Defining and Explaining Prosodic Phrasing: An Overview

Sun-Ah Jun
(jun@humnet.ucla.edu)

UCLA
Outline

• Defining Prosodic Phrasing
  – Syntax-based
  – Intonation-based

• Explaining Prosodic Phrasing
  – Domain of interpretation: sentence processing
  – Mark semantic properties: scope; wh- vs. yes/no-Q
  – deliver Information structure: focus

• Discussion and conclusion
What is prosody?

- Phonologically, it refers to:
  - suprasegmental categories or entities: tone, stress, length, pitch accent, etc.
  - a grouping of words in an utterance (=prosodic phrasing) and prominence relations within a group/among the groups

The child with asthma outgrew the condition last year.
(The child with asthma) (outgrew the condition last year)
(The child with asthma)(outgrew the condition)(last year)
Defining Prosodic Phrasing:
- prosodic phonology

• The relation between prosodic structure & syntactic structure

This is the cat that chased the rat that stole the cheese
(SPE)

syntactic structure ≠ prosodic structure

• Prosodic structure is distinct from the syntactic structure, but it is projected from syntactic structure, conforming to the properties of prosodic structure

=> A theory of prosodic structure or the Prosodic Hierarchy

Properties of prosodic structure

• (a) It consists of phonological categories of different types e.g. syllable (Syl: s), foot (Ft), prosodic word (PWd: w), Phonological Phrase (PPh), Intonational Phrase (I), utterance (U).

• (b) For any prosodic category, the sentence is exhaustively parsed into a sequence of such categories.

• (c) The prosodic categories are hierarchically ordered (in the order given above), and they are strictly organized into layers according to that hierarchy. => the Strict Layer Hypothesis

Strict Layer Hypothesis above the Word
How to define Phonological Phrase? --- in 80s

- **End-based theory** (Selkirk 1986)
  The Right/Left boundary of a prosodic constituent category C corresponds to the Right/Left boundary of morphosyntactic category X (lexical head or XP). The direction of edge is language specific.

- **Relation based theory** (Nespor and Vogel 1986)
  A Phonological phrase includes a syntactic head and elements on its non-recursive side that are not themselves syntactic heads.
  cf. c-command theory (Hayes 1989)
Ex. The domain of Italian RS rule (Nespor & Vogel 1986):

$$(\ldots V)\omega (Cv..)\omega$$

$$\Rightarrow V) (CCv$$

```
Ho visto tre colibri // molto scuri
```

$$\Rightarrow \{ \text{PPh} \} \{ \text{PPh} \} \{ \text{PPh} \}$$

‘I saw three very dark hummingbirds.’
Defining the Phonological Phrase
– since 90s

Syntax-phonology mapping is formalized in terms of ranked and violable constraints (OT: Prince and Smolensky 1993)

• Selkirk (1995)
  – *Constraints on prosodic domination* (=SLH)
    Layeredness, Headedness → inviolable
    Exhaustivity, Norrecursivity → violable
  – *Constraints on alignment of edges of constituents*
    ALIGN-XP, R/L: ALIGN (XP, R/L; P;R/L)
    (‘For each XP there is a P such that the right/left edge of XP coincides with the right/left edge of P’)

• Truckenbrodt (1996, 1999)
  – *WRAP-XP*
    : Each XP is contained in a phonological phrase.
So, the syntax-based approach ...

- tries to predict prosodic phrasing based on syntactic info.
- but cannot confirm if the domain defined is indeed a prosodic unit if no phonological rule exists related to the domain.

- Syntax is fixed but prosodic phrasing is not. Thus, syntax-based prosodic theories
  - have a limit in explaining optional and variable domains of phonological rules.
  - need to add constraints on various non-syntactic factors affecting prosodic phrasing.
Intonation based Prosodic Hierarchy => Intonational Phonology

- Define prosodic units based on intonation and duration
- Basic tone elements: High and Low
  - Intonation is a sequence of High and Low tonal targets. Syllables not associated with a tone get the surface f0 values by interpolation between the tonal targets.
- Tones are either “prominence marking” linked to a stressed syllable (=pitch accent) or “boundary marking” linked to the edge of a prosodic unit (=phrase accent or boundary tone).
- Korean: tones are marking the edge of a unit, small or large. No stress, so no pitch accent.
- Head prominence lg. (Eng) vs. edge prominence lg. (Kor)
5 types of pitch accent: $H^*$, $L^*$, $L+H^*$, $L^*+H$, $H+!H^*$
Prosodic structure defined by Intonation

Ex. *The child with asthma outgrew the condition last year*

One IP (L%), three ips (L-), 2 pitch accents in each ip.
Intonational Phonology of Korean

IP: Intonation Phrase
AP: Accentual Phrase
w: phonological word
s: syllable
T = H, when the AP-initial segment is aspirated or tense C or /h, s/;
Otherwise, T = L
%
: Intonation phrase boundary tone

AP tone pattern: medial two tones (H and L) may not be realized when AP is shorter than 4 syllables.
Example pitch track:
(Youngmi)(hates)(Younga). vs. (My older brother)(hates)(Younga)

<table>
<thead>
<tr>
<th>words</th>
<th>yEQminIN</th>
<th>yEQarIR</th>
<th>miwEheyo</th>
</tr>
</thead>
<tbody>
<tr>
<td>tones</td>
<td>L Ha</td>
<td>L Ha</td>
<td>L +H L%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>words</th>
<th>hyEQnimIN</th>
<th>yEQarIR</th>
<th>miwEheyo</th>
</tr>
</thead>
<tbody>
<tr>
<td>tones</td>
<td>H Ha</td>
<td>L Ha</td>
<td>L +H L%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Intonationally defined prosodic units can also be the domain of phonological rules (e.g., Hayes & Lahiri 1991, Jun 1993, 1998)

- Rules whose domain is an IP in Korean: Obstruent Nasalization, postlexical-Palatalization, and Spirantization

- Rules whose domain is an AP in Korean: Post Obstruent Tensing (POT) and Vowel Shortening
  - Ex. POT: /kuk/ ‘soup’ + /pap/ ‘rise’
    => [kuk\(p^*\)ap] ‘kind of soup’ (compound N)
  /tolmij\(\v\)k/ # /p\(\v\)\(\v\)iranik*\(\v\)a/ => [tolmij\(\v\)k \(p^*\)\(\v\)iranik*\(\v\)a]
    => [tolmij\(\v\)k] [p\(\v\)iranik*\(\v\)a]
  ‘seaweed’ ‘throw away’ => ‘Throw away seaweed’
Korean Post Obstruent Tensing rule
(Jun 1998)

Two prosodic phrases vs. One phrase
Mapping of prosodic units

Syntax-based

- Utterance
- Intonation phrase or major phrase
- Phonological phrase or minor phrase

Intonation-based

- Intonation phrase
- Intermediate phrase
- Accentual phrase
- Accentual phrase

Phonological word
Foot
Syllable / Mora
• **Phonetically**, the prosodic units are cued both at segmental and suprasegmental levels in both acoustic and articulatory dimensions, through the properties such as pitch, duration and intensity as well as by the movement of articulators.

  – Each prosodic unit/domain can be marked by

    • Domain final lengthening,
    • Domain initial strengthening,
    • f0 contour, amplitude, or duration, or
    • Pause at the boundary
Phrase final lengthening in French

French has Intonation Phr > Accentual Phr > Word

Data from Jun & Fougeron (2000); * = p < .05
Initial strengthening in Korean

VOT duration (ms) of /pʰ/ in 3 prosodic positions (Jun 1993)
Initial strengthening in French

Differences in Linguopalatal Contact of $[\text{n}]$ in French
as a function of Prosodic Position

Electro-palatogram (EPG) data from Keating et al. (2003)
Defining prosodic phrasing: Summary

- “Syntax-based” proposal aims at predicting prosodic phrasing, but cannot diagnose the domain if there is no phonological rule applying in the domain.

- “Intonation-based” proposal can diagnose/determine prosodic phrasing, but cannot predict prosodic phrasing.

- So, we need feedback from both approaches.
Explaining Prosodic Phrasing
= why do we prosodify an utterance?

In addition to serving as the domain of phonological rules, prosodic phrasing cues:

- syntactic/semantic grouping
  - Domain of interpretation in sentence processing
- semantic properties
  - Scope, wh- vs. yes/no-question
- information structure
  - focus
1. Domain of Interpretation

- Intonation phrase is the domain of meaning and interpretation. Items within the intonation phrase are interpreted together, but items across the phrase are not. (e.g., Marcus & Hindle 1990, Schafer 1997)

- Intonation phrase influences sentence processing (e.g., Kjelgaard & Speer, 1999, Schafer et al, 2000)

Ex. *When Roger leaves the house is dark.*

=> strong prosodic break after *leaves* facilitates ‘is’ reading.
Phrasing affects attachment
(Implicit Prosody Hypothesis, Fordor 1998, 2002)

• In NP1+NP2+RC structure, RC can attach high (modify NP1) or low (modify NP2).

  ex. *Someone shot the servant of the actress who was on the balcony* (Cuetos & Mitchell 1988)

• The length of RC influences attachment: the longer the RC, the more high attachment.

  EX. *Someone shot the servant of the actress who was on the balcony of the blue house*”

=> For high attachment, a prosodic phrase boundary comes before the attachee, RC.
2. Phrasing determines Scope

- English: “… not ... because ...” surface order
- [Because > Not] scope is achieved by
  - Phrase boundary between ‘not’ and ‘because’
  - pause is ok after ‘not’ phrase
- [Not > Because] scope is achieved by
  - Weaker or no phrase boundary between ‘not’ ‘because’
  - pause is bad after ‘not’ phrase
- Korean has a reverse surface order but has the same relation about scope and prosodic phrasing
Because > Not

John did not hit her (because she was yelling)’

<table>
<thead>
<tr>
<th>words</th>
<th>John</th>
<th>didn't</th>
<th>hit</th>
<th>Mary</th>
<th>because</th>
<th>she</th>
<th>was</th>
<th>yelling</th>
</tr>
</thead>
<tbody>
<tr>
<td>tones</td>
<td>H*</td>
<td>!H*</td>
<td>!H*</td>
<td>L-</td>
<td>H*</td>
<td>!H*</td>
<td></td>
<td>L-L%</td>
</tr>
</tbody>
</table>

Phrase boundary (pause is ok)
Not > Because

John hit her (for other reason)’

<table>
<thead>
<tr>
<th>words</th>
<th>John</th>
<th>didn't</th>
<th>hit</th>
<th>Mary</th>
<th>because</th>
<th>she</th>
<th>was</th>
<th>yelling</th>
</tr>
</thead>
<tbody>
<tr>
<td>tones</td>
<td>L+H*</td>
<td>(L-)</td>
<td></td>
<td>H*</td>
<td>L-H%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Weaker boundary pause is bad
Korean: Because > Not

‘I did not wait (because Youngi was crying)’

<table>
<thead>
<tr>
<th>words</th>
<th>gloss</th>
<th>tones</th>
</tr>
</thead>
<tbody>
<tr>
<td>j´Niga</td>
<td>u´s´</td>
<td>kida\idJin</td>
</tr>
<tr>
<td>Youngi-Nom.</td>
<td>cry-bec</td>
<td>wait-CT</td>
</tr>
<tr>
<td>L+H</td>
<td>Ha</td>
<td>L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L</td>
</tr>
</tbody>
</table>

AP boundary after ‘because’ pause is ok
Korean: Not > Because
‘I waited (for other reason)’

<table>
<thead>
<tr>
<th>words</th>
<th>j´Niga</th>
<th>u´s´</th>
<th>kida\idJin</th>
<th>anass´</th>
</tr>
</thead>
<tbody>
<tr>
<td>gloss</td>
<td>Youngi-Nom.</td>
<td>cry-Bec</td>
<td>wait-CT</td>
<td>not</td>
</tr>
<tr>
<td>tones</td>
<td>L+H</td>
<td>Ha</td>
<td>L</td>
<td>+H</td>
</tr>
</tbody>
</table>

One AP
No phrase boundary after ‘because’
Pause is bad
3. Wh-question vs. Y/N-question

- In both Korean and Japanese
  - Wh-question: no prosodic boundary after wh-word
  - Y/N-question: a prosodic boundary after wh-word
  (Maekawa 1991, Jun & Oh 1996)

Ex.  \[\text{nugu-riŋ moim issøjo?]} \\
     \text{who/anyone-with a meeting to have}

Wh-question: “With whom do you have a meeting?”
  => (nuguraŋ moim issøjo) -- see pitch track
Yes/No-question: “Do you have a meeting with anyone?”
  => (nuguraŋ) (moim issøjo) -- see pitch track
4. Focus and Prosodic phrasing

- In head-prominence languages (e.g., English, German, Greek), pitch accent marks a focused word (i.e., receives Nuclear pitch accent or sentence stress) and following words are deaccented.
- In edge-prominence languages, phrasing marks focus:
  - Focused word begins a new phrase, and following words are dephrased or produced in reduced pitch range
  - e.g., Korean, French, Japanese
  - Focused word ends a prosodic phrase, and following words are not affected - e.g., Chimwiini, Chichewa.
- Thus, phrasing in edge-prominence languages functions like pitch accent in head-prominence languages.
Deaccenting after focused word: English
Dephrasing after focused word: Korean

Younga tomorrow **evening** dumpling soup will eat

<table>
<thead>
<tr>
<th>words</th>
<th>Younga-ga</th>
<th>neil</th>
<th>jonyoge</th>
<th>mandukkug-ul</th>
<th>mongnunde</th>
</tr>
</thead>
<tbody>
<tr>
<td>tones</td>
<td>L Ha</td>
<td>LH+</td>
<td>L+ Ha</td>
<td>L +H</td>
<td>L+</td>
</tr>
</tbody>
</table>

| Hz            | 120       | 120  | 120     | 120           | 120       |

| ms            | 400       | 800  | 1200    | 1600          | 2000      |

**Focus**  Dephrased

**AP boundary**
Mary put the book in the BOX.

This sentence is a felicitous answer to the following wh-questions:

Where did Mary put the book? → [PP] focus
What did Mary do? → [VP] focus
What happened? → [S] focus

• Focus Projection principle
  (Selkirk 1984, 1995; Rochemont 1986, 1998)
  Focus on an internal argument can project its focus to its head. Focus on a head can project its focus to its head phrase recursively.
Problem: Not all internal arguments can project focus

*Ex. What did the butler do?*

*The butler [offered the president some COFFEE]_FOC*

*The butler [offered the PRESIDENT some coffee]_FOC*

- Valduvi & Engdahl (1996): Information Packaging Theory
  - argument structure and the type of argument should be considered.

- Chung, Kim & Sells (2003) adopted Valduvi & Engdahl’s theory and claimed that only theme argument may project focus to its head phrases in Korean.
Experiment: VP-focus and Narrow-focus in Korean (Kim et al. 2006, Jun et al. 2006)

i. Examined how VP-focus and NP(narrow)-focus in Korean are realized phonologically (phrasing) and acoustically (f0 and duration).

ii. Whether its realization is sensitive to the argument type (theme or oblique) and the word order.

iii. Whether Focus Projection theory can account for Korean data.
Method

• Data: Sentences containing an oblique argument (I.O/LOC) and a theme argument (D.O) in two different word orders.

Set 1: Subj+IO+DO+Verb vs. Subj+DO+IO+Verb
Set 2: Subj+LOC+DO+Verb vs. Subj+DO+LOC+Verb

➢ 8 sentences in each construction * 2 word order * 2 sets = 32 target sentences (plus 32 fillers)

➢ Wh-questions were used to elicit VP-focus and Narrow focus

➢ transcribed phrasing & tone patterns using K-ToBI (Jun 2000); measured f0 peak and duration of each word; also measured duration of word-initial syllable
Experiment procedures

1. A sentence was visually presented on the computer monitor.
2. Subjects read each sentence twice (elicited neutral focus) ex. *Youngho fed carrots to a horse.*
3. The sentence disappeared from the monitor.
4. Subjects heard a wh-question prompting the VP-focus, ex. *What did Youngho do?*
5. Subjects responded to the VP-focus question by supplying the sentence they had read.
6. Subjects heard a wh-question prompting narrow-focus on theme or goal. ex. *What did Youngho feed to a horse?*, or *What did Youngho feed carrots to?*
Results: Phrasing

(a) **Neutral**: each word formed one Accentual Phrase.

(b) **VP focus**: regardless of word order, an Intonation Phrase (IP) is often inserted before each argument within a VP, but more often before the VP-initial argument (70%). Dephrasing within VP was found very rarely.

(c) **Narrow focus**: post-focus string tends to be dephrased or produced in a reduced pitch range. An IP boundary is also inserted before (60%) or after (30%) or both before and after (20%) focused argument.
No. of IP boundaries in % at the left edge of each argument in VP focus condition

![Graph showing the percentage of IP boundaries at the left edge for different subjects and conditions.]
No. of IP boundaries (in %) before & after a focused argument in the **Narrow Focus** condition

![Bar chart showing the percentage of IP boundaries before and after a focused argument.](chart.png)
Results: f0 peak and duration

- **F0 peak**
  i. In **VP-focus**, f0 is raised for both arguments and sometimes raised even for the verb. The VP-initial argument was raised more.
  ii. In **narrow focus**, f0 is locally raised for the focused word and lowered for all words after focus.

- **Duration**
  i. In **VP-focus**, the word-initial syllable duration was significantly lengthened at VP-initial word.
  ii. In **narrow focus**, the duration was lengthened for narrowly focused word.

- **Sum**: For VP-initial word: Narrow focus ≈ VP-focus
  For VP-medial word: Narrow focus > VP-focus
F0 Peak: **DO(focus) + IO**
(data from males)

![Graph showing F0 peak data for different focus conditions](image-url)

- **Neutral**
- **VP-focus**
- **Narrow focus on Wd2**

Word 1  Word 2 (DO)  Word 3 (IO)  Word 4

NF/VP>Neut
F0 Peak: IO(focus) + DO
(data from males)
DO + IO(focus)

- Same f0 peak pattern for LOC data and female data
IO + DO (focus)

Graph showing comparison of neutral, VP-focus, and narrow focus on word 3 across different conditions.

- Neutral
- VP-focus
- Narrow focus on Wd3

Conditions:
- NF/VP > Neut
- NF > VP/Neut

Words:
- Word 1
- Word 2 (IO)
- Word 3 (DO)
- Word 4
Wd-initial Duration of Wd2 & Wd3 (ms)

- Neutral
- VP-foc
- Narrow foc

<table>
<thead>
<tr>
<th>Duration (ms)</th>
<th>IO (wd2)</th>
<th>DO (wd3)</th>
<th>DO (wd2)</th>
<th>IO (wd3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>120</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>140</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Duration (ms)</th>
<th>LOC(wd2)</th>
<th>DO(wd3)</th>
<th>DO(wd2)</th>
<th>LOC(wd3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>120</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>140</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Summary & Discussion of focus exp.

• In sum, no clear effect of argument type or word order in VP focus realization
  – In **VP-focus, the whole VP is affected**, and the VP-initial argument was more prominent than the VP-medial one.

• Implications to Selkirk’s and other syntax-based focus projection principles
  – Does the focus projection rule apply to Korean as in English?
    • Not really. In Korean, all arguments within the domain of focus should be prominent. If the VP-initial word is the only prominent one, it is interpreted as narrow focus on that word.
Why the difference?

- In English,
  1. In default prosody, the last word is most prominent.
  2. A focused word gets Nuclear Pitch Accent (NPA) and post-focus items are deaccented. Thus, a focused word (narrow or VP) is the rightmost prominent word in a phrase.
  3. Pre-focus words may or may not get pitch accent.

- In Korean,
  1. In default prosody, the first word is most prominent.
  2. A focused word is the most prominent and post-focus words are dephrased. Thus, a focused word (narrow) is the rightmost prominent word in a phrase.
  3. Pre-focus words still form an Accentual Phrase, thus prominent, though not as prominent as the focused word.
• Büring (2003, 2007) claims that we don’t need the Focus Projection theory. Focus prosody is determined by the interaction between the default prosody of the language and the ‘Focus-Prominence’ constraint (a constituent containing an F-mark is maximally prominent in its domain).

• Is there default prosody?
• In a head-prominence language, focus is realized with pitch accent (PA), thus a ‘head’ is easy to find. Then, do we know the edge of the domain whose head is PA? Is there a domain of PA?
Discussion: default prosody

- What is the ‘default’ prosody? Out-of-the-blue reading?
- Prosody (pitch accent and phrasing) in out-of-the-blue reading varies across speakers and even across repetitions of the same speaker.
  - optional pitch accent, phrasing and pause
- But it’s surely not random.
- Some prosodic features vary less than others.
  - pitch range reduction after focus
  - focused word is the most prominent word.
  - a large syntactic/semantic boundary is marked in prosody.
- In Japanese attachment data, the most common phrasing, (RC)(NP1 NP2), happened 65%. This grouping was achieved either at AP level or IP level. (Jun & Koike 2003)
Discussion: the domain of pitch accent

• Is there a domain of pitch accent in English?
  - Beckman & Edwards (1990) found no consistent evidence.
• Duration data suggest there might be one.
  - Wightman et al. (1992) found three levels higher than Word.
  - Shilman (2007) found that, with the same pitch accent distribution, the duration of *fugitive* below is:
    RC-VP condition > NP-VP condition > Adj-N condition.
    • Adj-N:  *Fleeing from the police, the* fugitive *pilot flew his plane into a mountainous area.*
    • NP-VP:  *Fleeing from the police, a* fugitive *piloted a plane into a mountainous area.*
    • RC-VP:  *Someone who had become a* fugitive *piloted a plane into a mountainous area.*
• RC-VP showed an ip boundary. Adj-N boundary is a word boundary. Is NP-VP boundary a boundary of accent domain?
Future research

• Need production and perception data of phrasing and focus. Also, need to explain optionality and variation.
  – German et al. (2006) found that, contrary to Selkirk (1984) and Schwarzchild (1999)’s predictions, English speakers prefer to place an accent on the NP or the verb than on the preposition. They propose to add a new constraint, *AccPrep and constraint-based stochastic model. This can explain inter- and intra-speaker variation in accent data.

• More research on intonation and semantics.
  Ex. Not all pitch accents mean focus, and not all focused words have pitch accent (cf. German et al. 2006, Büring’s ornamental PA, 2006, 2007)

• More languages (more than West Germanic)
• Need to improve the ways to elicit focus data
Conclusion

• Prosody (pitch accent and prosodic phrasing) is deeply involved in all linguistic representations.

• Studying interfaces between prosody and phonology/syntax/semantics will guide us to find a better way to predict and explain prosodic phrasing => more workshops like WPSI!

• Cross-linguistic data need to be examined.
Thank you!
References (selected)


