

Prosody in Dialog

Predicting Accent based on Information
Structure and context

Defining Prosodic Features in English

- Pitch Accent
 - Local Maxima or Minima in f_0
 - Variable depending on discourse
- Lexical Stress
 - Stress results in longer, slightly higher amplitude and pitch
 - Part of lexical item definition
 - Can be accented or unaccented

Information Structure

- Theme
 - The shared knowledge
 - Information previously discussed
 - Doesn't necessarily contain accented elements
- Rheme
 - The new information
 - Receives Pitch Accent
 - Always contains accented element
- FOCUS
 - Semantically salient parts of the utterance
 - Some kind of
- Background
 - Discourse neutral information (e.g. function words)

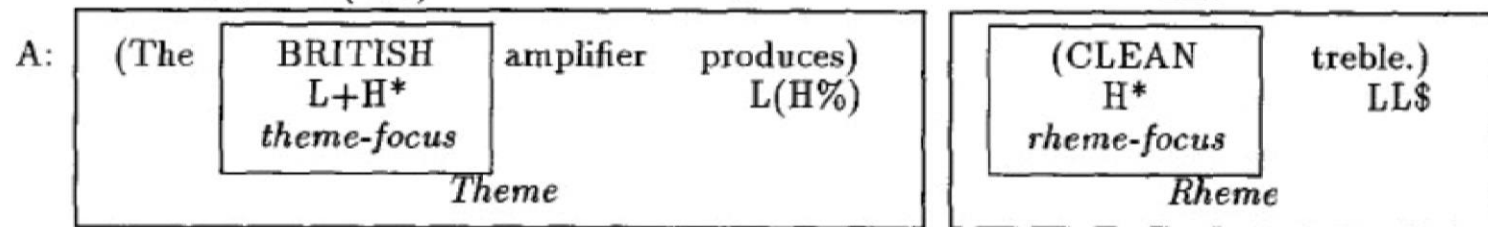
ToBI:

Tone and Break Indices

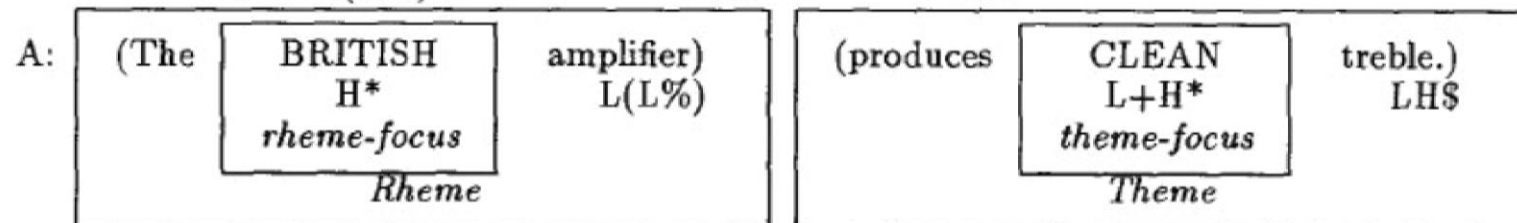
Label	Definition	Environment Found in
H	High Target Pitch	-
L	Low Target Pitch	-
H*	H centered on stressed syllable	Focus in Rheme
L*	L centered on stressed syllable	Focus in Rheme
L+H*	Rising pitch centered on H	Focus in Theme
L*+H	Rising pitch centered on L	Focus in Theme
H*+L	Falling pitch centered on H	Focus in Rheme
H+L*	Falling pitch centered on L	Focus in Rheme
!H	Relative Down step	Natural Loss of Intensity
%	Phrasal Boundary	End of sentence or phrase

Examples

(3) Q: I know the AMERICAN amplifier produces MUDDY treble,
 (But WHAT) (does the BRITISH amplifier produce?)
 L+H* L(H%) H* LL\$



(4) Q: I know the AMERICAN amplifier produces MUDDY treble,
 (But WHAT) (produces CLEAN treble?)
 L+H* L(H%) H* LL\$



Hirschberg (1990)

- Used more for reciting monologues not dialogue
- No Theme/Rheme tier
- Concept of Givenness
 - Based on context words (open vs. closed classed words)
- Used Syntax and Focus
 - Local Focus: resets when moving to new paragraph
 - Global Focus: remains relevant for entire monologue

Prevost (1996)

- Theme/Rheme used
- Developed an algorithm instead of heuristics for pitch accent assignment

DElist	Stack containing most recent knowledge seen in discourse
ASet	Set of alternatives
RSet	Set of alternatives filtered based on information in DElist and
CSet	Properties that will result in contrastive focus
Props	Grammatical properties of candidate, ranked by relevance

Example of Algorithm

(20) a. Describe the x4.
 b. The X4
 L+H* L(H%)
 is a SOLID-state AMPLIFIER.
 H* H* LL\$
 It COSTS EIGHT HUNDRED DOLLARS,
 H* H* H* H* LL%
 and PRODUCES
 H*
 ONE hundred watts-per-CHANNEL.
 H* H* LL\$
 It was PRAISED by STEREOFUOL,
 H_c !H_c LH%
 an AUDIO JOURNAL,
 H* H* LH%
 but was REVEILED by AUDIOFAD,
 H_c !H_c LH%
 ANOTHER audio journal.
 H* LL\$

(21) a. Describe the x5.
 b. The X5 is a TUBE amplifier.
 L+H_c L(H%) H_c LL\$
 IT costs NINE hundred dollars,
 L+H_c L(H%) H_c LH%
 produces TWO hundred watts-per-channel.
 H_c LH%
 and was praised
 by Stereofool AND Audiofad.
 H_c LL\$

Information State in GoDIS

PRIVATE

Information known only to speaker

SHARED

Shared Information between participants

PRIVATE

AGENDA

Immediate actions (Stack)

PLAN

Long term goal

BEL

Set of beliefs speaker has

SHARED

COM

Shared commitments (shared knowledge)

QUD

Question Under Discussion (Stack)

LU

The Latest Utterance

Assigning Focus

- QudTR:
 - A rule that looks at the question at the top of the stack and based on that assigns Rheme status to candidate answers and Theme status to the information found in the question.
- Background / Focus:
 - ComFB Rule: shared commitments
 - If there is shared commitment that semantically parallels something in the utterance the contrasting information will be assigned Focus
 - DomFB Rule: Domain (information found in question)
 - If ComFB doesn't find anything, contrasting information found in the same domain is assigned focus.

Example Dialogue

- (3) S1: Hello, how can I help you?
U1: What is the price of a flight from Paris to London on April fifth?
S2: What class did you have in mind?
U2: I don't know.
S3: BUSINESS class costs ONE THOUSAND euro.
ECONOMY class costs FIVE HUNDRED euro.

References

- Hirschberg, Julia. "Accent and discourse context: Assigning pitch accent in synthetic speech." *Proceedings of AAAI*. 1990.
- Kruijff-Korbayová, Ivana, et al. "Producing contextually appropriate intonation in an information-state based dialogue system." *Proceedings of the tenth conference on European chapter of the Association for Computational Linguistics-Volume 1*. Association for Computational Linguistics, 2003.
- Larsson, Staffan, et al. "GoDiS: an accommodating dialogue system." *Proceedings of the 2000 ANLP/NAACL Workshop on Conversational systems-Volume 3*. Association for Computational Linguistics, 2000.
- Prevost, Scott. "An information structural approach to spoken language generation." *Proceedings of the 34th annual meeting on Association for Computational Linguistics*. Association for Computational Linguistics, 1996.