

Referring Expression Generation & (Surface) Realiser

Key Questions

- Can and should a pronoun, such as *he*, *she*, or *it*, be used to refer to an entity?
- Should a proper name, such as *The Caledonian Express* or *Glasgow*, be used?
- Should a full or reduced noun phrase, such as *the Aberdeen train*, *the train on platform 12*, *the express about to depart from platform 12*, or simply *the train*, be used?

Different Referring Approaches

- Definite

The train is about to leave.

This train will leave before ours.

That train will leave before ours.

These trains will leave before ours.

- Indefinite

A train is about to leave.

Some trains will leave before ours.

When to use a pronoun?

- Method 1:
 - **If** the intended referent was last mentioned in the previous sentence, **then** use a pronoun
- Method 2:
 - **If** the intended referent was last mentioned in the previous sentence **and** no other entities within the same sentence share the same grammatical properties, **then** use a pronoun

Realiser Aspects

- Structure Realiser
 - Maps internal structure of a text specification into specific structural resources
 - e.g. paragraphs and sections
 - Resources provided by mark-up language
- Linguistic Realiser
 - Maps phrase specification into specific words and syntactic constructs
 - Provided by target language

Example

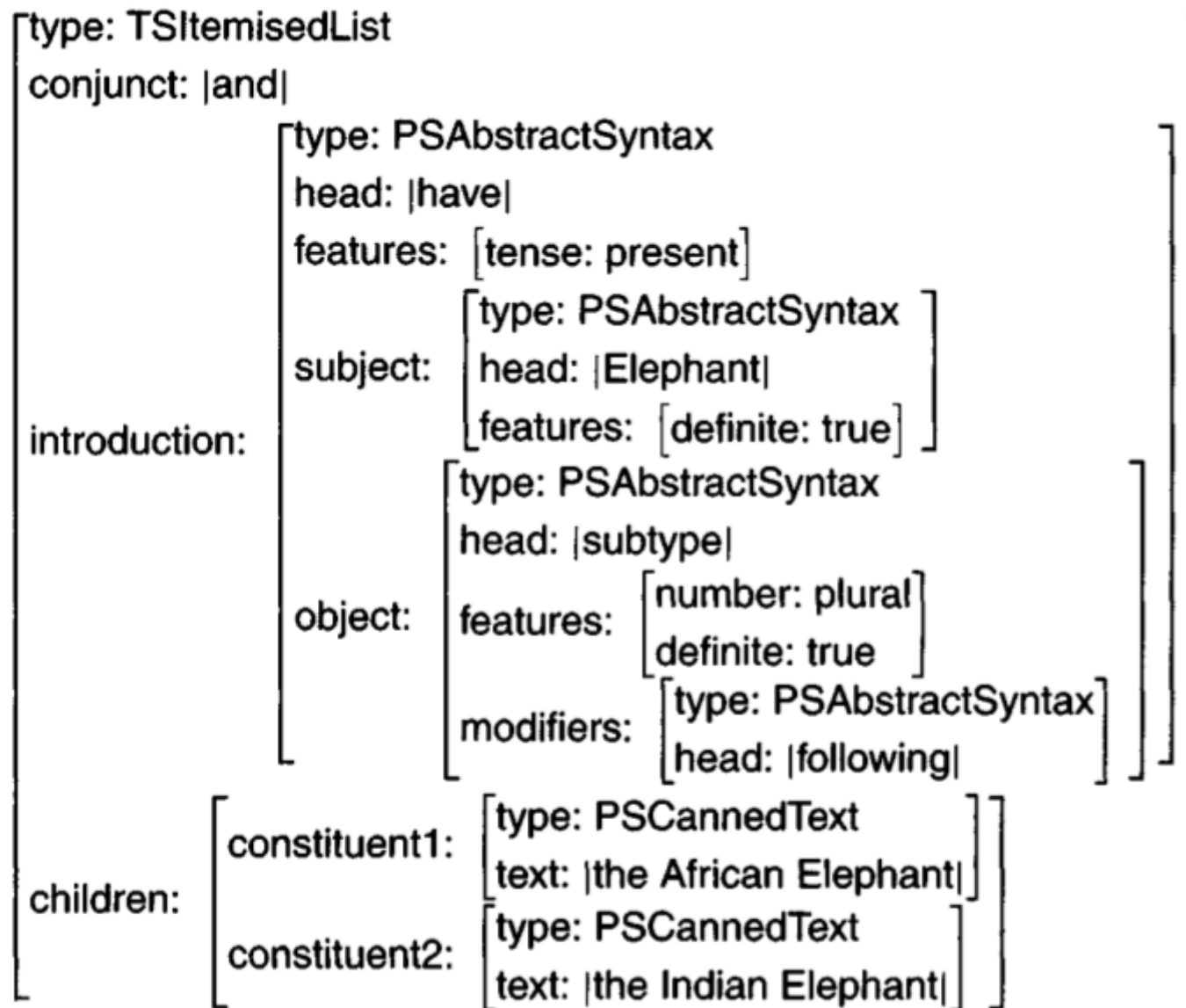


Figure 6.1 A simple PEBA text specification.

Example

```
The Elephant has the following subtypes:
<ul>
<li> the African Elephant; and
<li> the Indian Elephant.
</ul>
```

Figure 6.2 A surface form with mark-up annotations for the PEBA text.

Two
interpretation:

```
The Elephant has the following subtypes:
\begin{itemize}
\item the African Elephant; and
\item the Indian Elephant.
\end{itemize}
```

Figure 6.4 The logical structure specification in \LaTeX form.

End result:

```
The Elephant has the following subtypes:

• the African Elephant; and
• the Indian Elephant.
```

Figure 6.3 The PEBA text as displayed by the presentation system.

Input Example I

Input for the sentence:
“March had some rainy days”

Tool: KPML (Komet-Penman
Multilingual)

Type of Input: SPL (Sentence
Planning Language)

```
(S1 / generalized-possession
:tense past
:domain (N1 / time-interval
:lex march
:determiner zero)
:range (N2 / time-interval
:number plural
:lex day
:determiner some
:property-ascription
(A1 / quality :lex rainy)))
```

Figure 6.12 An input to KPML, from which KPML produces *March had some rainy days*.

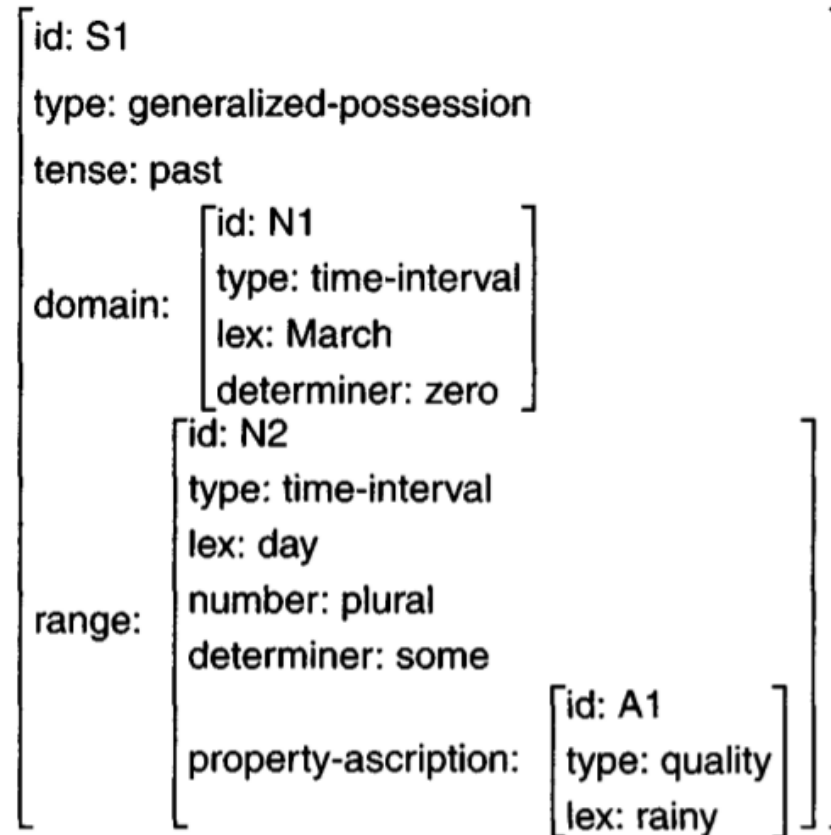


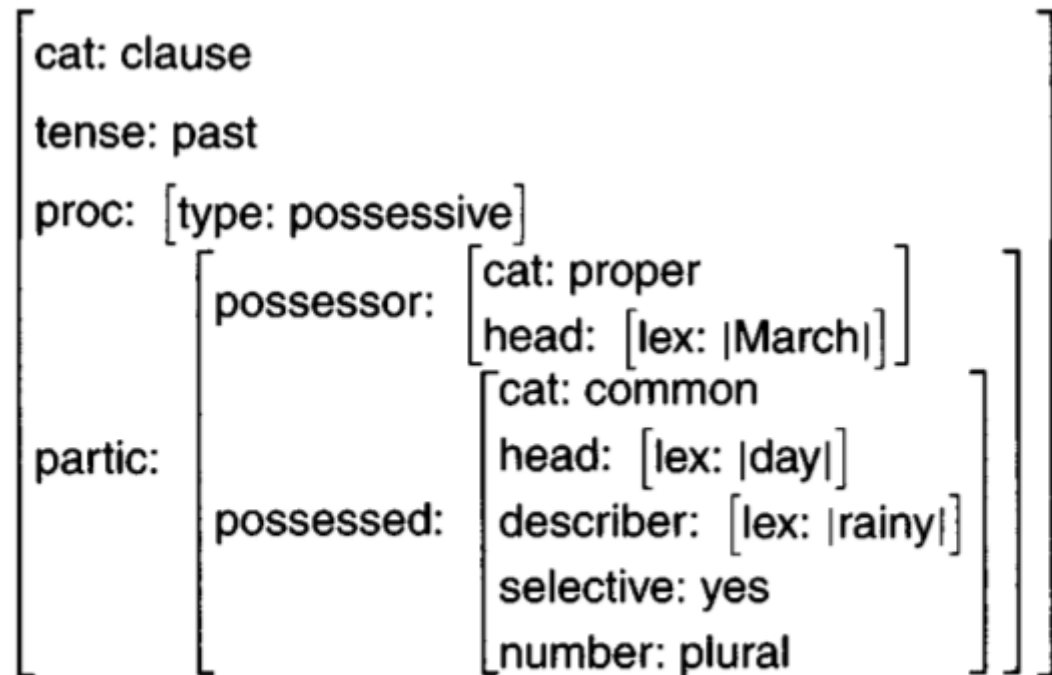
Figure 6.13 An AVM representation of the structure in Figure 6.12.

Input Example II

Input for the sentence:
"March had some rainy days"

Tool: SURGE (Systematical
Unification Realisation
Grammar of English)

Type of Input: FD (Functional
Description; Attribute Value
Pairs)



```
((cat clause)
 (proc ((type possessive)))
 (tense past)
 (partic ((possessor ((cat proper) (head ((lex "March"))))
 (possessed ((cat common) (head ((lex day))
 (describer ((lex rainy))
 (selective yes) (number plural))))))))))
```

Figure 6.21 An AVM representation of the structure in Figure 6.20.

Figure 6.20 An input to SURGE, from which SURGE produces *March had some rainy days*.

Input Example III

Input for the sentence:
"March had some rainy days"

Tool: RealPro

Type of Input: DsyntS (Deep
Syntactic Structure)

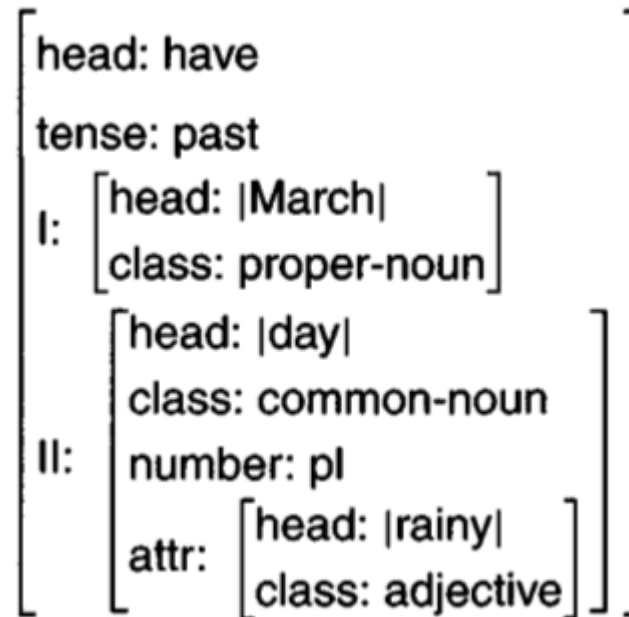


Figure 6.30 An AVM representation of the structure in Figure 6.29.

```
HAVE1 [tense:past]
  (I March [class:proper-noun]
   II day [class:common-noun number:pl]
    (ATTR rainy [class:adjective]))
```

Figure 6.29 A DSyntS for *March had some rain days*.

References

- Article: Choosing words in computer-generated weather forecasts, Reiter et al. (2005)
- Article: Statistical Natural Language Generation from Tabular Non-textual Data, Mahapatra et al. (2016)
- Book: Building Natural Language Generation Systems, Dale, Reiter (2000)