Generics and Habituals and Textual Inference

Graham Katz

Department of Linguistics
Georgetown University

Semantics for Textual Inference
LSA Summer Institute
Boulder, CO July 10-11, 2011
We all know how to treat simple entailments

(1)  a. Every student ate beans for lunch. Jack is a student.
    b. Jack ate lunch.

Generic sentences and Habitual sentences present difficulties:

- Relationship between individual (or individual event) and generalization

(2)  a. Students like beans.
    b. Jack likes beans.

- Relationship between individual event and habit

(3)  a. Jack ate beans for lunch in those days.
    b. Jack ate beans for lunch on Tuesday.
We all know how to treat simple entailments

\[(1)\]  
\[
a. \quad \text{Every student ate beans for lunch. Jack is a student.} \\
\]
\[
b. \quad \text{Jack ate lunch.} \\
\]

- **Generic sentences** and **Habitual sentences** present difficulties:
  - Relationship between individual (or individual event) and generalization
    \[(2)\]  
    \[
a. \quad \text{Students like beans.} \\
    \]
    \[
b. \quad \text{Jack likes beans.} \\
    \]
  - Relationship between individual event and habit
    \[(3)\]  
    \[
a. \quad \text{Jack ate beans for lunch in those days.} \\
    \]
    \[
b. \quad \text{Jack ate beans for lunch on Tuesday.} \\
    \]
Generics in RTE

Interpretation of **bare plural subjects** in RTE3:

(4)  
a. CSIRO researchers have established that culturing marine plants, oysters or other filtering organisms in farm discharge water can significantly improve the water quality.
b. Marine vegetation improves water quality.

(5)  
a. It didn’t happen because the cream of England’s thugs was smoking pot which is easily and legally available in the Netherlands.
b. Drugs in Holland are easily bought.

(6)  
a. Many of the Vikings who travelled to Scotland, and other parts of Europe were traders or peaceful settlers looking for land to farm.
b. Vikings landed in Scotland.
Habituals in RTE

Relationship between individual events and **habituals** in RTE:
Generalization follows from multiple event descriptions:

(7)  
   a. The Miracle Field, which is a field specifically designed for children with special needs, is the third field of its kind built in the last year in the Chicago area with funds from Chicago White Sox Charities.
   
b. Chicago White Sox Charities support children with special needs.

Inference from habit to single event:

(8)  
   a. Lauren landed her first job performing in the national touring company of the musical "Dreamgirls". Loraine would often step in for her sister in "Dreamgirls", in which they played the role of "Mimi Marquez".
   
b. Loraine took part in the musical "Dreamgirls".
Additional Problem: Temporal interpretation

(9)  
  a. Jack got up early. He worked hard all morning and ate a late lunch.
  b. John worked hard before he ate lunch.

(10)  
  a. Jack got up early. He worked for IBM and ate lunch early.
  b. Jack worked for IBM before he ate lunch.

To determine inferences that can be drawn, genericity of sentence must be known.
Problems **generic sentences** raise for practical reasoning

- **Semantics of Generics**
  - Quasi-quantificational effects
    - \( Dogs \text{ bark} \implies Fido \text{ barks} \)
  - Defeasibility
    - (Fido might not bark)

- **Identifying Generics:**
  - In most languages genericity is not (necessarily) overtly marked
    - \( Peter \text{ only ate oranges for breakfast.} \)
Generic sentences and habitual sentences ("characterizing sentences") denote "propositions which does not express specific episodes or isolated facts, but instead report a kind of general property, that is, report a regularity which summarizes a group of particular episodes or facts" (Krifka, Pelletier, Carlson, Ter Meulen, Chierchia, and Link 1995)

Characterizing sentences:

    b. Tweety flies.

Episodic sentences:

(12) a. Birds flew by my window last week.
    b. Tweety is flying by my window.
Characterizing sentences are often in the present tense with a bare plural subject. But not always:

(13) a. John smokes a cigarette after dinner.
    b. I am a pipe smoker.
    c. This book is readable.
    d. A professor drinks whisky.
    e. Every professor drinks whisky.
    f. Gold sells for $400 per ounce.

(14) a. John smoked a pipe
    b. John will smoke (when he grows up)
    c. John has been a pipe smoker.
Lexical and Habitual Characterizing Sentences

Characterizing sentences come in two sorts: *Habitual* and *Lexical* Characterizing Sentences:

(15)  
\[
\begin{align*}
\text{a.} & \quad \text{John knows French} \\
\text{b.} & \quad \text{Fido is a dog} \\
\text{c.} & \quad \text{Fido likes meat} \\
\text{d.} & \quad \text{Dogs have four legs}
\end{align*}
\]

<table>
<thead>
<tr>
<th>Characterizing Sentences</th>
<th>Eventive Verb</th>
<th>Stative Verb</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>Fido barks</em></td>
<td><em>Fido likes meat</em></td>
</tr>
<tr>
<td></td>
<td>(HABITUAL)</td>
<td>(LEXICAL)</td>
</tr>
<tr>
<td>Episodic sentences</td>
<td><em>Fido walked across my lawn</em></td>
<td><em>Fido is in this cage</em></td>
</tr>
</tbody>
</table>
In English, one of the major distinctions between characterizing sentences and episodic sentences is the interpretation of bare plural subjects.

- **Bare plural subjects in characterizing sentences express (quasi) universal interpretations.**

  (16) Dogs are mammals.
  a. ≈ All dogs are mammals
  b. ≠ There are dogs that are mammals

- **Bare plurals in episodic sentences express existential interpretations.**

  (17) Dogs walked across my lawn.
  a. ≠ All dogs walked across my lawn.
  b. = There are dogs that walked across my lawn.
### Bare Plural Interpretation

<table>
<thead>
<tr>
<th></th>
<th>Eventive Verb</th>
<th>Stative Verb</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Characterizing</strong></td>
<td><em>Dogs bark</em> (HABITUAL)</td>
<td><em>Dogs like meat</em> (LEXICAL)</td>
</tr>
<tr>
<td>sentences</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Episodic</strong></td>
<td><em>Dogs walked across my lawn</em></td>
<td><em>Dogs are in this cage</em></td>
</tr>
<tr>
<td>sentences</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Characterizing:** (aka “individual-level”)

(18)  
- b. Politicians are honest.  
- c. Sugar cubes dissolve in water.

**Episodic:** (aka “stage-level”)

(19)  
- a. Cats ate our lunch.  
- b. Politicians were at the party.  
- c. Sugar cubes fell in my tea.
How many is enough for a bare plural in a generic?

(20) a. Snakes are reptiles
b. Telephone books are thick books
c. Mammals give birth to live young
d. Shoplifters are prosecuted in criminal courts
e. Mosquitos carry the paramecium that causes yellow fever
f. White sharks attack bathers

\[ all \text{ snakes} \]
\[ those \text{ of large modern cities} \]
\[ only \text{ adult fertilized females} \]
\[ most \text{ are not even caught} \]
\[ very \text{ few do} \]
\[ only \text{ a tiny minority} \]
Other issues: **Exceptions** and **Non-monotonicity**

Which elements count?

(21)  
a. Ducks lay eggs.  
   Only adult fertilized females
b. Ducks have colorful feathers.  
   Only adult males

(22)  
  a. #Ducks are males.  
  b. #Ducks are females.

(23-b) doesn’t follow from (23-a)

(23)  
  a. Italians are good fencers.  
  b. Italians are fencers.

Unlike the monotonically increasing quantified cases:

(24)  
  a. Most/All/Some Italians are good fencers  ⇒
  b. Most/All/Some Italians are fencers
Characterizing Sentences express “gnomic” or law-like propositions.

(25) 
   a. Sugar dissolves in water.
   b. Objects are attracted to one another with a force proportional to the product of their masses and inversely proportional to the distance between them.
   c. Bishops move diagonally.

(Note - this doesn’t mean they are “timeless”)

(26) 
   a. John works for the Coca-Cola company.
   b. Apples cost 99 cents a kilo.
Carlson’s G/G’ Operators

Carlson (1977): gnomic statements are about individuals (objects or kinds) not stages:

(27)  a. John walked to school
      \[\exists x^S [R(x^S, john) \land \text{walk-to-school}(x^S)]\]
      b. John likes school.
         like-school(john)

G operator converts stage-level to individual-level predicate:

(28)  John walks to school
       G(\text{walk-to-school})(john)

G’ operator to convert individual-level predicates to kind-level predicates

(29)  a. Grad students like school
       G'(\lambda x \ [\text{like}(x,\text{school})])(\text{grad-students})
      b. Grad students walk to school
         G'(G(\text{walk-to-school}))(\text{grad-students})
Generic Operator as Q-Adverb

Classic relational analysis (Carlson 1989; Schubert and Pelletier 1989; Diesing 1992): Generic sentences contain **GEN** operator, a covert, modalized quantificational adverb:

\[(30)\]
\[
\begin{align*}
\text{a. } & \text{Cars have wheels (} \approx \text{ Cars usually/generally have wheels)} \\
\text{b. } & \text{GEN}_x [\text{car}(x)] \exists y [\text{wheel}(y) \land \text{has}(x,y)]
\end{align*}
\]

**Initial Consequence**: Provides straightforward account of: Existential/Universal ambiguity of Bare Plurals

\[(31)\]
\[
\begin{align*}
\text{a. } & \text{Cars are parked outside.} \\
\text{b. } & \exists x [\text{car}(x)] [\text{parked-outside}(x)]
\end{align*}
\]

The quantificational variability problems referred to above also appear in relation to restricting Q-adverbs (von Fintel 1994)
Habitual Semantics

Habituals appear to report lawlike “patterns of activity” (as generics report lawlike quantificational relations)

(32) Peter goes bowling on Tuesday nights.

**Means:** If it is Tuesday night (and nothing odd happens), Peter will go bowling

Capacity readings:

(33) a. This machine crushes oranges.
    b. Mary handles the mail from Antarctica.
    c. Peter drinks beer.

No (or not much) actual activity has to take place.
Kratzer (1995) suggested: Habituals are just generics (sentences with a GEN operator that binds an event argument)

(34)  
  a. When John eats fish, he throws up.  
  b. GEN_{e,e'} [eat-fish(e,john)]  \exists e' [C(e,e') \land throw-up(e',john)]

(35)  
  a. John drinks beer.  
  b. GEN_{e} [in(John,e) \land C(e)] [drink-beer(e,john)]

Happy prediction: Habituals are stative!
Habitual sentences in general pattern like lexical statives:

- **Appear in simple present tense**

  (36)  
  a. Peter knows French.
  b. Peter works in the garden.

- **Induce epistemic reading of modal verb *must***

  (37)  
  a. Peter must know French (...he got the table so quickly)
  b. Peter must work in the garden (...his tomatoes are so nice)
     (...or he’ll go crazy)

- **Induce simultaneous reading with when clause:**

  (38)  
  a. When Mary arrived in town, Steve owned a red Ferrari. *simultaneous*
  b. When Mary arrived in town, Steve bought a red Ferrari. *sequential*

  (39) When Mary arrive in town, Steve worked in the garden (now he has given up gardening).
**Practical Problem**: Determine event sequence from text:

*A man walked into a bar in the District last night. The bar was mostly empty, he worked there and the few customers were drinking, so nobody noticed when he went into a back room. At closing time, four employees of the bar were found tied up. They were all male Hispanic immigrants who lived in the building.*

Narrative interpretation (Partee 1984; Kamp and Reyle 1993; Hitzeman, Moens, and Grover 1995)

- Temporal adverbials locate “Narrative Time”
- Event sentences position events and relocate “Narrative Time”
- State sentences position states and leave “Narrative Time” alone

To correctly provide a temporal interpretation of a text we need to determine the stativity of each clause
Event sentences “move” narrative time, State sentences do not:

(40) a. Stefan woke up. He got out of bed and put on a sweater. Then he had breakfast.
b. Stefan woke up. It was cold. He had a fever. His book lay on the bedside table.

With overt then narrative time does move:

(41) a. It was cold and nobody was in the square. Then it was warm. Sometime later the square filled.
b. Rufus was a bright undergraduate. Then he was a motivated graduate student. Then he was an overworked assistant professor. And finally he was a disgruntled professor.

Habituals have much the same character:

(42) Peter woke up late Tuesday morning. He lectured Tuesdays and Thursdays, so he needed to prepare his lecture. He had a coffee and went to the office.

(43) Peter lectured Monday, Wednesday and Friday. Then he lectured Tuesdays and Thursdays. And much later he lectured only on Wednesdays.
Rimell (2004) notes contrast between modified and unmodified habituals:

(44)  
   a. Mary drinks a beer when she is at Dempsey’s Pub.  
   b. Mary drinks beer when she is at Dempsey’s Pub.

(45)  
   a. Mary drinks beer.  
   b. *Mary drinks a beer

Both should have the fine interpretation:

(46) \[ \text{GEN}_e \left[ C(e) \right] \exists x \left[ \text{beer}(x) \land \text{drink}(\text{mary},x) \right] \]

Contrast between “relational habitual” and “simple habitual”

(47)  
   a. *Writing an angry letter is a bad habit.  \text{not a habitual}  
   b. Writing letters is a good habit.  \text{simple habitual}  
   c. Writing an angry letter when someone insults you is a bad habit.  \text{relational habitual}

Relational habituals are quantificational in a way simple habituals are not.
Discourse can set up context for a quantificational adverb, but NOT for a relational habitual:

(48) Mary and Bob arrive at a party. Mary immediately goes to greet her friend, while Bob stays and chats with the host. The host offers Bob a drink and asks what Mary would like.

(49) Bob: Well, I’m not really sure, but she usually drinks a beer.

(50) Bob: ??Well, I’m not really sure, but she drinks a beer.

(51) Bob: Well, I’m not really sure, but she drinks beer.

We need a semantics for a habitual operator that is different from the Q-adverb-like GEN.
Habitual Marking Cross-linguistically

Dahl (1995) cross-linguistic questionnaire:

(52) [Q: What your brother DO right now? (= What activity is he engaged in?)
Answer by someone who can see him:] He WRITE letters.

(53) [Q: What your brother usually DO after breakfast?
A:] We WRITE letters

(54) [A: My brother works at an office. B: What kind of work he DO?
A:] We WRITE letters

Results from 76 languages:
- 24 Languages mark relational habituals
- Only 7 mark simple habituals with same marking
Claim: Simple habitual predicates are characterizing stative predicates derived from activities by a frequency-adverb-like operator HAB (Bennett and Partee 1978; Katz 1995; Dobrovie-Sorin 2003)

(55)  
(a) John drove a truck for an hour.  
(b) John is driving a truck.  
(c) John drove a truck for a living.  
   John drives a truck

(56)  
(a) John works in the garden.  
(b) ??John plants a garden  
(c) ??John recognizes a criminal.

Relational habitual predicates are derived via the Q-adv-like GEN operator

(57) John drinks a beer when he is thirsty.

HAB (≈ Carlson’s G) ≠ GEN (≈ Carlson’s G’)

Graham Katz
Generics & Habituality
Frequency adverbials are different from quantificational adverbials (although both quantify).

(58)  
   a.  Peter often smokes.  
   b.  Peter smokes often.  

(59)  
   a.  When he goes to the movies, Peter frequently uses the restroom.  
   b.  When he goes the the movies, Peter uses the restroom frequently.  

In English:

- Frequency adverbial associated with sentence final position
- Quantificational adverbial associated with VP initial position

- **Frequency adverbs**: frequently, regularly, often, daily
- **Quantificational**: always, usually, never, seldom, rarely
Frequency/Quantificational Ambiguity

Significant lexical overlap (many adverbs *ambiguous* between frequency and quantificational use)

(60)  
   a. John often writes letters.  
   b. John writes letters often.

(61)  
   a. When she goes to the beach, Stefanie often puts suncream on.  
   b. When she goes to the beach, Stefanie puts suncream on often.

Frequency adverbials tend to get used as quantificational adverbials (but not the other way around)

(62)  
   a. *John eats beans usually.  
   b. John regularly eats beans.

- Frequency adverbs: specify distribution of events in a time interval (how many of them there are, how close together, how evenly distributed, etc.) \( F-Adv(\phi,t) \)
- Quantificational adverbs: specify the quantificational relationship among two event predicates: \( Q-adv(\phi,\psi) \)
Frequency adverbs can be interpreted as generalized quantifiers over time

\[(63) \quad \text{[daily]} = \lambda P \forall \text{day}(t) \rightarrow \exists e [P(e) \land \forall e' [P(e') \land \tau(e') \subseteq t] \rightarrow e' = e]\]

Typically frequency adverbs are quite vague:

\[(64) \quad \text{[frequently]} = \lambda P \exists d \forall e [P(e) \rightarrow \exists e' [P(e') \land \text{distance}(e,e') < d]]\]

We can be pedantic:

\[(65) \quad \text{[regularly]} = \lambda P \exists d \forall e [P(e) \rightarrow \exists e' [P(e') \land \text{distance}(e,e') = d]]\]
Quantificational vs. Frequency

Frequency adverb is quantificational over times; Q-adverb relates events:

(66)  
\( a. \) In 1993, Marc ate beans at The Rib Palace regularly.
\( b. \) In 1993, Marc usually ate BEANS at The Rib Palace.

(67)  
\( a. \) regularly(\( \lambda \, e \, [\text{eat}(e,\text{Marc},\text{beans}) \wedge \text{at}(e, \text{The Rib Palace}) \wedge e \subseteq 1993] \))
\( b. \) usually(\( e \, [\text{eat}(e,\text{Marc},\text{beans})] \wedge \tau(e) \subseteq 1993, \text{at}(e, \text{The Rib Palace})] \))

Quantificational adverbs always “scope” over frequency adverbs:

(68)  
\( a. \) Peter sometimes exercises regularly.
\( b. \) sometimes(C(e),regularly(\( \lambda \, e \, [\text{exercise}(e,\text{Peter})] \)))

(69)  
\( a. \) When he is on vacation, Peter sometimes exercises regularly.
\( b. \) sometimes(on(t,peter,vacation),regularly(\( \lambda \, e \, [\text{exercise}(e,\text{Peter}) \wedge e \subseteq t] \)))
GEN is a quantificational adverb and HAB is a frequency adverb (Bennett and Partee 1978; Farkas and Sugioka 1983; Rimell 2004; Boneh and Doron 2008)
How do we determine if the HAB operator is present?
Habitual Marking

In some languages habituals are marked:
Czech vá (Filip 1994)

(70) a. V sobotu Honza sedává v hospodě
   “On Saturday John usually sits in the pub”

(71) a. V sobotu Honza sedí v hospodě
   “On Saturday John sits/is sitting in the pub”

In English we have to guess (although we have some markers, such as used to or habitually or these days)

(72) a. John worked in the yard
    b. John used to work in the yard
    c. John works in the yard.

Sometimes we get stuck without a cue:

(73) a. If I were to go to Germany, I would visit you
   (cf. If I were to eat meat, I would order steak.)
Cues for Habituals

Usually in English we have a number of weak cues, however:

- **Tense:** *He eats breakfast* vs. *He ate breakfast*
- **Aspect:** *He has eaten breakfast, He is eating breakfast*
- **Subject NP:** *Blacksmiths ate breakfast*
- **Object NP:** *He drank beer vs. He drank a beer*

Are there enough cues to guess reliably?
Supervised Habituals/Episodic Classification

Task: Distinguish **habitual** use of verb from **episodic** (Joint work with Thomas Mathew (Mathew and Katz 2009))

- Choose verbs to annotate which are not lexically stative.
  
  (74)  
  a. John *likes* oranges.  
  b. John *ate* oranges.

- Annotate verb in context as to whether it’s use is habitual or episodic
  
  (75)  
  a. John rarely ate fruit. He just ate oranges.  
  b. John didn’t eat much at breakfast. He just ate oranges.

- Identify **features** which might be relevant to classification
  
  (76)  
  a. John ate an orange *yesterday*.  
  b. John eats oranges *every day*.

- Use Machine Learning techniques to classify
  
  - Decision Tree
  - Naive Baysian
Heuristics for Annotator:

- Identify if clause describes single (or a group events). Single event descriptors are episodic.
- Determine whether adding the adverb *usually* changes meaning significantly. If not, clause is classified as habitual.

Annotation:

- 1,816 clauses from Penn Treebank
  - Wall Street Journal (newspaper text)
  - Brown Corpus of Contemporary English (fiction, journalism, wide range)
- 72 distinct verbs
  - Chosen randomly
  - NOT lexical statives

Results:

- Relatively simple task:
  - Double annotation of 100 clauses ⇒ 98% agreement
- Distribution: 80% episodic and 20% habitual.
- Highly skewed by verb
Many verbs appear mostly as episodics (57 of 74)

<table>
<thead>
<tr>
<th>Verb</th>
<th>Tokens</th>
<th>Habitual (%)</th>
<th>Episodic (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>report</td>
<td>115</td>
<td>0.87</td>
<td>99.13</td>
</tr>
<tr>
<td>cite</td>
<td>54</td>
<td>1.85</td>
<td>98.15</td>
</tr>
<tr>
<td>add</td>
<td>53</td>
<td>3.77</td>
<td>96.23</td>
</tr>
<tr>
<td>drop</td>
<td>25</td>
<td>4.00</td>
<td>96.00</td>
</tr>
<tr>
<td>tell</td>
<td>24</td>
<td>4.17</td>
<td>95.83</td>
</tr>
<tr>
<td>stop</td>
<td>18</td>
<td>5.56</td>
<td>94.44</td>
</tr>
<tr>
<td>pick</td>
<td>34</td>
<td>5.88</td>
<td>94.12</td>
</tr>
<tr>
<td>say</td>
<td>26</td>
<td>7.69</td>
<td>92.31</td>
</tr>
<tr>
<td>set</td>
<td>24</td>
<td>8.33</td>
<td>91.67</td>
</tr>
<tr>
<td>ask</td>
<td>28</td>
<td>10.71</td>
<td>89.29</td>
</tr>
<tr>
<td>give</td>
<td>45</td>
<td>11.11</td>
<td>88.89</td>
</tr>
<tr>
<td>come</td>
<td>207</td>
<td>11.59</td>
<td>88.41</td>
</tr>
<tr>
<td>hand</td>
<td>17</td>
<td>11.76</td>
<td>88.24</td>
</tr>
<tr>
<td>move</td>
<td>68</td>
<td>11.76</td>
<td>88.24</td>
</tr>
<tr>
<td>laugh</td>
<td>8</td>
<td>12.50</td>
<td>87.50</td>
</tr>
</tbody>
</table>
Habitual-skewed Verbs

- Few verbs appear mostly as habituals—only these 12
- (By design we included no verbs that were lexically stative)

<table>
<thead>
<tr>
<th>Verb</th>
<th>Tokens</th>
<th>Habitual (%)</th>
<th>Episodic (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>keep</td>
<td>31</td>
<td>51.61</td>
<td>48.39</td>
</tr>
<tr>
<td>remain</td>
<td>22</td>
<td>54.55</td>
<td>45.45</td>
</tr>
<tr>
<td>refer</td>
<td>9</td>
<td>55.56</td>
<td>44.44</td>
</tr>
<tr>
<td>apply</td>
<td>8</td>
<td>62.50</td>
<td>37.50</td>
</tr>
<tr>
<td>alert</td>
<td>3</td>
<td>66.67</td>
<td>33.33</td>
</tr>
<tr>
<td>be</td>
<td>3</td>
<td>66.67</td>
<td>33.33</td>
</tr>
<tr>
<td>create</td>
<td>3</td>
<td>66.67</td>
<td>33.33</td>
</tr>
<tr>
<td>deny</td>
<td>3</td>
<td>66.67</td>
<td>33.33</td>
</tr>
<tr>
<td>persist</td>
<td>4</td>
<td>75.00</td>
<td>25.00</td>
</tr>
<tr>
<td>maintain</td>
<td>1</td>
<td>83.33</td>
<td>16.67</td>
</tr>
<tr>
<td>act</td>
<td>7</td>
<td>85.71</td>
<td>14.29</td>
</tr>
<tr>
<td>require</td>
<td>13</td>
<td>92.31</td>
<td>7.69</td>
</tr>
</tbody>
</table>
Most-skewed verbs were left out.

- Top 25% (17) verb stems skewed towards the episodic category and the top 25% of the verb stems skewed towards the habitual category.
- Classification sample: 1,052 clauses (57 distinct verb stems) Of them: 73.1% episodic and 26.9% habitual
  Majority class baseline: 73.1%
Features for Predicting Habituality/Episodicity

Tense
- **Present**: She sings
- **Past**: She sang
- **Infinitive**: We expect her to sing

Aspect
- **Progressive**: She is singing
- **Perfect Aspect**: She has sung

Temporal Adverbial
- **Quantificational Temporal Adverbial**: She sang every night
- **Specific Temporal Adverbial**: She sang yesterday

Subject NP
- **Bare-plural Subject**: Slaves sang
- **Definite Subject**: Sonia sang

Object NP
- **Absent Object**: She ate
- **Bare-plural Object**: She sang spirituals
- **Definite Object**: She sang “Strange Fruit”

Prepositional Phrase
- **at-PP**: She sang at home
- **in-PP**: She sang in prison
- **on-PP**: She sang on Monday
Distribution of Features by Category: Habituals

- % to Feature measures how frequently the feature indicates the category
- % to Category measures how frequently a feature is used in the category

Habitual features

- Present tense is habitual 84% of time, but only 50% of habituals in present
- Presence of a quantifying temporal adverb indicates habitual 91%, but only 30% of habituals have one
- 80% of habitual sentences have indefinite object, but indefinite object indicates habitual only 34% of the time.
Distribution of Features by Category: Episodics

- % to Feature measures how frequently the feature indicates the category
- % to Category measures how frequently a feature is used in the category

Episodics
- Past tense and presence of a definite subject provide a good indication on episodicity and are also frequently used in episodics.
- Presence of specific temporal adverbials is a rare but good indicator of episodicity
Supervised Classification

We used two standard machine-learning algorithms (Naive Bayesian and Decision Tree) to classify sentences (10-fold cross validation; 1,052 items)

<table>
<thead>
<tr>
<th></th>
<th>Habitual Precision</th>
<th>Habitual Recall</th>
<th>Episodic Precision</th>
<th>Episodic Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision Tree</td>
<td>84.3%</td>
<td>60.6%</td>
<td>86.8%</td>
<td>95.8%</td>
</tr>
<tr>
<td>Naive Bayes</td>
<td>81.7%</td>
<td>62.7%</td>
<td>87.3%</td>
<td>94.8%</td>
</tr>
<tr>
<td>Baseline</td>
<td>0%</td>
<td>0%</td>
<td>73%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Measures of Classification Success:

- **Precision**: Hearer perspective—Ratio of clauses classified as habitual (episodics) that were intended (by speaker) as habituals (episodics)
- **Recall**: Speaker perspective—Ratio of clauses intended as habitual (episodic) correctly classified as habituals (episodics) by hearer.
On basis of simple superficial cues, classification is relatively simple

Information not being used:

- Lexical knowledge
- Discourse cues

(77)  

a. She didn’t like fruit. She only ate oranges.  
b. She didn’t have much to eat. She only ate oranges.
Wild speculation: Habitual Marking is only obligatory when the grammatical system of the language doesn’t provide enough cues for a speaker to guess.

- Obligatory determiners provide clear markers of preferred accomplishment interpretation:
  *John drank the beer* ⇒ likely non-habitual

- Obligatory progressive provide clear marking of preferred ongoing event interpretation:
  *John is drinking beer* ⇒ likely non-habitual

(Language with no determiners, no plural, and no progressive might need overt morphological marking (Bybee, Perkins, and Pagliuca 1994))
Results:

- Habituals are readily detected in English from surface cues

Future:

- Is there a correlation between use of habitual marking in general and difficulty in determining habituality?
The habitual/episodic distinction is crucial to drawing the correct inferences (even if we don’t see it much of the time).

(78) a. Jimmy Dean usually didn’t move without jeans
    b. Jimmy Dean generally danced without pants. (contradiction)
    c. Jimmy Dean once danced without pants. (no contradiction).

(79) a. Betty Ford ate a grapefruit for breakfast.
    b. Betty Ford ate a bowl of oatmeal for breakfast.
       (contradiction on habitual reading, no contradiction on episodic)
**Bennett, M. and B. Partee (1978).**
Toward the logic of tense and aspect in English.

**Boneh, N. and E. Doron (2008).**
Habituality and the habitual aspect.

**Bybee, J. L., R. Perkins, and W. Pagliuca (1994).**
University of Chicago Press.

**Carlson, G. (1989).**
On the semantic composition of English generic sentences.

**Carlson, G. N. (1977).**
*Reference to Kinds in English.*
Ph. D. thesis, University of Massachusetts, Amherst, Amherst, MA.

**Dahl, O. (1995).**
The episodic/generic disinction in tense/aspect systems.

**Diesing, M. (1992).**
Bare plural subjects and the derivation of logical representations.

**Dobrovie-Sorin, C. (2003).**
Adverbs of quantification and genericity.

**Farkas, D. F. and Y. Sugioka (1983).**
Restrictive if/when clauses.
*Linguistics and Philosophy* 6, 225–258.

**Filip, H. (1994).**
Quantificational morphology.