Autosegmental-Metrical Theory I & II (AM)

Amalia Arvaniti (I)
Amalia Arvaniti & Janet Fletcher (II)

Aix Summer School on Prosody
6-9 September 2016
What is AM?

• AM (Autosegmental-Metrical theory of intonational phonology) is a phonological theory of intonational structure.

• Before we discuss principles of AM it is essential to clarify:
  – Terminology
  – Differences between AM and other theories
  – Confusion between AM and ToBI
What is intonation?

• Intonation is the linguistically controlled and pragmatically meaningful use of F0 that spans entire utterances

• F0 changes related to intonation are NOT associated with word meaning, but to the syntax and pragmatics of the utterance

• e.g. think of saying *my ex* as a response to
  – (i) who is that?
  – (ii) being surprised that your new partner has invited their ex to your birthday party
F0

- F0 is the main exponent of intonation
- F0 (fundamental frequency) is an acoustic property of the speech signal
- It is determined by the rate of vibration of the vocal folds (the number of times they open and close per second): F0 is higher the higher the rate of vibration
- F0 is measured in Hz (cycles per second)
- F0 is typically lower in the voices of men and highest in the voices of children, with women in between
- These ranges are determined both by physiology and social norms
F0 & pitch

• F0 gives rise to the percept of pitch
• Scales for pitch: ERB, mel, bark
• Semitones: NOT a perceptual scale
• Under 700 Hz, the relationship between F0 and pitch is linear, so we can refer to either and use Hz to measure both
Brief pit stop to have a look at PRAAT

- Scales
- Speckles, lines and microprosodic perturbation
- Algorithms for viewing and analysis
- Viewing and analysis range
- Praat pictures
F0 and intonation

• All languages use and modulate F0: it is part of voicing
• However, not all modulations of F0 are intonation
• DO NOT USE the term intonation when you mean pitch or F0
Uses of F0

• Languages use F0 lexically, post-lexically, and para-linguistically and sociolinguistically
• Lexical uses of pitch are referred to as tone and pitch accent
• Post-lexical uses of pitch are referred as intonation
• Para-linguistic uses (changes in pitch to show anger, boredom, surprise, excitement) often blur the issue, especially for intonation
• Changes in pitch range, pitch span and dynamism may also serve sociolinguistic functions
• These uses are not mutually exclusive
Lexical uses of F0: tone

- Changes of F0 change lexical meaning
- In tone languages most syllables are lexically specified for tone
  - **contour tone languages**: tones tend to be dynamic, such as rise, fall or fall-rise; tones can be combined with register (so a language may distinguish a high fall from a low fall) and/or with phonation differences; e.g. Mandarin, Cantonese, Thai
  - **level tone languages**: tones tend to be levels, such as high and low; most tone languages distinguish two or at most three level tones (high, low, mid); some have only a high tone (and absence of tone in the other syllables); Yoruba, Shona, Zulu, Luganda, Ibibio
- Phonetically, dynamic tones exist in all systems of lexical tone, either only phonetically or because they are phonologically specified that way (e.g. Mandarin, Cantonese, Thai)
Igbo narrative

Pitch (Hz)

Time (s)
Bole example

![Graph showing time (s) vs. pitch (Hz) with data points at 0, 1.078, 75, and 250.]
And a Mandarin example

![Graph showing pitch over time]
Lexical uses of F0: lexical pitch accent

• In languages with lexical pitch accent at most one syllable per word is lexically specified to carry tone
• A language may have only one type of pitch accent
  – in Classical Greek one syllable of each word had rising, falling or rising-falling pitch; e.g. φως “light” [pʰóòs] δος “give (imperative)”[dós]
  – in Swedish, there are two accents: both are falls but one is timed early with respect to the accented syllable, while the other is timed late
• Almost all words of the language may have pitch accent (Swedish, Serbian, Classical Greek) or only some (Japanese: about 30%)
• In some of these languages, pitch accent plays the same role as stress in English
• In other languages, e.g. Swedish, both pitch accent and stress exist
Tone and lexical pitch accent

• Some argue that there is no essential difference from a phonological perspective between the two phenomena
  – Gooden et al. (2009) suggest that it is simply a matter of how dense the lexical specifications for tone

• Hybrid systems from creole languages, such as Papiamentu (Remijsen and van Heuven, 2005, Phonology) support this idea
What is intonation?

• Intonation is the linguistically controlled and pragmatically meaningful use of F0 that spans entire utterances

• F0 changes related to intonation are NOT associated with word meaning, but to the syntax and pragmatics of the utterance
Some questionable statements about intonation

- Intonation is difficult or unsystematic
- Intonation shares patterns across languages based on biological characteristics (cf. Ohala (1983); cf. Gussenhoven’s (2004) frequency code)
- Tone languages do not have intonation (correct for some but only in a very narrow sense)
- Intonation has to do with affect or emotion
- Particular tunes have particular meanings; e.g. questions end in a pitch rise and statements in a fall
The status of intonation

• Intonation is part of the linguistic system, specifically part of a language’s phonology
• This point has not always been addressed and by some it has even been refuted, so that in other models intonation is practically equated to F0
• Many models of intonation
  – British School
  – the level system(s) of the American structuralists
  – INTSINT
  – Xu’s Penta
  – Fujisaki’s model

attempt to faithfully reproduce/model F0 contours and to visually represent the course of F0
An example from INSINT

- M = mid
- S = same
- H = High(er)
- L = Low(er)
- U = up
- T = top
- D = down
- B = Bottom
The British School

Well 'make up your \textit{mind}.
And others

Where are you going?

Tókhiya lá he?

Lakota dictionary

He bought apples, peaches, pears, and oranges.
I’ll have two pencils, a black pen, and some ink.
We went to Paris, Brussels, Amsterdam, and London.
I saw Esther, Jane, Neil, and Susan.

Good evening. How do you do? I hope your mother is feeling well today, Miss Johnson. The weather is very nice. Would you like to go out for a walk?
AM

• AM is a phonological theory of intonational structure that connects stress and phrasing with intonation

• The aim of AM is NOT to faithfully present F0 contours – that’s what pitch tracks are for

• AM is not a transcription system for intonation

• AM relies on an abstract phonological representation coupled with pitch tracks ↓

Amalia Arvaniti, AixPros 2016
AM and ToBI: what ToBI is not

• ToBI is not a transcription system for intonation (or prosody) either

• **ToBI is not the IPA of intonation**

• Such annotation systems do exist e.g. INTSINT (International Transcription System for Intonation)

• ToBI is not one of them

• So, one cannot “learn ToBI” as one can learn the IPA Alphabet

• **THERE IS NO TOBI POLICE** that will stop you from employing solutions suitable for the language you analyze; e.g.

  *I cannot use ToBI because my language does not have stars*

• **NO LANGUAGE HAS STARS!**

• Stars are a notational device
What is ToBI?

- **ToBI** = **T**ones and **B**reak **I**ndices
- A name given to a family of systems designed for the prosodic annotation of spoken corpora
- Family of annotation systems: they share basic assumptions about aims and about intonation and prosody
- Aim of ToBI systems: to create searchable databases and extract regularities from the data
- Assumptions about intonation and prosody: AM conception of how intonation and prosody are structured
- The first such system, called ToBI, was designed for English (now MAE ToBI for Mainstream American English)
Some reassurance

• (Repeat) There is NO ToBI police
• There is no obligation to use MAE ToBI labels for other languages
• There is no obligation to use the same number of tiers or the same types of tiers as MAE ToBI
Why a phonological model of intonation?

• Intonation is part of a language’s grammar
• Representations must
  – capture useful generalizations about the system
  – capture what is contrastive in the system
  – have predictive value
• These properties allow a speaker or a learner to
  – use the appropriate melody in a given situation
  – process F0 contours as instances of the same melody independently of their exact shape
An illustration of variability & gestalts

From Arvaniti & Ladd, 2009, *Phonology*

Where (does s/he) live?

Where could they be speaking to me from?

Amalia Arvaniti, AixPros 2016
An illustration of variability & gestalts

Where (does s/he) live?

Where could they be speaking to me from?

From Arvaniti & Ladd, 2009, *Phonology*
An illustration of overspecification

From Arvaniti & Ladd, 2009, *Phonology*

Where (does s/he) live?

Where could they be speaking to me from?
Recap of slides 26, 27, 28

• Parts of the tune stay put and others move about or change size
• This suggests that
  – Tunes are not gestalts: we cannot stretch or shrink them, accordion-like, to fit utterances of varying lengths
  – We can describe them by specifying the F0 of each syllable, but doing so means we miss significant generalizations: what do they tunes depicted in the previous slides have in common and what is responsible for the differences between them?
Breakthrough

• Pierrehumbert’s thesis (1980): *The phonetics and phonology of English intonation* (with lots of ideas from the British School, the autosegmentalists like Goldsmith and Leben working on tone, and Mark Liberman and his work on stress and intonation)

• Bob Ladd (1996) dubbed the theory AM for *autosegmental metrical theory of intonational phonology*
• **Autosegmental-Metrical Theory of Intonational Phonology**

• Why this mouthful of a term?

• Intonation consists of discrete elements called tones, L(ow) and H(igh), which are concatenated in various configurations and associated with metrical structure

• **Autosegmental**
  – Tones are autosegments

• **Metrical (prosodic)**
  – Tones associate with structural positions in the metrical structure: heads and boundaries of constituents
Tones

• The primitives of intonation are **tones**: low (L) and high (H)

• Tones are represented as a string of **autosegments**: LHLHLH
  – They are independent of vowels and consonants (though they have to co-occur with them to be realized)
  – They are independent of each other
  – They may form bitonal (or even tritonal) configurations
Tones as autosegments

• Etung
  
  étùmbá  bìsônjé  ékúé  òbô
  
  * pot  * wife  * forest  * arm

How can we account for these patterns, especially the fact that in some words every syllable has a different tone, in others they all have the same tone, and in yet others, some syllables have a complex tone (e.g. last syllable of arm)?

(from Gussenhoven & Jacobs 1998)
Tones as autosegments

• Etung
  étìmbá  bìsnjé  ékúé  òbô
  *pot*  *wife*  *forest*  *arm*

• Etung has a number of tonal melodies, among them H LH HLH and LHL

  étìmbá  bìsnjé  ékúé  òbô
  \[\text{\texttt{H L H L H L H L HL}}\]

(from Gussenhoven & Jacobs 1998)
Autosegments

• Basic principle about autosegments
• The relationship between tones and segments (Tone Bearing Units or TBUs) is NOT one-to-one
• The number of TBUs and tones may match
• We can have more tones than TBUs
• We can have more TBUs than tones
Underspecification in AM

• In the AM model, tones do not exhaustively represent the course of F0

• Tones represent the *linguistically relevant* points of an F0 contour – they are assumed to be contrastive

Circled elements: elements likely to make it into an AM representation
What are the linguistically relevant points?

• The linguistically relevant points reflect pitch modulations that co-occur with
  • prominent syllables
  or
  • phrasal boundaries
Tones and metrical structure

• The association of tones with prominent syllables and phrasal boundaries presupposes the existence of a **metrical structure**

• This structure provides the necessary information:
  – prominence relations among constituents (heads)
  – position of phrasal (and other constituent) boundaries

• Note that a language need not have stress to have intonation but it will need to have phrasing (cf. Korean)
Metrical structure and phrasing

• For an AM 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 accentual phrase [α or AP]
    • Used in analyses of Japanese, Korean, French
  – The intermediate phrase [ip]
    • akin to the phonological phrase of other models
    • Used in analyses of English, Greek, Italian
    • Not used in, e.g., ToDI (Dutch ToBI-like system)
  – The intonational phrase [IP]
    • Generally accepted
Criteria for phrasing

• Accentual Phrase
  – A prosodic constituent roughly intermediate between a prosodic word and a phrase
  – Defined by a pitch contour (e.g. by initial and final phrasal tones)

• Intermediate Phrase
  – Sense of juncture, but no pause or sense of completion
  – Pitch movements that are simpler and less extreme compared to those in IPs

• Intonational Phrase
  – Often followed by pause
  – Found at the end of a turn
  – Can form an utterance on its own
  – (Typically) no sandhi with following utterance
  – REMEMBER TO TREAT SANDHI WITH CAUTION
Association by tone type

• Association with the heads of metrical constituents (roughly, stressed syllables):
  – pitch accents, e.g. L*, L+H* ...

• Association with boundaries:
  – phrase accents, e.g. L-, H-, LH-, associate with IP boundaries (typically right boundary)
  – boundary tones, e.g. L%, HL%, associate with IP boundaries (typically right boundary)
Pitch accents

• Pitch accents represent pitch movements that associate with metrical heads (stressed syllables): formally, pitch accents associate with specific constituents and percolate down to the constituent’s head
• Pitch accents are flagged by * e.g. L*
• If bitonal, one element is usually starred, e.g. L+H*
Pitch accents and stress revisited

• Recall Fry: stressed syllables have high or rising pitch

• Why did Fry conclude that?
  – Pitch accents (which are often, though not always rising) associate with stressed syllables
  – When you hear a pitch rise on a syllable that could be stressed, you assume it is

• But pitch accents are NOT a correlate of stress – pitch accents are prominence-cueing pitch movements
Phrase accents

• Phrase accents are indicated by -, e.g. L-
• They associate with intermediate phrase (ip) boundaries
• Not all models of AM accept the presence of phrase accents
• If they do not fulfil a purpose, they are not required
• They can be monotonal (e.g. L-) or bitonal (e.g. L+H-)
Boundary tones

- Boundary tones are indicated by %, e.g. H%
- They associate with Intonational Phrase (IP) boundaries
- English does not have bitonal BTs, but bitonal and multitonal BTs do exist, e.g. in Korean, which has LHLH% and HLHL% (among other complex boundary tones)
Stars, percentages and crosses

- *, - and % are diacritics that indicate the function of a tone
- In bitonal accents the * indicates the **metrically stronger** tone of the accent and typically the one that co-occurs with the stressed vowel
- When tonal events consist of two (or more) tones, a + may be used to indicate their connection: L+H% or LH%
- Jun & Fletcher (2014) have suggested that + be used only if the tones are independent of each other:
  - L+H* is a weak L tone followed by a strong H tone
  - LH* is a rise
- This practice is not yet widely accepted however
- Brackets are meant to show that the * is related to the unit rather than a particular tone, e.g. (L+H)*
Connecting Autosegmental and Metrical

• The tones are associated, by means of association lines, with structural positions in metrical structure:
  – heads of constituents (informally stresses)
  – phrasal boundaries

[[John loves Mary]IP]
RECAP

• AM is an abstract phonological model to represent the contrastive elements of an intonational system.
• These elements are a string of tonal autosegments.
• They associate with structural positions in metrical structure:
  – heads of constituents (informally stresses)
  – phrasal boundaries
Phonetic realization

• AM includes a theory of how its abstract representations are phonetically realized
• This is known as text-tune association
• Phonological tones are realized as tonal targets defined along two dimensions:
  • **Scaling** = their F0
  • **Alignment** = their temporal position with respect to segmental landmarks
• What are the **segmental landmarks**?
• It depends on the type of intonational event
Phonetic realization of pitch accents

• Pitch accents co-occur with stressed syllables
• Pitch accents are typically realized on the vowel (most likely Tone Bearing Unit or TBU)
• Some time-lag is possible, especially with bitonal pitch accents
Bitonal accents

• Bitonal accents in different languages have been shown to have different types of alignment

  1. One tone shows stable alignment, the other precedes or follows by a given amount of time

  2. The tones align independently of each other (i.e. with its own segmental landmark)
Boundary tones

• Boundary tones are typically realized on the last TBU

• So, a boundary tone like H% will likely be realized as a rise on the last vowel of a phrase

• Bitonal and multitonal boundary tones are also possible; attested, e.g., in Korean
Text-tune association

NOTE: the AM analysis of the top example is based primarily on meaning, rather than solely on pitch accent shape; alternatives are possible.
Phrase accents

• Phrase accents were originally said to be ‘floating tones’ with inconsistent realization between the last PA (nuclear accent) and the boundary tone; H* L- L%

• Beckman & Pierrehumbert (1986) proposed that in English phrase accents associate with ips (not everyone agrees and not all analyses rely on phrase accents)
Name a mammal

platypus is a mammal

Time (s)
0 1.224
Pitch (Hz)
0 375

Amalia Arvaniti, AixPros 2016
Phrase accents & secondary association

• Phrase accents appear to be a valid category with consistent realization in many languages, occupying the area between the last pitch accent and the boundary tone.

• In a number of languages, they may also show “stress-seeking” behavior, co-occurring with stressed vowels; this is analysed as secondary association in Grice, Ladd & Arvaniti (2000) – following a proposal by Pierrehumbert & Beckman (1988).
An example: Greek polar questions

In “On the BLUE balcony?” (left), the stressed syllable of [bal’kon] is high, while in “on the blue BALCONY?” it is low because it is in focus, and the rise moves to the last syllable of the question.
Underspecification in AM

- Since only the linguistically significant points are specified, the phonological representation is sparse.
- Crucially, the representation remains unspecified in phonetics as well: the F0 contour, beyond the tonal targets, is determined by interpolation, not by tones.
- In other words, there are typically fewer tonal targets to realize than there are TBUs.
An illustration of interpolation

Where (does s/he) live?

Where could they be speaking to me from?
Tonal crowding

• The flip side of underspecification is that in many instances there may be more tones than TBUs
• This is known as tonal crowding
• Languages have a variety of ways of dealing with tonal crowding
  – Deleting or truncating tones
  – Undershooting (compressing) tones
  – Using, by preference, a different melody when there is tonal crowding
• No language of those investigated uses one of these methods exclusively
Text-tune association: Polish calling melodies
Text-tune association: Polish calling melodies

The rise is the reflex of a pitch accent
The fall-rise is the reflex of phrasal tones

Arvaniti, Žygis & Jascula, *Phonetica*, to appear
Tonal crowding: Polish calling melodies

The L is indispensable in the routine call tune. The L is truncated under extreme tonal crowing in the urgent call, suggesting it is an optional element.
Connecting AM with phonetics

[[John loves Mary]ip]IP
An example from Polish

Prosodic Structure

Intonational Phrase IP

intermediate phrase ip

Prosodic Word w

(Foot) F

Syllable

Phoneme Tier

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

Amalia Arvaniti, AixPros 2016
An example from Polish

Intonational Phrase IP

intermediate phrase ip

Prosodic Word w

(Foot) F

Syllable

Tone Tier

Phoneme Tier

Data from Arvaniti et al., Phonetica, to appear
An example from Polish

LH* pitch accent
!H- phrase accent
H% boundary tone

Data from Arvaniti et al., Phonetica, to appear
An example from Polish

Intonational Phrase IP
intermediate phrase ip
Prosodic Word w
(Foot) F
Syllable
Tone Tier
Phoneme Tier

LH* pitch accent
!H- phrase accent
H% boundary tone

Mister Doctor Wabuda!

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

Amalia Arvaniti, AixPros 2016
Consequences of having phonological structure

• If we accept that intonation is a phonological phenomenon, then certain additional principles follow with respect to phonetic realization
  – No two languages need have the same intonational units
  – No two languages need have the same representation for similar melodies
    • cf. [k] is /k/ French and /g/ in English
  – Languages will show dialectal and contextual variation, including neutralization between two tonal categories
  – The connection between meaning and intonation is arbitrary (and has to be investigated)
  – Contrasts in the system (intonational elements) are posited on the basis of meaning

• So, intonation is not variable and unsystematic any more than segmentals are
Problems with too much focus on scaling and alignment

- Phonetic realization has been operationalized as stable alignment and scaling of tonal targets, resulting in
  - little tolerance for variability in the phonetics of intonation
  - the overlooking other parameters
Typical reaction to variability

- Differences among the three Greek accents relate to:
  - Peak alignment
  - Scaling of the peak
  - Slope of drop after the peak
  - Slope is of interest as all three accents are said to be followed by L-L% phrasal tones

From Katsika & Arvaniti, 2016, LabPhon
How is this typically dealt with in AM?

• Typical response: conduct control study to examine how systematic the difference is
  a. If unsystematic, ignore
  b. If systematic, incorporate into phonological representation; e.g. L+H*+L

• But is phonetic variation unsystematic unless predicted by phonological representation?
• The answer is NO; see VOT example in slide 73
These are distributions of VOT for three stops as produced by American English speakers. The fact that the distributions overlap, especially for the velars, does not stop us from positing two phonemes in English for each place of articulation.
Intonation meaning in AM

• No agreement so far as to how intonational meaning works
• Pierrehumbert & Hirschberg (1990) propose that intonational meaning is derived compositionally from the meaning of each element in a tune (plus the semantics and pragmatics of the other components of the utterance)
• Their analysis is not perfect but it is possibly the best we have so far
• Important to consider meaning when conducting analysis, including, at a minimum, elements such as sentence modality and focus
Representations and meaning

• Phonological representations of intonation need not always be phonetically transparent
• They need to capture what is contrastive
• For example: if a pitch accent is a rise it does not have to be presented as L+H*, unless it can be shown that L+H* and H* convey different meaning in the language under analysis; if they do not, the simpler H* analysis should be preferred
• Consider: we use /p/, /t/, /k/ in English to phonologically represent the language’s voiceless stops, even though we know that their most likely phonetic realization is [pʰ], [tʰ], [kʰ] respectively
Intonational analysis: Keeping things simple

• Analysis should be as simple as possible until this proves untenable
• Bottom up & top down approach to analysis
• Analysis should take into account possible sources of variation
  – General prosodic context
  – Segmental context & microprosodic variation
  – Tonal context, coarticulation and crowding
  – Variety-specific preferences for crowding resolution
  – Neutralization
• Decisions about contrasts should be guided by meaning
• Tonal representations need not be
  – Surface faithful
  – Similar to those in other languages if this is not warranted by the system
Take-home message

- Intonation is one of the uses of F0
- Intonation is part of phonology and like the rest of phonology subject to context and dialect variation (at least!)
- AM is a theory of the phonology of intonation
- Its principles can be (and have been) used to analyze languages that are typologically distinct
- This can be successfully accomplished if we understand that
  - intonation is not just F0
  - intonation is only one component of prosody
  - Intonation interacts with the other prosody components (phrasing, stress, rhythm) in ways that cannot be ignored during the analysis of intonation itself
ENJOY THE REST OF THE SCHOOL!