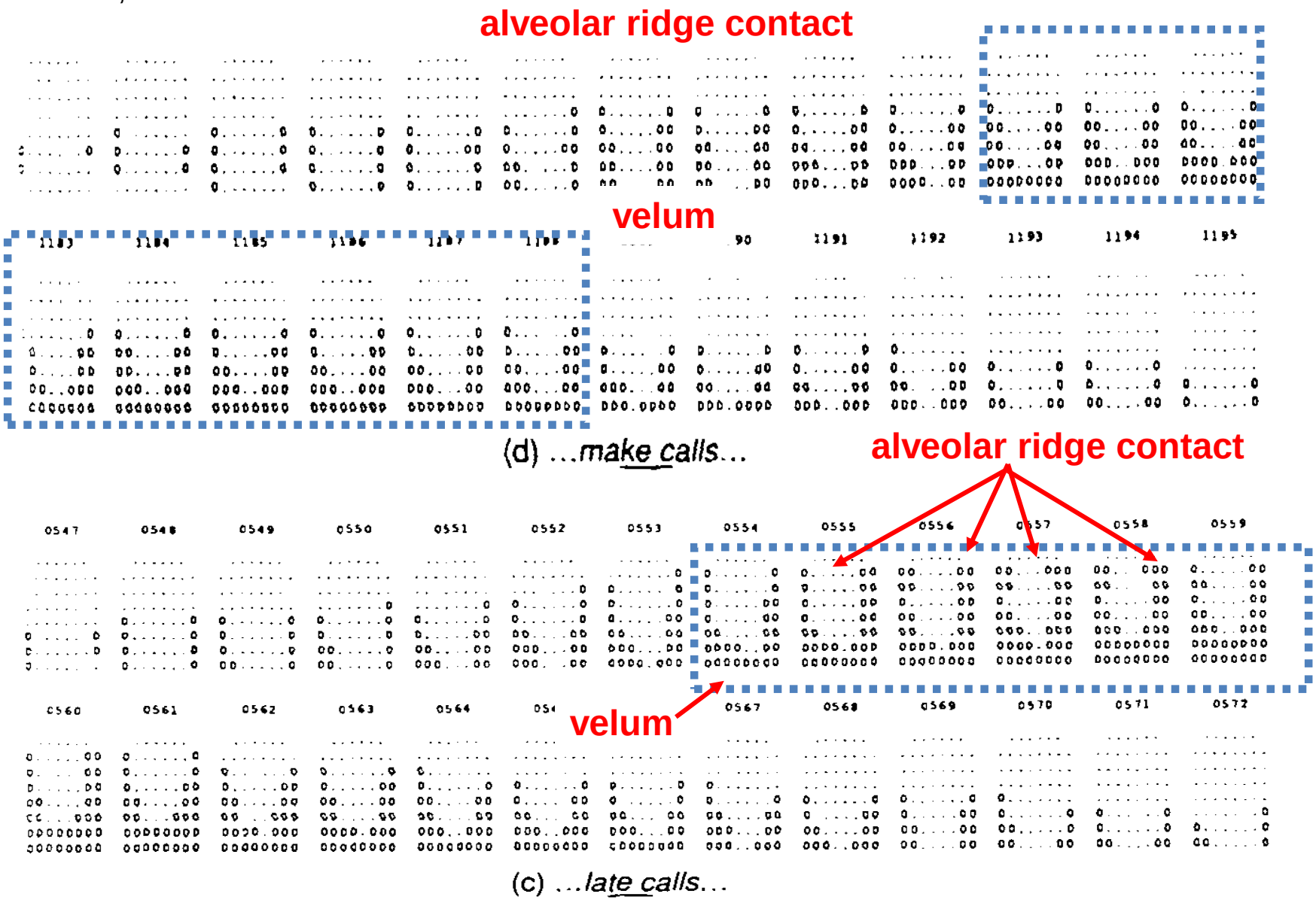
# Sandhi and phrasing

- Many phonological phenomena (assimilations, lenitions, elisions, etc.) including stress rules like the Rhythm Rule are affected by phrasing: they may occur only across or within certain phrasal constituents
- Thus, **sandhi** changes of this sort act as cues to prosodic phrasing and help listeners parse the segmental string
- Flapping in AE applies within a foot but not across feet
  - Don't hi**t** it to Joey! - the /t/ of "hit" is flapped ([ɾ]) in American English
  - Don't hit! It's Joey's turn! - the same /t/ is not flapped

# Sandhi and levels of phrasing

- Sandhi has been often used as a prosodic constituency diagnostic (e.g. Nespor & Vogel 1986)
- Many sandhi phenomena turned out to be
  - Affected by intonation rather than phrasing
    - Deletion/coalescence in hiatus is blocked in Greek if the second word is in focus (Baltazani 2006)
  - Variable
    - Alveolar stop assimilation in English (*bad guys* does not turn into *baguys;* Nolan 1992)

From Nolan, 1992




Circles = tongue-palate contact; dots = lack of contact; in *late calls* [t] is not deleted as phonology predicts: it is reduced and overlaps with the following [k]

# Sandhi and levels of phrasing

- Caution is needed when considering sandhi as a criterion

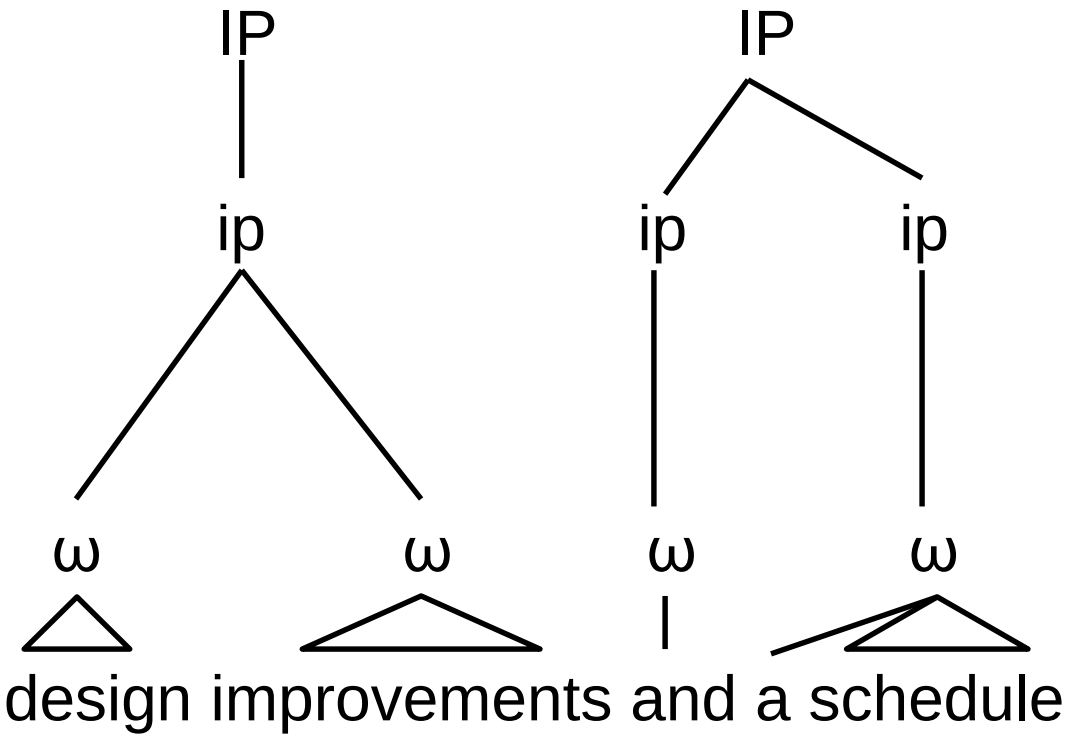
# Additional cues to phrasing

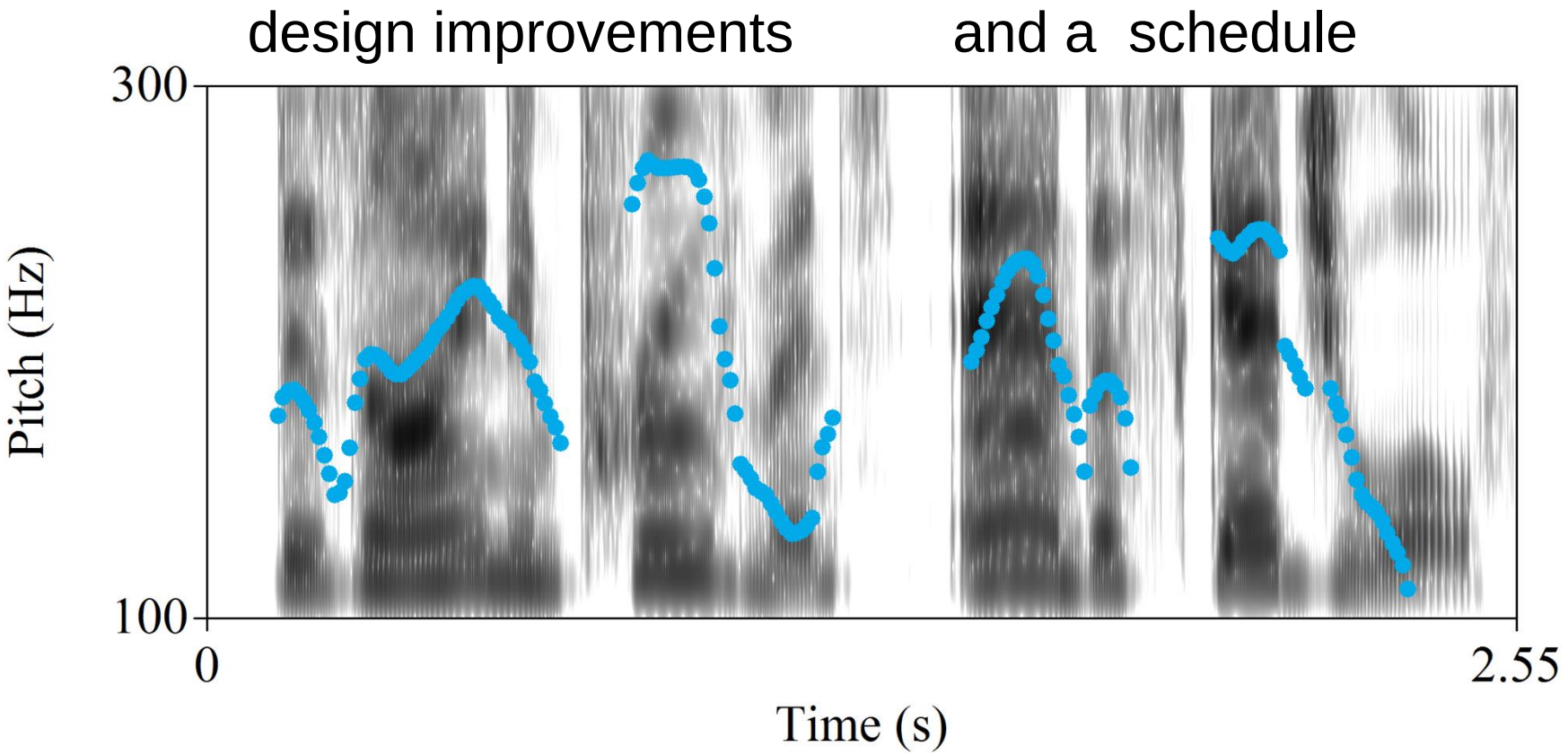
- Intonational melodies apply over phrasal constituents (thereby providing listeners with more cues to the presence of phrasal boundaries)
- At prosodic boundaries, short silences can and may be inserted
- The higher the constituent, the stronger the boundary is felt to be and the more likely it is for a pause to be inserted
- Left edges of phrasal constituents in particular show articulatory strengthening

# Representation of phrasing

- There are many models of representation:
  - bracketed grids, binary trees, n-ary trees
- For an analysis the *minimum* needed is
  - To know (or decide) if the language has stress
  - To know (or decide) on levels of phrasing
- For English, Beckman & Pierrehumbert (1986) recognize
  - The intermediate phrase [ip]
    - akin to the phonological phrase of other models
  - The intonational phrase [IP]

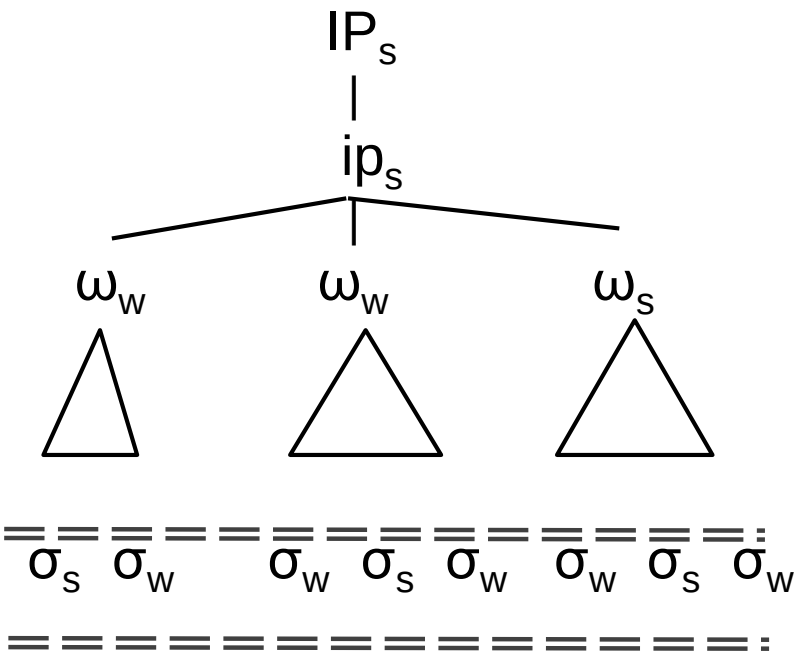
# An example from MAE ToBI




design improvements
and a schedule
300
100
Pitch (Hz)
0
2.55
Time (s)

# An example from Polish

**Prosodic Structure**




Intonational Phrase IP

intermediate phrase ip

Prosodic Word w

(Foot) F

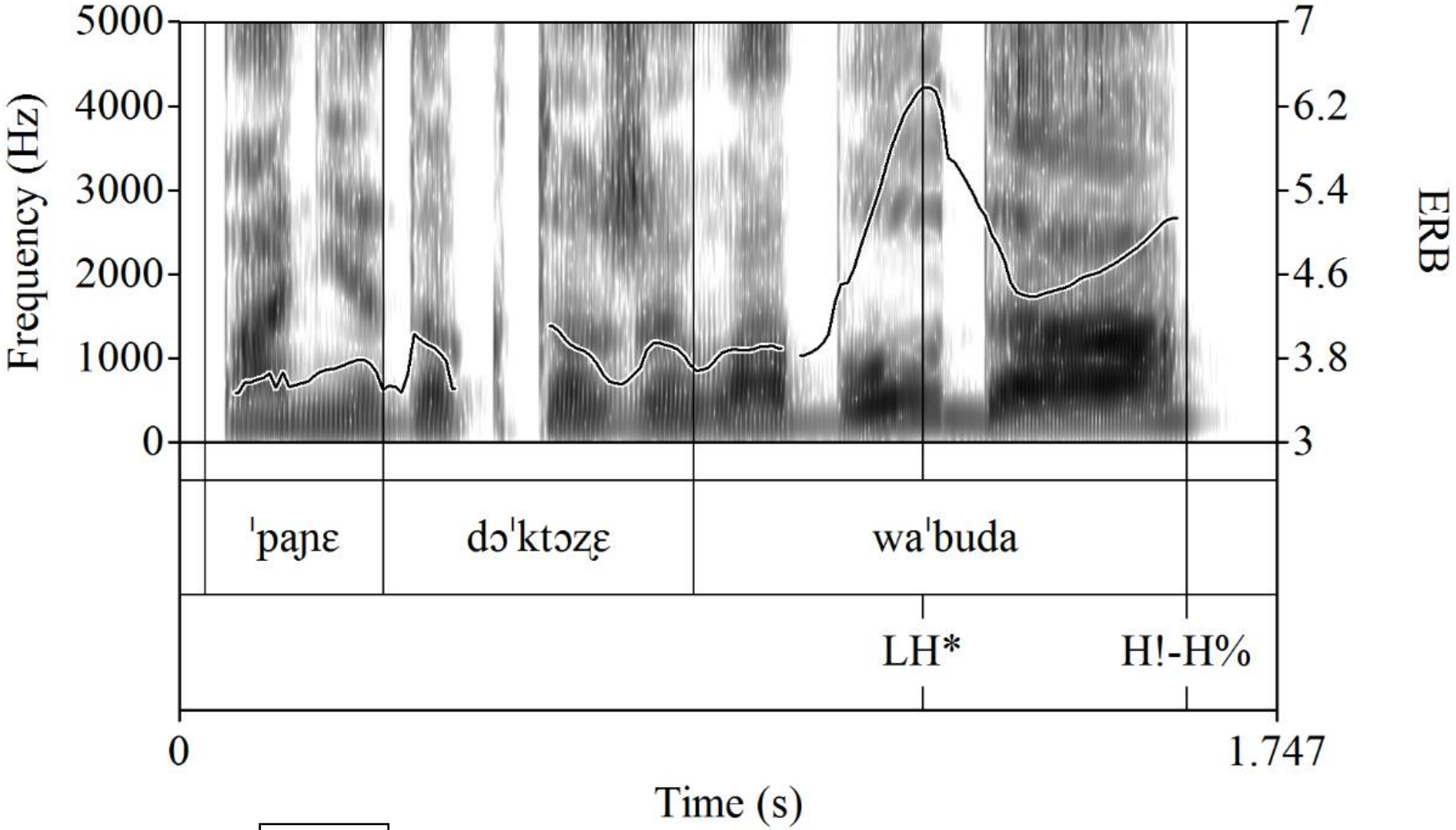
Syllable

ˈp a ɲ ɛ d ɔk ˈt ɔ ʐ ɛ w a ˈb u d a — Phoneme Tier

*Mister Doctor Wabuda!*



Data from Arvaniti et al., *Phonetica*, to appear


5000
4000
3000
2000
1000
0
Frequency (Hz)
7
6.2
5.4
4.6
3.8
3
ERB
ˈpaɲɛ
dɔˈktɔzɛ̨
waˈbuda
LH*
H!-H%
0
1.747
Time (s)

# Intonation: come back at 11:00!