Syntax in Language Production:

An Approach Using Tree-Adjoining Grammars

by Fernanda Ferreira (1999)

Gerald Schoch
May 28, 2011
Outline

- Introduction
  - Production of Language and Syntax

- Tree-Adjoining Grammar
  - Elementary trees and operations

- Aspects of Syntactic Production – captured with TAG

- Model of Syntactic Production – based on TAG
  - Implications for incrementality
Lemmas are retrieved and assigned grammatical functions (subject, object, ...)

Serial order of phrases
Order of elements within any given phrase
Inflectional processing
Incrementality is assumed!
Simone was eating tuna yesterday.

- Decisions about word order

- Constraints:
  - *Eating*: requires appropriate subject and object
  - Subject before and object after verb
  - *Yesterday*: beginning or end of the sentence
  - *Tuna*: object or subject (requires passive)
Introduction

• Syntactic information for these decisions:
  consulted *quickly* and *efficiently*

  How is this speed and efficiency accomplished?

• Why active form rather than passive?
• How to manage agreement between *to be* and *Simone*?

  How are these decisions made?
How do speakers make syntactic decisions?

• Considering psychological mechanisms underlying the ability to combine words to form appropriate sentences

→ Approach: Tree-Adjoining Grammar (TAG)
Grammar:
- set of objects
- set of operations for object manipulation

Objects: *elementary trees*
- Primitive *syntactic units* consisting of
  - Lexical head
  - argument(s) licensed by the head
Two types of elementary trees:

- **Auxiliary** tree:
  - Root node **identical** to one of the non-terminal nodes

  ➔ Recursion
Two types of elementary trees:

- **Initial** trees:
  - All elementary trees that are **not auxiliaries**
  - Do **not** permit recursion
• **Substitution**
  - attaching one elementary tree to **bottom node** of another one

• **Restriction:** root node **matches** bottom node
• Adjoining:

• inserting elementary tree inside another one
● Primitive syntactic objects (*elementary trees*)
  
  ● retrieved as single chunk
  
  ● Containing all *dependency* relations
e.g., relation between *head* as verb and its *arguments*
  
  ● Information about *sorts* of further *syntactic entities*
e.g., NP needed for subject position
  
  ● Operations: *substitution* and *adjoining*
Outline

- Introduction
  - Production of Language and Syntax

- Tree-Adjoining Grammar
  - Elementary trees and operations

- Aspects of Syntactic Production – captured with TAG

- Model of Syntactic Production – based on TAG
  - Implications for incrementality
Syntactic Production

- Using TAG to describe syntactic production
  - Lexical influences on syntactic form
  - Syntactic priming
  - Subject-verb agreement

- Implications for the assumption of the **incrementality** of language production
Using TAG to describe syntactic production

- Lexical influences on syntactic form
- Syntactic priming
- Subject-verb agreement

Implications for the assumption of the *incrementality* of language production
Tom quoted Mary.
Mary was quoted by Tom.

- Same idea, expressed differently

What factors influence the decision to choose one of these structures during the on-line production?
Syntactic Production - Lexical Influences

- Syntactic form influenced by availability* of concepts
  - More available concepts tend to be subject
  - Rest of the structure is adjusted appropriately

→ quote: if agent (Tom) is more available than the patient (Mary), agent is in subject position

(*) ”available”: concepts that are more prototypical, more concrete, more animate, generally more activated
Syntactic Production - Lexical Influences

- Example:
  - **Patient**: highly **available** (topic)
  - Production system begins working on it
    - Principle of **incrementality**!

- Grammatical encoder:
  - **first thing** it can do: **entity** = **subject**

- Few options for encoding the rest:
  - subject – verb – object

- Patient = Subject → overall structure **passive**:
  - *Mary was quoted by Tom*.
Lexical Influences - TAG

Can TAG describe this more precisely?

- Propositional representation of the idea:

quote(Tom: agent, Mary: patient, PAST)

- Assuming MARY as highly available, it can immediately be syntactically encoded
Concept **QUOTE** constrains encoder to select

- an **elementary tree** headed by *quote*

And

- the information that patient *Mary* has already been **encoded** as **subject** and **requires** passive
Substitution:

- Principle of **incrementality**: substitution at *earliest position possible* → subject position
Principle of incrementality:

- Insertion of NP *Mary*: phonological encoder begins to work, converting syntactic structure into suitable output

- Syntactic encoder still works on the remaining parts

- Syntactic representation done: phonological representation is nearly complete
Syntactic Production

- Using TAG to describe syntactic production
  - Lexical influences on syntactic form
  - Syntactic priming
  - Subject-verb agreement

- Implications for the assumption of the **incrementality** of language production
Syntactic Production - Syntactic Priming

- Tendency to repeat a particular syntactic form

- Example:
  
  Speaker just described a transitive action using passive

  ➔ Subsequent transitive event is likely to be passive too
  (Bock, 1986):

  The referee was punched by one of the fans.

  The church is bring struck by lightning.
Implications of these results:

- **Challenging extreme forms** of incremental production

- **Point** during production where the entire syntactic form of a sentence can be influenced by its prior presentation

  - If a syntactic structure is simply built up in little bits, immediately converted into phonological units:

    when is a syntactic representation available to be primed?
Assumption because of syntactic priming effect:

- **Point in syntactic encoding** where a **large chunk** of syntactic structure is **simultaneously available**

- Explanation with a model based on TAG
Syntactic Priming - TAG

- **Availability** of verb
  - → **availability** of entire clause's **overall syntactic form**
    - active/passive, preposition/double-object dative, …

- Syntactic Priming **independent** of **semantic** content
  - Expected on model:
    - elementary **tree headed by verb** may **not include internal content** of any arguments in the tree

- Only thing that may be primed:
  - number, configuration, max. projection labels of verb's arguments
Syntactic Priming - TAG

- TAG-based model provides an account of SP effect:
  - **Elementary trees** can be **primed**

- Prediction:
  - not just **clausal trees** (i.e., trees headed by verbs) may be **primed**, but **other structures** as well
    - e.g., ADJ before N
      - (testing not possible in English: strict word order)

- Surface order, tested in Dutch (picture description task):
  - *A ball is on the table.* vs. *On the table is a ball.*
    - Expected by TAG: each order with own elementary tree (although both headed by *is*)
Further concept in TAG: “families”
- **Clusters** of related elementary trees, i.e.:
  - **Ditransitive** elementary trees including NP + PP as post-verbal arguments: *He gave a ball to the cat.*
  - **Variations** on the same basic tree headed by the same lemma (i.e., same verb with different tenses, aspects)

- Priming would occur across similar trees
- Similarity relations captured with “families”
Syntactic Production

- Using TAG to describe syntactic production
  - Lexical influences on syntactic form
  - Syntactic priming
  - Subject-verb agreement

- Implications for the assumption of the \textit{incrementality} of language production
Subject-Verb Agreement

- Agreement between *Subject* and *Verb*, e.g.
  - *The report* and *to have* or *to be* (number)

- Agreement errors in sentence completing experiments:
  - More errors with phrases like
    
    *The report of the destructive fires* (PP)

    - as with phrases like

    *The report that they controlled the fires* (relative clause)
The report of the destructive fires

- *report* takes PP as argument
- elementary tree for NP includes the PP
- *fires* part of same elementary tree headed by *report*

The report that they controlled the fires

- relative clause merely modifier of *report*
- not in the same elementary tree
- *fires* in different elementary tree (head: *control*)
- inserted by substitution
More agreement errors with the PP-construction
  e.g., *The report of the fires are* ...

Head and local noun part of the same structure
  - Simultaneously available
    (in contrast to the relative clause construction!)
  - Plural feature of *fires* could end up on head noun

Explanation for more agreement errors
Outline

- Introduction
  - Production of Language and Syntax

- Tree-Adjoining Grammar
  - Elementary trees and operations

- Aspects of Syntactic Production – captured with TAG

- Model of Syntactic Production – based on TAG
  - Implications for incrementality
Production of Syntax – based on TAG

- Critical assumptions of the TAG model:

  - **Syntactic structure** built up by **primitive syntactic templates**
  
  - Each template **based** on a **single lexical item**
  
  - Templates **retrieved** when its **head** is **activated**
  
  - Head: template's **only** primitive lexical **content**
  
  - Other material: inserted by a **operation**
  
  - Other lexical items: **bound** to **appropriate** syntactic **positions**
  
  - **Incrementality**: insertion at the **earliest** possible **point**
Example:

*The dog bit a flower.*

Propositional representation

```
event: BITE(def/1/agent/topic: DOG;
    indef/1/patient: FLOWER;
    past)
```
Syntactic Production based on TAG - Example

Activated first: DOG (topic)

Retrieval of lemma for DOG (sg/def)

Agent: checked off as grammatically encoded

NP placed in syntactic buffer, awaiting retrieval of clausal tree

event: BITE(DOG, FLOWER)
Syntactic Production based on TAG - Example

Assumed as next activated: verb

Retrieval of BITE (past)

Active form: agent has been already encoded

event: BITE(DOG, FLOWER)
Syntactic Production based on TAG - Example

NP (the dog) in syntactic buffer

Incrementality: NP in the leftmost NP slot

The dog encoded as subject

event: BITE(DOG, FLOWER)
First entity of sentence encoded

Piece of utterance (S+V)
sent for phonological encoding

Retrieval of lemma for FLOWER
(sg/indef)

Indefinite NP structure

Inserted in the last remaining NP slot

event: BITE(DOG, FLOWER)
Grammatical encoding of the sentence is complete!

event: BITE(DOG, FLOWER)
Another example:

- advantages of assuming only a **moderate degree** of incrementality

event: PUT (def/1/agent: MAN; def/1/theme: BODY; def/1/location/topic: TRUNK; Past)

Idea:

“a particular trunk was the location in which a singular male placed a body”
Syntactic Production based on TAG – Example 2

Available first: TRUNK (topic)

LOCATION checked off as encoded

NP placed in syntactic buffer

event: PUT(MAN, BODY, TRUNK)
Syntactic Production based on TAG – Example 2

- Assumption: lemma for PUT becomes available

- LOCATION encoded first:
  - Retrieval of two lemmas (and trees) for PUT
  - Active and passive!

- Lexical semantics of *put*:
  - LOCATION is not allowed to be subject

(N.B.: *contain* allows this: *The trunk contains the body.*)

(event: PUT(MAN, BODY, TRUNK))
Syntactic Production based on TAG – Example 2

- Two trees available in parallel
- Wait for another argument to be encoded

<table>
<thead>
<tr>
<th>Agent</th>
<th>Patient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incrementality</td>
<td>Incrementality</td>
</tr>
<tr>
<td>Leftmost position of <strong>active</strong> structure</td>
<td>Leftmost position of <strong>passive</strong> structure</td>
</tr>
</tbody>
</table>

- The structure that is not chosen loses its activation

event: PUT(MAN, BODY, TRUNK)
Syntactic Production based on TAG – Example 2

- Example 2: moderate degree of incrementality

- With extreme degree of incrementality:
  
  - System would not wait for the verb
  - **Nominal** entities *immediately* made into **subjects**

  -> **Ungrammatical** utterances, e.g.

  *The trunk was put the body by the man*
Syntactic Production based on TAG

- Syntactic encoding not necessarily a serial process
  
  - All structures compatible with a lemma are activated at one time
  
  - As more information available: competing lemmas drop out until one structure is left when encoding is complete
  
  - Two nominal lemmas equally available: speaker might be disfluent
Syntactic Production based on TAG - Conclusion

- Utterances: generated from **propositional representations**

- Concepts: **differentially** activated

- Topic: **most available** concept
  - Most affinity for **subject** position

- Verb:
  - Determines verb **lemma** (active/passive, dative, …) and
  - Retrieval of **elementary tree(s)**
Syntactic Production based on TAG - Conclusion

- As grammatical encoding enfolds:
  - Remaining of one activated clausal elementary tree
    - Determines form of the sentence
  - Elementary trees others than clausal trees:
    - Must be inserted into clausal tree
    - Order: determined by availability
Conclusion

- Tree-Adjoining Grammar

- TAG for capturing aspects of Syntactic Production

- Model for Syntactic Production based on TAG
  - Incrementality
  - Propositional representations
  - Different activations of concepts
  - Simultaneously available trees

Fernanda Ferreira and Paul Engelhardt (2006). Syntax and Production