Comma intonation and the syntax-phonology interface

Shigeto Kawahara
kawahara@rci.rutgers.edu
Rutgers University, the State University of New Jersey

September 19th, 2008
1 Introduction

2 Experiment 1
   • Introduction
   • Method
   • MaP vs. MiP
   • IntP vs. MaP
   • Utterance vs. IntP
   • Discussion

3 Experiment II
   • Introduction
   • Method
   • Results
   • Discussion

4 General discussion
   • Summary
   • General discussion

5 References
   References75
Prosodic hierarchy

- Prosodic levels define domains of (i) phonological processes and distribution of boundary tones, and (ii) regulate phonetic articulations.
- Prosodic hierarchy above the Prosodic Word (Selkirk, 1986, *et seq*):

```
  Utterance
  | Intonational Phrase
  | Major Phrase (a.k.a. Phonological / Intermediate phrase)
  | Minor Phrase (a.k.a. Accentual Phrase)
```

**Figure**: Prosodic levels above Word
The issue

The issue: cross-linguistic variations in prosodic levels.

- Korean may not have an intermediate phrase (Jun, 1993) (cf. Jun, 2006).
- “[L]anguages vary in the number of prosodic units above the Word, ranging from one to three” (Jun, 2005).
The issue: cross-linguistic variations in prosodic levels.

- The ToBI framework is a general labeling guideline for intonation (Silverman, Beckman, Pitrelli, & Ostendorf, 1992, *et seq*).
- It attempts to find and define prosodic levels exclusively based on observations.
- The framework, however, is not a predictive theory—it admits language-specific gaps, and it in principle admits any number of levels.
But...

- Do we want to admit cross-linguistic variations in the number of prosodic levels?
- Especially if prosodic structures serve as “a general organizing principle for the phonology” (Hayes, 1995, 82).
From a UG perspective

An edge-based theory of the syntax-phonology interface postulates that phonological edges coincide with syntactic XP edges (Selkirk, 1986):

ALIGN(Syntactic phrase, Phonological phrase) (McCarthy & Prince, 1993).

- We do not want to give up the universality of prosodic hierarchy so easily.
- We can use syntax as a guide to find and define prosodic levels (Selkirk, 2005).
Japanese prosodic hierarchy

- This talk fills one gap in the Japanese prosodic hierarchy.
- Previous studies did not posit an Intonational Phrase, a level above a Major Phrase (MaP) (Pierrehumbert & Beckman, 1988).

Figure: Prosodic levels postulated for Japanese

<table>
<thead>
<tr>
<th>Utterance</th>
<th>Domain of downtrend (see below)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MaP (Intermediate Phrase)</td>
<td>1. Contains one accent</td>
</tr>
<tr>
<td>MiP (Accentual Phrase)</td>
<td>2. Initial LH rise</td>
</tr>
</tbody>
</table>
The lack of an IntP contradicts Selkirk’s (2005) proposal:

ALIGN(CommaP, IntP):
An IntP is aligned with a syntactic CommaP.

A CommaP includes an epithet, a parenthetical phrase, etc. (Potts, 2003).

I assume that a CommaP also includes a syntactic clause following Selkirk (2005), although a CommaP and a clause may form a larger natural class.

To simplify a bit, a [+comma] feature conveys an independent and complete speech act (Potts 2003; Selkirk 2005).
John, an English teacher, had lunch at an Indian restaurant:

1. John had lunch at an Indian restaurant.
2. John is an English teacher.

If the CommaP exists in Japanese, we would expect that Japanese has an IntP by virtue of Align(CommaP, IntP).

Rather than giving up the universality of the prosodic hierarchy, we should look at the intonation of CommaPs in Japanese.
The roadmap

Experiment 1: The experiment shows that in multiple-clause constructions (gapping and coordination), clause edges coincide with IntP boundaries (Kawahara & Shinya, 2008).

Experiment 2: The experiment shows that in nominal parenthetical constructions, the left—and right, for some speakers—edge of parentheticals coincides with IntP boundaries.

Conclusions:

1. The IntP does exist in the Japanese prosodic hierarchy.
2. The syntax-phonology mapping is mediated by general principles, not by construction-specific rules.
Rationale

- Since IntP boundaries are observed at clause edges (Ladd, 1980; Nespor & Vogel, 1986), we should look at multiple-clause constructions.
- Japanese gapping and coordination:

  Coordination:   S O V, S O V, S O V  
  Gapping:        S O (V), S O (V), S O V
<table>
<thead>
<tr>
<th>Muramatsu-wa</th>
<th>aomori-o</th>
<th>aruki</th>
</tr>
</thead>
<tbody>
<tr>
<td>Muramatsu-TOP</td>
<td>Aomori-ACC</td>
<td>walk</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Iwasaki-wa</th>
<th>Yamanashi-o</th>
<th>aruki</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iwasaki-TOP</td>
<td>Yamanashi-ACC</td>
<td>walk</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Morishitawa</th>
<th>Miyazaki-o</th>
<th>aruita.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morishita-TOP</td>
<td>Miyazaki-ACC</td>
<td>walk</td>
</tr>
</tbody>
</table>
Proposal and strategy

Proposed structure:

Utterance

IntP  IntP  IntP

Clause 1 Clause 2 Clause 3

Strategy

sentence [...]

clause [...]

VP(XP) [...]

Noun(X°) [...]

Level 4

Level 3

Level 2

Level 1

Utterance

IntP

MaP

MiP
Stimuli

- All words: 4-mora, accent on the 2nd mora (i.e. LHLL contour). The first two moras contain only sonorants.
- Two constituency lengths: S(hort) vs. L(ong). S=One noun; L=Two nouns (Noun1-GEN Noun2).
- To compare distinct syntactic edges, four pairs of gapping vs. coordination:
  - Short SUB Short OBJ (SS)
  - Short SUB Long OBJ (SL)
  - Long SUB Short OBJ (LS)
  - Short DAT Short ACC (Dative)
An example of a comparison

SS coordination:

Muramatsu-wa aomori-o... aruki...
Muramatsu-TOP Aomori-ACC walked

s[\text{NP}] v\text{P}[

LS gapping:

Morioka-no aniyome-wa Muramatsu-o...
Morioka-GEN brothers wife-TOP Muramatsu-ACC

s[\text{NP}[\text{NP}]] v\text{P}[]
Recording & Measurement

- Single-clause filler items mixed with multi-clause target items.
- Four female native speakers of Japanese, all from Tokyo and the surrounding areas.
- Two recording sessions on separate days. No minimal pairs in the same recording session. 6 repetitions in total. Order randomized between each repetition.
- $F_0$ valleys and peaks measured with PitchWorks (Scicon R&D).
Data analysis

- Pitch range normalization (Truckenbrodt, 2004).

\[
\text{NormalizedValue} = \frac{(\text{Mean}_{\text{Highest}}) - \text{OriginalValue}}{(\text{Mean}_{\text{Highest}}) - (\text{Mean}_{\text{Lowest}})}
\]

- Relativizes each pitch value to that speaker's pitch range.
- Reduces inter-speaker variability and allows us to avoid drastic Bonferronization.
Two major phonetic correlates

- The amount of pitch reset (Ladd, 1988)

![Figure: The amount of pitch reset](image1)

- The amount of initial rise (Ladd, 1990)

![Figure: The size of initial rise](image2)
**MaP vs. MiP**

 ALIGNL(\(XP, \text{MaP}\)): A MaP is left-aligned with a VP (\(XP\)) boundary (Selkirk, 2000; Selkirk & Tateishi, 1991).

*Figure: A schematic illustration of Selkirk & Tateishi 1991*
Further evidence from our data

An example of the comparison:

Muramatsu-wa aomori-o... aruki...
Muramatsu-TOP Aomori-ACC walked

Morioka-no aniyome-wa Muramatsu-o...
Morioka-GEN brothers wife-TOP Muramatsu-ACC

Other comparisons include dative gapping vs. SL gapping; dative coordination vs. SL coordination.
Example pitch tracks

**Figure:** Pitch reset across a VP boundary
Statistical comparison

**Figure:** A statistical comparison of downtrend

The amount of downtrend is smaller across a VP boundary ($t(97) = 5.30, p < .001$).
Pitch reset: IntP vs. MaP

Figure: The intonation of LS coordination
Pitch reset: IntP vs. MaP

Figure: A statistical comparison of pitch reset

Pitch reset is larger across a clause boundary than across a VP boundary ($t(197) = 5.36, p < .001$).
Initial rise: comparison

Figure: The intonation of LS coordination
The amount of initial rise

![Bar chart showing comparison of initial rises](image)

**Figure**: A statistical comparison of initial rises

Initial rises are larger clause-initially than VP-initially ($t(198) = 2.89, p < .01$).
Obligatory pauses

Figure: A spectrogram of SS coordination

Pauses appear clause-finally, but not (necessarily) phrase-finally.
Distribution of creaky vowels

Figure: Spectrograms of subject-final and clause-final vowels
Distribution of creaky vowels

<table>
<thead>
<tr>
<th>Speaker J</th>
<th>MaP-final</th>
<th>IntP-final</th>
<th>Spaker N</th>
<th>MaP-final</th>
<th>IntP-final</th>
</tr>
</thead>
<tbody>
<tr>
<td>Creaky</td>
<td>0</td>
<td>276</td>
<td>Creaky</td>
<td>4</td>
<td>241</td>
</tr>
<tr>
<td>Non-Creaky</td>
<td>288</td>
<td>12</td>
<td>Non-Creaky</td>
<td>308</td>
<td>71</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Speaker R</th>
<th>MaP-final</th>
<th>IntP-final</th>
<th>Spaker Y</th>
<th>MaP-final</th>
<th>IntP-final</th>
</tr>
</thead>
<tbody>
<tr>
<td>Creaky</td>
<td>0</td>
<td>268</td>
<td>Creaky</td>
<td>0</td>
<td>287</td>
</tr>
<tr>
<td>Non-Creaky</td>
<td>306</td>
<td>38</td>
<td>Non-Creaky</td>
<td>288</td>
<td>1</td>
</tr>
</tbody>
</table>
Final lowering

Figure: Final lowering: Comparing SS gapping and SS coordination
Figure: Final lowering quantified: SS & SL
Final lowering

**Figure:** Final lowering quantified: LS & Dative
Summary

Properties of an IntP:

1. Large pitch reset.
2. Large initial rise.
3. Creaky vowels and pause.
4. Final lowering.
A possible reaction: What you mean by an IntP is an Utterance, as in (a). I will now show Utterances and IntPs are different, as in (b).

Figure: Two possible models
Initial rise

Figure: SS coordination—a comparison of initial rises

The sentence-initial rises are larger than sentence-internal rises.
The sentence-initial rises are larger than sentence-internal rises (C1-C2:
\[ t(394) = 13.73, p < .001; \] C2-C3: \[ t(394) = 4.99, p < .001; \] The difference b/w the differences:
\[ t(394) = 5.86, p < .001). \]
Clause-initial Hs

Figure: The heights of clause initial Hs

H1-H2: $t(394) = 26.67, p < .001$; H2-H3: $t(394) = 5.96, p < .001$; The difference b/w the differences: $t(394) = 13.74, p < .001$.
Clause-initial Hs

- The height of clause-initial Hs generally decline.
- An Utterance, not an IntP, is the domain of declination.
- The slope between H1 and H2 is steeper than the slope between H2 and H3.
- An Utterance-initial H is particularly higher.
Verb rises

- We have observed that clause-final Hs undergo final lowering.
- We have also observed that utterance-initial Hs undergo an extra boost.

- We thus may expect extra lowering utterance-finally.
- We can look at initial rises in verbs in coordination sentences to test this prediction.
Verb rises

Figure: Verb rises

The first and second rises involve a small rise or a flat contour, whereas the third rise is actually a fall.
Verb rises

C1-C2: $t(41) = 3.83, p < .001$; C2-C3: $t(41) = 8.96, p < .001$; The difference b/w the differences: $t(41) = 3.59, p < .001$.

Figure: Verb rises
An Utterance has a boundary H% tone on its penultimate MiP whereas an IntP does not (cf. Venditti, Maekawa, & Beckman, 2008).

This boundary H% tone exists for all speakers, but the realization of this H tone differs between speakers.
Speakers N and Y show an extra peak on the case particle of the final object (penultimate MiP of the Utterance).
For Speakers J and R, the object peaks of the final clause are higher than or as high as the object peaks of the second clause.
Between-speaker comparison

Figure: A comparison of clause-final peaks of all speakers

1 Pitch reset: MiP < MaP < IntP.
2 Initial rise: (MiP <) MaP < IntP < Utterance.
3 Creakiness and pauses at IntP-edges.
4 Final lowering: Present IntP-finally, stronger Utterance-finally.
5 Final H% tone: Utterance-finally.
Discussion

- When we look at multi-clause sentences, we do observe a level between a MaP and an Utterance, contra previous theories.
- A syntactic clause (CommaP) corresponds to an IntP in Japanese.
- Guided by syntax and a general syntax-phonology mapping principle, we discovered the new prosodic level.
ALIGN(Syntactic Cat, Phono Cat) constraints are perhaps not construction-specific.

**Restrictiveness argument:** It results in a more restrictive theory if constraints can refer to syntactic category types (e.g. $X^0, XP$), not sentence types (e.g. gapping, coordination).

**Empirical argument:** The same types of syntactic categories pattern prosodically in the same way.
If alignment constraints are not construction-specific, then \( \text{ALIGN}(\text{CommaP}, \text{IntP}) \) should not care what type of CommaP it is.

We should look at some other constructions (which are not clauses) → Nominal parentheticals.
NP-Parentheticals in Japanese

- A NP parenthetical is a side-remark, often introduced by connective phrases like *iwaba* ‘so-called’, *iwayuru* ‘so-called’, and *tsumari* ‘that is’.
- Aemono \( p_a r [iwayuru gomayogoshi] \) ‘aemono, so-called mixed salad’
- These phrases do not need to contain tense, hence NP/nominal parentheticals.
Stimuli

- Sets of sentences that have the same number of moras and accent placement. The first and the second words are always accented.
- The second words followed either a noun boundary, a VP boundary or a parenthetical boundary.

1. \( \text{XXXX}_{\text{Noun}}[\text{XX’XX}] \) (Noun boundary condition)
2. \( \text{X’XXX}_{\text{VP}}[\text{XX’XX}] \) (VP boundary condition)
3. \( \text{X’XXX}_{\text{Par}}[\text{XX’XX}] \) (Parenthetical condition)

- Four sets with different lexical items, using words with only sonorants in the first two moras.
### An example set

#### Noun boundary condition

<table>
<thead>
<tr>
<th>Candy seller’s screen kitchen broke</th>
</tr>
</thead>
<tbody>
<tr>
<td>ame’ya-no ami’do-no daidokoro-ga kowarema’shita</td>
</tr>
</tbody>
</table>

#### VP boundary condition

<table>
<thead>
<tr>
<th>Leak seller-NOM candy-seller’s suger-mix-ACC took out.</th>
</tr>
</thead>
<tbody>
<tr>
<td>nira’ya-ga ame’ya-no zarameae-o mochida’shita</td>
</tr>
</tbody>
</table>

#### Parenthetical boundary condition

<table>
<thead>
<tr>
<th>Brother’s wife so-called sister-in-law took</th>
</tr>
</thead>
<tbody>
<tr>
<td>ani’yome iwa’yuru girinoyome-o tsuretei’itta</td>
</tr>
</tbody>
</table>

Kawahara  (Rutgers) Comma intonation September 19th, 2008 53 / 74
Pitch reset 1

Figure: Pitch reset across N^0
Pitch reset across VP

Figure: Pitch reset across VP
Pitch reset 3

Figure: Pitch reset across a parenthetical boundary
Pitch reset by set

Figure: Pitch reset by set
Pitch reset by speaker

Figure: Pitch reset by speaker

Kawahara (Rutgers)
Summary

- In all of the lexical sets for all speakers, the pitch reset showed the following pattern: Across Par > Across VP (XP) > Across Noun (X^0).

- Planned contrast analyses:
  - Noun-VP \( (t(318) = 5.38, p < .003) \)
  - VP-Par \( (t(318) = 4.51, p < .001) \).
Initial rise

**Figure:** Initial rises at various positions

Higher prosodic categories show larger initial rises.
Initial rise

Figure: Initial rise by set
Initial rise

![Bar graphs showing initial rise by speaker.](image)

**Figure:** Initial rise by speaker
Discussion

- All speakers had the smallest rise noun-initially (NP vs. VP; \(t(318) = 5.38, p < .001\)).
- Speakers M and R had larger rise in Par-initial positions than VP-initial positions \((t(158) = 4.63, p < .001)\).
- Speakers K and O had larger rise in VP-initial positions than Par-initial positions \((t(158) = 10.1, p < .001)\).
- Nevertheless these speakers showed a stronger reset across a parenthetical boundary than across a VP-boundary: these speakers boost parenthetical initial %Ls as well.
Interim summary

- The left edges of nominal parentheticals show properties different from those at VP edges or noun edges, both in terms of pitch reset and initial rise.
- Left edges of nouns ($X^0$), VPs (XP), and nominal parentheticals correspond to distinct prosodic levels, which we may call MiP, MaP and IntP.
Figure: Pause: Speaker K
Speakers O and R always inserted a pause before the parentheticals, and variably after the parentheticals.
Creakiness: Speaker K

Figure: The distribution of creaky vowels: Speaker K
Creakiness: Speaker M

Figure: The distribution of creaky vowels: Speaker M
Prosodic structure: Speaker K

Figure: Prosodic structure of Speaker K
Prosodic structure: Speaker M

Figure: Prosodic structure of Speaker M
Discussion

- All speakers show evidence for ALIGN-L(CommaP, IntP).
- Speaker M—and sometimes Speakers O and R—in addition show evidence for ALIGN-R(CommaP, IntP).

- The second type of prosodic parsing is an example of both edge alignment (de Lacy, 2003; Samek-Lodovici, 1998).
Figure: General model of syntax-phonology mapping
Discussion

- The effects of this phrasing on processing (Hirotani, 2005; Warren, 1999)?
- Evidence from spontaneous utterances?
- Other phonetic cues for different levels?
  - Domain-edge articulatory strengthening (Fougeron & Keating, 1997).
Summary

- I defended the universality of the prosodic hierarchy i.e. the putative lack of the IntP in Japanese is due to the fact that previous studies focused on lower categories.

- Syntax-phonology mapping is demonstrably constrained by universal principles.

- Without the guide of syntax and ALIGN(CommaP, IntP) (Selkirk 2005), we would not have known what to look for.

- Theories provide guidelines for what to look in phonetic studies. Phonetic researches can contribute to theoretical debates.
References


Ito, J., & Mester, A. (2008). The onset of the prosodic word. In S. Parker...


