Using Dependency Grammars in guiding templatic Natural Language Generation

Ariel Gutman
relgu@google.com

Anton Ivanov
aii@google.com

Jessica Kirchner
jkirchner@google.com

Google

Introduction

Notwithstanding recent advances, neural NLG is still error-prone. To ensure **high-quality** messages in a **multilingual** & **wide-coverage** NLG system, human-authored NLG templates are still the easiest way to go.

To alleviate efforts of the template authors, we propose a templatic system enriched **by dependency relations**. This allows easy integration of **grammatical regularities** using a simple grammatical formalism, while at the same time maintaining **maximum flexibility** of the templates, which can combine static and dynamic elements.

System components

- **Lexical features of dynamic content,** e.g.
  AGR [NUMBER, GENDER, PERSON]
  DET [DEFINITENESS, DECLENSION]
- **Lexical & POS constraints:**
  e.g. for nouns, set the PERSON feature to third.
- **Selection of lexical forms**, according to grammatical constraints and **markedness** of forms.
- **Dependency analysis** using the Universal Dependencies framework

- **Hierarchy of attribute types**
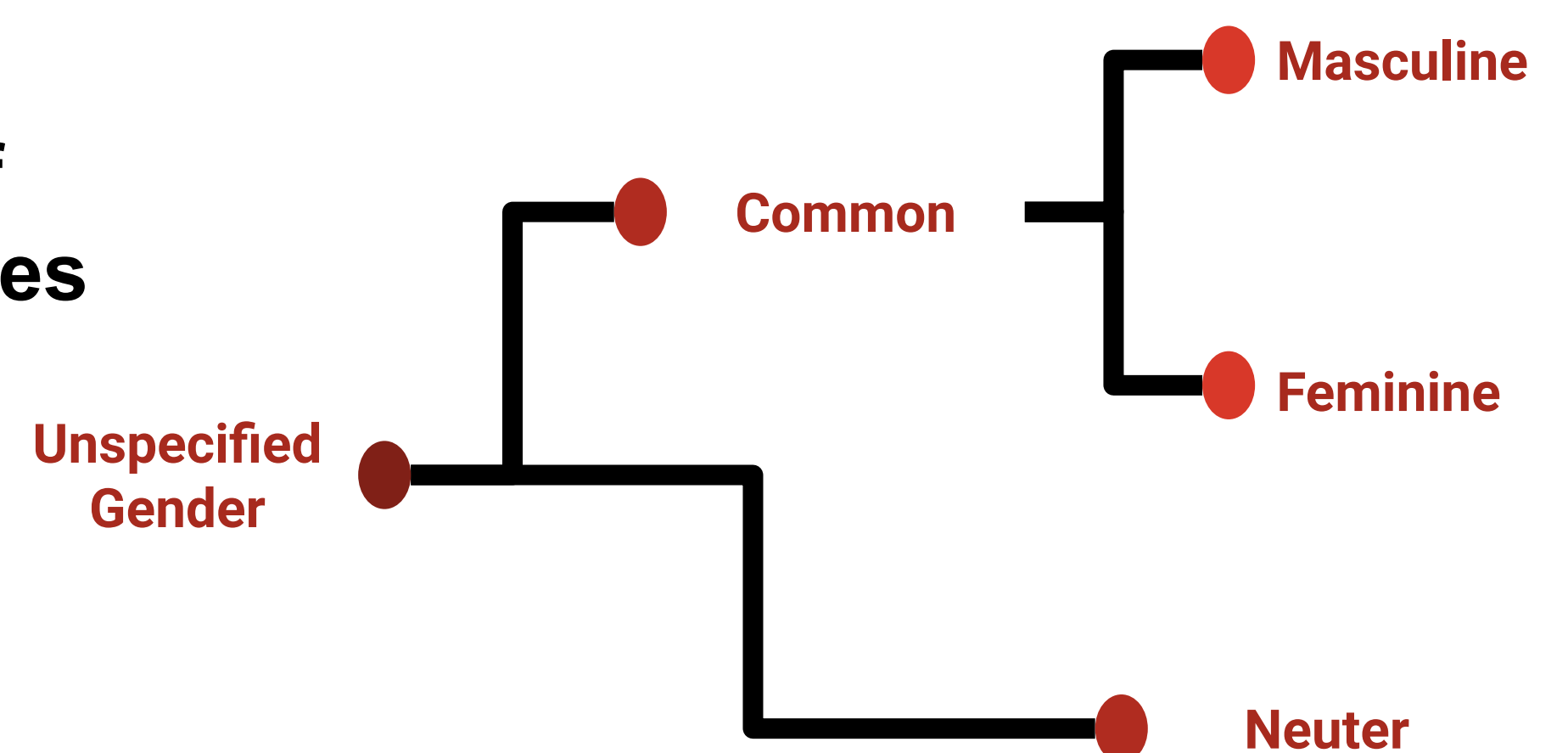
Feature unification across **dependency arcs**
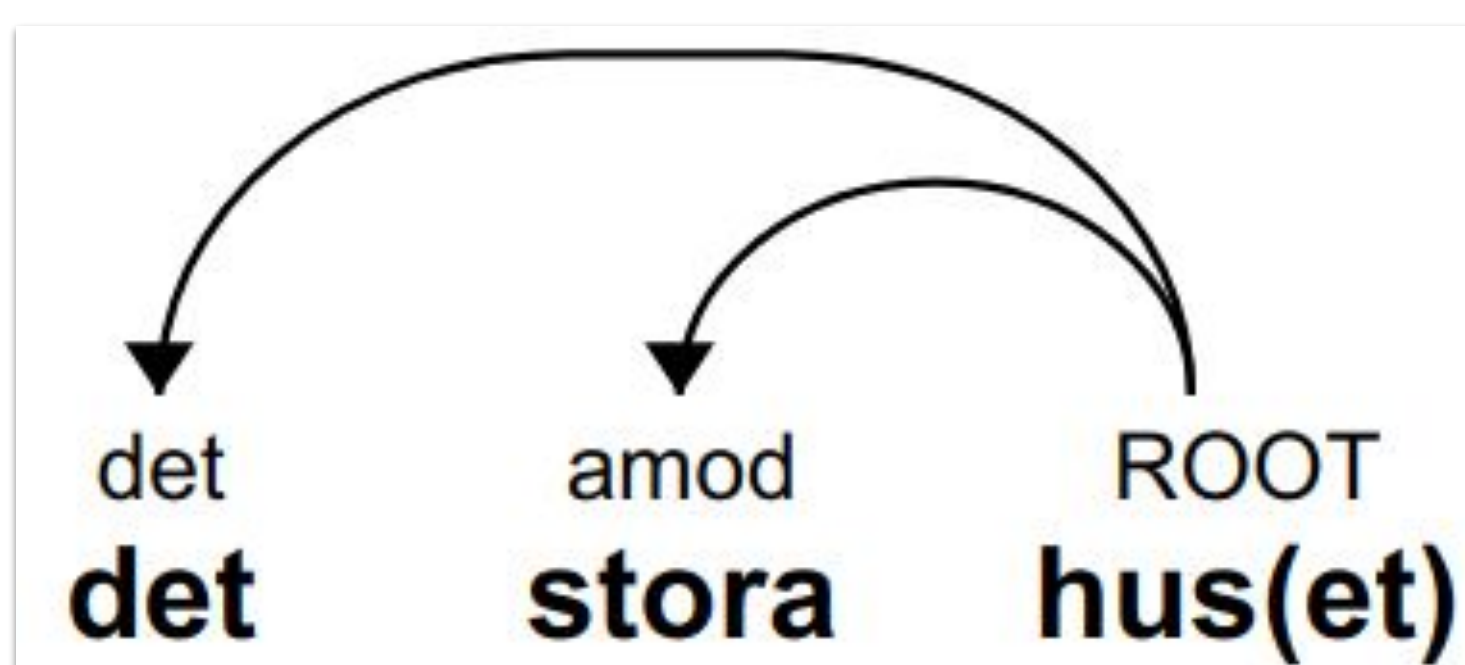
Examples and Challenges of Templates with Dependency Annotation

Dependency annotation · **Static text** · **Sub-tree's root**
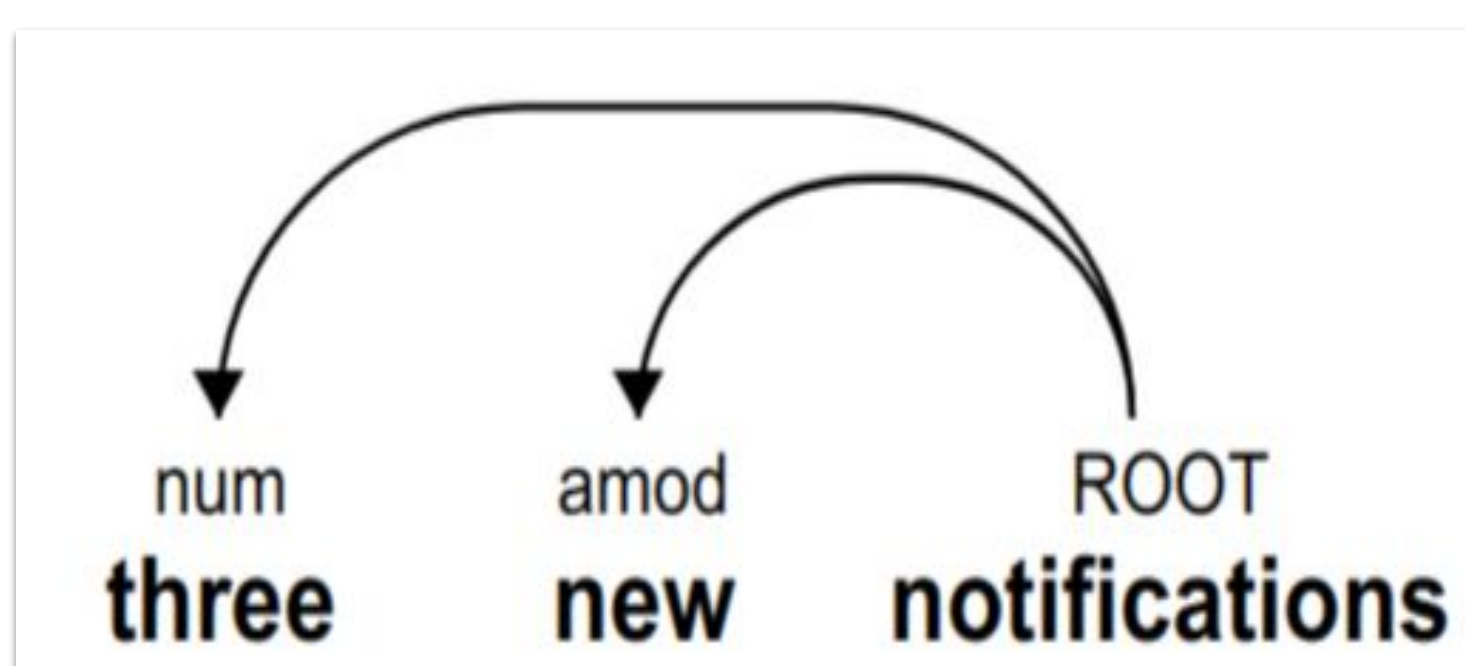
(English) (num:$number) new (root:NOTIFICATION)
(French) (num:$number) (amod:NOUVEAU) (root:NOTIFICATION)
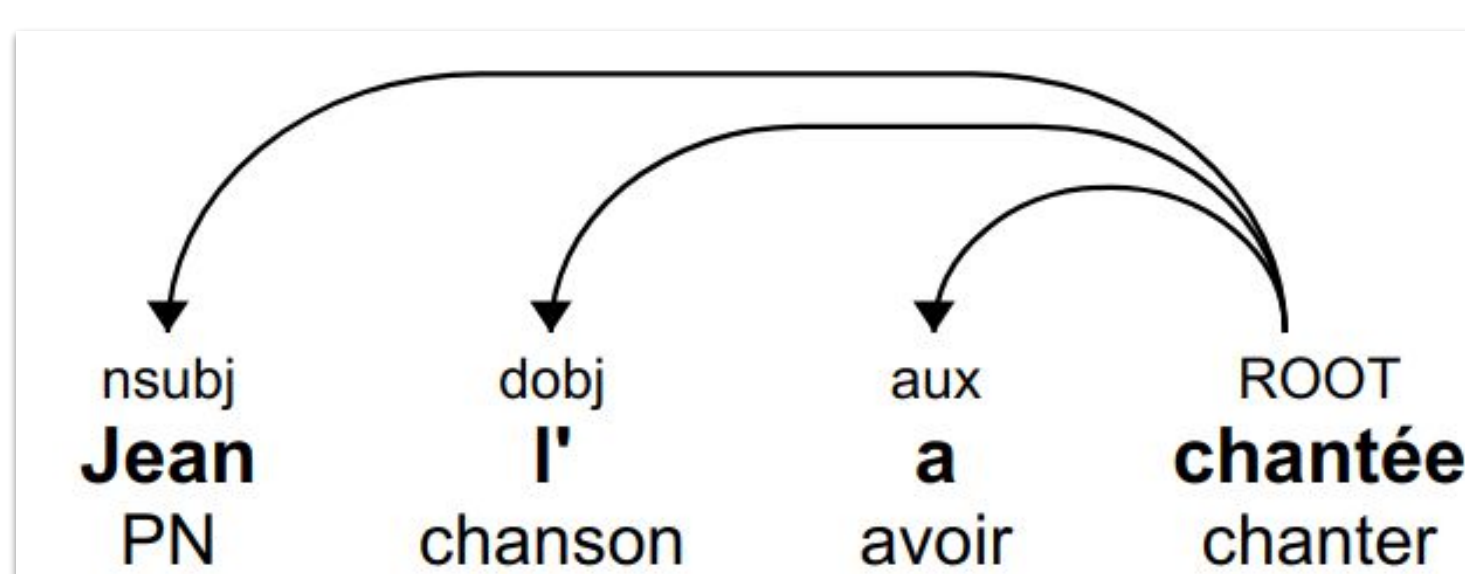
Dynamic content · **Inflecting lexemes**

(Danish/Swedish) (det:$article) (amod:BIG) (root:HOUSE)
Difference between Swedish and Danish is lexical, for the definite article:
- Swedish has feature DECLENSION strong ⤳ Select noun form *huset*
- Danish has feature DECLENSION weak ⤳ Select noun form *hus*
- Both have DEFINITENESS definitive ⤳ Select adjective form *stora*

(French) (nsubj:$agent) (dobj:$pronoun) (aux:AUX) (root:$verb)
- Subject agreement features flow from the subject to the auxiliary (**nsubj** ⤳ **root** ⤳ **aux**), while being parked as **covert agreement features** in the participle.
- Object agreement flows through the **dobj** relation under some conditions.
- Selection of auxiliary verb possible through the **aux** relation.

Advantages of the System

- Reuse of template structure for multiple languages: Dependency parses abstract away from language-specific details (at least for similar languages)
- Hybrid templates simplify system design and template creation: No need to parse static or irrelevant parts of template, making it much easier to get a system up and running, and to add new templates.

Select references

Ariel Gutman, Alexandros A. Chaaraoui, and Pascal Fleury. 2018. Crafting a lexicon of referential expressions for NLG applications. In: Ilan Kernerman and Simon Krek (eds.), *Proceedings of the LREC 2018 Workshop "Globalex 2018 – Lexicography & WordNets"*.

Richard Kittredge and Igor A. Mel'Cuk. 1983. Towards a Computable Model of Meaning-Text Relations Within a Natural Sublanguage. In *Proceedings of the Eighth International Joint Conference on Artificial Intelligence (IJCAI-83)*. 657–659.

François Lareau and Leo Wanner. 2007. *Towards a Generic Multilingual Dependency Grammar for Text Generation*. In Tracy Holloway King and Emily M. Bender (eds.), *Proceedings of the GEAF 2007 Workshop*. CSLI Publications, Stanford. 203–223.

Ryan McDonald, Joakim Nivre, Yvonne Quirmbach-Brundage, Yoav Goldberg, Dipanjan Das, Kuzman Ganchev, Keith Hall, Slav Petrov, Hao Zhang, Oscar Täckström, Claudia Bedini, Núria Bertomeu Castelló, and Jungmee Lee. 2013. Universal Dependency Annotation for Multilingual Parsing. In *Proceedings of ACL 2013*.

Ivan A. Sag, Thomas Wasow, and Emily M. Bender. 2003. *Syntactic Theory: A Formal Introduction*. Second edition. CSLI publications, Stanford.

Nikolai Trubetzkoy. 1931. *Die phonologischen Systeme*. Travaux du Cercle Linguistique de Prague 4. 96-116.