Overview

• What are auxiliaries?
• General properties of auxiliaries
• Lexical type/lexical entries for auxiliaries
• NICE properties (lexical rules)
What Auxiliaries Are

- Sometimes called “helping verbs,” auxiliaries are little words that come before the main verb of a sentence, including forms of *be, have, do, can, could, may, might, must, shall, should, will,* and *would*

- They tend to be involved in the expression of time, necessity, possibility, permission, and obligation, as well as such things as negation, affirmation, and questioning
Some Basic Facts about Auxiliaries

• They are optional
  
  Pat tapdanced. Pat can tapdance. Pat is tapdancing.

• They precede any non-auxiliary verbs
  
  *Pat tapdance can. *Pat tapdancing is.

• They determine the form of the following verb
  
  *Pat can tapdancing. *Pat is tapdance.

• When they co-occur, their order is fixed
  
  Pat must be tapdancing. *Pat is musting tapdance.

• Auxiliaries of any given type cannot iterate
  
  *Pat could should tapdance.
A Little History

• Chomsky’s first book, *Syntactic Structures* (1957), contained a detailed analysis of the English system of auxiliary verbs

• It showed how formal analysis could reveal subtle generalizations

• The power of Chomsky’s analysis of auxiliaries was one of the early selling points for transformational grammar
  • Especially, his unified treatment of auxiliary *do*

• So it’s a challenge to any theory of grammar to deal with the same phenomena
Two Approaches to Analyzing Auxiliaries

• Treat auxiliaries as a special category, and formulate specialized transformations sensitive to their presence

• Assimilate their properties to existing types as much as possible, and elaborate the lexicon to handle what is special about them

• We adopt the latter, treating auxiliaries as a subtype of $srv-lxm$
Consequences of Making $auxv-lxm$ a Subtype of $srv-lxm$

- Auxiliaries should express one-place predicates
- Auxiliaries should allow non-referential subjects (dummy *there*, *it*, and idiom chunks)
- Passivization of the main verb (the auxiliary’s complement) should preserve truth conditions
- Are these borne out?
Why call auxiliaries verbs?

• *be, have,* and *do* exhibit verbal inflections (tense, agreement)

• *be, have,* and *do* can all appear as main verbs (that is, as the only verb in a sentence)
  • Their inflections are the same in main and auxiliary uses
  • *be* exhibits auxiliary behavior, even in its main verb uses

• Modals (*can, might, will*, etc.) don’t inflect, but they occur in environments requiring a finite verb with no (other) finite verb around.
What’s special about auxiliaries?

- Unlike other subject-raising verbs we have looked at, their complements aren’t introduced by *to*.
- The modals and *do* have defective paradigms.
- There are restrictions on the ordering and iterability of auxiliaries.
- They have a set of special characteristics known as the NICE properties.
## Some Type Constraints

<table>
<thead>
<tr>
<th>TYPE</th>
<th>FEATURES/CONSTRAINTS</th>
<th>IST</th>
</tr>
</thead>
</table>
| verb-lxm | SYN: \[
SYN
HEAD
\text{verb} / \text{AUX}
\]
ARG-ST: \[
\text{ARG-ST}
\langle \text{HEAD nominal}, \ldots \rangle
\]
SEM: \[
\text{SEM}
\text{MODE prop}
\] | infl-lxm |
| srv-lxm | ARG-ST: \[
\text{ARG-ST}
\langle 1, \text{SPR} \langle 1 \rangle, \text{COMPS} \langle \rangle \rangle
\] | verb-lxm |
| ic-srv-lxm | ARG-ST: \[
\text{ARG-ST}
\langle X, \text{VP} \langle \text{INF} + \text{INDEX s} \rangle \rangle
\]
SEM: \[
\text{SEM}
\text{RESTR} \langle \text{ARG s} \rangle
\] | srv-lxm |
| auxv-lxm | SYN: \[
\text{SYN}
\text{HEAD} \text{AUX} +
\] | srv-lxm |
A Lexical Entry for *be*

\[
\langle \text{be}, \langle x, \langle \text{syn} \ \text{head} \ \text{pred} + \rangle \rangle \rangle
\]

\[
\langle \text{arg-st} \ \text{auxv-lxm} \rangle
\]

\[
\langle \text{sem} \ \text{index} \rangle
\]

\[
\langle \text{restr} \rangle
\]
The Entry for *be*, with Inherited Information

![Diagram of syntactic and semantic structures for the verb "be".]
Entry for *have*

- Note the FORM restriction on the complement VP
- What accounts for the analogous FORM restriction on verbs following *be*?
Lexical Entry for a Modal

• What inflectional lexical rules apply to this lexeme?
Accounting for the Basic Facts Cited Earlier

- **Optionality of auxiliaries:**
  As raising verbs, their subjects and complements go together.

- **Auxiliaries precede non-auxiliary verbs:**
  Auxiliaries are heads, and complements follow heads in English.

- **Auxiliaries determine the form of the following verb:**
  This is built into their lexical entries.

- **When auxiliaries co-occur, their order is fixed:**
  Different explanations for different combinations; see next slide.

- **Non-iterability of auxiliaries:**
  Ditto.
Accounting for Restrictions on Order and Iterability

• **Order**
  - Modals are finite, and all auxiliaries take non-finite complements. Hence, modals must come first.
  - Stative verbs (like *own*) don’t have present participles, and auxiliary *have* is stative. Hence, *Pat is having tapdanced*.

• **Iterability**
  - Auxiliary *be* is also stative, so *Pat is being tapdancing*.
  - Modals must be finite, and their complements must be base, so *Pat can should tapdance*.
  - *Pat has had tapdanced* can be ruled out in various ways, e.g. stipulating that auxiliary *have* has no past participle.
Sketch of Chomsky’s Old Analysis

\[ S \to NP \text{ AUX } VP \]
\[ \text{AUX} \to T(M)(\text{PERF})(\text{PROG}) \]

NP

Chris

AUX

past  could  have\text{+en}  be\text{+ing}

VP

eat
How this Analysis Handles the Basic Facts

• Optionality of auxiliaries:
  Stipulated in the phrase structure rule (with parentheses)

• Auxiliaries precede non-auxiliary verbs:
  Built into the phrase structure rule, with AUX before VP

• Auxiliaries determine the form of the following verb:
  Inflections are inserted with the auxiliaries and moved onto
  the following verb transformationally.

• When auxiliaries co-occur, their order is fixed:
  Stipulated in the phrase structure rule for AUX

• Non-iterability of auxiliaries:
  Ditto.
The two analyses assign very different trees

- *could have been VP, have been VP, and been VP* are all constituents
- *could have been* is not a constituent

- *could have been VP, have been VP, and been VP* are not constituents
- *could have been* is a constituent
Ellipsis and Constituency

• Consider:

*Pat couldn’t have been eating garlic, but Chris could have been*  
*Pat couldn’t have been eating garlic, but Chris could have*  
*Pat couldn’t have been eating garlic, but Chris could*

• On the nested analysis, the missing material is a (VP) constituent in each case

• On the flat analysis, the missing material is never a constituent

• This argues for our analysis over the old transformational one.
Our Analysis of Auxiliaries So Far

• Auxiliaries are subject-raising verbs

• Most basic distributional facts about them can be handled through selectional restrictions between auxiliaries and their complements (that is, as ARG-ST constraints)

• Auxiliaries are identified via a HEAD feature AUX, which we have not yet put to use
# Descriptive Summary of the NICE Properties

<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Negation</strong></td>
<td>Sentences are negated by putting <em>not</em> after the first auxiliary verb; they can be reaffirmed by putting <em>too</em> or <em>so</em> in the same position.</td>
</tr>
<tr>
<td><strong>Inversion</strong></td>
<td>Questions are formed by putting an auxiliary verb before the subject NP.</td>
</tr>
<tr>
<td><strong>Contraction</strong></td>
<td>Auxiliary verbs take negated forms, with <em>n’t</em> affixed.</td>
</tr>
<tr>
<td><strong>Ellipsis</strong></td>
<td>Verb phrases immediately following an auxiliary verb can be omitted.</td>
</tr>
</tbody>
</table>
Negation (and Reaffirmation)

- Polar adverbs (sentential *not, so, and too*) appear immediately following an auxiliary
  
  *Pat will not leave*
  *Pat will SO leave*
  *Pat will TOO leave*

- What about examples like *Not many people left*?

- What happens when you want to deny or reaffirm a sentence with no auxiliary?
  
  *Pat left*
  *Pat did not leave*
  *Pat did TOO leave*
The Auxiliary *do*

- Like modals, *do* only occurs in finite contexts:
  *Pat continued to do not leave*

- Unlike modals, *do* cannot be followed by other auxiliaries:
  *Pat did not have left*
The $ADV_{pol}$-Addition Lexical Rule

\[
\begin{aligned}
\text{pi-rule} & \quad \left[ \begin{array}{c}
\text{SYN} \\
\text{HEAD} \\
\text{POL} \\
\text{AUX}
\end{array} \right] \\
\text{ARG-ST} & \quad \left[ \begin{array}{c}
\langle 1 \rangle \\
\oplus \\
\Lambda
\end{array} \right] \\
\text{SEM} & \quad \left[ \begin{array}{c}
\text{INDEX} \\
\langle s_1 \rangle
\end{array} \right]
\end{aligned}
\]

\[
\begin{aligned}
\text{INPUT} & \quad \left[ \begin{array}{c}
\langle X, \\
\text{POL} + \\
\text{AUX} +
\end{array} \right] \\
\text{ARG-ST} & \quad \langle 1 \rangle \\
\text{SEM} & \quad \langle \text{INDEX} \langle s_1 \rangle \rangle \\
\end{aligned}
\]

\[
\begin{aligned}
\text{SYN} & \quad \left[ \begin{array}{c}
\text{HEAD} \\
\text{POL} +
\end{array} \right] \\
\text{VAL} & \quad \left[ \begin{array}{c}
\text{SPR} \\
\langle Z \rangle
\end{array} \right]
\end{aligned}
\]

\[
\begin{aligned}
\text{OUTPUT} & \quad \left[ \begin{array}{c}
\langle Y, \\
\text{POL} + \\
\text{AUX} +
\end{array} \right] \\
\text{ARG-ST} & \quad \langle 1 \rangle \\
\text{SEM} & \quad \langle \text{INDEX} \langle s_2 \rangle \rangle \\
\end{aligned}
\]

\[
\begin{aligned}
\text{SYN} & \quad \left[ \begin{array}{c}
\text{HEAD} \\
\langle \text{INDEX} \langle s_2 \rangle \rangle \\
\text{RESTR} \\
\langle [\text{ARG} \langle s_1 \rangle] \rangle
\end{array} \right] \\
\end{aligned}
\]

\[
\begin{aligned}
\text{SEM} & \quad \langle \text{INDEX} \langle s_2 \rangle \rangle \\
\end{aligned}
\]
What does the type *pi-rule* mean?

- It maps words to words (hence, “post-inflectional”)
- It preserves MOD values, HEAD values as a default, and (like other lexical rule types) SEM values as a default

\[
\text{INPUT} \quad \left\langle / \left[ \begin{array}{c}
\text{syn} \\
\text{sem}
\end{array} \right] \right\rangle \\
\text{OUTPUT} \quad \left\langle / \left[ \begin{array}{c}
\text{syn} \\
\text{sem}
\end{array} \right] \right\rangle
\]
Why doesn’t $\text{ADV}_{pol}$-Addition LR mention VAL?

\[
\begin{align*}
\text{INPUT} & \quad \langle X, \rangle \\
\text{ARG-ST} & \quad \langle 1 \rangle \oplus A \\
\text{SEM} & \quad \left[\text{INDEX} \ s_1\right] \\
\text{OUTPUT} & \quad \langle Y, \rangle \\
\text{ARG-ST} & \quad \langle 1 \rangle \oplus \left[\text{INDEX} \ s_2\right] \\
\text{SEM} & \quad \left[\text{INDEX} \ s_2\right]
\end{align*}
\]
What is the role of these indices?

\[
\begin{align*}
\text{INPUT} & \quad \langle X, \rangle \\
\text{ARG-ST} & \quad \langle \text{INDEX } s_1 \rangle \\
\text{SEM} & \quad \langle \text{INDEX } s_1 \rangle \\
\text{OUTPUT} & \quad \langle Y, \rangle \\
\text{ARG-ST} & \quad \langle \text{INDEX } s_2 \rangle \\
\text{SEM} & \quad \langle \text{INDEX } s_2 \rangle
\end{align*}
\]
Which *nots* does the rule license?

Andy must *not* have been sleeping? ✓
Andy must have *not* been sleeping? ✗
Andy must have been *not sleeping*? ✗
*Kleptomaniacs cannot* *not* steal. ✓
*Kleptomaniacs cannot* *not* steal. ✗
Negation and Reaffirmation: A Sample Tree

S

NP

VP

Leslie

V

did

ADV$_{pol}$

so

eat the whole pizza
Inversion

• Yes-no questions begin with an auxiliary: 
  *Will Robin win?*

• The NP after the auxiliary has all the properties of a subject
  • Agreement:  *Have they left?*  vs.  *Has they left?*
  • Case:  *Have them left?*
  • Raising:  *Will there continue to be food at the meetings?*

• What happens if you make a question out of a sentence without an auxiliary? 
  *Robin won*
  *Did Robin win?*
The Inversion Lexical Rule

\[
\begin{align*}
\text{INPUT} & \quad \left( \langle W, \text{arg-st} A, \text{sem} \text{[MODE prop]} \rangle, \text{syn} \text{[val SPR [X]]} \right) \\
\text{OUTPUT} & \quad \left( \langle Z, \text{arg-st} A, \text{sem} \text{[MODE ques]} \rangle, \text{syn} \text{[head INV +] \text{val SPR [ ]}} \right)
\end{align*}
\]
How the Rule Yields Inverted Order

\[ pi\text{-rule} \]

INPUT \[ \langle W, \text{SYN} \rangle \]
- ARG-ST \[ A \]
- SEM \[ \text{SEM} \]
- VAL \[ \text{VAL} \]
- HEAD \[ \text{HEAD} \]

\[ \text{FORM} \text{ fin} + \]
\[ \text{SPR} \langle X \rangle \]

OUTPUT \[ \langle Z, \text{SYN} \rangle \]
- ARG-ST \[ A \]
- SEM \[ \text{SEM} \]
- VAL \[ \text{VAL} \]
- HEAD \[ \text{HEAD} \]

\[ \text{INV} + \]
\[ \text{SPR} \langle \rangle \]

...plus the ARP
The Feature INV

• What is the INV value of inputs to the Inversion LR?
  • Perhaps surprisingly, the input is [INV +]
  • Word-to-word rules (pi-rules) have default identity of HEAD features, and no INV value is given on the input

• Then what work is the feature doing?
  • It’s used to mark auxiliaries that can’t or must be inverted
    \[
    \text{You better watch out } \quad \text{vs. } \quad \text{*Better you watch out}
    \]
    \[
    \text{I shall go } \quad (\text{shall } \sim \text{‘will’}) \quad \text{vs. } \quad \text{Shall I go? } \quad (\text{shall } \sim \text{‘should’})
    \]
Other Cases of Inversion

• Inversion is not limited to questions
  • Preposed negatives: *Never have I been so upset!*
  • Conditionals: *Had we known, we would have left.*
  • Exclamations: *May your teeth fall out!*

• Does our rule account for these?
  • No. Our rule’s output says [MODE ques]. And each construction has slightly different idiosyncrasies.

• How might we extend our analysis to cover them?
  • Define a type of inversion lexical rules, sharing certain properties, but with some differences.
Inversion: A Sample Tree

S
  /\  /
 V NP VP
 /     /
Did Leslie eat the entire pizza?
Contraction

• There are several types of contraction in English, but we’re only talking about words ending in *n’t

• It may seem like just *not said fast, but there’s more to it
  • Only finite verbs can take *n’t:
    *Terry must haven’t seen us

• There are morphological irregularities:
  won’t, not *willn’t %shan’t, not *shalln’t
  mustn’t pronounced mussn’t
don’t pronounced doesn’t, not dewn’t
  *amn’t
The Contraction Lexical Rule

\[
p_{i-rule}:
\begin{align*}
\text{INPUT} & \quad \left\langle 2, \begin{bmatrix}
\text{SYN} & \text{HEAD} \\
\text{ARG-ST} & \text{SEM} \\
\end{bmatrix}
\begin{bmatrix}
\verb FORM \\ 
\verb AUX + \\
\verb POL - \\
\end{bmatrix}
\right\rangle \\
\text{OUTPUT} & \quad \left\langle \text{F}_{\text{NEG}}(2), \begin{bmatrix}
\text{SYN} & \text{VAL} \\
\text{ARG-ST} & \text{SEM} \\
\end{bmatrix}
\begin{bmatrix}
\verb POL + \\
\verb SPR \left\langle X \right\rangle \\
\end{bmatrix}
\right\rangle
\end{align*}
\]
Most of the work is in the semantics

Why?
What does POL do?

\[
\begin{align*}
\text{pi-rule} & \\
\text{INPUT} & \langle 2, \text{ARG-ST } [\text{B}] \rangle \\
\text{SEM} & \langle \text{INDEX } s_1, \text{RESTR } [\text{A}] \rangle \\
\text{OUTPUT} & \langle F_{\text{NEG}(2)}, \text{ARG-ST } [\text{B}] \rangle \\
\text{SEM} & \langle \text{INDEX } s_2, \text{RESTR } \langle \text{RELN} \text{ not } s_2, \text{SIT } s_2 \rangle \oplus [\text{A}] \rangle
\end{align*}
\]

\*We can’tn’t stop
\*They won’t TOO mind
Contraction: Sample Tree

S

NP

Leslie

V

wouldn’t

VP

eat the entire pizza
Ellipsis

• Ellipsis allows VPs to be omitted, so long as they would have been preceded by an auxiliary

  *Pat couldn’t have been watching us, but Chris could have been watching us.*

• Unlike the other NICE properties, this holds of all auxiliaries, not just finite ones.

• What is the elliptical counterpart to a sentence with no auxiliary?

  *Whenever Pat watches TV, Chris watches TV Whenever Pat watches TV, Chris does*
The Ellipsis Lexical Rule

\[
\begin{align*}
\text{INPUT} & \quad \left\langle 1, \left[ \text{auxv-lxm} \right. \right. \\
& \quad \left. \text{ARG-ST} \left\langle 2 \right\rangle \oplus \text{A} \right]\rightangle \\
\text{OUTPUT} & \quad \left\langle 1, \left[ \text{dervv-lxm} \right. \right. \\
& \quad \left. \text{ARG-ST} \left\langle 2 \right\rangle \right]\rightangle
\end{align*}
\]

• Note that this is a derivational LR (\textit{d-rule}) -- that is, lexeme-to-lexeme

• This means that SYN and SEM are unchanged, by default
Ellipsis: A Sample Output

\[
\begin{array}{c}
\langle \text{could} , \\
\text{ARG-ST} \\
\text{SEM} \\
\rangle
\end{array}
\]

\[
\begin{array}{c}
\text{auxv-lxm} \\
\text{SYN} \\
\text{SEM} \\
\end{array}
\]

\[
\begin{array}{c}
\langle \text{could} , \\
\text{ARG-ST} \\
\text{SEM} \\
\rangle
\end{array}
\]

\[
\begin{array}{c}
\langle \text{could} , \\
\text{ARG-ST} \\
\text{SEM} \\
\rangle
\end{array}
\]

\[
\begin{array}{c}
\langle \text{could} , \\
\text{ARG-ST} \\
\text{SEM} \\
\rangle
\end{array}
\]

\[
\begin{array}{c}
\langle \text{could} , \\
\text{ARG-ST} \\
\text{SEM} \\
\rangle
\end{array}
\]
Ellipsis: A Sample Tree

S
  /\   \\
NP  VP
  |    |
Kim V    VP
  |    |
could V    VP
    |    |
   have V    VP
    |    |
   been V    VP
    |    |
attending the conference
Semantics of Ellipsis

S

NP   VP

Kim   could

What is the SEM value of the S node of this tree?

\[
\begin{bmatrix}
\text{INDEX} & s_1 \\
\text{MODE} & \text{prop} \\
\text{RESTR} & \begin{bmatrix}
\text{RELN} & \text{name} \\
\text{NAME} & \text{Kim} \\
\text{NAMED} & i \\
\text{SIT} & s_1 \\
\text{ARG} & s_2 
\end{bmatrix}
\end{bmatrix}
\]

Note: $s_2$ has to be filled in by context.
Infinitival to Revisited

• VP Ellipsis can occur after *to*:

   We didn’t find the solution, but we tried *to*.

• This is covered by our Ellipsis LR if we say *to* is [AUX +].

• Since AUX is declared on type *verb*, it follows that *to* is a verb.
do Revisited

• Chomsky’s old analysis: in sentences w/o auxiliaries...
  • Tense can get separated from the verb in various ways
    • Negation/Reaffirmation inserts something between Tense and the following verb
    • Inversion moves Tense to the left of the subject NP
    • Ellipsis deletes what follows Tense
  • When this happens, *do* is inserted to support Tense

• Our counterpart:
  • NICE properties hold only of auxiliaries
  • *do* is a semantically empty auxiliary, so negated, reaffirmed, inverted, and elliptical sentences that are the semantic counterparts to sentences w/o auxiliaries are ones with *do*. 
Summary

• Our analysis employs straightforward mechanisms
  • Lexical entries for auxiliaries
  • 3 new features (AUX, POL, INV)
  • 4 lexical rules

• We handle a complex array of facts
  • co-occurrence restrictions (ordering & iteration)
  • the NICE properties
  • auxiliary *do*
  • combinations of NICE constructions