

Structural Grammars†

R. B. Lees, Research Laboratory of Electronics,
Massachusetts Institute of Technology, Cambridge, Massachusetts

We adopt the view that the grammar of a language is a predictive theory which isolates the grammatical sentences of that language by means of immediate constituent analyses, morphophonemic conversions, and grammatical transformations. A sample grammatical analysis is given for the development of the verb phrase in German independent clauses. Simple rules are given for converting the verb phrase as a sequence of personal affixes, various auxiliaries, and the main verb into passive, future, or conditional clauses, and then introducing word boundaries, choosing the proper auxiliaries, arranging the word-order, and finally mapping the resulting morpheme sequence into the correct sequence of words in the independent clause.

ANY reasonably general, mechanized program for translating texts from one language into another can avoid dealing with each and every sentence as a completely new and arbitrary sequence of dictionary items only if it provides each source-language sentence with a grammatical analysis.

Traditional notional or semantic-based grammatical descriptions are useless for this purpose, since an analysis using such a grammar can be carried out only if the meanings of all of the constituents of the sentence are given. These meanings cannot be assumed: one of the main purposes of a syntax program is to aid in determining them so that they can be transferred, i.e., translated, into the appropriate target-language equivalents. Furthermore, contemporary descriptive linguistic grammatical practice is also faulty, especially when it is to be employed in a machine program; for, while the descriptive linguist no longer designates sentence constituents by means of meaning-labels but refers exclusively to their perceptible shapes, the description is still largely ad hoc — each particular grammatical category is designated by an arbitrary stigma or mark

of class membership and must be devised differently for each language. Moreover, descriptive sketches are deficient in their presentation of the syntax in that they are either fragmentary or else require very complicated, arbitrary, and often-repeated rules for specifying the constituent structure of even fairly simple sentences. This is largely the result of assuming that all sentences of a natural language are describable in terms of an immediate-constituent analysis or branching-diagrams.

N. Chomsky ⁽¹⁾ has described a theory of language which avoids these difficulties by relaxation of requirements on a grammar to the weaker position of satisfying some evaluation procedures (instead of requiring a discovery or decision procedure), introduction of natural canons of simplicity or elegance, statement in terms of a set of expansion rules for generating all grammatical utterances, and, above all, introduction of a level of grammatical transformations. These grammatical transformations convert the constituent-structures of a set of the most central sentences (i.e., basic, nonderived sentence types, such as affirmative assertions) into the derived structures of a more complex, less central, and infinitely extendible set of sentences.

† This work was supported in part by the U.S. Army (Signal Corps), the U.S. Air Force (Office of Scientific Research, Air Research and Development Command), and the U.S. Navy (Office of Naval Research); and in part by the National Science Foundation.

1. Chomsky, N., "The Logical Structure of Linguistic Theory", Preliminary Draft, M.I.T., 1956, 713 + li pp.

Following certain suggestions of Chomsky and Lukoff⁽²⁾ a scheme has been constructed as an illustration of a small, isolated portion of such a grammar for German. The scheme is intended to generate all verb phrases of independent clauses, active and passive, subject to the following limitations:

a) The device generates several types of verb phrase which would occur only rarely in natural speech, not for any clearly grammatical reason, but simply because they are too long or clumsy. Three types generated are probably only semigrammatical, containing two past participles in ge-. In addition, several very

long, but not obviously excluded, types will not be generated.

b) There is no provision for conforming the affixes of the finite verb to those of the accompanying noun phrases in the sentence, or for adjusting the selection between particular verbphrase morphemes and other morphemes external to the verb phrase, such as subject, object, or adverbial, or between the verb and the separable prefix. (This last selection would devolve upon the lexicon.)

c) No provisions are made to generate impersonal constructions, zu- infinitives, nominalized verb phrases, dependent clauses, or other verbal constructions.

d) The rules for generating the proper allomorphic shapes of the stems and affixes are only suggested by reference to a few examples, since a complete listing of morpheme spellings would be as long as the lexicon.

2. Chomsky, N. and Lukoff, F., "Construction of the German Verb Phrase", Mechanical Translation Group Memo, Aug. 12, 1955, Research Laboratory of Electronics, M.I.T.

GLOSSARY OF SYMBOLS

Af	Any affix or connected sequence of affixes of the set Ps, sbj, pst, I, G	M_{ord}^{ob}	Obligatory mapping which rearranges the word-order, placing non-finite verb forms at end in correct order
Af₁	Affix of subject nominal	M_W^{ob}	Obligatory mapping which introduces word boundaries at proper places
Af₂	Affix of object nominal	N₁	Subject nominal
Aux	Auxiliary verb stem, <u>hab</u> or <u>sei</u>	N₂	Object nominal
D	Any post-verbal objects, adverbials, predicate nominals or adjectivals	Ps	Any personal affix
G	Affix of past participle	pst	Past tense affix
I	Affix of infinitive	Q	Any St + Af or sequence of these
M	Modal stem, <u>könn</u> , <u>müss</u> , etc.	S	Sentence
M_{Af}^{ob}	Obligatory mapping which places the affixes after the appropriate stem	sbj	Subjunctive morpheme
M_{Aux}^{ob}	Obligatory mapping which selects the proper auxiliary stem	sep	Any separable prefix
M_{DI}^{ob}	Obligatory mapping which replaces a participle with an infinitive in the so-called "Double Infinitive" construction	St	Any stem of the set V, M, W, or Aux
M_{ge}^{ob}	Obligatory mapping which introduces the special participle of <u>werden</u> after another participle	t₂	2nd person plural personal affix
		t₃	3rd person singular personal affix
		T_P^{op}	Optional transformation of kernel sentences producing passive sentences

T_{W}^{OP}	Optional transformation of kernel sentences or passive sentences producing future and conditional sentences	Y	Any string
V	Any verb stem = either V_h or V_s	Z	Any St + Af
V_h	Any verb stem which takes <u>haben</u> as auxiliary	+	Symbol of grammatical concatenation
V_s	Any verb stem which takes <u>sein</u> as auxiliary	/	Word boundary
W	Stem of the verb <u>werden</u>	*	Sentence boundary
X	Any St(+ Af)	=	The grammatical rule "Rewrite the foregoing as:"
		()	Optionally present
		{ }	Alternatively present

DEVELOPMENT OF THE VERB PHRASE

1. PHRASE-STRUCTURE RULE to yield verb phrases of kernel sentences

$$S = N_1 + Af_1 (+pst)(+sbj) + Ps (+ \left\{ \begin{array}{l} (M+I) + M + I (+Aux+G) \\ (Aux+G)(+M+I)(+M+I) \\ M + I + Aux + G + M + I \end{array} \right\}) + V(+D)(+sep+)*$$

(Abbrev. : N_1 Subject nominal	N_2 Object nominal
Af_1 Subject nominal affix	Af_2 Object nominal affix
pst Past tense morpheme	sep Separable verb prefix
sbj Subjunctive morpheme	S Sentence
Ps Personal affix	St V, M, W, Aux, hab, or sei
M Modal stem	Af Ps, sbj, pst, I, G, or any connected sequence of these
I Infinitive affix	D Objects, adverbials, predicate nominals, adjectives, etc.)
Aux Auxiliary verb stem	
G Past participle affix	
V Verb stem	

2. Optional GRAMMATICAL TRANSFORMATIONS to yield non-kernel sentences

a. Passive transformation:

$$T_P^{OP}: N_1 + Af_1 (+Af) + Ps (+X) + V + N_2 + Af_2 (+D)(+sep+)$$

$$= N_2 + Af_2 (+Af) + Ps (+X) + W + G + V(+D)(+von+N_1+Af_1)(+sep+)$$

b. Werden transformation to yield future and conditional phrases:

$$T_W^{OP}: \left\{ \begin{array}{l} pst + sbj \\ (+sbj) + Ps \end{array} \right\} + St = \left\{ \begin{array}{l} pst + sbj \\ (+sbj) + Ps \end{array} \right\} + W + I + St$$

3. Obligatory MAPPINGS to yield proper word-order, word boundaries, and auxiliary selections

a. Word boundary:

$$M_{W}^{ob}: \left\{ \begin{array}{c} \text{St} \\ \text{D} \\ \text{von} \\ \text{Af}_1 \\ \text{Af}_2 \end{array} \right\} + = \left\{ \begin{array}{c} \text{St} \\ \text{D} \\ \text{von} \\ \text{Af}_1 \\ \text{Af}_2 \end{array} \right\} /$$

b. Affixation:

$$M_{Af}^{ob}: \text{Af} + \text{St} = \text{St} + \text{Af}$$

c. Auxiliary selection:

$$M_{Aux}^{ob}: \left\{ \begin{array}{l} \text{Aux} + \text{Af} / \left\{ \begin{array}{c} \text{V}_h \\ \text{M} \end{array} \right\} = \text{hab} + \text{Af} / \left\{ \begin{array}{c} \text{V}_h \\ \text{M} \end{array} \right\} \\ \text{Aux} + \text{Af} / \left\{ \begin{array}{c} \text{V}_s \\ \text{W} \end{array} \right\} = \text{sei} + \text{Af} / \left\{ \begin{array}{c} \text{V}_s \\ \text{W} \end{array} \right\} \end{array} \right.$$

d. Word order:

$$M_{ord}^{ob}: \left\{ \begin{array}{l} X / (Y/) Z / V + \text{Af} / (Q/) * = X / (Y/) V + \text{Af} / (Q/) Z * \\ X / (Y/) Z / \left\{ \begin{array}{c} \text{D} / (\text{sep}+) \\ \text{sep}+ \end{array} \right\} (Q/) * = X / (Y/) \left\{ \begin{array}{c} \text{D} / (\text{sep}+) \\ \text{sep}+ \end{array} \right\} (Q/) Z * \end{array} \right.$$

(where X = St(+Af) + Ps; Y = any string; Z = St + Af;

Q = St + Af or any sequence of these)

e. Double Infinitive:

$$M_{DI}^{ob}: M + G / * = M + I / *$$

f. Special participle:

$$M_{ge}^{ob}: V + G / W + G = V + G / \text{worden}$$

MORPHOPHONEMIC RULES

a. Personal endings:

$$Ps = \left\{ \begin{array}{c} e \\ st \\ t_3 \\ en \\ t_2 \end{array} \right\}$$

$$\text{St} + X + \left\{ \begin{array}{c} e \\ t_3 \end{array} \right\} = \text{St} + X$$

pst = te

sbj = e

b. Stems:

M = könn, müss, woll, soll, dürf, mög

W = werd

V_h = sing, hör, mach, trag, geb, ...

V_s = geh, fahr, bleib, werd, ...

c. Separable prefixes:

sep + = an, auf, ein, hin, vor, ...

d. Infinitives and participles:

I = en (special rules for verbs like tadeln where I = n)

sing + G = gesungen

hör + G = gehört

studier + G = studiert

etc.

e. Past and subjunctive stems:

sing + te = sang

sang + e = sänge

f. General morphophonemic rules:

e + e = e

g. Finite verb affixes:

sing + e	= singe	fahr + e	= fahre	hör + e	= höre
sing + st	= singst	fahr + st	= fährst	hör + st	= hörst
sing + t ₃	= singt	fahr + t ₃	= fährt	hör + t ₃	= hört
sing + en	= singen	fahr + en	= fahren	hör + en	= hören
sing + t ₂	= singt	fahr + t ₂	= fahrt	hör + t ₂	= hört
sing + e	= singe	fahr + e	= fahre	hör + e	= höre
sing + e + st	= singest	fahr + e + st	= fahrest	hör + e + st	= hörst
sing + e	= singe	fahr + e	= fahre	hör + e	= höre
sing + en	= singen	fahr + en	= fahren	hör + en	= hören
sing + e + t ₂	= singet	fahr + e + t ₂	= fahret	hör + e + t ₂	= horet
sang	= sang	fuhr	= fuhr	hör + te	= hörte
sang + st	= sangst	fuhr + st	= fuhrst	hör + te + st	= hörtest
sang	= sang	fuhr	= fuhr	hör + te	= hörte
sang + en	= sangen	fuhr + en	= fuhren	hör + ten	= hörten
sang + t ₂	= sangt	fuhr + t ₂	= fuhrte	hör + te + t ₂	= hörtest
sänge	= sänge	führe	= führe		
sänge + st	= sängest	führe + st	= führest		
sänge	= sänge	führe	= führe		
sängen	= sängen	führen	= führen		
sänge + t ₂	= sänget	führe + t ₂	= führet		

A SAMPLE DERIVATION

VP

- (1.) $N_1 + Af_1 + pst + sbj + Ps + Aux + G + V + N_2 + Af_2 + D + sep + *$
- (2. a.) $N_2 + Af_2 + pst + sbj + Ps + Aux + G + W + G + V + D + von + N_1 + Af_1 + sep + *$
- (2. b.) $sbj + Ps + W + I + Aux + G + W + G + V + D + von + N_1 + Af_1 + sep + *$
- (3. a.) $N_2 + Af_2 / pst + sbj + Ps + W / I + Aux / G + W / G + V / D / von / N_1 + Af_1 / sep + *$
- (3. b.) $W + pst + sbj + Ps / Aux + I / W + G / V + G / D$
- (3. c.) $N_2 + Af_2 / \underbrace{W + pst + sbj + Ps}_X / \underbrace{sei + I / W + G}_Y / \underbrace{V + G}_Z / \underbrace{D / von / N_1 + Af_1}_D / sep + *$
- (3. d.) $\underbrace{W + pst + sbj + Ps}_X / \underbrace{sei + I}_Y / \underbrace{W + G}_Z / \underbrace{D / von / N_1 + Af_1}_D / sep + \underbrace{V + G}_Q / *$
- (3. d.) $\underbrace{W + pst + sbj + Ps}_X / \underbrace{sei + I}_Z / \underbrace{D / von / N_1 + Af_1}_D / sep + \underbrace{V + G / W + G}_Q / *$
- (3. d.) $\underbrace{W + pst + sbj + Ps}_X / \underbrace{D / von / N_1 + Af_1}_D / sep + \underbrace{V + G / W + G / sei + I}_Q / *$
- (3. f.) $/ \text{worden} /$
- (4. a.) $W + pst + sbj + t_3$
- (4. a.) $W + pst + sbj$
- (4. a.) $W + te + e$
- (4. b.) $/ \text{werd} + te + e / D / von / N_1 + Af_1 / sep + \text{trag} + G /$
- (4. c.) $/ \text{ein} + \text{trag} + G /$
- (4. d.) $/ \text{ein} + \text{getragen} / \text{worden} / \text{sein} / *$
- (4. e.) $/ \text{würde} + e /$
- (4. f.) $N_2 + Af_2 / \text{würde} / D / von / N_1 + Af_1 / \text{eingetragen} / \text{worden} / \text{sein} / *$
- Let $N_2 + Af_2 = \text{Das} / \text{Geld}$
- $D = \text{früher} / \text{auf} / \text{meinem} / \text{Konto}$
- $N_1 + Af_1 = \text{der} / \text{Gesellschaft}$
- $\text{Das} / \text{Geld} / \text{würde} / \text{früher} / \text{auf} / \text{meinem} / \text{Konto} / \text{von} / \text{der} / \text{Gesellschaft} / \text{eingetragen} /$
 $\text{worden} / \text{sein} / *$