Prosodic cues of focus marking in Hungarian: do they exist?

Katalin Mády

RIL HAS, Budapest

April 17, 2012

Overview

1. Introduction: intonation models, tone sequence model of intonation
2. Production experiment: focus realisation in controlled sentences.
3. Perception experiment: naturalness judgements for answers with matching and mismatching focus types.
4. Conclusions

Research areas of prosody

Traditional term: suprasegmentals. But: interaction with segmental features, thus misleading.

Prosody research is typically concerned with:

- intonation (macroprosody such as speech melody, prosodic phrasing),
- stress and accent (prominence on word and utterance level),
- speech rhythm and timing (sequences of strong and weak syllables and their duration),
- interaction between prosodic phrasing and segment realisations (microprosody such as phrase-final devoicing, phrase-initial strengthening etc.).

Intonation models

Intonation models have been developed based on following aspects:

- contour-based vs. tone-based,
- symbolic vs. parametric,
- perceptual vs. mathematic,
- single-layer vs. superpositional.

Models in Hungarian linguistics

Linguistically relevant: symbolic and perceptual models:

- Contour-based: goes back to the British school. It has been adapted to Hungarian by Fónagy & Magdics (1967). Extensive enhancement by Varga 1994, 2002.

- Tone-based: based on autosegmental phonology in which the assignment of tones to certain units is assumed. Several aspects have been investigated for Hungarian (Kornai & Kálmán 1988, Vogel & Kenesei, Varga 2008 etc.) No complete adaption so far.

- Metrical: degree of prominence is defined by a metrical grid that is based on the syntactic and information structure of the sentence (Hunyadi 2002).

Intonational phonology

- Other terms: tone sequence model, autosegmental-metrical phonology.
- Originates from Asian tone languages, adapted to Swedish by Bruce (1976), for English by Pierrehumbert (1980) and Ladd (1996). Relevance of tones were shown not only for lexical accent, but for sentence-level intonation.
- Instrumentalised via the intonation transcription system ToBI (**To**nal and **B**reak **I**ndices), Silverman et al. (1992).
- Growing importance in linguistics in the past 30 years.

Basic units of ToBI

Intonation contours (tunes) can be deconstructed into H(igh) or L(ow) target tones. They are assigned to pitch accents (*) and boundaries (%).
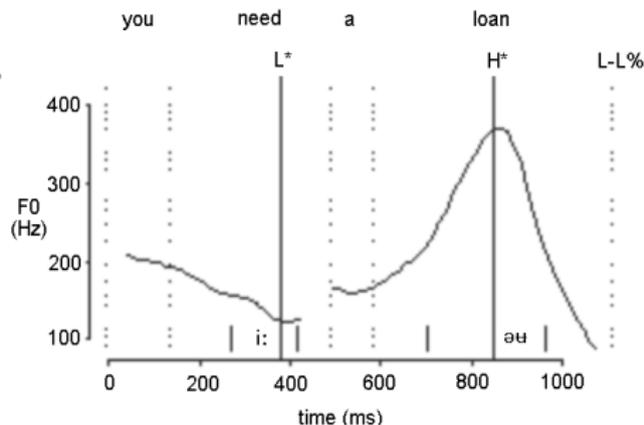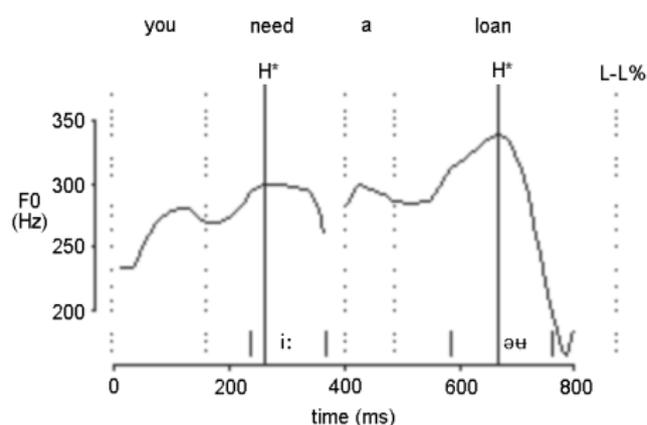
Pitch accent types

Changes in the fundamental frequency (f0):

- a local pitch maximum H* or minimum L*,
- potential combination with a preceding leading tone (L+) or a following trailing tone (+H).
- usual domain: syllable of primary lexical stress, typically a CVC sequence.
- Labels are phonological and not phonetic, i.e. only perceptually distinctive accent types are distinguished.

N.B. In Intonational Phonology, *stress* refers to lexical stress (a potential domain for prominence), whereas *accent* means prosodic prominence on the sentence level.

Pitch accent types

Pitch accent contour over CVC sequence:
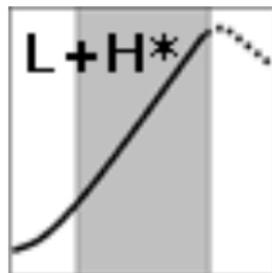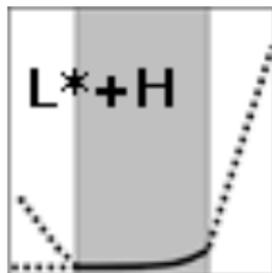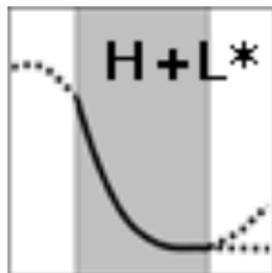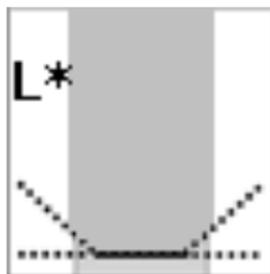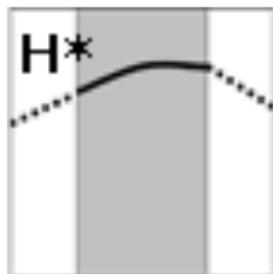
Boundary tones

ToBI is used for American English, but systems for many other languages have been developed since then.

- Original ToBI: highest prosodic level is intonational phrase (IP), and there is an intermediary phrase (ip) above word level, marked by a phrase accent (H- or L-).
- Boundary tones appear at the end of intonational phrase.
- Intermediary phrase and thus phrase accent is not relevant for all languages (e.g. Serbian).
- Dutch ToDI: highest level is utterance that can include one or more intonational phrases. The system also contains IP-initial boundary tones.

Acoustic parametrisation

Emphasis can be expressed by categorical and gradual means:

- higher f0 maximum,
- lower f0 minimum,
- larger f0 range,
- steeper slope,
- later peak aligment, e.g. an L+H* accent rather than an H* type one,
- preference of certain accent types over others (such as H*+L over H+L* in European Portuguese),
- longer segment duration.

N.B. According to most studies on Indo-European languages, intensity is relevant on word level, but not on sentence level.

Research questions

Goals of the present study:

- How is implicite or explicite contrast marked on the focussed element itself?

- What impact does focussing have on the topic?

- In what way does focussing influence the postverbal part of the sentence?

Material

Two target sentences:

sentence 1:

A	lányom	Németországba	ment	munkát keresni.
the	daughter-my	to Germany	went	job+ACC search-to

'My daughter went to Germany to look for a job.'

sentence 2:

Marianna	Máltán	nyaral	Mónival.
Marianna	Malta-on	has-holiday	Móni-with

'Marianna is having summer holiday with Móni on Malta.'

Potentially focussed words: *NÉmetországba* and *MÁLtán*.

Choice of nasal consonants in stressed syllable: voiced segments whose start and end can easily be detected in acoustic signal.

Focus types

Contexts preceding the sentence to read:

1. **broad focus:** (1): *Mi újság veletek? Mi van Erikával?* 'How is it going? How is Erika?'(2): *Mi újság nálatok? Hogy van a kolléganőd?* 'How are you (plural)? And how is your colleague?'

2. **narrow focus:** (1): *Végül hol keres állást Erika?* 'So where is Erika looking for a job?' (2): *Végül a kolléganőd és a barátnője hol töltik a szabadságukat?* 'So where did your colleague and her friend decide to go for holiday?'

3. **contrastive focus:** *Erika Pesten keres állást, ugye?* 'Erika is looking for a job in Pest, right?' (2): *A kolléganőd és a barátnője végül Svédországba mentek, ugye?* 'Your colleague and her friend went to Sweden in the end, didn't they?'

Recordings

- Speakers: 7 native speakers of Hungarian, all students (6 females, 1 male).
- Location: sound-treated room at the Department of Phonetics, RIL, HAS.
- Task: to read question in brackets silently first and then read target sentence aloud.
- Part of a larger experiment with other stimuli.
- 6 repetitions of each sentence in randomised order (mixed with distractors).
- Intended: head-mounted microphone. Reality: built-in microphone of our laptop. (Unforeseen incompatibility between recording software and the Mac OS.)

Prosodic cues on focus and topic

Categorical cues:

- distribution of pitch accent types,
- occurrence of deaccented content words.

Gradual cues (acoustic parameters) within word-initial CVC sequence:

- f0 maximum,
- f0 minimum,
- f0 range,
- f0 slope,
- aligment of target tone within accented syllable,
- duration of f0 rise or fall,
- duration of accented syllable.

Prosodic cues: postfocal part

Categorical cues:

- distribution of pitch accent types,
- occurrence of deaccented content words.

Gradual cues:

- overall f0 fall,
- overall f0 mean.

Details of measurements

$\rightarrow$ F0 was measured in semitones, because distances in Hertz express different pitch distances for male and female speakers. E.g. the interval between 100 Hz and 200 Hz is perceptually identical with the interval between 200 Hz and 400 Hz (one octave each).

$\rightarrow$ Duration of accented syllable was normalised to the entire sentence length.

$\rightarrow$ Intensity was not measured, because it is only reliable if the distance between source (speaker's lips) and microphone is constant (i.e. it requires a head-mounted microphone).

Statistical analysis

χ^2 test for comparison of accent type distributions: are the frequency patterns of accent types in broad, narrow and contrastive condition identical?

Mixed models for analysis of parametric measures where differences are computed independently of variation between speakers and sentences (multiple random factors possible).

Repeated-measures multivariate analysis of variances (RM MANOVA) for parametric measures: differences *within* each speaker are compared. As apposed to RM ANOVA, no sphericity (i.e. equality of variances for each condition) is required. Independent tests for two target sentences.

Accent distribution

Statistics: χ^2 test. Comparison of focus types for sentence 1 and 2.

Focus: large majority of H+L* accents in all contexts (around 80%).
About 8% H*+L accents and L* accents, 2% H* accents. Differences not
significant, sentences 1 and 2 do not differ.

Topic: Deaccented in ~70%, accents starting low (L*, L+H*, L*+H) in
about 10% of all cases each. Excepction: L* occurred only once before
contrastive focus. Differences not significant, sentences 1 and 2 do not
differ.

Verb: Verb was deaccented in 6 (93% of all) cases. Differences not
significant. Sentences 1 and 2 do not differ.

Postverbal element: sentence 1 and 2 differ: deaccentuation about 2/3 in
sentence 1, 1/3 in sentence 2. Accents mostly H+L*. Significantly less
deaccentuations in broad focus.

Prosodic cues of focus

No significant difference for f0 maximum, f0 minimum, f0 range, slope, segment duration, and duration of f0 fall.

Only the alignment of the f0 peak differed between broad and narrow focus on the one hand and contrastive focus on the other ($p = 0.03$).

High variation between speakers. Weak tendencies for lower f0 minimum (3 speakers), shorter duration for f0 falls (4 speakers).

Diverging tendencies for f0 range and segment duration (some speakers show the expected trend, some the opposite, others none).

Prosodic cues of focus

No significant difference for f0 maximum, f0 minimum, f0 range, slope, segment duration, and duration of f0 fall.

Only the alignment of the f0 peak differed between broad and narrow focus on the one hand and contrastive focus on the other ($p = 0.03$).

High variation between speakers. Weak tendencies for lower f0 minimum (3 speakers), shorter duration for f0 falls (4 speakers).

Diverging tendencies for f0 range and segment duration (some speakers show the expected trend, some the opposite, others none).

Peak alignment
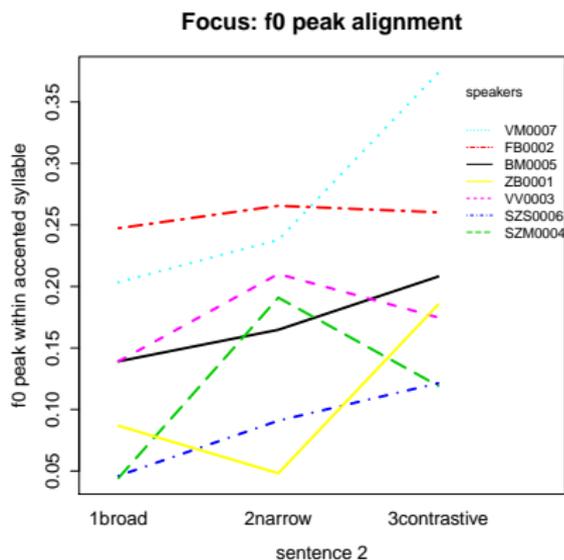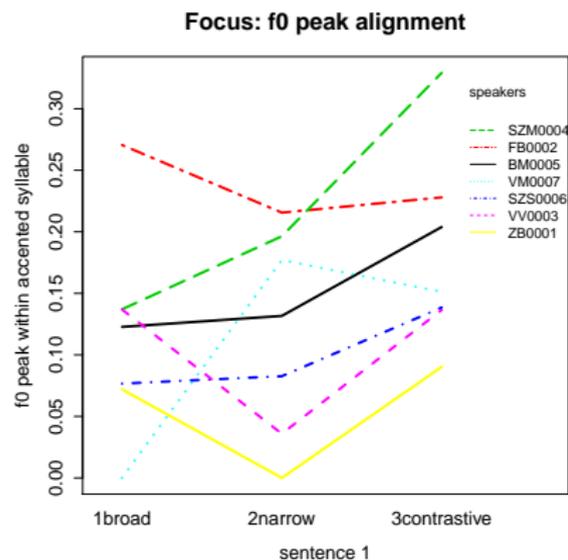
Separate analysis for sentences: difference only significant in sentence 2
(RM MANOVA for sentence 1: $p = 0.06$, sentence 2: $p = 0.02$).

Prosodic cues of topic

No significant difference for f0 maximum, f0 minimum, f0 range, slope, peak alignment. $p < 0.1$ for minimum alignment and duration of f0 rise, but only in sentence 2.

Highly significant difference for duration of stressed syllable, but only in sentence 2. (RM MANOVA for sentence 1: $p = 0.99$, sentence 2: $p = 0.007$.) Duration: narrow $<$ broad $<$ contrastive focus (in 5 speakers).

Topic: duration of stressed syllable
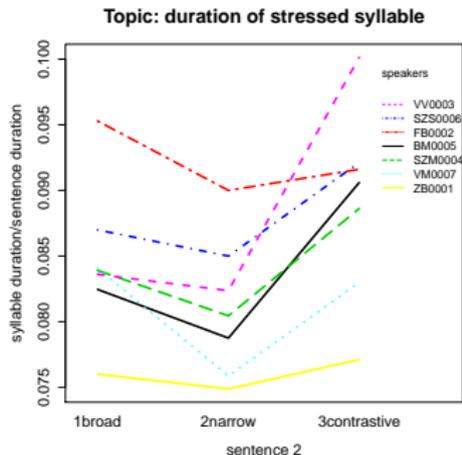
Prosodic cues of postverbal part

No significant difference for overall f0 range or for mean f0.

3 typical examples:

Summary and discussion of production experiment

- Pitch accents show a homogenous pattern: a rising (L*+H or L+H*) accent on topic, a falling one (mostly H+L*) on focus. Same observation by Mycock (2000) and Surányi, Ishihara, Schubö (in press).
- F0 and duration parameters do not typically show any tendency from broad over narrow to contrastive focus.
- Alignment of f0 maximum in focus contributes to focus distinction, but tendencies differ in the 2 target sentences.
- Duration of stressed syllable in topic is affected by focus type, but only in sentence 2, narrow < broad < contrastive focus ⇒ not quite the expected tendency.
- Verb is almost never accented in any sentence and condition (supports Kálmán & Nádasdy 1994).
- High inter-speaker variation for all parameters (also observed by Baumann et al. 2006).

Material

- 2 sentences, 3 conditions: broad, narrow and contrastive focus.
- 1 matching and 2 mismatching conditions, e.g. question requires answer with broad focus, answer has broad/narrow/contrastive focus.
- 18 combinations altogether: 1 matching and 2 mismatching answers to each question.
- 21 subjects, one occurence of each utterance.
- Praat's perception experiment modul, part of a larger experiment with other stimuli.
- Statistics: non-parametric Kruskal-Wallis test.

Setup

Instruction:
You will see a question on top of the screen, and you will listen to a reply. Please indicate how natural the answer sounds to the question.

5: very natural

1: very unnatural.

Stimuli

broad focus:
Mi újság veletek? Mi van Erikával?
Mi újság nálatok? Hogy van a kolléganőd?

narrow focus:
Végül hol keres állást Erika?
*Végül a kolléganőd és a barátnője hol töltik a szabadságukat?*

contrastive focus:
Erika Pesten keres állást, ugye?
*A kolléganőd és a barátnője végül Svédországba mentek, ugye?*

Broad focus

Question: *How are you? How is Erika/your colleague?*

Match: broad focus
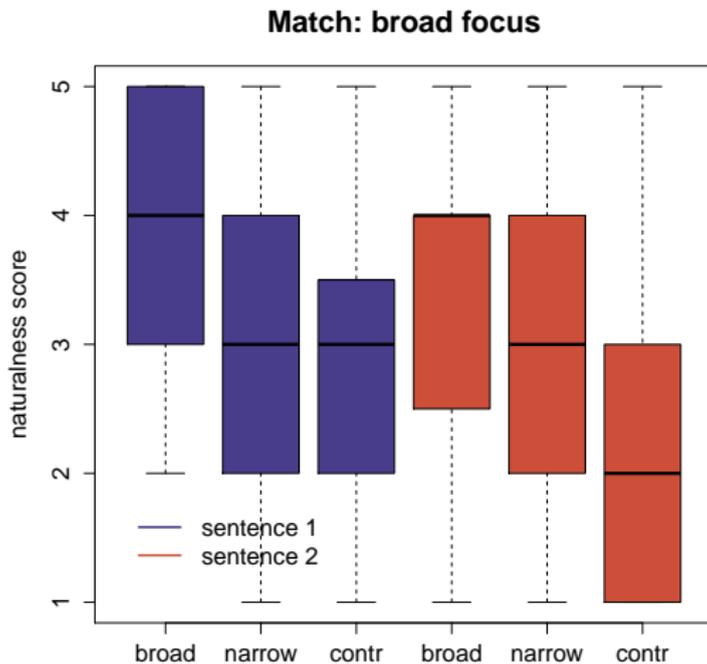
Narrow focus

Question: *Where is she looking for a job?/Where are they having holiday?*

Match: narrow focus
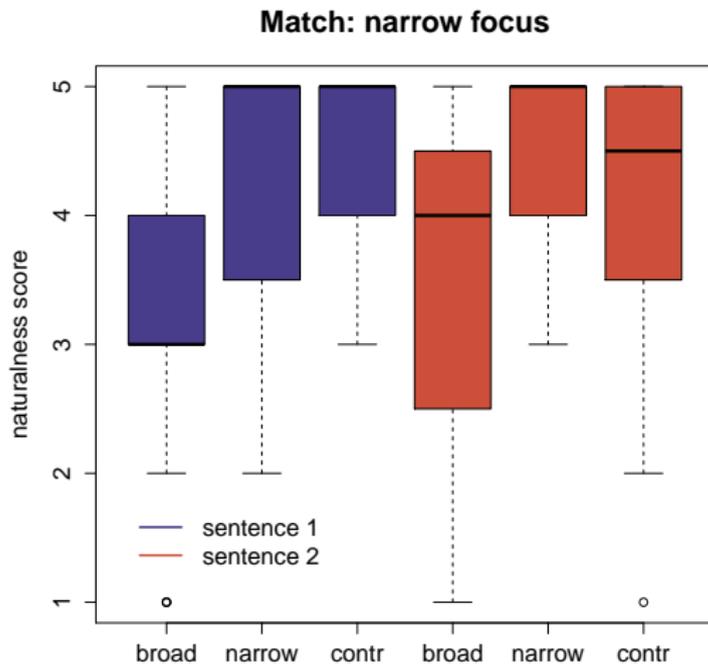
Contrastive focus

Question: *Is she looking for a job in Budapest?/Are they having holiday in Sweden?*

Match: contrastive focus
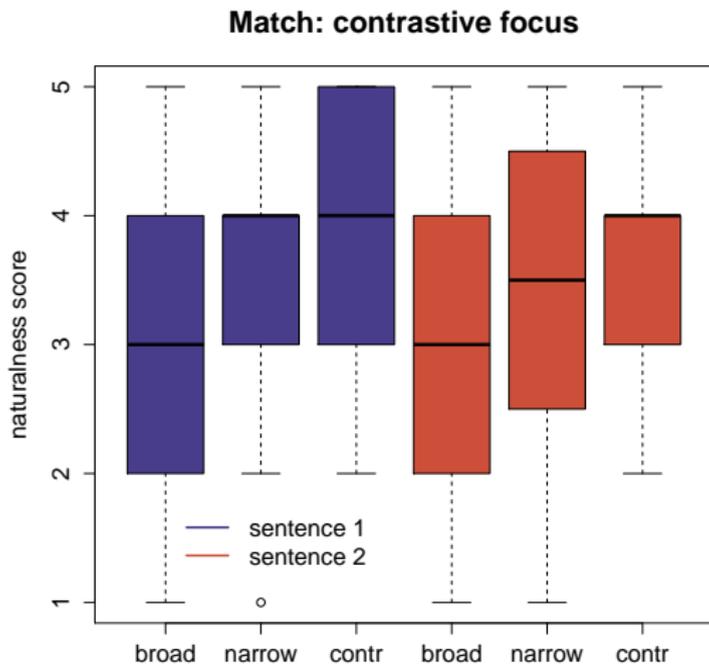
Interpretation

Hard job. . .

Differences between matching and mismatching answers are throughout significant.

But: Naturalness of matching answer is scored around 4, non-matching ones around 3. Is this difference really relevant?

Best naturalness scores for matching answers with narrow focus.

Occurrence of focus types in spontaneous speech

Information structure was labelled in a maptask dialogue of about 13 minutes.

Occurrence of focus types:

broad	91
narrow	99
contrastive	18

Contrastive focus far less frequent than narrow focus.

Broad focus is usually not an answer to *What happened?*, but to more specific questions (*Where should I go next?*), or it is part of longer monologues without a preceding question.

⇒ narrow focus is presumably the most natural type of question and answer pairs.

Discussion and conclusions

Prosodic cues of focus marking in Hungarian: do they exist?

- In sentence 1, none of the cues was clearly associated with focus (narrow, contrastive, or both).
- In sentence 2, (close-to) significant differences can be summarised in this way: f0 reaches its maximum earlier on the topic and remains high longer on the focus.
- Speakers differ wrt gradualness: peak is aligned latest in contrastive focus for some speakers, in narrow focus for others.
- Difference cannot be expressed by categorical means: an earlier rise on topic and a later fall on focus would be described as a sequence of L+H* H*+L accents as opposed to broad focus L*/L*+H H+L*, but this is not the case: the most frequent pattern is 0 H+L* (0: deaccented topic).

Future work

- Testing the perceptual relevance of peak alignment in acoustically manipulated and/or synthesised material.
- A comparison with European Portuguese data where categorical differences (H+L* vs. H*+L) were found.
- To capture the difference between sentence 1 and sentence 2: why are differences more prevalent in sentence 2?

Thanks to:

OTKA PD101050
Zsuzsanna Bárkányi
Mária Gósy and her colleagues
Balázs Surányi and his colleagues
Krisztián Tronka