# Towards a Head-Driven account of information structure: evidence from Japanese and Korean

Yo Sato

University of Hertfordshire, UK

Paris LingLunch, Dec 2012

# Outline

## The main points

- Head-driven conception of information structure for Japanese and Korean, where predicates determine which arguments to be focused / topicalised: but why?

## The main points

- Head-driven conception of information structure for Japanese and Korean, where predicates determine which arguments to be focused / topicalised: but why?
- This aspect affects the syntax of 'particles' (case-markers and topic-marker) in these languages, i.e.

## The main points

- Head-driven conception of information structure for Japanese and Korean, where predicates determine which arguments to be focused / topicalised: but why?
- This aspect affects the syntax of 'particles' (case-markers and topic-marker) in these languages, i.e.
  - which particles to use, in particular, case-marker or topic-marker (they cannot co-occur)
  - whether to use one at all (elide it or not)

## The main points

- Head-driven conception of information structure for Japanese and Korean, where predicates determine which arguments to be focused / topicalised: but why?
- This aspect affects the syntax of 'particles' (case-markers and topic-marker) in these languages, i.e.
    - which particles to use, in particular, case-marker or topic-marker (they cannot co-occur)
    - whether to use one at all (elide it or not)
- In this talk I focus on focushood and case-markers and their ellipsis, with evidence that shows:

## The main points

- Head-driven conception of information structure for Japanese and Korean, where predicates determine which arguments to be focused / topicalised: but why?
- This aspect affects the syntax of 'particles' (case-markers and topic-marker) in these languages, i.e.
  - which particles to use, in particular, case-marker or topic-marker (they cannot co-occur)
  - whether to use one at all (elide it or not)
- In this talk I focus on focushood and case-markers and their ellipsis, with evidence that shows:
  - the likelihood that information structure considerations are a key contributive factor to case-marking
  - the overt/zero marking varies from predicate to predicate

## The main points

- Head-driven conception of information structure for Japanese and Korean, where predicates determine which arguments to be focused / topicalised: but why?
- This aspect affects the syntax of 'particles' (case-markers and topic-marker) in these languages, i.e.
  - which particles to use, in particular, case-marker or topic-marker (they cannot co-occur)
  - whether to use one at all (elide it or not)
- In this talk I focus on focushood and case-markers and their ellipsis, with evidence that shows:
  - the likelihood that information structure considerations are a key contributive factor to case-marking
  - the overt/zero marking varies from predicate to predicate
  - and cross-dialectal and cross-linguistic variations

## The main points

- Head-driven conception of information structure for Japanese and Korean, where predicates determine which arguments to be focused / topicalised: but why?
- This aspect affects the syntax of 'particles' (case-markers and topic-marker) in these languages, i.e.
  - which particles to use, in particular, case-marker or topic-marker (they cannot co-occur)
  - whether to use one at all (elide it or not)
- In this talk I focus on focushood and case-markers and their ellipsis, with evidence that shows:
  - the likelihood that information structure considerations are a key contributive factor to case-marking
  - the overt/zero marking varies from predicate to predicate
  - and cross-dialectal and cross-linguistic variations
- I also indicate, on the way, that this study is rich with general implications (i.e. not just about Japanese and Korean!)

## Some Japanese data

- Japanese: 'Tarô is laughing'

| | | |
|---|---|---|
| Tarô-ga | waratteru | (case-marker) |
| Tarô-wa | waratteru | (topic-marker) |
| Tarô-$\phi$ | waratteru | (ellipsis) |

## Some Japanese data

- Generally, an NP seems to require case-marking to be focused:
  '(Who's laughing?) Taro is laughing.'

|          |           |                |
|----------|-----------|----------------|
| Tarô-ga  | waratteru | (case-marker)  |
| Tarô-wa  | waratteru | (topic-marker) |
| Tarô-$\phi$ | waratteru | (ellipsis)  |

## Some Japanese data

- Generally, an NP seems to require case-marking to be focused: '(Who's laughing?) Taro is laughing.'

  | | | | |
  |---|---|---|---|
  | ✓ | Tarô-ga | waratteru | (case-marker) |
  | | Tarô-wa | waratteru | (topic-marker) |
  | | Tarô-$\phi$ | waratteru | (ellipsis) |

## Some Japanese data

- Generally, an NP seems to require case-marking to be focused:
  '(Who's laughing?) Taro is laughing.'

  | ✓ | Tarô-ga | waratteru | (case-marker) |
  |-----|---------|-----------|---------------|
  | ??? | Tarô-wa | waratteru | (topic-marker) |
  | | Tarô-$\phi$ | waratteru | (ellipsis) |

## Some Japanese data

- Generally, an NP seems to require case-marking to be focused:
  '(Who's laughing?) Taro is laughing.'

  | | | |
  |---|---|---|
  | ✓ | Tarô-ga | waratteru (case-marker) |
  | ??? | Tarô-wa | waratteru (topic-marker) |
  | ?? | Tarô-$\phi$ | waratteru (ellipsis) |

## Some Japanese data

- Generally, an NP seems to require case-marking to be focused:
  '(Who's laughing?) Taro is laughing.'

  | ✓ | Tarô-ga | waratteru | (case-marker) |
  |---|---------|-----------|---------------|
  | ??? | Tarô-wa | waratteru | (topic-marker) |
  | ?? | Tarô-$\phi$ | waratteru | (ellipsis) |

- However, focushood also seems to depend on:

## Some Japanese data

- Generally, an NP seems to require case-marking to be focused:
  '(Who's laughing?) Taro is laughing.'

  |      |           |            |                 |
  |------|-----------|------------|-----------------|
  | ✓    | Tarô-ga   | waratteru  | (case-marker)   |
  | ???  | Tarô-wa   | waratteru  | (topic-marker)  |
  | ??   | Tarô-$\phi$ | waratteru | (ellipsis)     |

- However, focushood also seems to depend on:
  - case, as accusative seems exempt from the above constraint
    E.g. '(What did you lose?) I lost my wallet'

    Saifu-{✓o/✓$\phi$/?wa} nakushichatta

## Some Japanese data

- Generally, an NP seems to require case-marking to be focused:
  '(Who's laughing?) Taro is laughing.'

  | ✓ | Tarô-ga | waratteru | (case-marker) |
  |---|---------|-----------|---------------|
  | ??? | Tarô-wa | waratteru | (topic-marker) |
  | ?? | Tarô-$\phi$ | waratteru | (ellipsis) |

- However, focushood also seems to depend on:
  - case, as accusative seems exempt from the above constraint
    E.g. '(What did you lose?) I lost my wallet'

    Saifu-{✓o/✓$\phi$/?wa} nakushichatta

  - more subtle argument structure difference, e.g. unaccusative
    E.g. '(Who came out?) Taro came out.'

    Taro-{✓ga/✓$\phi$/?wa} detekita

## Taking stock...

- Two curious facts:
  - case-markers, though not dedicated for focus marking, have effects on focus, even without context (context cannot override their focal effect)

# Taking stock...

- Two curious facts:
    - case-markers, though not dedicated for focus marking, have effects on focus, even without context (context cannot override their focal effect)
    - but not always: the effects vary depending on the argument structure of a head

## Taking stock...

- Two curious facts:
    - case-markers, though not dedicated for focus marking, have effects on focus, even without context (context cannot override their focal effect)
    - but not always: the effects vary depending on the argument structure of a head
- Case-marking appears to be motivated by information structure consideration, but only indirectly, via argument structure

## Taking stock...

- Two curious facts:
  - case-markers, though not dedicated for focus marking, have effects on focus, even without context (context cannot override their focal effect)
  - but not always: the effects vary depending on the argument structure of a head
- Case-marking appears to be motivated by information structure consideration, but only indirectly, via argument structure
- Focus articulation specification for verbs provides the required flexibility

## Taking stock...

- Two curious facts:
  - case-markers, though not dedicated for focus marking, have effects on focus, even without context (context cannot override their focal effect)
  - but not always: the effects vary depending on the argument structure of a head
- Case-marking appears to be motivated by information structure consideration, but only indirectly, via argument structure
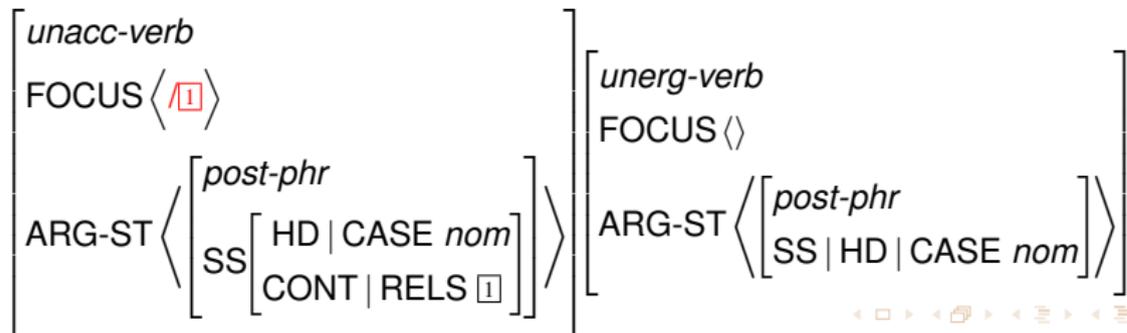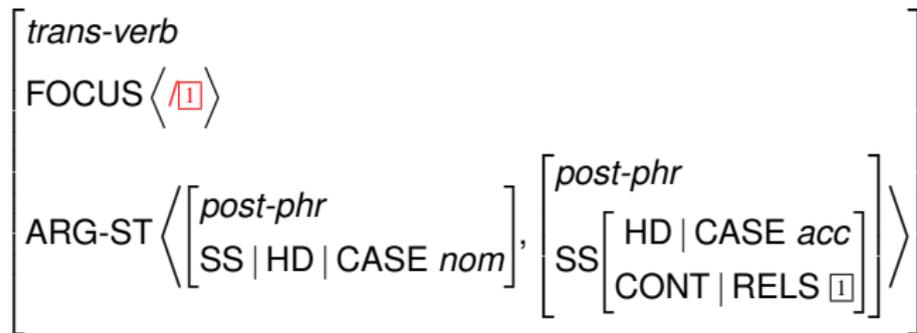- Focus articulation specification for verbs provides the required flexibility
- We will see some HPSG formalisation first, then move on to some relevant cross-linguistic data to show general implications

# Outline

## Implicit focus for a verb

- Implicit focus: can be focused without explicit focalisation (but the focalisation is optional)

$$
\begin{bmatrix}
\textit{trans-verb} \\
\text{FOCUS} \left\langle /\boxed{1} \right\rangle \\
\\
\text{ARG-ST} \left\langle \begin{bmatrix} \textit{post-phr} \\ \text{SS} \mid \text{HD} \mid \text{CASE } \textit{nom} \end{bmatrix}, \begin{bmatrix} \textit{post-phr} \\ \text{SS} \begin{bmatrix} \text{HD} \mid \text{CASE } \textit{acc} \\ \text{CONT} \mid \text{RELS } \boxed{1} \end{bmatrix} \end{bmatrix} \right\rangle
\end{bmatrix}
$$

$$
\begin{bmatrix}
\textit{unacc-verb} \\
\text{FOCUS} \left\langle /\boxed{1} \right\rangle \\
\\
\text{ARG-ST} \left\langle \begin{bmatrix} \textit{post-phr} \\ \text{SS} \begin{bmatrix} \text{HD} \mid \text{CASE } \textit{nom} \\ \text{CONT} \mid \text{RELS } \boxed{1} \end{bmatrix} \end{bmatrix} \right\rangle
\end{bmatrix}
\begin{bmatrix}
\textit{unerg-verb} \\
\text{FOCUS} \left\langle \right\rangle \\
\\
\text{ARG-ST} \left\langle \begin{bmatrix} \textit{post-phr} \\ \text{SS} \mid \text{HD} \mid \text{CASE } \textit{nom} \end{bmatrix} \right\rangle
\end{bmatrix}
$$

# How focus articulation is shaped

- Focus projection: vertically percolated, conditioned on:
    - case-marking
    - verb's implicit focus specification
- Base case: mother inherits head daughter's focus value (list) (cf. Selkirk's 'vertical' projection rule)

# How focus articulation is shaped

- Focus projection: vertically percolated, conditioned on:
    - case-marking
    - verb's implicit focus specification
- Base case: mother inherits head daughter's focus value (list) (cf. Selkirk's 'vertical' projection rule)
- Special case: if not an implicit focus, overt case-marking takes the role of explicit focus-marking
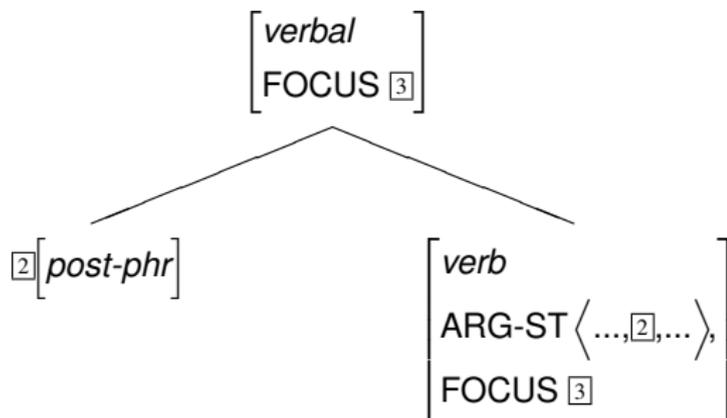
# How focus articulation is shaped

- Focus projection: vertically percolated, conditioned on:
    - case-marking
    - verb's implicit focus specification
- Base case: mother inherits head daughter's focus value (list) (cf. Selkirk's 'vertical' projection rule)
- Special case: if not an implicit focus, overt case-marking takes the role of explicit focus-marking
- Corollary: otherwise (for implicit foci) case-marking does not do anything

## Focus Projection

- Base:

$$
\begin{bmatrix} verbal \\ \text{FOCUS } \boxed{3} \end{bmatrix}
$$

$\boxed{2}\begin{bmatrix} post\text{-}phr \end{bmatrix}$
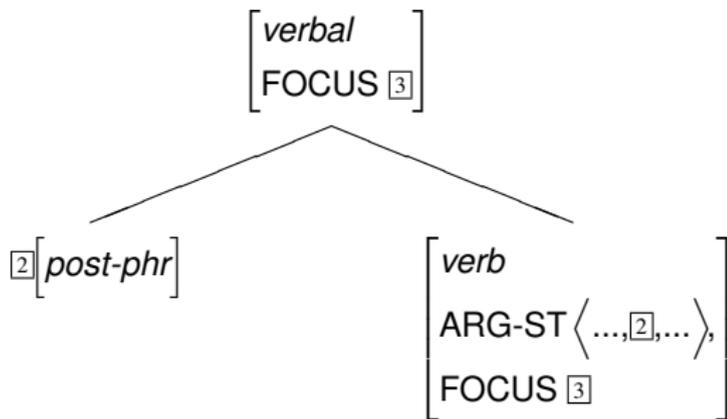$\qquad\begin{bmatrix} verb \\ \text{ARG-ST } \left\langle ...,\boxed{2},... \right\rangle, \\ \text{FOCUS } \boxed{3} \end{bmatrix}$
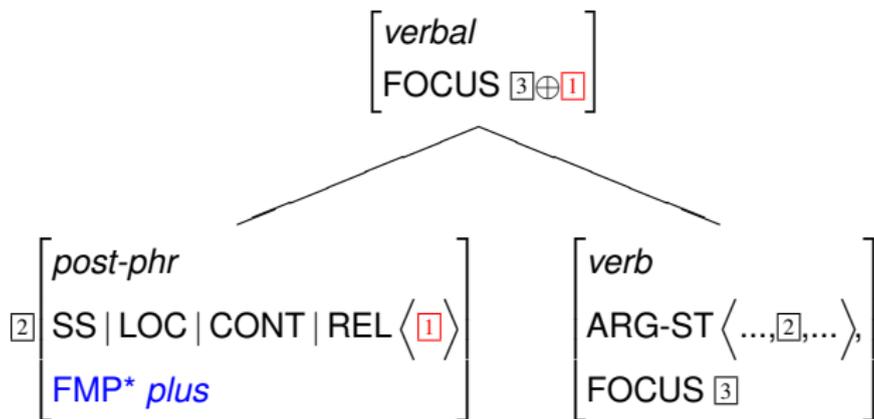
## Focus Projection

- Base:

## Focus Projection

- If the NP (content) is not an implicit focus in the focus list, in this case $/\boxed{1} \notin \boxed{3}$
- AND it is overtly marked, then the NP is added to the focus list

$$\begin{bmatrix} verbal \\ \text{FOCUS } \boxed{3} \oplus \boxed{1} \end{bmatrix}$$

$$\boxed{2}\begin{bmatrix} post\text{-}phr \\ \text{SS | LOC | CONT | REL } \left\langle \boxed{1} \right\rangle \\ \text{FMP* } plus \end{bmatrix} \qquad \begin{bmatrix} verb \\ \text{ARG-ST } \left\langle ...,\boxed{2},... \right\rangle, \\ \text{FOCUS } \boxed{3} \end{bmatrix}$$
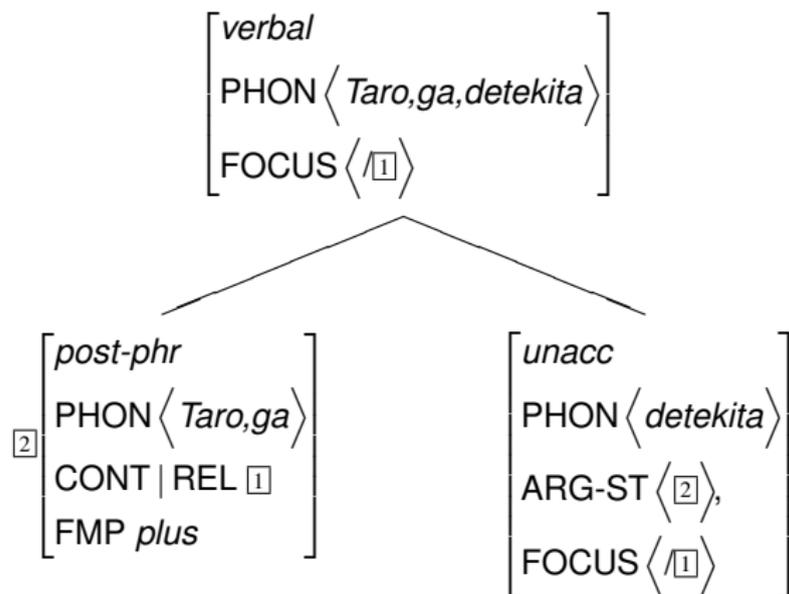
\* FMP = focus marking potential, here simply whether there is an overt marker or not

Sato

# Focus and case-marking: unaccusative / unergative

- To recap: unaccusative case-marking ellipsis OK, unergative not in a focused context, e.g.:
  ✓ '*Taro detekita.*' vs. × '*Taro waratta.*'
- We are saying this is because (only) for the latter *ga* is obligatory to receive focus

# Focus and case-marking: unaccusative

With overt marking

$$
\begin{bmatrix}
\textit{verbal} \\
\text{PHON} \left\langle \textit{Taro,ga,detekita} \right\rangle \\
\text{FOCUS} \left\langle \boxed{1} \right\rangle
\end{bmatrix}
$$

$$
\boxed{2}
\begin{bmatrix}
\textit{post-phr} \\
\text{PHON} \left\langle \textit{Taro,ga} \right\rangle \\
\text{CONT} \mid \text{REL} \ \boxed{1} \\
\text{FMP } \textit{plus}
\end{bmatrix}
\qquad
\begin{bmatrix}
\textit{unacc} \\
\text{PHON} \left\langle \textit{detekita} \right\rangle \\
\text{ARG-ST} \left\langle \boxed{2} \right\rangle, \\
\text{FOCUS} \left\langle \boxed{1} \right\rangle
\end{bmatrix}
$$

# Focus and case-marking: unaccusative

Marking dropped

$$
\begin{bmatrix}
\textit{verbal} \\
\text{PHON} \left\langle \textit{Taro,detekita} \right\rangle \\
\text{FOCUS} \left\langle /\boxed{1} \right\rangle
\end{bmatrix}
$$

$$
\boxed{2}\begin{bmatrix}
\textit{post-phr} \\
\text{PHON} \left\langle \textit{Taro} \right\rangle \\
\text{CONT} \mid \text{REL} \ \boxed{1} \\
\text{FMP } \textit{minus}
\end{bmatrix}
\qquad
\begin{bmatrix}
\textit{unacc} \\
\text{PHON} \left\langle \textit{detekita} \right\rangle \\
\text{ARG-ST} \left\langle \boxed{2} \right\rangle, \\
\text{FOCUS} \left\langle /\boxed{1} \right\rangle
\end{bmatrix}
$$

Sato
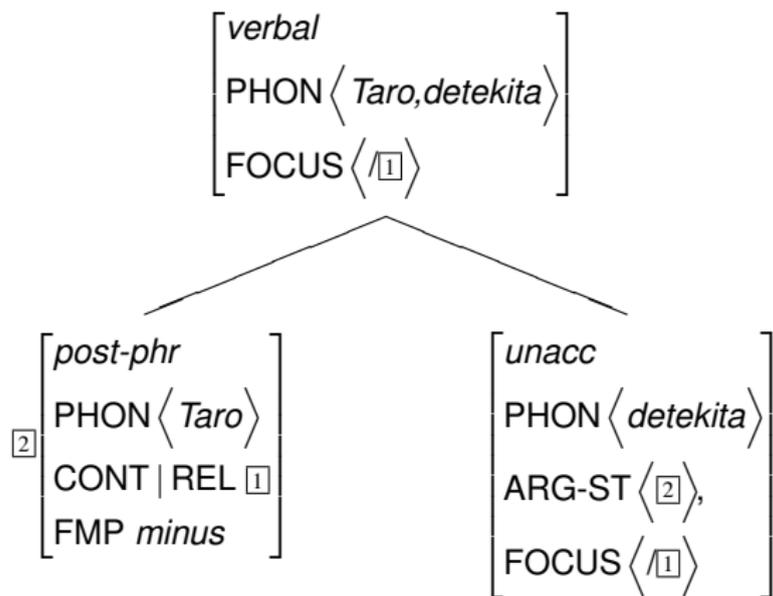
# Focus and case-marking: unergative

With overt marking

$$
\begin{bmatrix}
\textit{verbal} \\
\text{PHON} \left\langle \textit{Taro,ga,waratteru} \right\rangle \\
\text{FOCUS} \left\langle \boxed{1} \right\rangle
\end{bmatrix}
$$

$$
\boxed{2}
\begin{bmatrix}
\textit{post-phr} \\
\text{PHON} \left\langle \textit{Taro,ga} \right\rangle \\
\text{CONT} \mid \text{REL} \left\langle \boxed{1} \right\rangle \\
\text{FMP } \textit{plus}
\end{bmatrix}
\qquad
\begin{bmatrix}
\textit{unerg} \\
\text{PHON} \left\langle \textit{waratteru} \right\rangle \\
\text{ARG-ST} \left\langle \boxed{2} \right\rangle, \\
\text{FOCUS} \left\langle \right\rangle
\end{bmatrix}
$$
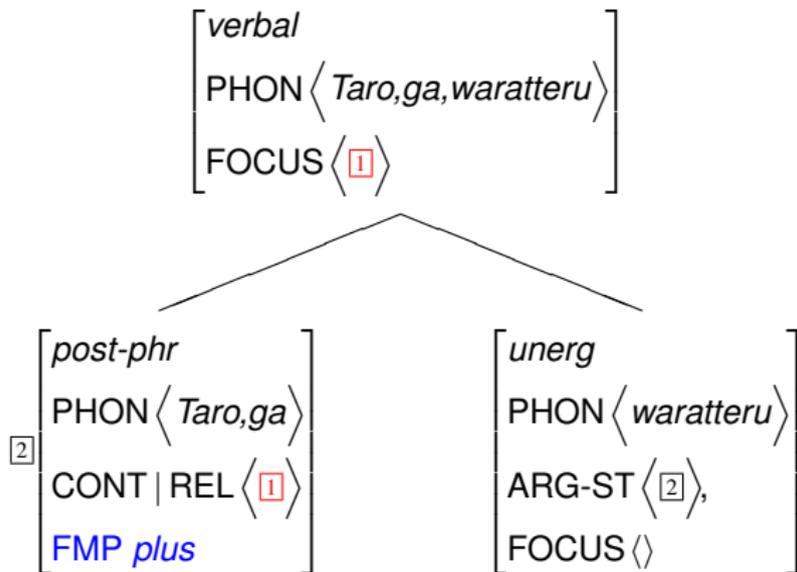
# Focus and case-marking: unergative

Marking dropped
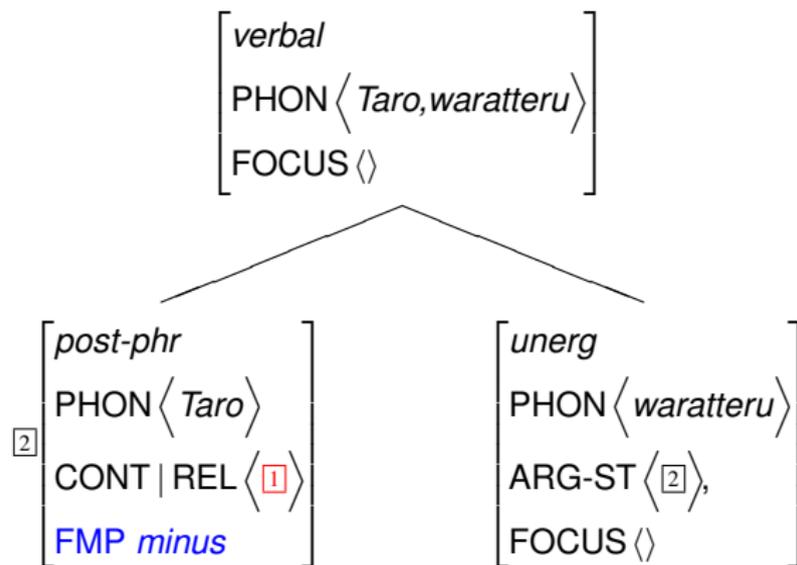
$$
\begin{bmatrix}
\textit{verbal} \\
\text{PHON}\ \langle \textit{Taro,waratteru} \rangle \\
\text{FOCUS}\ \langle\rangle
\end{bmatrix}
$$

$$
\boxed{2}\begin{bmatrix}
\textit{post-phr} \\
\text{PHON}\ \langle \textit{Taro} \rangle \\
\text{CONT}\ |\ \text{REL}\ \langle \boxed{1} \rangle \\
\text{FMP}\ \textit{minus}
\end{bmatrix}
\qquad
\begin{bmatrix}
\textit{unerg} \\
\text{PHON}\ \langle \textit{waratteru} \rangle \\
\text{ARG-ST}\ \langle \boxed{2} \rangle, \\
\text{FOCUS}\ \langle\rangle
\end{bmatrix}
$$

# Focus and case-marking: transitive

- Object (acc) argument specified as an implicit focus, so eventually:

## Focus and case-marking: transitive

- Object (acc) argument specified as an implicit focus, so eventually:
- E.g. '*Taro*-{*ga/ϕ*} *Jiro*-{*o/ϕ*} *ijimeteru*'
  where 1*taro_rel* and 2*jiro_rel*

  | Taro-ga | Jiro-o | ijimeteru | }FOC⟨ 1, /2 ⟩ |
  |---------|--------|-----------|---------------|
  | Taro-ga | Jiro   | ijimeteru |               |

  | Taro | Jiro-o | ijimeteru | }FOC⟨ /2 ⟩ |
  |------|--------|-----------|------------|
  | Taro | Jiro   | ijimeteru |            |

## In a nutshell...

- A phrase, if it is an implicit focus (as specified by its head) does not have to be case-marked to be focused, whereas

## In a nutshell...

- A phrase, if it is an implicit focus (as specified by its head) does not have to be case-marked to be focused, whereas
- if not an implicit focus, must be case-marked to be focused

# Outline

# Narrow (argument) focus

- The same mechanism applicable to the predicate itself

## Narrow (argument) focus

- The same mechanism applicable to the predicate itself

$$
\begin{bmatrix}
\textit{imp-focus-not-on-pred} \\
\text{FOCUS} \langle \rangle \\
\text{ARG-ST} \left\langle \begin{bmatrix} \textit{post-phr} \\ \text{SS} \,|\, \text{HD} \,|\, \text{CASE } \textit{nom} \end{bmatrix} \right\rangle
\end{bmatrix}
$$

# Narrow (argument) focus

- The same mechanism applicable to the predicate itself

$$
\begin{bmatrix}
\textit{imp-focus-not-on-pred} \\
\text{FOCUS} \langle \rangle \\
\text{ARG-ST} \left\langle \begin{bmatrix} \textit{post-phr} \\ \text{SS} | \text{HD} | \text{CASE } \textit{nom} \end{bmatrix} \right\rangle
\end{bmatrix}
\quad \boxed{1}
\begin{bmatrix}
\textit{imp-focus-on-pred} \\
\text{FOCUS} \left\langle \cancel{\boxed{1}} \right\rangle \\
\text{ARG-ST} \left\langle \begin{bmatrix} \textit{post-phr} \\ \text{SS} | \text{HD} | \text{CASE } \textit{nom} \end{bmatrix} \right\rangle
\end{bmatrix}
$$
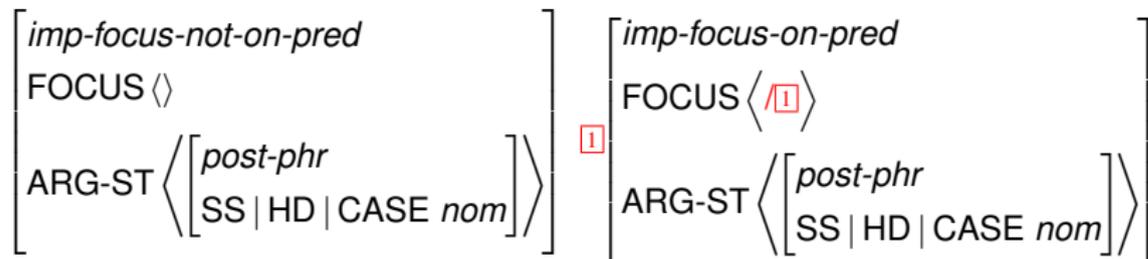
# Narrow (argument) focus

- The same mechanism applicable to the predicate itself

- This will allow for expressing 'sentence/argument focushood'

# Narrow (argument) focus

- The same mechanism applicable to the predicate itself

- This will allow for expressing 'sentence/argument focushood'
- Now, Japanese has another curious phenomenon: argument focus for case-marked subject NPs in statives

# Narrow (argument) focus

- The same mechanism applicable to the predicate itself


- This will allow for expressing 'sentence/argument focushood'
- Now, Japanese has another curious phenomenon: argument focus for case-marked subject NPs in statives
  - Copula (desu: 'be'):
    Watashi-no namae-ga Satô-desu
    (surface: 'My name is Sato', but implies: 'It is Satô that is my name')

# Narrow (argument) focus

- The same mechanism applicable to the predicate itself

- This will allow for expressing 'sentence/argument focushood'
- Now, Japanese has another curious phenomenon: argument focus for case-marked subject NPs in statives
  - Copula (desu: 'be'):
    Watashi-no namae-ga Satô-desu
    (surface: 'My name is Sato', but implies: 'It is Satô that is my name')
  - Adjective (yasashii: 'kind'):
    Tarô-ga yasashii-desu ('It is Tarô who is kind')

# Narrow (argument) focus

- The same mechanism applicable to the predicate itself

- This will allow for expressing 'sentence/argument focushood'
- Now, Japanese has another curious phenomenon: argument focus for case-marked subject NPs in statives
    - Copula (desu: 'be'):
      Watashi-no namae-ga Satô-desu
      (surface: 'My name is Sato', but implies: 'It is Satô that is my name')
    - Adjective (yasashii: 'kind'):
      Tarô-ga yasashii-desu ('It is Tarô who is kind')
    - Stative verb (shitteiru: 'know'):
      Tarô-ga sono-koto shittemasu ('It is Taro who knows about that')'
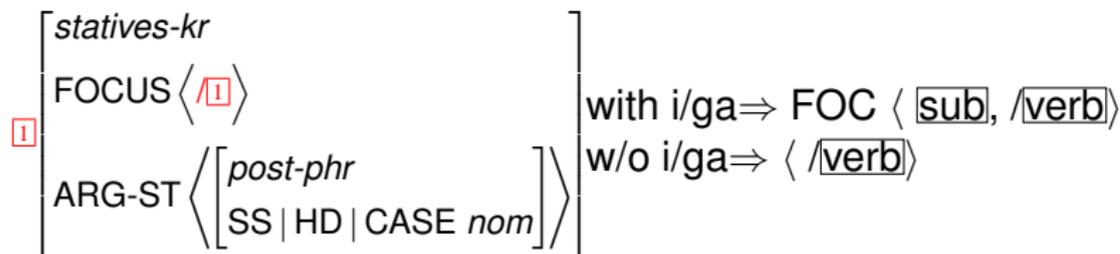
# Contrast with Korean

- This 'argument focus' effect of a case-marker appears absent in Korean
    - Che irŭm-i Sato-yeyo ('My name is Sato')
    - Hyeonsu-ga tajeon-heyo ('Hyeonsu is kind')
    - Hyeonsu-ga kŭ-geot chal al-ayo ('Hyeonsu knows about that')
- The contrast can be expressed by specifying/not specifying the predicate as an implicit focus

Sato

# Contrast with Korean

- This 'argument focus' effect of a case-marker appears absent in Korean
    - Che irŭm-i Sato-yeyo ('My name is Sato')
    - Hyeonsu-ga tajeon-heyo ('Hyeonsu is kind')
    - Hyeonsu-ga kŭ-geot chal al-ayo ('Hyeonsu knows about that')
- The contrast can be expressed by specifying/not specifying the predicate as an implicit focus

Eventual focus articulation
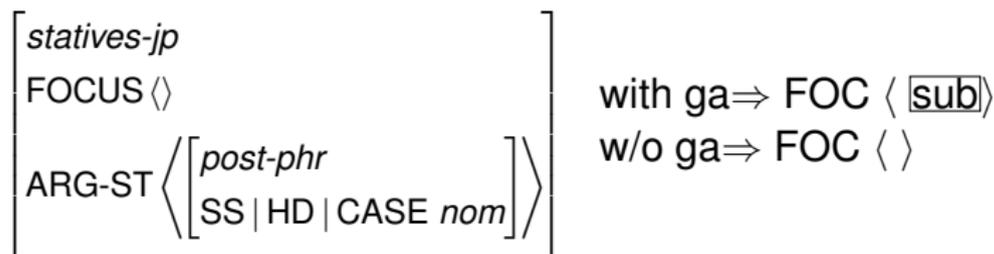
$$
\begin{bmatrix}
\textit{statives-jp} \\
\text{FOCUS} \langle \rangle \\
\text{ARG-ST} \left\langle \begin{bmatrix} \textit{post-phr} \\ \text{SS | HD | CASE } \textit{nom} \end{bmatrix} \right\rangle
\end{bmatrix}
$$

with ga $\Rightarrow$ FOC $\langle \boxed{\text{sub}} \rangle$
w/o ga $\Rightarrow$ FOC $\langle \ \rangle$

## Contrast with Korean

- This 'argument focus' effect of a case-marker appears absent in Korean

    - Che irŭm-i Sato-yeyo ('My name is Sato')
    - Hyeonsu-ga tajeon-heyo ('Hyeonsu is kind')
    - Hyeonsu-ga kŭ-geot chal al-ayo ('Hyeonsu knows about that')

- The contrast can be expressed by specifying/not specifying the predicate as an implicit focus

$$
\boxed{1}
\begin{bmatrix}
\textit{statives-kr} \\
\text{FOCUS} \left\langle /\boxed{1} \right\rangle \\
\text{ARG-ST} \left\langle
\begin{bmatrix}
\textit{post-phr} \\
\text{SS | HD | CASE } \textit{nom}
\end{bmatrix}
\right\rangle
\end{bmatrix}
\quad
\begin{array}{l}
\text{with i/ga} \Rightarrow \text{FOC } \langle \; \underline{\text{sub}}, \; /\underline{\text{verb}} \rangle \\
\text{w/o i/ga} \Rightarrow \langle \; /\underline{\text{verb}} \rangle
\end{array}
$$

## Dialectal contrast

- Kansai Japanese (spoken in Ôsaka & Kyôto) speakers report narrow focus effect for *ga* in unaccusatives as well for subject case marking
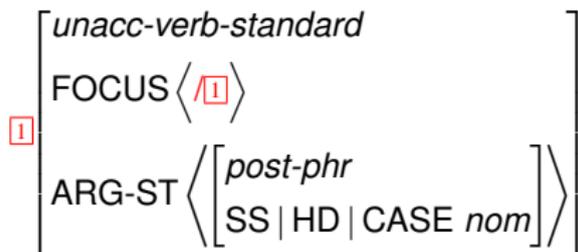
## Dialectal contrast

- Kansai Japanese (spoken in Ôsaka & Kyôto) speakers report narrow focus effect for *ga* in unaccusatives as well for subject case marking
  (What was it (say, an open day) like?) '– Many people turned up.'
  Tokyo: *Takusan hito*-{*ga/ϕ*} *kitetayo.*
  Kansai: *Takusan hito*-{*??ga/ϕ*} *kiteyatten.*

# Dialectal contrast

- Kansai Japanese (spoken in Ôsaka & Kyôto) speakers report narrow focus effect for *ga* in unaccusatives as well for subject case marking
  (What was it (say, an open day) like?) '– Many people turned up.'
  Tokyo: *Takusan hito-{ga/ϕ} kitetayo.*
  Kansai: *Takusan hito-{ ??ga/ϕ} kiteyatten.*
- This can be modelled by specifying the verb itself as an implicit focus for Tokyo Japanese but not in the Kansai variety, i.e.
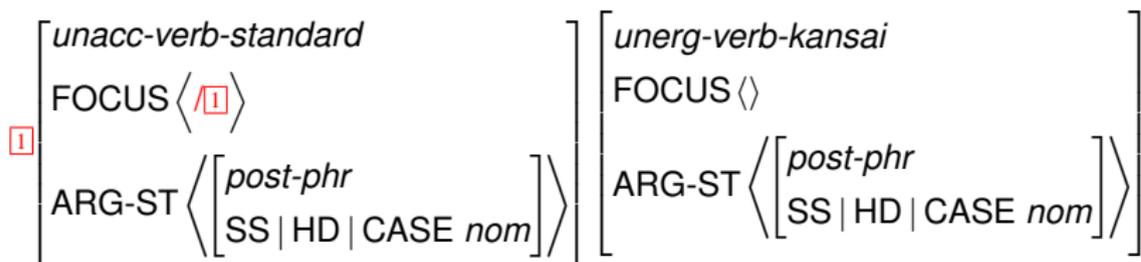
## Dialectal contrast

- Kansai Japanese (spoken in Ôsaka & Kyôto) speakers report narrow focus effect for *ga* in unaccusatives as well for subject case marking
  (What was it (say, an open day) like?) '– Many people turned up.'
  Tokyo: *Takusan hito-*{*ga/$\phi$*} *kitetayo.*
  Kansai: *Takusan hito-*{*??ga/$\phi$*} *kiteyatten.*
- This can be modelled by specifying the verb itself as an implicit focus for Tokyo Japanese but not in the Kansai variety, i.e.
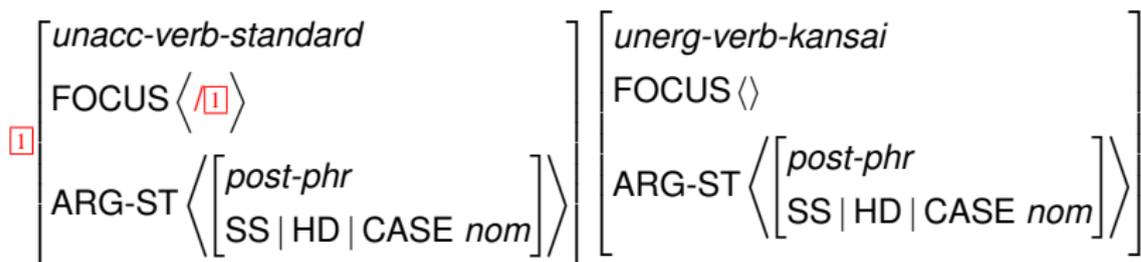
## Dialectal contrast

- Kansai Japanese (spoken in Ôsaka & Kyôto) speakers report narrow focus effect for *ga* in unaccusatives as well for subject case marking
  (What was it (say, an open day) like?) '– Many people turned up.'
  Tokyo: *Takusan hito-*{*ga/ϕ*} *kitetayo.*
  Kansai: *Takusan hito-*{*??ga/ϕ*} *kiteyatten.*
- This can be modelled by specifying the verb itself as an implicit focus for Tokyo Japanese but not in the Kansai variety, i.e.

$$
\boxed{1}
\begin{bmatrix}
\textit{unacc-verb-standard} \\
\text{FOCUS} \left\langle \boxed{/1} \right\rangle \\
\text{ARG-ST} \left\langle
\begin{bmatrix}
\textit{post-phr} \\
\text{SS} \mid \text{HD} \mid \text{CASE } \textit{nom}
\end{bmatrix}
\right\rangle
\end{bmatrix}
$$

## Dialectal contrast

- Kansai Japanese (spoken in Ôsaka & Kyôto) speakers report narrow focus effect for *ga* in unaccusatives as well for subject case marking
  (What was it (say, an open day) like?) '– Many people turned up.'
  Tokyo: *Takusan hito-*{*ga/φ*} *kitetayo.*
  Kansai: *Takusan hito-*{*??ga/φ*} *kiteyatten.*
- This can be modelled by specifying the verb itself as an implicit focus for Tokyo Japanese but not in the Kansai variety, i.e.

$$
\boxed{1}\begin{bmatrix} \textit{unacc-verb-standard} \\ \text{FOCUS}\,\big\langle\,\boxed{1}\,\big\rangle \\ \text{ARG-ST}\,\Big\langle\,\begin{bmatrix}\textit{post-phr}\\ \text{SS}\,|\,\text{HD}\,|\,\text{CASE }\textit{nom}\end{bmatrix}\,\Big\rangle \end{bmatrix}\begin{bmatrix} \textit{unerg-verb-kansai} \\ \text{FOCUS}\,\langle\,\rangle \\ \text{ARG-ST}\,\Big\langle\,\begin{bmatrix}\textit{post-phr}\\ \text{SS}\,|\,\text{HD}\,|\,\text{CASE }\textit{nom}\end{bmatrix}\,\Big\rangle \end{bmatrix}
$$

## Dialectal contrast

- Kansai Japanese (spoken in Ôsaka & Kyôto) speakers report narrow focus effect for *ga* in unaccusatives as well for subject case marking
  (What was it (say, an open day) like?) '– Many people turned up.'
  Tokyo: *Takusan hito-*{*ga*/$\phi$} *kitetayo.*
  Kansai: *Takusan hito-*{*??ga*/$\phi$} *kiteyatten.*
- This can be modelled by specifying the verb itself as an implicit focus for Tokyo Japanese but not in the Kansai variety, i.e.

$$\boxed{1}\begin{bmatrix} \textit{unacc-verb-standard} \\ \text{FOCUS}\left\langle \boxed{1} \right\rangle \\ \text{ARG-ST}\left\langle \begin{bmatrix} \textit{post-phr} \\ \text{SS | HD | CASE } \textit{nom} \end{bmatrix} \right\rangle \end{bmatrix} \begin{bmatrix} \textit{unerg-verb-kansai} \\ \text{FOCUS}\left\langle \right\rangle \\ \text{ARG-ST}\left\langle \begin{bmatrix} \textit{post-phr} \\ \text{SS | HD | CASE } \textit{nom} \end{bmatrix} \right\rangle \end{bmatrix}$$

- Eventually:
  Tokyo: FOC $\langle$ sub, verb $\rangle$ / Kansai: FOC $\langle$ sub $\rangle$

# Outline

# Predictability, probability and case-marking

- Today's point: conventional (predicate-dependent and language-dependent) nature of the relationship between information structure and case-marking

# Predictability, probability and case-marking

- Today's point: conventional (predicate-dependent and language-dependent) nature of the relationship between information structure and case-marking
- Two remaining issues:
    - How do the speaker know, or learn, this language-specific and implicit convention?
    - The implicit focus position should not always be arbitrary: universal patterns can be found e.g. object of a transitive verb

# Predictability, probability and case-marking

- Today's point: conventional (predicate-dependent and language-dependent) nature of the relationship between information structure and case-marking

- Two remaining issues:
    - How do the speaker know, or learn, this language-specific and implicit convention?
    - The implicit focus position should not always be arbitrary: universal patterns can be found e.g. object of a transitive verb

- Two potential sources for explanation
    - Aissen's (2003) constraint-based (optimality-oriented) view: economy vs. markedness, i.e. no case-marking for predictable positions
    - Jaeger's (2009) information-theoretic view: the more unpredictable probabilistically, the more overt syntax

# Predictability, probability and case-marking

- Today's point: conventional (predicate-dependent and language-dependent) nature of the relationship between information structure and case-marking

- Two remaining issues:
    - How do the speaker know, or learn, this language-specific and implicit convention?
    - The implicit focus position should not always be arbitrary: universal patterns can be found e.g. object of a transitive verb

- Two potential sources for explanation
    - Aissen's (2003) constraint-based (optimality-oriented) view: economy vs. markedness, i.e. no case-marking for predictable positions
    - Jaeger's (2009) information-theoretic view: the more unpredictable probabilistically, the more overt syntax

- Constraint-based view and probabilistic method can be combined

# Feature-conditioned probability

- Probabilistic methods do, if not provide explanation, indicate which features do affect case-marking

# Feature-conditioned probability

- Probabilistic methods do, if not provide explanation, indicate which features do affect case-marking
- Logistic regression computes the probability of a binary feature (e.g. presence of a case-marker) conditioned on other features (e.g. argument structure)

# Feature-conditioned probability

- Probabilistic methods do, if not provide explanation, indicate which features do affect case-marking
- Logistic regression computes the probability of a binary feature (e.g. presence of a case-marker) conditioned on other features (e.g. argument structure)
- Using dialect corpora, Sato and Nakagawa (2012) identifies the features that significantly contribute to case-marking, which are consistent with what we saw (case and argument structure)

# Feature-conditioned probability

- Probabilistic methods do, if not provide explanation, indicate which features do affect case-marking
- Logistic regression computes the probability of a binary feature (e.g. presence of a case-marker) conditioned on other features (e.g. argument structure)
- Using dialect corpora, Sato and Nakagawa (2012) identifies the features that significantly contribute to case-marking, which are consistent with what we saw (case and argument structure)
- Hypothesis: it is feature-conditioned probability that makes the learning of the convention possible
- Or a stronger one: different probabilities conditioned on different features *cause* conventions to arise

|  | Overall | Tokyo | Kansai | Sig (dialects) |
|---|---|---|---|---|
| nom | 52.31% | 58.92% | 50.43% | ** |
| acc | 47.04% | 45.87% | 48.28% | n/s |
| Sig (features) | * | ** | n/s | |
| acc/animate | 48.14% | 48.96% | 47.09% | n/s |
| acc/inanimate | 47.11% | 47.24% | 47.12% | n/s |
| Sig (features) | n/s | * | n/s | |
| nom/unacc | 51.43% | 53.25% | 49.02% | * |
| acc/unerg | 48.63% | 57.33% | 42.15% | *** |
| nom/trans | 71.95% | 74.43% | 70.24% | * |
| acc/trans | 50.77% | 50.31% | 51.37% | n/s |
| Sig (features) | ** | *** | ** | |

Figure : Probability of case-marking a la logistic regression (Sato and Nakagawa 2012)

# Outline

## Concluding remarks

- Main claim: case-marking becomes optional for a position at which the head predicate predicts focus, and mandatory for a position at which it does not

## Concluding remarks

- Main claim: case-marking becomes optional for a position at which the head predicate predicts focus, and mandatory for a position at which it does not
- which causes indirect mechanism to arise, where case-marking functions as a focus-marking surrogate

## Concluding remarks

- Main claim: case-marking becomes optional for a position at which the head predicate predicts focus, and mandatory for a position at which it does not
- which causes indirect mechanism to arise, where case-marking functions as a focus-marking surrogate
- More generally (and speculatively) on the explicit/implicit syntactic marking:

## Concluding remarks

- Main claim: case-marking becomes optional for a position at which the head predicate predicts focus, and mandatory for a position at which it does not
- which causes indirect mechanism to arise, where case-marking functions as a focus-marking surrogate
- More generally (and speculatively) on the explicit/implicit syntactic marking:
  - explicit marking occurs in more **unpredictable** places but

## Concluding remarks

- Main claim: case-marking becomes optional for a position at which the head predicate predicts focus, and mandatory for a position at which it does not
- which causes indirect mechanism to arise, where case-marking functions as a focus-marking surrogate
- More generally (and speculatively) on the explicit/implicit syntactic marking:
  - explicit marking occurs in more **unpredictable** places but
  - there are 'intermediate' cases that are subject to variation across languages/dialects **conventional** nature of explicit/implicit marking or where the boundary is

## Concluding remarks

- Main claim: case-marking becomes optional for a position at which the head predicate predicts focus, and mandatory for a position at which it does not
- which causes indirect mechanism to arise, where case-marking functions as a focus-marking surrogate
- More generally (and speculatively) on the explicit/implicit syntactic marking:
  - explicit marking occurs in more **unpredictable** places but
  - there are 'intermediate' cases that are subject to variation across languages/dialects **conventional** nature of explicit/implicit marking or where the boundary is
  - **conditional probability** may help to determine what the relevant features are and where the boundary is in a particular language

## Concluding remarks

- Main claim: case-marking becomes optional for a position at which the head predicate predicts focus, and mandatory for a position at which it does not
- which causes indirect mechanism to arise, where case-marking functions as a focus-marking surrogate
- More generally (and speculatively) on the explicit/implicit syntactic marking:
    - explicit marking occurs in more **unpredictable** places but
    - there are 'intermediate' cases that are subject to variation across languages/dialects **conventional** nature of explicit/implicit marking or where the boundary is
    - **conditional probability** may help to determine what the relevant features are and where the boundary is in a particular language

Thanks for your attention!