SAMPLE OUTPUT SENTENCES OF VARIOUS COMIT AND SNOBOL PROGRAMS FOR TESTING A CHINESE GENERATIVE GRAMMAR ARE PRESENTED. THE GRAMMAR CHosen FOR EXPERIMENTATION IS A PRELIMINARY VERSION OF A TRANSFORMATIONAL GRAMMAR. ALL OF THE COMIT PROGRAMS AND ONE OF THE SNOBOL PROGRAMS USE A LINEARIZED REPRESENTATION OF TREE STRUCTURES, WITH ADDITIONAL NUMERICAL SUBSCRIPTS FOR EASIER IDENTIFICATION OF CORRESPONDING PARENTHESES. A SNOBOL PROGRAM USES INSTEAD A LIST OF NODE NUMBERS, TOGETHER WITH IMMEDIATE ANCESTORS AND DESCENDANTS. A SET OF SUBROUTINES NOW BEING WRITTEN IN COMIT USES MODIFIED POLISH NOTATION, IN WHICH A SUBSCRIPT GIVES THE NUMBER OF NODES IMMEDIATELY DOMINATED BY A GIVEN ONE. ONE SUCCESSFUL COMIT PROGRAM FOR THE PHRASE STRUCTURE OF THE GRAMMAR IS ALSO EXTENDED TO A FEW TRANSFORMATION RULES, WHICH INVOLVE EMBEDDING. THIS ARTICLE APPEARS IN THE OHIO STATE UNIVERSITY RESEARCH FOUNDATION PROJECT ON LINGUISTIC ANALYSIS, REPORT NUMBER 10. (IT)
Project on Linguistic Analysis
Report No. 10

CHINESE GRAMMARS AND THE COMPUTER
AT THE OHIO STATE UNIVERSITY
Leroy F. Meyers and James Yang
Divisions of Mathematics and
Electrical Engineering

May 1965
NATIONAL SCIENCE FOUNDATION
Grant No. GN-174
BIBLIOGRAPHY


Chomsky, A. N., Aspects of the Theory of Syntax, in press.


______, "Relationship between Yes-No Questions and Answers," unpublished term paper.


______, "Conjoining and Deletion in Mandarin Syntax," paper read at the summer meeting of LSA, 1964.
Planning for testing a Chinese generative grammar on a high-speed digital computer began in May, 1964. The grammar chosen for experimentation is a preliminary version of a transformational grammar prepared by Anne Yue Hanhimoto. James H. Yang and Leroy F. Meyers have done the programming for it, using both COMIT and SNOBOL.

Since transformational grammars require manipulation of tree structures, an appropriate machine representation of trees is needed. For example, the tree:

\[
\begin{align*}
S & \rightarrow A \quad D \quad A \\
& \quad B \quad C \\
A & \quad D \quad A \\
& \quad G
\end{align*}
\]

can be linearized, using parentheses, as:

\[
( (S \ (A \ B \ C)) \ A \ D \ (A \ G) \ A \ F ) \ S.
\]

All of the COMIT programs and one of the SNOBOL programs have used this representation, with additional numerical subscripts for easier identification of corresponding parentheses, as in:

\[
( (1, S \ (2, A \ B \ C))_2, A \ D \ (3, A \ G) \ (3, A \ F) )_1, S.
\]

A SNOBOL program uses instead a list of node numbers, together with immediate ancestors and descendants:

1. 0 S 3 4 2 3 4 5
2. 1 A 2 2 6 7
3. 1 D 0 0
4. 1 A 5 1 8
5. 1 F 0 0
6. 2 B 0 0
7. 2 C 0 0
8. 4 0 0

For example, node 2 is dominated by node 1, is labeled A, and uses subrule number 2 to dominate the two nodes numbered 6 and 7.
A set of subroutines now being written in COMIT uses modified Polish notation, in which a subscript tells how many nodes are immediately dominated by a given one:

$$S_4 A_2 B C D A_1 G F$$

At present, Yang's successful COMIT program for the phrase structure portion of the grammar is being extended to take care of the first few transformation rules, which involve embedding. One difficulty with programming the embedding rules (besides their complexity) is that it is seldom possible to generate two sentences at random such that one can be embedded into the other.

Some sample output sentences are given on the next few pages. These are based on various slight modifications of the rules in Mrs. Hashimoto's grammar. The corresponding trees are also illustrated.
Sample Output of Yang Program 2 (Constituent Structure only; COMIT)

Sentence

\begin{align*}
&\text{(001) \ ZHE4 Yi1 GE4 GOU3 MOD DE XIANG3 WO3MEN} \\
&\text{\begin{cases} \\
&S + \ast ( / .1 , S + \ast ( / .4 , N + \ast ( / .7 , NPM + \ast ( / .9 , \\
&\text{DE} + \ast ( / .13 , D + \ast ( / .14 , \text{NUM} + \ast ) / .15 , \text{CL} + \ast ) / .17 , \text{MOD} + \ast ) / .19 , \\
&\text{DE} + \ast ( / .21 , \text{VP} + \ast ) / .23 , \text{VRM} + \ast ( / .25 , \text{NP} + \ast ) / .27 , \text{NV} + \ast ( / .29 , \text{PP} + \ast ) / .31 , \text{VP} + \ast ) / .33 , \\
&S + \ast \end{cases}}
\end{align*}

Tree description

\begin{align*}
&\begin{cases} \\
&S + \ast ( / .1 , S + \ast ( / .4 , N + \ast ( / .7 , NPM + \ast ( / .9 , \\
&\text{DE} + \ast ( / .13 , D + \ast ( / .14 , \text{NUM} + \ast ) / .15 , \text{CL} + \ast ) / .17 , \text{MOD} + \ast ) / .19 , \\
&\text{DE} + \ast ( / .21 , \text{VP} + \ast ) / .23 , \text{VRM} + \ast ( / .25 , \text{NP} + \ast ) / .27 , \text{NV} + \ast ( / .29 , \text{PP} + \ast ) / .31 , \text{VP} + \ast ) / .33 , \\
&S + \ast \end{cases}
\end{align*}

Rules and subrules used

\begin{align*}
&\begin{cases} \\
&S + \ast ( / .1 , S + \ast ( / .4 , N + \ast ( / .7 , NPM + \ast ( / .9 , \\
&\text{DE} + \ast ( / .13 , D + \ast ( / .14 , \text{NUM} + \ast ) / .15 , \text{CL} + \ast ) / .17 , \text{MOD} + \ast ) / .19 , \\
&\text{DE} + \ast ( / .21 , \text{VP} + \ast ) / .23 , \text{VRM} + \ast ( / .25 , \text{NP} + \ast ) / .27 , \text{NV} + \ast ( / .29 , \text{PP} + \ast ) / .31 , \text{VP} + \ast ) / .33 , \\
&S + \ast \end{cases}
\end{align*}

Tree
Sample Output of Meyers Program 1 (Constituent Structure only: COMIT)

\[
S \rightarrow \text{NOM} \cdot \text{VP} + \\
\text{NOM} \rightarrow \text{NP} + \\
\text{NP} \rightarrow \text{PP} \cdot \text{MEN} + \\
\text{PP} \rightarrow \text{WO*3} + \\
\text{VP} \rightarrow \text{VBM} + \\
\text{VBM} \rightarrow \text{DE} \cdot \text{VM} \cdot \text{NOM} + \\
\text{Rules used} \\
\text{VM} \rightarrow \text{AI*4} + \\
\text{NOM} \rightarrow \text{NP} + \\
\text{NP} \rightarrow \text{DET} \cdot \text{N} + \\
\text{DET} \rightarrow \text{NUM} \cdot \text{CL} + \\
\text{NUM} \rightarrow \text{-ER*4} + \\
\text{CL} \rightarrow \text{ZHANG*1} + \\
\text{N} \rightarrow \text{NNH} + \\
\text{NNH} \rightarrow \text{-CHE*1} +
\]

Sentence

OO1. WO3 MEN DE AI4 ER4 ZHANG1 CHE1.

Tree description

\[
eg \cdot 0 \cdot 0 \cdot 1 + \ast (S / .1 + \ast (NOM / .2 + \ast (NP / .3 + \ast (PP / .4 + \\
-WO*3 + PP*) / .4 + MEN + NP*) / .3 + NOM*) / .2 + \ast (VP / .5 \\
+ \ast (VBM / .6 + -DE + \ast (VM / .7 + -AI*4 + VM*) / .7 + \ast (NOM \\
/ .8 + \ast (NP / .9 + \ast (DET / .1) + \ast (NUM / .11 + -ER*4 + NUM*) \\
/ .11 + \ast (CL / .12 + ZHANG*1 + CL*) / .12 + DET*) / .10 + \ast (N \\
/ .13 + \ast (NNH / .14 + -CHE*1 + NNH*) / .14 + N*) / .13 + NP*) \\
/ .9 + NOM*) / .8 + VBM*) / .6 + VP*) / .5 + S*) / .1 +
\]

Rules

\[
S \cdot \text{NOM} \cdot \text{NP} \cdot \text{PP} \cdot \text{VP} \cdot \text{VBM} \cdot \text{NOM} \cdot \text{NP} \cdot \text{DET} \cdot \text{NUM} \cdot \text{CL} \cdot \text{N} \cdot \text{NNH}
\]
Sample Output of Meyers Program 1 (Constituent Structure only: COMIT)

Tree

1: S
  2: NOM
  3: NP
    4: PP men
      wǒ
  5: VP
    6: VBM
      -de
      7: VM
dài
  8: NOM
    9: NP
      10: DET
      11: NUM
dér
      12: CL
      13: N
      14: NNH
      zhāng
      chē
Sample Output from Meyers Program 2 (Constituent Structure; SNOBOL)

\[
S = \text{NOM} \quad \text{VP} \quad F \\
\text{NOM} = \text{NP} \\
\text{NP} = \text{PP} \\
\text{Rules} \\
\text{used} \\
\text{VP} = \text{VBNST} \\
\text{VBNST} = \text{ASP} \quad \text{VQUO} \quad \text{C} \\
\text{ASP} = \text{LE} \\
\text{VQUO} = \text{SHUO1} \\
F = \text{MA} \\
\text{(Hello, Newton!)}
\]

Sentence 1. \text{NI3} \quad \text{LE} \quad \text{SHUO1} \quad \text{C} \quad \text{MA} .

Tree description

Sample Output of Yang Program 7 (Constituent Structure and the First Transformational Rule; COMIT)

Sentence

\[
\{ \text{OUTPUT-OF-ETRI} ((((( \text{GOU3}))))( \text{BU}(\text{LE})(\text{C}((\text{GONG4}))))(\text{X}((\text{REN2}))(\text{MEN})))(((\text{LE})(\text{SHANG4})(\text{LAI2})(\text{YI1})(\text{TANG4}))))(\text{DE})))((\text{LE})) \}
\]

Tree description

\[
\{ \begin{align*}
\text{MOD} + * & (/ .1, \text{s} + * (/ .6, \text{NOM} + * (/ .8, \text{NP} + * (/ .10, \text{N} + * (/ .16, \text{NNH} + \text{GOU}^*5) *)), / .16, \text{NNH} + *)), / .10, \text{N} + *) / .8, \text{NP} + *) / .6, \text{NOM} + * (/ .2, \text{VP} + \text{BU} + * (/ .3, \text{VBNSTB} + * (/ .13, \text{ASP} + \text{LE} + *)), / .13, \text{ASP} + * (/, .4, \text{VTR} + \text{C} + * ((/ .5, \text{VTA} + * (/, .14, \text{VTAB} + \text{GONG}^*4 + *) / .14, \text{VTAB} + *)), / .5, \text{VTA} + *)), / .4, \text{VTR} + * (/, .7, \text{NOM} + * (/, .9, \text{NPM} + *) + * (/, .11, \text{N} + * (/, .15, \text{NH} + \text{REN}^*2 + *)), / .15, \text{NH} + \text{MEN} + *)), / .11, \text{N} + *)) / .9, \text{NPM} + * (/, .17, \text{MOD} + * (/, .1, \text{s} + * (/, .2, \text{VP} + * (/, .3, \text{VBNSTB} + * (/, .8, \text{ASP} + \text{LE} + *)), / .8, \text{ASP} + * (/, .9, \text{VMO} + \text{SHANG}^*4 + *) / .9, \text{VMO} + * (/, .10, \text{DIR} + \text{LAI}^*2 + *)), / .10, \text{DIR} + * (/, .7, \text{FPC} + * (/, .12, \text{NUM} + \text{YI}^*1 + *)), / .12, \text{NUM} + * (/, .13, \text{CLV} + \text{TANG}^*4 + *) / .13, \text{CLV} + *)), / .7, \text{FPC} + *) / .3, \text{VBNSTB} + *)), / .2, \text{VP} + *), / .1, \text{s} + \text{DE} + *)), / .17, \text{MOD} + *)), / .7, \text{NOM} + *)), / .3, \text{VBNSTB} + *) / .2, \text{VP} + *))), / .12, \text{F} + \text{LE} + *)), / .12, \text{F} + *), / .1, \text{s} + \text{NOM} + *) \end{align*} \}
\]
Sample Output of Yang Program 7 (Constituent Structure and the First Transformational Rule; COMIT)

```
1:S
   6:NOM
   2:VP
   12:F
   8:NP
   3:VBNSTB
   10:N
   13:ASP
   14:VTAB
   7:NOM
   16:NNH
   11:N
   1:S
   9:NPM
   17:MOD
      15:NH
      2:VP
      11:N
      1:S
dev
   8:ASP
   9:VMO
   10:DIR
   7:FPC
   5:VTA
   4:VTR
   17:MOD
   12:NUM
   13:CLV
   6:NOM
   2:VP
   12:F
   8:ASP
   9:VMO
   10:DIR
   7:FPC
   5:VTA
   4:VTR
   17:MOD
   12:NUM
   13:CLV
   6:NOM
   2:VP
   12:F
   8:ASP
   9:VMO
   10:DIR
   7:FPC
   5:VTA
   4:VTR
   17:MOD
   12:NUM
   13:CLV
   6:NOM
   2:VP
   12:F
   8:ASP
   9:VMO
   10:DIR
   7:FPC
   5:VTA
   4:VTR
   17:MOD
   12:NUM
   13:CLV
   6:NOM
   2:VP
   12:F
   8:ASP
   9:VMO
   10:DIR
   7:FPC
   5:VTA
   4:VTR
   17:MOD
   12:NUM
   13:CLV
   6:NOM
   2:VP
   12:F
   8:ASP
   9:VMO
   10:DIR
   7:FPC
   5:VTA
   4:VTR
   17:MOD
   12:NUM
   13:CLV
```
Sample Output of Meyers Program 3 (Constituent Structure only; SNOBOL)

1. 0 S.1 2. 2 3
2. 1 NOM.0 1. 4
3. 1 VP.5 1. 5
4. 2 NP.1 2. 6 7
5. 3 VEM.2 3. 8 9 10
6. 4 PP.1 1. 11
7. 4 MEN.0 0.
8. 5 EMP.1 1. 12
9. 5 VM.1 1. 13
10. 5 NOM.0 1. 14

Tree

description

12. 8 HA03.0 0.
13. 9 AI4.0 0.
14. 10 NP.0 2. 15 16
15. 14 DET.2 2. 17 18
16. 14 N.1 1. 19
17. 15 D.1 1. 20
18. 15 CL.1 1. 21
19. 16 NNH.1 1. 22
20. 17 NA4.0 0.
21. 18 ZHANG1.0 0.
22. 19 CHE1.0 0.

Sentence 1. NI3 MEN HA03 AI4 NA4 ZHANG1 CHE1.
Sample Output of Meyers Program 3 (Constituent Structure only; SNOBOL)