INFLECTIONAL ISLANDS

Sally Rice & John Newman

University of Alberta

CANADA
defective verb paradigms
inflectional islands
the conceit of the lemma
words-in-context (WICs)
defective verb paradigms
STRUCTURE OF THE TALK

I. describe what we mean by “inflectional islands”

II. survey examples from published literature and our own queries of the BNC

III. suggest implications for linguistic theory, lexicography, typology, and psycholinguistic research
STRUCTURE OF THE TALK

I. describe what we mean by “inflectional islands”

II. survey examples from published literature and our own queries of the BNC

III. suggest implications for linguistic theory, lexicography, typology, and psycholinguistic research
children tend to use uninflected verb roots before inflected forms
verb inflections are mastered on a verb-by-verb basis
generalization is gradual
initially, particular verbs “strand” inflections

adults use particular inflected forms of individual verbs on a register-specific basis
verb inflections adhere to verbs on a verb-by-verb basis
particularization is gradual
eventually, inflections “strand” particular verbs
An English Verb Paradigm: SUBJ x TAM

<table>
<thead>
<tr>
<th></th>
<th>INF</th>
<th>PRES</th>
<th>PAST</th>
<th>PROG</th>
<th>PERF</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.SG</td>
<td>I need to go</td>
<td>I go</td>
<td>I went</td>
<td>I am/was going</td>
<td>I have/had gone</td>
</tr>
<tr>
<td>2</td>
<td>you need to go</td>
<td>you go</td>
<td>you went</td>
<td>you are/were going</td>
<td>you have/had gone</td>
</tr>
<tr>
<td>3.SG</td>
<td>s/he/it need to go</td>
<td>s/he/it goes</td>
<td>s/he/it went</td>
<td>s/he/it is/was going</td>
<td>s/he/it has/had gone</td>
</tr>
<tr>
<td>1.PL</td>
<td>we need to go</td>
<td>we go</td>
<td>we went</td>
<td>we are/were going</td>
<td>we have/had gone</td>
</tr>
<tr>
<td>3.PL</td>
<td>they need to go</td>
<td>they go</td>
<td>they went</td>
<td>they are/were going</td>
<td>they have/had gone</td>
</tr>
</tbody>
</table>
Frequency Distribution in BNC<sub>all</sub>

**GO**

<table>
<thead>
<tr>
<th></th>
<th>INF</th>
<th>PRES</th>
<th>PAST</th>
<th>PROG</th>
<th>PERF</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.SG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.SG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.PL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.PL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Frequency Distribution in BNC

**GO**

<table>
<thead>
<tr>
<th></th>
<th>INF</th>
<th>PRES</th>
<th>PAST</th>
<th>PROG</th>
<th>PERF</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.SG</td>
<td>6 %</td>
<td>3 %</td>
<td>2 %</td>
<td>4 %</td>
<td>0 %</td>
</tr>
<tr>
<td>2</td>
<td>2 %</td>
<td>10 %</td>
<td>0 %</td>
<td>3 %</td>
<td>0 %</td>
</tr>
<tr>
<td>3.SG</td>
<td>13 %</td>
<td>4 %</td>
<td>10 %</td>
<td>15 %</td>
<td>2 %</td>
</tr>
<tr>
<td>1.PL</td>
<td>6 %</td>
<td>0 %</td>
<td>1 %</td>
<td>2 %</td>
<td>0 %</td>
</tr>
<tr>
<td>3.PL</td>
<td>2 %</td>
<td>6 %</td>
<td>1 %</td>
<td>2 %</td>
<td>3 %</td>
</tr>
</tbody>
</table>
## Frequency Distribution in BNCall

### GO

<table>
<thead>
<tr>
<th></th>
<th>INF</th>
<th>PRES</th>
<th>PAST</th>
<th>PROG</th>
<th>PERF</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.SG</td>
<td>6 %</td>
<td>3 %</td>
<td>2 %</td>
<td>4 %</td>
<td>0 %</td>
</tr>
<tr>
<td>2</td>
<td>2 %</td>
<td>10 %</td>
<td>0 %</td>
<td>3 %</td>
<td>0 %</td>
</tr>
<tr>
<td>3.SG</td>
<td>13 %</td>
<td>4 %</td>
<td>10 %</td>
<td>15 %</td>
<td>2 %</td>
</tr>
<tr>
<td>1.PL</td>
<td>6 %</td>
<td>0 %</td>
<td>1 %</td>
<td>2 %</td>
<td>0 %</td>
</tr>
<tr>
<td>3.PL</td>
<td>2 %</td>
<td>6 %</td>
<td>1 %</td>
<td>2 %</td>
<td>3 %</td>
</tr>
</tbody>
</table>
Frequency Distribution in BNCall

GO

![Chart showing frequency distribution for different verb forms in BNCall](chart.png)
lemmas

- argument structure(s)
- syntactic constructions
- lexical meaning

inflected forms

“have a life of their own”  
Thompson & Hopper 2001:44

words in context (WICs)

- distribution patterns (usage)
- collocations & N-grams
- pragmatic associations
- incipient grammaticalization & idiomaticization
VVB - present, imperative
  go

VVZ - 3SG.present
  goes

VVI - infinitive
  go

VVD - past
  went

VVG - progressive
  going

VVN - perfect
  gone
conversation
fiction
news
academic writing
conversation
### Another English Verb Paradigm

#### THINK

<table>
<thead>
<tr>
<th></th>
<th>INF</th>
<th>PRES</th>
<th>PAST</th>
<th>PROG</th>
<th>PERF</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1.SG</strong></td>
<td>I need to think</td>
<td>I think</td>
<td>I thought</td>
<td>I am/was thinking</td>
<td>I have/had thought</td>
</tr>
<tr>
<td><strong>2</strong></td>
<td>you need to think</td>
<td>you think</td>
<td>you thought</td>
<td>you are/were thinking</td>
<td>you have/had thought</td>
</tr>
<tr>
<td><strong>3.SG</strong></td>
<td>s/he/it needs to think</td>
<td>s/he/it thinks</td>
<td>s/he/it thought</td>
<td>s/he/it is/was thinking</td>
<td>s/he/it has/had thought</td>
</tr>
<tr>
<td><strong>1.PL</strong></td>
<td>we need to think</td>
<td>we think</td>
<td>we thought</td>
<td>we are/were thinking</td>
<td>we have/had thought</td>
</tr>
<tr>
<td><strong>3.PL</strong></td>
<td>they need to think</td>
<td>they think</td>
<td>they thought</td>
<td>they are/were thinking</td>
<td>they have/had thought</td>
</tr>
</tbody>
</table>
Frequency Distribution in BNC

<table>
<thead>
<tr>
<th></th>
<th>INF</th>
<th>PRES</th>
<th>PAST</th>
<th>PROG</th>
<th>PERF</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.SG</td>
<td>80%</td>
<td>93%</td>
<td>82%</td>
<td>65%</td>
<td>75%</td>
</tr>
<tr>
<td>2</td>
<td>10%</td>
<td>2%</td>
<td>2%</td>
<td>8%</td>
<td>11%</td>
</tr>
<tr>
<td>3.SG</td>
<td>7%</td>
<td>0%</td>
<td>5%</td>
<td>9%</td>
<td>10%</td>
</tr>
<tr>
<td>1.PL</td>
<td>2%</td>
<td>2%</td>
<td>7%</td>
<td>12%</td>
<td>0%</td>
</tr>
<tr>
<td>3.PL</td>
<td>1%</td>
<td>3%</td>
<td>4%</td>
<td>6%</td>
<td>4%</td>
</tr>
</tbody>
</table>
### Frequency Distribution in BNC

**THINK**

<table>
<thead>
<tr>
<th></th>
<th>INF</th>
<th>PRES</th>
<th>PAST</th>
<th>PROG</th>
<th>PERF</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.SG</td>
<td>80%</td>
<td>93%</td>
<td>82%</td>
<td>65%</td>
<td>75%</td>
</tr>
<tr>
<td>2</td>
<td>10%</td>
<td>2%</td>
<td>2%</td>
<td>8%</td>
<td>11%</td>
</tr>
<tr>
<td>3.SG</td>
<td>7%</td>
<td>0%</td>
<td>5%</td>
<td>9%</td>
<td>10%</td>
</tr>
<tr>
<td>1.PL</td>
<td>2%</td>
<td>2%</td>
<td>7%</td>
<td>12%</td>
<td>0%</td>
</tr>
<tr>
<td>3.PL</td>
<td>1%</td>
<td>3%</td>
<td>4%</td>
<td>6%</td>
<td>4%</td>
</tr>
</tbody>
</table>
Frequency Distribution in BNCcc

**THINK**

(Subject x TAM)

<table>
<thead>
<tr>
<th>VVB</th>
<th>VVI</th>
<th>VVD</th>
<th>VVG</th>
<th>VVN</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="chart.png" alt="Chart with data distribution for different TAMs (Subject x TAM)" /></td>
<td><img src="chart.png" alt="Chart with data distribution for different TAMs (Subject x TAM)" /></td>
<td><img src="chart.png" alt="Chart with data distribution for different TAMs (Subject x TAM)" /></td>
<td><img src="chart.png" alt="Chart with data distribution for different TAMs (Subject x TAM)" /></td>
<td><img src="chart.png" alt="Chart with data distribution for different TAMs (Subject x TAM)" /></td>
</tr>
</tbody>
</table>

- **1s**: Yellow
- **2**: Maroon
- **3s**: Pink
- **1p**: White
- **3p**: Purple
Frequency Distribution in BNCcc

**THINK**

- I think... (93%)
- I don’t think... (70%)
- I thought... (82%)
- I was thinking... (28%)
- I would have thought... (39%)
Hongyin Tao’s (2001, 2003) Spoken Corpus (CSAE) Results

<table>
<thead>
<tr>
<th></th>
<th>REMEMBER</th>
<th>FORGET</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>that</em> complement</td>
<td>19%</td>
<td>4%</td>
</tr>
<tr>
<td>gerundive complement</td>
<td>6%</td>
<td>1%</td>
</tr>
<tr>
<td>infinitival complement</td>
<td>1%</td>
<td>14%</td>
</tr>
<tr>
<td>non-complement</td>
<td>74%</td>
<td>79%</td>
</tr>
<tr>
<td>1st subject</td>
<td>55%</td>
<td>61%</td>
</tr>
<tr>
<td>2nd subject</td>
<td>14%</td>
<td>4%</td>
</tr>
<tr>
<td>3rd subject</td>
<td>3%</td>
<td>6%</td>
</tr>
<tr>
<td>null subject</td>
<td>28%</td>
<td>29%</td>
</tr>
</tbody>
</table>
Hongyin Tao’s (2001, 2003) Spoken Corpus (CSAE) Results

<table>
<thead>
<tr>
<th></th>
<th>REMEMBER</th>
<th>FORGET</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>that</em> complement</td>
<td>19%</td>
<td>4%</td>
</tr>
<tr>
<td>gerundive complement</td>
<td>6%</td>
<td>1%</td>
</tr>
<tr>
<td>infinitival complement</td>
<td>1%</td>
<td>14%</td>
</tr>
<tr>
<td>non-complement</td>
<td>74%</td>
<td>79%</td>
</tr>
<tr>
<td>1st subject</td>
<td>55%</td>
<td>61%</td>
</tr>
<tr>
<td>2nd subject</td>
<td>14%</td>
<td>4%</td>
</tr>
<tr>
<td>3rd subject</td>
<td>3%</td>
<td>6%</td>
</tr>
<tr>
<td>null subject</td>
<td>28%</td>
<td>29%</td>
</tr>
</tbody>
</table>
Hongyin Tao’s (2001, 2003) Spoken Corpus (CSAE) Results

<table>
<thead>
<tr>
<th></th>
<th>REMEMBER</th>
<th>FORGET</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>that complement</strong></td>
<td>19%</td>
<td>4%</td>
</tr>
<tr>
<td>gerundive complement</td>
<td>6%</td>
<td>1%</td>
</tr>
<tr>
<td>infinitival complement</td>
<td>1%</td>
<td>14%</td>
</tr>
<tr>
<td>non-complement</td>
<td>74%</td>
<td>79%</td>
</tr>
<tr>
<td>1st subject</td>
<td>55%</td>
<td>61%</td>
</tr>
<tr>
<td>2nd subject</td>
<td>14%</td>
<td>4%</td>
</tr>
<tr>
<td>3rd subject</td>
<td>3%</td>
<td>6%</td>
</tr>
<tr>
<td>null subject</td>
<td>28%</td>
<td>29%</td>
</tr>
</tbody>
</table>
Hongyin Tao’s (2001, 2003) Spoken Corpus (CSAE) Results

<table>
<thead>
<tr>
<th></th>
<th>I remember</th>
<th>I forget</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>that</em> complement</td>
<td>19%</td>
<td>4%</td>
</tr>
<tr>
<td>gerundive complement</td>
<td>6%</td>
<td>1%</td>
</tr>
<tr>
<td>infinitival complement</td>
<td>1%</td>
<td>14%</td>
</tr>
<tr>
<td>non-complement</td>
<td>74%</td>
<td>79%</td>
</tr>
<tr>
<td>1st subject</td>
<td>55%</td>
<td>61%</td>
</tr>
<tr>
<td>2nd subject</td>
<td>14%</td>
<td>4%</td>
</tr>
<tr>
<td>3rd subject</td>
<td>3%</td>
<td>6%</td>
</tr>
<tr>
<td>null subject</td>
<td>28%</td>
<td>29%</td>
</tr>
</tbody>
</table>
In spoken English, REMEMBER and FORGET are de facto discourse particles or epistemic stance predications; moreover, “complement-taking is actually a marginal feature”

Tao 2003:75.
THE INFLECTIONAL ISLAND HYPOTHESIS

Rice & Newman 2005

• uneven distribution of inflection

• verbs (and verb classes) have “weighted” inflectional profiles

• weightings may be universal (experientially motivated)

• inflectional categories are lexically & pragmatically meaningful
  (and not just part of grammatical house-keeping or concord relationships)

• especially “weighty” inflected verbs (WICs) may idiomaticize and grammaticalize
STRUCTURE OF THE TALK

I. describe what we mean by “inflectional islands

II. survey examples from published literature and our own queries of the BNC

III. suggest implications for linguistic theory, lexicography, typology, and psycholinguistic research
Looking for **Islands** (Stranded Verbs)

sought BNC with Mark Davies’ corpus tool:
Variation in English Words and Phrases: http://view.byu.edu

de-lemmatized the verb (re-inflectionalized it)
downloaded 100 hits each for every verb matching a BNC tag

factored in genre/register
Casual Conversation (4.2M sub-corpus)

tracked subject & TAM distribution
coded each hit for subject, tense, complement type, etc.

examined inflectional “skew”
Some Classic Stranded Verbs
(inflectional islands)

<table>
<thead>
<tr>
<th>Category</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>MODALS</td>
<td>rumour</td>
</tr>
<tr>
<td>IMPERSONALS</td>
<td>rid</td>
</tr>
<tr>
<td>WEATHER VERBS</td>
<td>allow</td>
</tr>
</tbody>
</table>
rumour

RUMOUR
frequency per million

RUMOUR (SUBJ x TAM)
**rumour**

**it BE rumoured to V...**

- 100M BNC
- 273 hits
- 2.8 (freq per M)
- not in casual conversation

---

**Subjects of (BE) RUMOURED [VVN]**

- dummy it/there
- personal
- inanimate
- corporations
- animate

**Complements of (BE) RUMOURED [VVN]**

- inf
- that S
- S
- other (as, for)
- none
rid
allow

VVB-base
VVZ-3sg.pres
VVI-inf
VVD-past
VVG-prog
VVN-perf part

RID
frequency per million

1/2 get rid of

ALLOW
frequency per million

2/3 not be allowed to VP
2/3 be allowed (to have) NP
Some Emerging Stranded Verbs
(inflectional islands)

think
know
mean
want
*say
**think**

The image shows a bar chart and a line graph for the word **think**.

### Line Graph
- **THINK frequency per million**
- **Rank across BNC-cc**
- The x-axis represents different TAM forms: VVB, VVZ, VVI, VVD, VVG, VVN.
- The y-axis shows frequency per million.
- The graph indicates the frequency of **think** across different TAM forms.

### Bar Chart
- **THINK (Subject x TAM)**
- The chart shows the distribution of TAM forms for **think**.
- The y-axis represents different TAM forms: VVB, VVZ, VVI, VVD, VVG, VVN.
- The x-axis represents the percentage from 0% to 100%.
- The chart includes five TAM forms: 1s, 2, 3s, 1p, 3p.
- The TAM forms are color-coded:
  - **1s** (Yellow)
  - **2** (Purple)
  - **3s** (Pink)
  - **1p** (Light Purple)
  - **3p** (Dark Purple)
- The chart visualizes the distribution of TAM forms across different TAM types for **think**.
**KNOW**

**KNOW** frequency per million

**Rank across BNC-cc**

**you know?**

**I don't know**

**I dunno**

**KNOW** (Subj x TAM)

- **1s**
- **2**
- **3s**
- **1p**
- **3p**
- **INDET/UNSPEC**

**VVB-base**

**VVZ-3sg.pres**

**VVI-inf**

**VVD-past**

**VVG-prog**

**VVN-perf part**
mean

**MEAN**

*frequency per million*

![Bar chart showing frequency per million for different verb forms: VVB, VVZ, VVI, VVD, VVG, VVN. The chart illustrates the mean frequency per million for each form across different TAMs.](image)

**MEAN**

*(Subject x TAM)*

![Bar chart showing the mean frequency per million for different verb forms across subject and TAM combinations: 1s, 2, 3s, 1p, 3p, indet.](image)
want

VVB-base
VVZ-3sg.pres
VVI-inf
VVD-past
VVG-prog
VVN-perf part

WANT
frequency per million

I want to, if you want
(what) do you want...?

WANT
(Subj x TAM)

1s 2 3s 1p 3p
say

VVB-base
VVZ-3sg.pres
VVI-inf
VVD-past
VVG-prog
VVN-perf part

**SAY**

frequency per million

<table>
<thead>
<tr>
<th>Rank across BNC-cc</th>
<th>VVB</th>
<th>VVZ</th>
<th>VVI</th>
<th>VVD</th>
<th>VVG</th>
<th>VVN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**SAY**

(Subj x TAM)

- 1s
- 2
- 3s
- 1p
- 3p
- unspec
High frequency constructions are more likely to undergo semantic/pragmatic and phonological change over time.

Bybee 1985, Bybee & Hopper 2001
STRUCTURE OF THE TALK

I. describe what we mean by “inflectional islands”

II. survey examples from published literature and our own queries of the BNC

III. suggest implications for linguistic theory, lexicography, typology, and psycholinguistic research
USAGE-BASED GRAMMARS
LEMMA x GENRE

INFLECTED FORMS
IMPLICATIONS OF USAGE-BASED APPROACHES TO GRAMMAR

a new starting point for linguistic analysis

put lemmas aside (as done earlier with syntactic rule in favor of constructions)

substitute words-in-context or WICs (intersection of genre, register, & inflection)

a new (lower) level of linguistic generalization

find the “hierarchy of lower-level structures...[that] specify the actual array of subcases and specific instances that support and give rise to the higher-level generalization”

RWL, Concept, Image, & Symbol, 1991:281-282
WICs

locus of grammaticalization

active in borrowings & morphological realignment

spawn psychological associations, induce priming effect

WICs are relevant for speakers....why not for linguists??
WICs

locus of grammaticalization

active in borrowings & morphological realignment

spawn psychological associations, induce priming effect

normalize suppletion & polysynthesis
A Typical Dene (Athapaskan) Verb Paradigm

<table>
<thead>
<tr>
<th>sit.IMPF</th>
<th>SG</th>
<th>DU</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>thida</td>
<td>th7ke</td>
<td>deth7ltth’ i</td>
</tr>
<tr>
<td>2</td>
<td>th8da</td>
<td>thuhke</td>
<td>dumttth’ i</td>
</tr>
<tr>
<td>3</td>
<td>theda</td>
<td>heheke</td>
<td>d4mtth’ i</td>
</tr>
</tbody>
</table>
Another Dene (Athapaskan) Verb Paradigm

<table>
<thead>
<tr>
<th>go.IMPF</th>
<th>SG</th>
<th>DU</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>hessa</td>
<td>h7t’ 1s</td>
<td>h7d4m</td>
</tr>
<tr>
<td>2</td>
<td>h8gha</td>
<td>huh/1s</td>
<td>huhd4m</td>
</tr>
<tr>
<td>3</td>
<td>hegha</td>
<td>he/1s</td>
<td>hed4m</td>
</tr>
</tbody>
</table>
Another Dene (Athapaskan) Verb Paradigm

PERF forms are different again....so which is chosen as the head word?

<table>
<thead>
<tr>
<th>go.IMPF</th>
<th>SG</th>
<th>DU</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>hessa</td>
<td>h7't' 1s</td>
<td>h7d4m</td>
</tr>
<tr>
<td>2</td>
<td>h8gha</td>
<td>huh/1s</td>
<td>huhd4m</td>
</tr>
<tr>
<td>3</td>
<td>hegha</td>
<td>he/1s</td>
<td>hed4m</td>
</tr>
</tbody>
</table>
Thank you.

Sally Rice
sally.rice@ualberta.ca

John Newman
john.newman@ualberta.ca