

# The Building Blocks of Spoken Language

Suwei Wu & Kasper Kok, PhD candidates  
 Dept. of Language and Communication  
 Faculty of Arts, VU

## Introduction

In Construction Grammar, human's grammar system is composed of grammatical patterns (form-meaning pairing) stored in human mind. However, most of previous studies focus on peripheral or author-invented patterns like *John sneezed the napkin off the table*.

This research examines grammatical patterns in actual language use in a top-down approach.

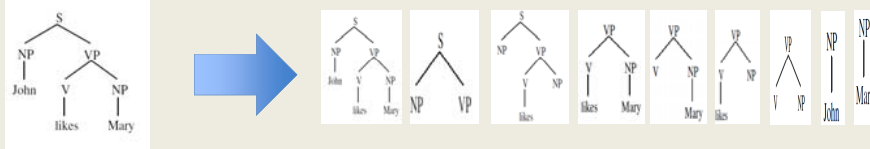
## Research questions

➤What are the frequent grammatical patterns in spoken language?

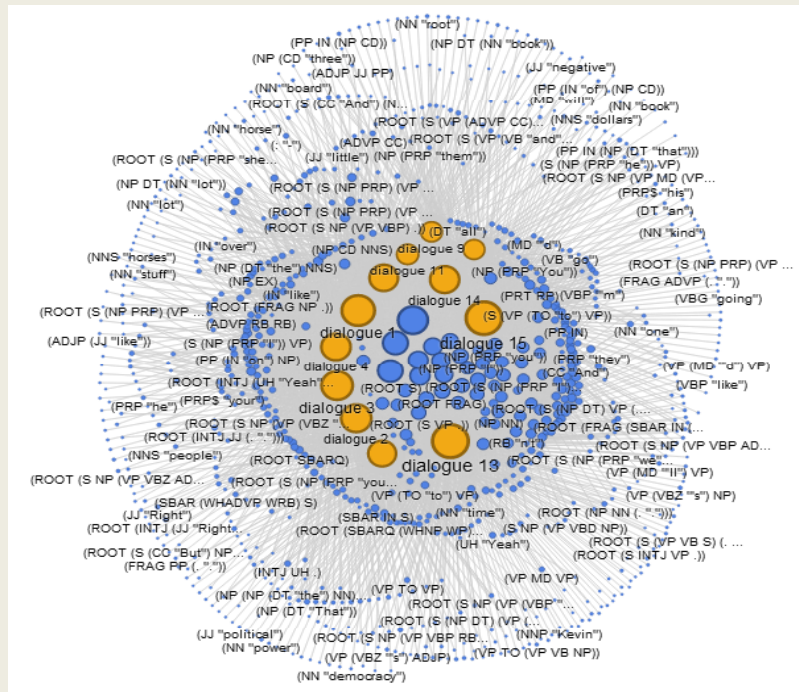
This is the first part of my PhD project, in which I examine how the framework of Construction Grammar might provide insights into spoken language as both an audio and visual phenomenon — in terms of how we use gesture with speech.

## Procedure and results

- 1) Annotate a spoken corpus (Santa Barbara Corpus of Spoken American English) with Stanford Parser, like Figure 1 extract all fragments with their frequency (Figure 2);



- 2) Identify the link between the frequency, fragments and the dialogues, as shown in Graph 1.



Graph 1: Tree fragments in 12 dialogues of a corpus. Node weights correspond to token frequency.  
**Orange nodes:** Dialogues (conversations, lectures and task-related talks)  
**Blue nodes:** Frequent constructions in these dialogues; the size of the ball represents the frequency of the construction.

## Results

- 1) Grammrtil patterns have various degrees of token frequencies.
- 2) There appears to be a correlation between the token frequency of the tree fragments and the number of dialogues in which they occur (bigger nodes are centrally positioned)
- 3) Frequently observed tree fragments are generally more abstract than peripheral tree fragments.

## Future directions and open questions

- 1) Comparison of differnt genres (e.g., lectures, novels, everyday conversation)
- 2) Different degrees of granularity
- 3) Investigating the relation of token frequency and centrality to gesture occurrence and form
- 4) What kind of statistics would help to confirm the results obtained so far?

## Acknowledgements:

- Supervisors: Alan Cienki and Mike Hannay
- Funding: China Scholarship Council and VU joint scholarship