Complexity in grammar

Komplexität im Lexikon: Jackendoff (1975)

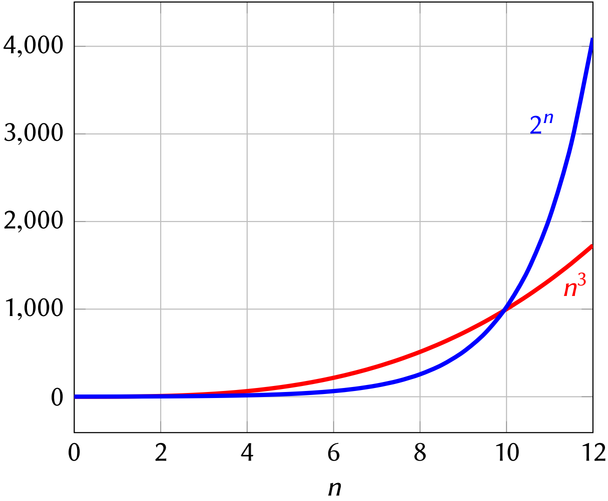
Timm Lichte

HHU Düsseldorf

WS 2015/2016, 11.11.2015

Kritik der algorithmischen Komplexitätsbegriffe

Berwick & Weinberg (1984); Pollard (1996)



Diese Sitzung

Jackendoff, Ray. 1975. Morphological and Semantic Regularities in the Lexicon. Language 51(3). 639–671.

(2011, Quelle: Wikipedia)

Intro

The starting point of the **Lexicalist Hypothesis**, proposed in Chomsky's 'Remarks on nominalization' (1970), is the rejection of the position that a nominal such as ***Bill's decision*** to go is derived transformationally from a sentence such as ***Bill decided to go***.

Rather, Chomsky proposes that the nominal is generated by the base rules as an NP, no S node appearing in its derivation. His paper is concerned with the consequences of this position for the syntactic component of the grammar.

The present paper will develop a more highly articulated theory of the lexical treatment of nominals, show that it is independently necessary, and extend it to a wide range of cases other than nominalizations. (S.639)

Überblick

1. “Levels of adequacy in description”
   - observational/descriptive/explanatory adequacy
2. “Formulation of two preliminary theories”
   - transformational/empoverished-entry/full-entry theory
3. “Which theory?”
4. “Separate morphological and semantic rules”
5. “Other applications”
   - Präfixverben, Komposita, kausative Verben, Idiome
6. “The cost of refering to redundancy rules”
7. “Creativity in the lexicon and its implications”

Levels of adequacy of grammars/theories

(following Chomsky 1965)

1. **Observational adequacy**
   - correct enumeration of the set of ~~sentences~~ lexical items in a language
2. **Descriptive adequacy**
   - relationships, sub-regularities, and generalizations among lexical items of the language
   - Beispiel: *decide* and *decision* are related.
   - Beispiel: *decide* is more 'basic' than *decision*.
3. **Explanatory adequacy**
   - Why the chosen relationships in the description?

   ⇒ **evaluation measures:**
   - typically length of grammar
   - here "independent information content"

Komplexität!

Formulation of two preliminary theories

Transformational theory (TG)

- *John decided to go* → *John's decision to go*
- contra **Lexicalist Hypothesis** (=no transformation between word forms in syntax)

Impoverished-entry theory (IET)

- *decide* has a full entry; *decision* has an impoverished entry.
- **redundancy rules** expand impoverished entries during lexical insertion

Beispiele für Einträge für *decide* und *decision*

(2) $\begin{bmatrix} \text{784} \\ \text{/decīd/} \\ +\text{V} \\ +[\text{NP}_1\underline{\quad\quad}\ \textit{on}\ \text{NP}_2] \\ \text{NP}_1\ \textsc{decide on}\ \text{NP}_2 \end{bmatrix}$

(4) $\begin{bmatrix} \text{375} \\ \text{derived from 784} \\ \quad\text{by rule 3} \end{bmatrix}$

Impoverished-entry theory (IET)

- *decide* has a full entry; *decision* has an impoverished entry.
- **redundancy rules** expand impoverished entries during lexical insertion

Beispiel einer Redundanzregel

$$(3)\quad \left[\begin{array}{l} x \\ /y + \text{ion}/ \\ +\text{N} \\ +[\text{NP}_1\text{'}s\ \underline{\hphantom{xxxx}}\ (\text{P})\ \text{NP}_2] \\ \text{ABSTRACT RESULT OF ACT} \\ \quad \text{OF NP}_1\text{'S } Z\text{-ING NP}_2 \end{array}\right] \leftrightarrow \left[\begin{array}{l} w \\ /y/ \\ +\text{V} \\ +[\text{NP}_1\ \underline{\hphantom{xxxx}}\ (\text{P})\ \text{NP}_2] \\ \text{NP}_1\ Z\ \text{NP}_2 \\ \\ \end{array}\right]$$

[…] the two-way arrow may be read as the **symmetric relation** ‘is lexically related to’. The rule thus can be read: ‘A lexical entry *x* having such-and-such properties is related to a lexical entry *w* having such-and-such properties. (S. 642)

Formulation of two preliminary theories

Full-entry theory (FET)

- *decide* and *decision* have fully specified lexical entries.

Beispiel für *decision*

(5) $\left[\begin{array}{l} /\text{decīd} + \text{ion}/ \\ +\text{N} \\ +[\text{NP}_1\text{'}\textit{s} \_\_\_ \textit{ on } \text{NP}_2] \\ \text{ABSTRACT RESULT OF ACT OF} \\ \quad \text{NP}_1\text{'S DECIDING NP}_2 \end{array}\right]$

- The redundancy rule plays no part in the derivation of sentences.
- "Rather, the redundancy rule plays a role in the **information measure** for the lexicon. It designates as redundant that information in a lexical entry which is **predictable** by the existence of a related lexical item; redundant information will not be counted as independent." (S.643)

Formulation of two preliminary theories

Full-entry theory (FET)

(6) (Information measure)
Given a fully specified lexical entry *W* to be introduced into the lexicon, the independent information it adds to the lexicon is

  (a) the information that *W* exists in the lexicon, i.e. that *W* is a word of the language; plus
  (b) all the information in *W* which cannot be predicted by the existence of some redundancy rule *R* which permits *W* to be partially described in terms of information already in the lexicon; plus
  (c) the cost of referring to the redundancy rule *R*.

⇒ Reihenfolge der Worteinfügungen ins Lexikon ist entscheidend für “information measure”.

Formulation of two preliminary theories

Full-entry theory (FET)

⇒ Reihenfolge der Worteinfügungen ins Lexikon ist entscheidend für "information measure".

> Given 2 in the lexicon, now let us add 5. Since its lexical entry is completely predictable from 2 and redundancy rule 3, its cost is the information that a word exists plus the cost of referring to 3, which is presumably less than the cost of all the information in 5. Thus the cost of adding the pair *decide-decision* is the information that two words exist, plus the total information of the entry 2, plus the cost of referring to redundancy rule 3.

> [...] if we add *decision* first, then *decide*, we arrive at a different sum: the information that two words exist, plus the information contained in 5, plus the cost of referring to redundancy rule 3 (operating in the opposite direction). This is more than the previous sum, since 5 contains more information than 2 [...]. (S.644)

Formulation of two preliminary theories

Full-entry theory (FET)

(7) (Information content of the lexicon)
Given a lexicon L containing n entries, $W_1, \ldots, W_n$, each permutation P of the integers $1, \ldots, n$ determines an order A_p in which $W_1, \ldots, W_n$, can be introduced into L. For each ordering A_p, introduce the words one by one and add up the information specified piecemeal by procedure 6, to get a sum S_p. The independent information content of the lexicon L is the least of the $n!$ sums S_p, plus the information content of the redundancy rules.

(8) (Full-entry theory evaluation measure)
Of two lexicons describing the same data, that with a lower information content is more highly valued.

Which theory?

Affigierung ist unregelmäßig:

- *aggression*, *retribution*, *fission*
- **aggress*, **retribute*, **fiss*

Transformational theory:

- Obligatheitsmarkierung (EXCEPTION-Merkmal):

$$^{*}\textit{fiss}_{\text{EXC}} \rightarrow \textit{fissation}$$

- “ […] it claims that English would be simpler if **fiss* were a word, since one would not have to learn that it is exceptional.” (S. 646)

Which theory?

Affigierung ist unregelmäßig:

- *aggression*, *retribution*, *fission*
- **aggress*, **retribute*, **fiss*

Impoverished-entry theory:

- Option 1: Annahme von Pseudo-Wörtern

  $^{*}\textit{retribute}_{[-\text{ Lexical Insertion}]} \leftrightarrow \textit{retribution}$

- Option 2: Einbettungsansatz

(9)
$$\begin{bmatrix} 511 \\ \text{derived by rule 3 from} \\ \quad \begin{bmatrix} /\text{retribūt}/ \\ +\text{V} \\ +[\text{NP}_1 \;\_\_\_\; \textit{for} \;\; \text{NP}_2] \\ \text{NP}_2 \;\; \text{RETRIBUTE} \;\; \text{NP}_2 \end{bmatrix} \end{bmatrix}$$

Which theory?

Affigierung ist unregelmäßig:

- *aggression*, *retribution*, *fission*
- **aggress*, **retribute*, **fiss*

Impoverished-entry theory:

- Option 1: Annahme von Pseudo-Wörtern

  **retribute*$_{[-\text{ Lexical Insertion}]}$ $\leftrightarrow$ *retribution*
- Option 2: Einbettungsansatz
  - Problem:

    {*aggression*, *aggressive*, *aggressor*},

    {*aviation*, *aviator*},

    {*retribution*, *retributiv*}

    **aggress*, **aviat*, **retribute*

  (a) Redundanz/fehlende Generalisierung

  (b) arbiträre Auswahl einer Basisform (qua Regelsequenz)

Which theory?

Full-entry theory (FET)

> Note that 6b, the measure of non-redundant information in the lexical entry, is cleverly worded so as to depend on the existence of redundant information **somewhere** in the lexicon, but not necessarily on the existence of related lexical entries. (S. 648)

> …
>
> (6b) all the information in *W* which cannot be predicted by the existence of some redundancy rule *R* which permits *W* to be partially described in terms of information already in the lexicon; plus
>
> …

- *perdition* $\leftrightarrow_3$ [**perdite*]
  - ⇒ *perdition* komplexer als *damnation* (wegen *damn*)
- {*aggression*, *aggressive*, *aggressor*} $\leftrightarrow_3$ [**aggress*]
  - ⇒ Komplexität gleich; keine willkürliche Basisform

Separate morphological and semantic rules

Morphologische und semantische Redundanzregeln sind nicht immer deckungsgleich:

govern+ment:

1. "group that Z-s"
2. "act/process of Z-ing"

Deshalb: unterschiedliche Behandlung im "information measure" (Reformulierung von 6b)

Other applications

Präfixverben

(17)	*transist	transmit	transfer	*transcede	*transcur
	persist	permit	prefer	precede	*precur
	consist	commit	confer	concede	concur
	assist	admit	*affer	accede	*accur
	subsist	submit	suffer	succeed	*succur
	desist	*demit	defer	*decede	*decur
	insist	*immit	infer	*incede	incur

Redundanzregel

$$(18)\quad \begin{bmatrix} /x = y/ \\ +\mathrm{V} \end{bmatrix} \leftrightarrow \left\{ \begin{matrix} \begin{bmatrix} /x/ \\ +\mathrm{Prefix} \end{bmatrix} \\ \begin{bmatrix} /y/ \\ +\mathrm{Stem} \end{bmatrix} \end{matrix} \right\}$$

Komposita

(21) a. garbage man, iceman, milkman, breadbasket, oil drum
b. snowman, gingerbread man, bread crumb, sand castle
c. bulldog, kettledrum, sandstone, tissue paper

Redundanzregeln

(22) $$\begin{bmatrix} /[_{\mathrm{N}}x]\ [_{\mathrm{N}}y]/ \\ +\mathrm{N} \end{bmatrix} \leftrightarrow \left\{ \begin{matrix} \begin{bmatrix} /x/ \\ +\mathrm{N} \end{bmatrix} \\ \begin{bmatrix} /y/ \\ +\mathrm{N} \end{bmatrix} \end{matrix} \right\}$$

Komposita

(21) a. garbage man, iceman, milkman, breadbasket, oil drum
b. snowman, gingerbread man, bread crumb, sand castle
c. bulldog, kettledrum, sandstone, tissue paper

Redundanzregeln

(23) a. $$\begin{bmatrix} +\mathrm{N} \\ Z \text{ THAT CARRIES } W \end{bmatrix} \leftrightarrow \left\{ \begin{matrix} \begin{bmatrix} +\mathrm{N} \\ Z \end{bmatrix} \\ \begin{bmatrix} +\mathrm{N} \\ W \end{bmatrix} \end{matrix} \right\}$$

b. $$\begin{bmatrix} +\mathrm{N} \\ Z \text{ MADE OF } W \end{bmatrix} \leftrightarrow \left\{ \begin{matrix} \begin{bmatrix} +\mathrm{N} \\ Z \end{bmatrix} \\ \begin{bmatrix} +\mathrm{N} \\ W \end{bmatrix} \end{matrix} \right\}$$

c. $$\begin{bmatrix} +\mathrm{N} \\ Z \text{ LIKE A } W \end{bmatrix} \leftrightarrow \left\{ \begin{matrix} \begin{bmatrix} +\mathrm{N} \\ Z \end{bmatrix} \\ \begin{bmatrix} +\mathrm{N} \\ W \end{bmatrix} \end{matrix} \right\}$$

Other applications

Komposita

(24) a. blueberry, blackberry
  b. cranberry, huckleberry
  c. gooseberry, strawberry

exocentric compunds:

- *redhead*, *blackhead*, *redwing*, *yellow jacket*, *redcoat*, *greenback*, *bigmouth*, *big top*

Other applications

Kausativverben

(28) a. The door opened.
  b. Bill opened the door.

(29) a. The window broke.
  b. John broke the window.

(30) a. The coach changed into a pumpkin.
  b. Mombi the witch changed the coach from a handsome young man into a pumpkin.

(35) a. Bees swarmed in the garden.
  We sprayed paint on the wall.
  b. The garden swarmed with bees.
  We sprayed the wall with paint.

Other applications

Kausativverben

(29) a. The window broke.
     b. John broke the window.

Lexikalische Einträge

(32) a. $\begin{bmatrix} /\text{bræk}/ \\ +\text{V} \\ +[\text{NP}_1 \;\_\_\_\_] \\ \text{NP}_1 \;\; \textsc{break} \end{bmatrix}$

b. $\begin{bmatrix} /\text{bræk}/ \\ +\text{V} \\ +[\text{NP}_2 \;\_\_\_\_\; \text{NP}_1] \\ \text{NP}_2 \;\; \textsc{cause} \;\; (\text{NP}_1 \;\; \textsc{break}) \end{bmatrix}$

Other applications

Kausativverben

(29) a. The window broke.
b. John broke the window.

Redundanzregeln

(33) a. $\begin{bmatrix} /x/ \\ +\mathrm{V} \end{bmatrix} \leftrightarrow \begin{bmatrix} /x/ \\ +\mathrm{V} \end{bmatrix}$

b. $\begin{bmatrix} +\mathrm{V} \\ +[\mathrm{NP}_1 \;\_\_\_\_] \\ \mathrm{NP}_1 \;\; W \end{bmatrix} \leftrightarrow \begin{bmatrix} +\mathrm{V} \\ +[\mathrm{NP}_2 \;\_\_\_\_\; \mathrm{NP}_1] \\ \mathrm{NP}_2 \;\; \textsc{cause} \;\; (\mathrm{NP}_1 \;\; W) \end{bmatrix}$

Idiome (und Partikelverben)

(37) a. $\begin{bmatrix} \text{NP}_1\ [_{\text{VP}}\ [_{\text{V}}\text{kik}]\ [_{\text{NP}}\ [_{\text{Art}}\text{ðə}]\ [_{\text{N}}\text{bukət}]]]] \\ \text{NP}_1\ \text{DIE} \end{bmatrix}$

b. $\begin{bmatrix} \text{NP}_1\ [_{\text{VP}}\ [_{\text{V}}\text{giv}]\ [_{\text{NP}}\ [_{\text{N}}\text{hel}]]\ [_{\text{PP}}\ [_{\text{P}}\text{to}]\ \text{NP}_2]] \\ \text{NP}_1\ \text{YELL AT}\ \text{NP}_2 \end{bmatrix}$

c. $\begin{bmatrix} \text{NP}_1\ [_{\text{VP}}\ [_{\text{V}}\text{tǣk}]\ \text{NP}_2\ [_{\text{PP}}\ [_{\text{P}}\text{to}]\ [_{\text{NP}}\ [_{\text{N}}\text{tæsk}]]]]] \\ \text{NP}_1\ \text{CRITICIZE}\ \text{NP}_2 \end{bmatrix}$

=> reguläre Syntax, idiomatische Semantik

The cost of refering to redundancy rules

- *refuse*, *refusal*, **refusion*
- *confuse*, **confusal*, *confusion*

(40) The cost of referring to redundancy rule R in evaluating a lexical entry W is $I_{R,W} \times P_{R,W}$, where $I_{R,W}$ is the amount of information in W predicted by R, and $P_{R,W}$ is a number between 0 and 1 measuring the regularity of R in applying to the derivation of W.

The sum of the actual uses and the non-uses is the number of POTENTIAL uses of R. $P_{R,W}$ should be near zero when the number of actual uses of R is close to the number of potential uses; $P_{R,w}$ should be near 1 when the number of actual uses is much smaller than the number of potential uses; and it should rise monotonically from the former extreme to the latter.

Creativity in the lexicon and its implications

> The accepted view of the lexicon is that it is simply a repository of learned information. Creativity is taken to be a product of the phrase-structure rules and transformations. (S. 667)

> Lexical redundancy rules are learned from generalizations observed in already known lexical items. Once learned, they make it easier to learn new lexical items: we have designed them specifically to represent what new independent information must be learned. However, after a redundancy rule is learned, it can be used generatively, producing a class of partially specified possible lexical entries. (S. 668)

> We have thus abandoned the standard view that the lexicon is memorized and only the syntax is creative. In its place we have a somewhat more flexible theory of linguistic creativity. Both creativity and memorization take place in both the syntactic and the lexical component. [...] However, the normal mode for syntactic rules is creative, and the normal mode for lexical rules is passive. (S. 668)

Zusammenfassung

- Full-entry theory mit Redundanzregeln
- neues Informationsmaß (als Evaluationsmaß für Erklärungsadequatheit)
  - $\Rightarrow$ Beschreibungskomplexität

Kritikpunkte:

- Redundanzregeln unbeschränkt: /y/+*or* ↔ /y/+*ion*
- Informationsmaß abhängig von optimaler Ordnung: $n!$

[1] Berwick, Robert C. & Amy S. Weinberg. 1984. *The grammatical basis of linguistic performance: Language use and acquisition*. Cambridge, MA: MIT Press.

[2] Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: The MIT Press.

[3] Jackendoff, Ray. 1975. Morphological and semantic regularities in the lexicon. *Language* 51(3). 639–671. `http://www.jstor.org/stable/412891`.

[4] Pollard, Carl. 1996. The nature of constraint grammar. Paper presented at the 11th Pacific Conference of Language, Information and Computation (PACLIC).