The parse tree / constituent tree distinction in a computational grammar of Hebrew (HeGram)

Petter Haugereid[1,2,3], Nurit Melnik[2] and Shuly Wintner[3]

Department of Linguistic, Literary and Aesthetic Studies,
University of Bergen, Norway

Department of Literature, Language and the Arts,
The Open University of Israel, Israel

Department of Computer Science,
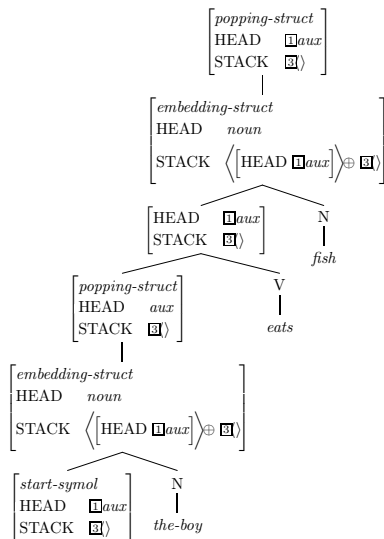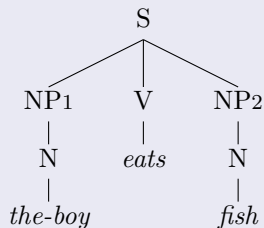University of Haifa, Israel

Constituent tree / parse tree

Transitive sentence

(1) ha-yeled oxel dag
     the-boy   eats fish

The boy eats fish.

Constituent tree

Popping structure

$$
\begin{bmatrix}
\text{HEAD} & \boxed{1} \\
\text{STACK} & \boxed{2} \\
\text{ARGS} & \left\langle \begin{bmatrix} \text{STACK} & \left\langle \begin{bmatrix} \text{HEAD} & \boxed{1} \end{bmatrix} \right\rangle \oplus \boxed{3} \end{bmatrix} \right\rangle
\end{bmatrix}
$$

Embedding structure

$$
\begin{bmatrix}
\text{HEAD} & \boxed{2} \\
\text{STACK} & \left\langle \begin{bmatrix} \text{HEAD} & \boxed{1} \end{bmatrix} \right\rangle \oplus \boxed{3} \\
\text{ARGS} & \left\langle \begin{bmatrix} \text{HEAD} & \boxed{1} \\ \text{STACK} & \boxed{3} \end{bmatrix}, \begin{bmatrix} \text{HEAD} & \boxed{2} \end{bmatrix} \right\rangle
\end{bmatrix}
$$
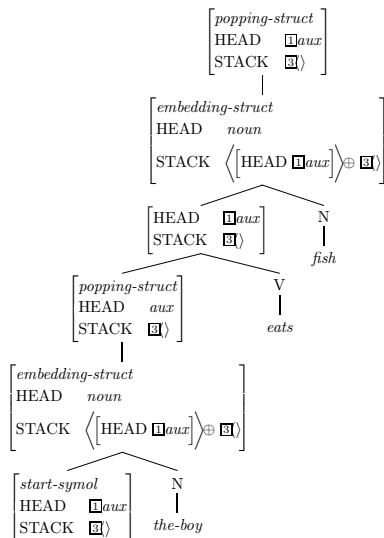
Constituent tree / parse tree

Subordinate clause complement

(2) ani   xoshev she-hu   lomed
    I   think   that-he studies

I think that he studies.

Constituent tree

$$\begin{bmatrix} \textbf{mrs} \\ \text{LTOP} \quad \boxed{h1}\, h \\ \text{INDEX} \quad \boxed{e2}\, e \\ \\ \text{RELS} \quad \left\langle \begin{bmatrix} \textbf{pron\_rel} \\ \text{LBL} \quad \boxed{h3}\, h \\ \text{ARG0} \quad \boxed{x4}\, x \end{bmatrix}, \begin{bmatrix} \textbf{pronoun\_q\_rel} \\ \text{LBL} \quad \boxed{h5}\, h \\ \text{ARG0} \quad \boxed{x4} \\ \text{RSTR} \quad \boxed{h6}\, h \\ \text{BODY} \quad \boxed{h7}\, h \end{bmatrix}, \begin{bmatrix} \textbf{\_xoshev\_v\_1\_rel} \\ \text{LBL} \quad \boxed{h8}\, h \\ \text{ARG0} \quad \boxed{e2} \\ \text{ARG1} \quad \boxed{x4} \\ \text{ARG2} \quad \boxed{e9}\, e \end{bmatrix}, \begin{bmatrix} \textbf{subcompl\_c\_rel} \\ \text{LBL} \quad \boxed{h10}\, h \\ \text{ARG0} \quad \boxed{e9} \end{bmatrix}, \right. \\ \\ \left. \begin{bmatrix} \textbf{pron\_rel} \\ \text{LBL} \quad \boxed{h11}\, h \\ \text{ARG0} \quad \boxed{x12}\, x \end{bmatrix}, \begin{bmatrix} \textbf{pronoun\_q\_rel} \\ \text{LBL} \quad \boxed{h13}\, h \\ \text{ARG0} \quad \boxed{x12} \\ \text{RSTR} \quad \boxed{h14}\, h \\ \text{BODY} \quad \boxed{h15}\, h \end{bmatrix}, \begin{bmatrix} \textbf{\_lomed\_v\_1\_rel} \\ \text{LBL} \quad \boxed{h10} \\ \text{ARG0} \quad \boxed{e9} \\ \text{ARG1} \quad \boxed{x12} \end{bmatrix} \right\rangle \\ \\ \text{HCONS} \quad \left\langle \begin{bmatrix} \textbf{qeq} \\ \text{HARG} \quad \boxed{h6} \\ \text{LARG} \quad \boxed{h3} \end{bmatrix}, \begin{bmatrix} \textbf{qeq} \\ \text{HARG} \quad \boxed{h14} \\ \text{LARG} \quad \boxed{h11} \end{bmatrix} \right\rangle \end{bmatrix}$$

Constituent tree

Constituent tree with 4 embeddings
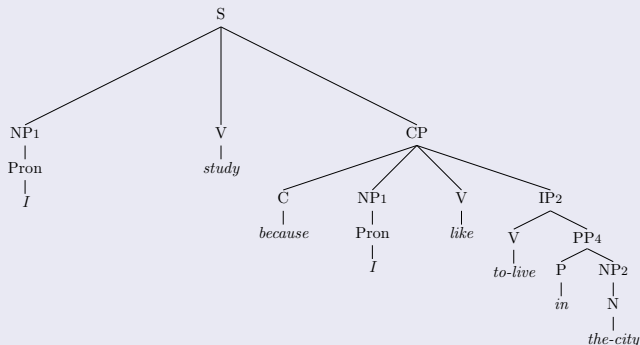
(3) ani lomed ki   ani ohev lagur   ba-ir
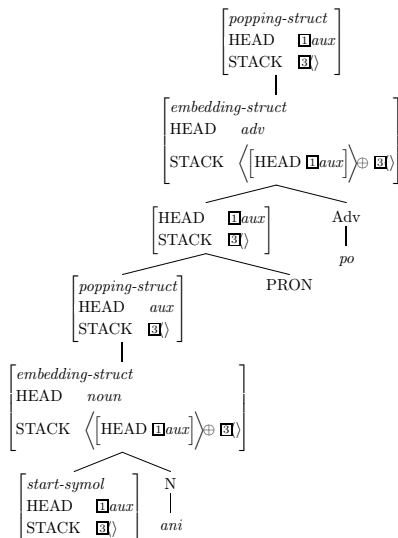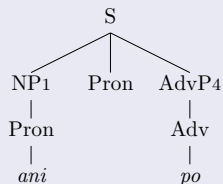   I   study because   I   like to.live in.the-city

     I study because I like to live in the city.

Constituent tree / parse tree

Empty copula

(4) ani po
     I   here

I am here.

Constituent tree

```
                S
        ┌───────┼────────┐
      NP₁      Pron     AdvP₄
       │                  │
      Pron               Adv
       │                  │
      ani                 po
```

$$
\begin{bmatrix} popping\text{-}struct \\ \text{HEAD} & \boxed{1}\,aux \\ \text{STACK} & \boxed{6}\rangle \end{bmatrix}
$$

$$
\begin{bmatrix} embedding\text{-}struct \\ \text{HEAD} & adv \\ \text{STACK} & \left\langle \begin{bmatrix} \text{HEAD } \boxed{1}aux \end{bmatrix} \right\rangle \oplus \boxed{6}\rangle \end{bmatrix}
$$

$$
\begin{bmatrix} \text{HEAD} & \boxed{1}\,aux \\ \text{STACK} & \boxed{6}\rangle \end{bmatrix}
$$

Adv

po

PRON

$$
\begin{bmatrix} popping\text{-}struct \\ \text{HEAD} & aux \\ \text{STACK} & \boxed{6}\rangle \end{bmatrix}
$$

$$
\begin{bmatrix} embedding\text{-}struct \\ \text{HEAD} & noun \\ \text{STACK} & \left\langle \begin{bmatrix} \text{HEAD } \boxed{1}aux \end{bmatrix} \right\rangle \oplus \boxed{6}\rangle \end{bmatrix}
$$

$$
\begin{bmatrix} start\text{-}symol \\ \text{HEAD} & \boxed{1}\,aux \\ \text{STACK} & \boxed{6}\rangle \end{bmatrix}
$$

N

ani

ani

Constituent tree / parse tree

Wh-question with empty copula

(5) eifo   xaifa
     where Haifa

Where is Haifa?

Wh-question

```
               S
         _____|_____
        |      |      |
     AdvP4    Pron    NP1
        |             |
       Adv            PN
        |             |
      where         Haifa
```

$$\begin{bmatrix} \textit{popping-struct} \\ \text{HEAD} \quad \boxed{1}\textit{aux} \\ \text{STACK} \quad \boxed{3}\rangle \end{bmatrix}$$

$$\begin{bmatrix} \textit{embedding-struct} \\ \text{HEAD} \quad \textit{noun} \\ \text{STACK} \quad \left\langle \left[\text{HEAD } \boxed{1}\textit{aux}\right]\right\rangle \oplus \boxed{3}\rangle \end{bmatrix}$$

$$\begin{bmatrix} \text{HEAD} \quad \boxed{1}\textit{aux} \\ \text{STACK} \quad \boxed{3}\rangle \end{bmatrix} \qquad \begin{matrix} \text{N} \\ \textit{Haifa} \end{matrix}$$

PRON

$$\begin{bmatrix} \textit{popping-struct} \\ \text{HEAD} \quad \textit{aux} \\ \text{STACK} \quad \boxed{3}\rangle \end{bmatrix}$$

$$\begin{bmatrix} \textit{embedding-struct} \\ \text{HEAD} \quad \textit{adv} \\ \text{STACK} \quad \left\langle \left[\text{HEAD } \boxed{1}\textit{aux}\right]\right\rangle \oplus \boxed{3}\rangle \end{bmatrix}$$

$$\begin{bmatrix} \textit{start-symol} \\ \text{HEAD} \quad \boxed{1}\textit{aux} \\ \text{STACK} \quad \boxed{3}\rangle \end{bmatrix} \qquad \begin{matrix} \text{Adv} \\ \textit{where} \end{matrix}$$
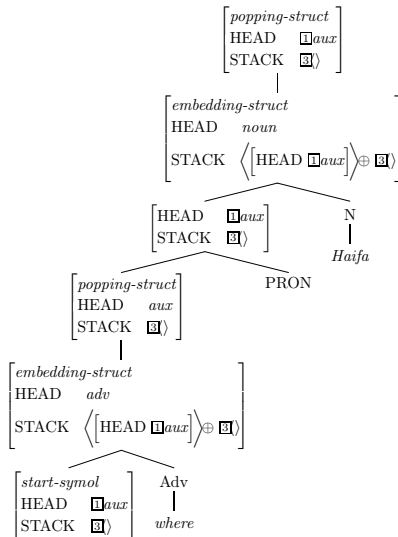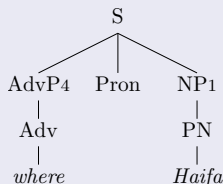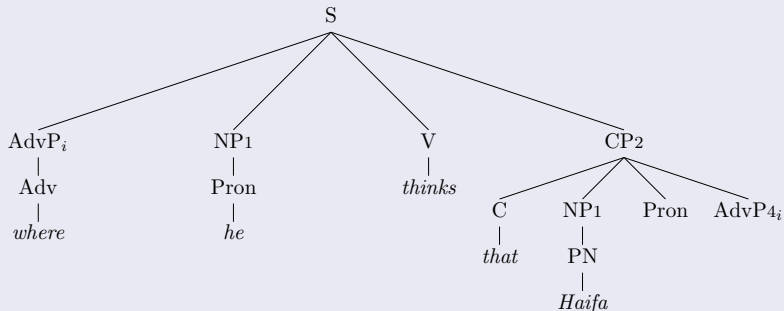
Constituent tree

Long distance dependency

(6) eifo   hu xoshev she-xaifa
where he thinks that-Haifa

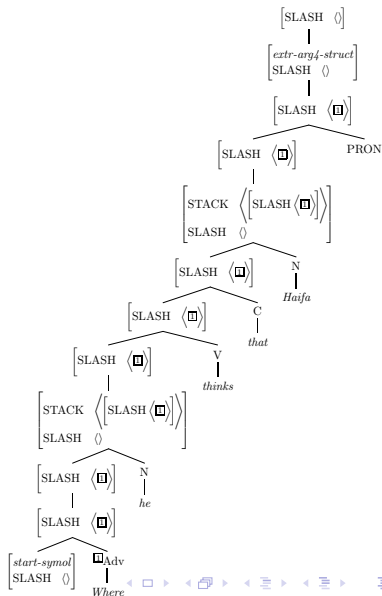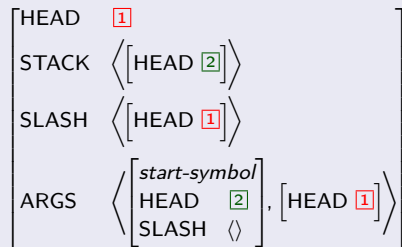Where does he think Haifa is?

Constituent tree

Long distance dependencies

Extraction structure

$$
\begin{bmatrix}
\text{SLASH} & \langle \rangle \\
\text{ARGS} & \left\langle \begin{bmatrix} \text{ARG} & \boxed{1} \\ \text{SLASH} & \langle \boxed{1} \rangle \end{bmatrix} \right\rangle
\end{bmatrix}
$$

Embedding filler structure

$$
\begin{bmatrix}
\text{HEAD} & \boxed{1} \\
\text{STACK} & \left\langle \begin{bmatrix} \text{HEAD} & \boxed{2} \end{bmatrix} \right\rangle \\
\text{SLASH} & \left\langle \begin{bmatrix} \text{HEAD} & \boxed{1} \end{bmatrix} \right\rangle \\
\text{ARGS} & \left\langle \begin{bmatrix} \textit{start-symbol} \\ \text{HEAD} & \boxed{2} \\ \text{SLASH} & \langle \rangle \end{bmatrix}, \begin{bmatrix} \text{HEAD} & \boxed{1} \end{bmatrix} \right\rangle
\end{bmatrix}
$$

HeGram (or The Haifa Hebrew Hi-quality HPSG Grammar, HaHeHiHG) – Data

Types	1,033
Rules	39
Inflectional rules	39
Lexical rules	0
Lexicon (handwritten)	248
Lexicon (automatically generated)	31,618