ወር እርስል

赊ት ስንስ ሳንስ
This is the malt that the cat that the dog
built.

This is the dog, that worried the cat, that Jack
rat, that ate the malt, that lay in the house that

or sentences more difficult to understand.

Linguistic complexity: what makes individual constructions

languages

the resources needed in order to process classes of

Computational complexity: the expressive power of (and

Complexity of natural languages

---

Natural Language Processing
are. are, the lower in the hierarchy the languages they generate generalizing all its languages. The more restricted the rules and the format of phrase-structure rules necessary for there exists a correspondence between the class of languages includes the languages in lower classes. In particular, the class of languages in one class properly the hierarchy, the more “complex” it is. ordered by their “complexity”. The higher the language is in a hierarchy of classes of languages, viewed as sets of strings, 

The Chomsky Hierarchy of Languages

-------------------------- Natural Language Processing --------------------------
The Chomsky Hierarchy of Languages
The Chomsky hierarchy of languages

Regular (type-3) languages:

**Grammar:** right-linear or left-linear grammars

**Rule form:** $A \rightarrow \alpha$ where $A \in V$ and $\alpha \in \Sigma^* \cdot V \cup \{\epsilon\}$.

**Computational device:** finite-state automata
The Chomsky hierarchy of languages

Computational device: push-down automata

Context-free grammars

Rule form: \( A \rightarrow a \) where \( A \in \sum \) and \( a \in \Sigma \)

Context-free (type-2) languages
Length of the input string (m) machines with a finite tape, linear-bounded by the
Computational device: Linear-bounded automata (Turing
Rule form: \( a \Rightarrow \gamma \) where \( \gamma \leftarrow a \)

Grammar: context-sensitive grammars

Language: context-sensitive (type-1) languages

The Chomsky hierarchy of languages
Computational device: Turing machines

Rule form: $a \rightarrow b$ where $a \neq e$

General Rewriting systems

Recursively-enumerable (type-0) languages

The Chomsky hierarchy of languages
language

It can shed light on questions of human processing of

It can help accept or reject linguistic theories

complexity of natural languages

The hierarchy represents some informal notion of the

Why is it interesting?

Where are natural languages located?
worried killed ate.

This is the malt that the rat that the cat that the dog
built.

rat, that ate the malt, that lay in the house that Jack
this is the dog, that worried the cat, that killed the

Competition vs. Performance

language? Is it context-free? How about Hebrew?
When viewed as a set of strings, is English a regular

What exactly is the question?

Where are natural languages located?
English is itself literally outside the range of such analyses.

As for context-free languages, "I do not know whether or not "English is not a regular language" "English is not a natural language" Chomsky (1957): Where are natural languages located?"
introduce a new kind of rule that will permit us to do so.

the significant aspects of language structure. We must
phrase structure grammars cannot possibly represent all
one language, namely English, we can be sure that
to represent an obviously significant generalization about
since there seem to be no way of using such PS rules

(Asman and Heny, 1976):
An introduction to the principles of transformational syntax

How not to do it
to generate English. Hence, the phrase structure analysis will not be sufficient.
there will always be an infinite number of such sentences. Grammar will not generate. In fact, because of recursion
will always be possible to construct a sentence that the
finite number of rules like (2.5), (2.52) and (2.54) it
in general, for any phrase structure grammar containing

Syntax (Peter Culicover, 1976):

How not to do it
transformational grammar (Grinder & Elgin, 1973)

...clearly the inadequacy of ... context-free phrase-structure
this well-known syntactic phenomenon demonstrates

*the girl kiss the boy
the girl saw the boy

How not to do it
sequences the girl kiss the boy and the girls kiss the boy. sequences the girl kiss the boy will also generate (incorrectly) the sequence the girl. In other words, any set of context-free rules that generate the symbol into the sequence the girls and that will expand the symbol by a kiss in the context of being immediately preceded by the symbol cannot write a context-free rule that will expand the symbol to the context in which it occurs. Thus, by definition, one defining characteristic of a context-free rule is that the

How not to do it
English is not a context-free language. Agreement is sufficient to guarantee the accuracy of the grammatical phenomenon of Subject-Predicate. How not to do it.
verb. (or even be immediately adjacent) to the present-tense
noun phrase that determines the agreement will precede
structure grammar, since we cannot guarantee that the
a simple agreement rule of the type given in a phrase-
English number agreement is literally impossible using
These examples show that to describe the facts of
Syntactic theory (Bach 1974):

How not to do it
correctly describable as a finite string of phrases. For the possible context is not
phrase-structure rules, cannot be adequately described even by context-sensitive
or patience permits... the distant type of agreement... this type of syntaxic dependency can extend as far as memory
of a noun phrase at some distance from it... this type in many cases the number of a verb agrees with that

A realistic transformational grammar (Bresnan, 1978):

How not to do it
The notion of "context-freeness"

What is the source for this confusion?

How not to do it
• Closure under homomorphism
• Closure under intersection
• The pumping lemma for regular languages

Proof techniques:

How to do it right
A \text{ white male hired another white male.}

A \text{ white male hired another white male.}

A \text{ white male hired another white male.}

The following is a sequence of grammatical English sentences:

\text{Center embedding: English is not a regular language}.
Since $L_{\text{reg}}$ is trans-regular, English is not a regular language. 

Since $L_{\text{reg}}$ is trans-regular, $L_{\text{reg}}$ is intersected with $L_{\text{reg}}$, namely $L_{\text{reg}}$, would be regular. 

Since $L_{\text{reg}}$ is a regular language, then if English were regular, 

Since the regular languages are closed under intersection, and 

$L_{\text{reg}}$ is regular, as it is defined by a regular expression, 

\[
\{ \text{hired another white male} \} \cap \{ \text{hired a white male whom a white male} \} = L_{\text{reg}}
\]

the regular set 

$L_{\text{reg}}$ is the intersection of the natural language English with 

$L_{\text{reg}}$ is not a regular language 

English is not a regular language
The cat the dog the rat the elephant admired but chased. Likes tuna fish.
\{\{q + a\} \in x \mid \exists xx\} \text{ the trans-context-free language}

or to \( q \), and all other words to \( e \), English can be mapped to

\text{English is not a regular language}

\text{Another construction:}

\text{either } S_3 \text{ or } S_4

\text{if } S_1 \text{ then } S_2
Some languages that are not context-free:

\[
\{ uPw \# uq \# v \}
\]

\[
\{ \ast \{ q + v \} \ni x \mid xx \}
\]

Languages under homomorphisms and under intersection with regular languages, and two closure properties: closure is not context-free is the pumping lemma for context-free languages. The common proof technique for showing that a language is English context-free?
we have wanted to let the kids help Hans paint the house~

we have wanted to let help paint

we have let the kids Hans-DAT the house-ACC have

mer de'child em Hans es hus mer

we helped Hans-DAT the house-ACC helped paint

mer em Hans es hus mer de'child em Hans-DAT the house-ACC have

jan sain das (jan said that)

Data from Swiss-German:

Natural Languages are not context-free
which is trans-context-free.

In half-astrichie
Jan sagt das, mer (d’chind)’n (en Hans)’n es hius hand welle

the above regular expression with Swiss-German yields

similarly for accusatives. Intersecting the language defined by
must equal the number of dative NPs (en Hans)’n, and
however, the number of verbs requiring dative objects (half-e)

verbs must precede accusative-taking verbs:

Dative NPs must precede accusative NPs, and dative-taking

Natural languages are not context-free.
Limitations on stack size?
The cat the dog the rat bit chased likes tuna fish
others?

Why are some sentences more difficult to understand than

Linguistic complexity
sentence.

Two grammars are strongly equivalent if they are weakly equivalent and, in addition, assign the same structure to each string. When grammars for natural languages are involved, we say that $G_1$ and $G_2$ are weakly equivalent if their string languages are identical.

If formal language theory, the natural equivalence relation on weak and strong generative capacity
capacity is probably not sufficient for natural languages. Even if context-free grammars are weakly sufficient for
formalisms is its ability to associate structure to strings.

The strong generative capacity (or power) of a linguistic weak and strong generative capacity