

GenDR

A Generic Deep Realizer with Complex Lexicalization

F. Lareau, F. Lambrey, I. Dubinskaite, D. Galarreta-Piquette, M. Nejat
Observatoire de linguistique Sens-Texte (OLST)

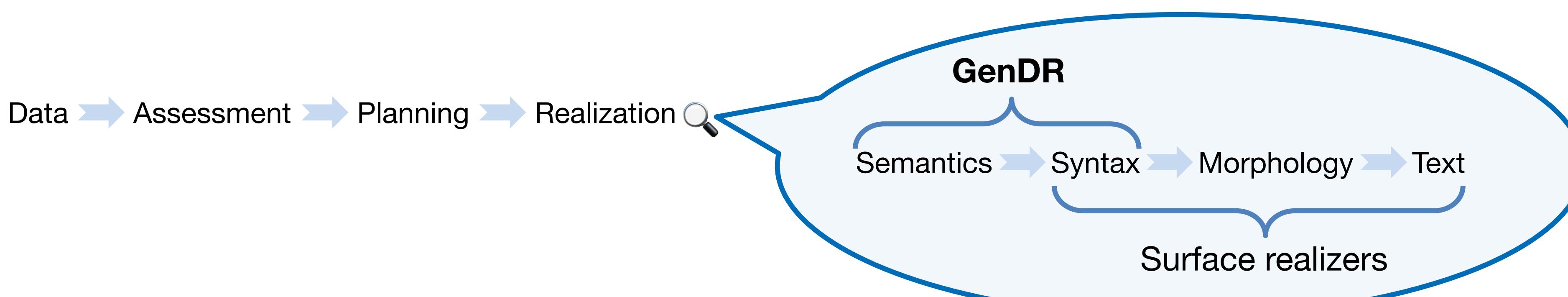
Université de Montréal

NLG

Salience-driven realization

Lexical resources

Step-by-step example

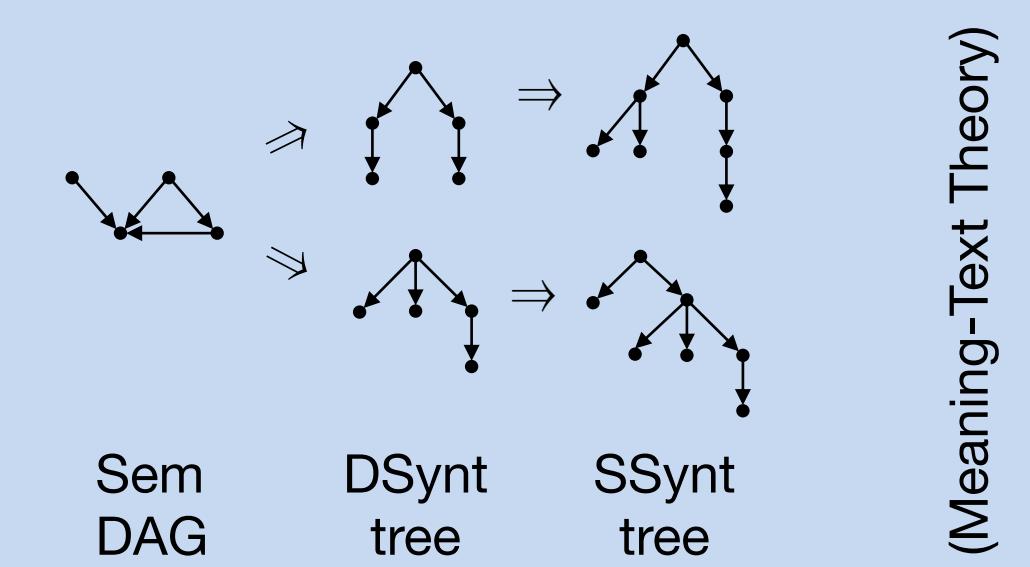


Which meaning is the most salient?

'small' →
 1 → 'cat' ← 1
 'lie' → 2 → 'sofa' ← 1
 'red' ←
 ≈ small(e_1 , x) ∧ cat(x) ∧ lie(e_2 , x , y) ∧ sofa(y) ∧ red(e_3 , y)
 The cat lying on the red sofa is small.
 The small cat is lying on the red sofa.
 The sofa on which the small cat is lying is red.

Semantics-syntactic interface

Arborization



Lexicalization

- Simple (lexemes)
- Template-based (full idioms)
- Bound (collocations)
- Class-based (numbers, etc.)
- Grammatical (function words)
- Fallback (unknown words)

Lexical dictionary

```

owe {
    dpos=V spos=verb
    gp = { // X owes Y to Z
        1=I 2=II 3=III // trivial diathesis
        I={dpos=N rel=subj}
        II={dpos=Num rel=dobj}
        III={dpos=N rel=iobj prep=to} } }

debt {
    dpos=N spos=noun
    gp = { // X's debt of Y to Z
        1=II 2=I 3=III // special diathesis
        I={dpos=Num rel=ncomp prep=of}
        II={dpos=N rel=poss case=GEN}
        III={dpos=N rel=ncomp prep=to} }
    lf={name=Oper1 value=have}
    lf={name=Oper13 value=owe}
    lf={name=Func2 value=amount_v_1}
    lf={name=Func2 value=stand_v_2}
    lf={name=Func2 value=total_v} }

```

Semantic dictionary

```

owe {
    lex=owe
    lex=debt }

```

Lexical coverage

English	1,500
French	1,500
Lithuanian	180
Persian	60

Lexical functions meta-dictionary

```

Oper1 { // e.g., X has a debt
    dpos = V
    gp = { 1=I L=II } } Arg 1 → Actant I
                                         Base → Actant II

Oper13 { // e.g., X owes a debt to Z
    dpos = V
    gp = { 1=I L=II 3=III } }

Func2 { // e.g., the debt amounts to Y
    dpos = V
    gp = { L=I 2=II } }

```

Lexical Functions (LFs)

Collocations tend to be instances of recurrent patterns across languages, e.g., *strong preference*, *gravely ill*, *intense flavour* and *win hands down* are all instances of the same pattern where a base is intensified by a syntactic modifier (the collocate).

What defines a collocation is the special relationship that exists between the base and the collocate it selects, which is modeled as a function. Intensification collocations are modeled with the *Magn* function: *Magn(preference)=strong*, etc.

Over the years, around 60 basic LFs have been identified, that combine to form a large number of complex LFs. GenDR implements ~26,000.

Semantics

Deep syntax

Surface syntax

