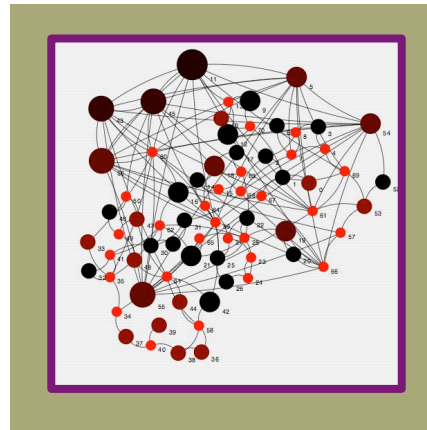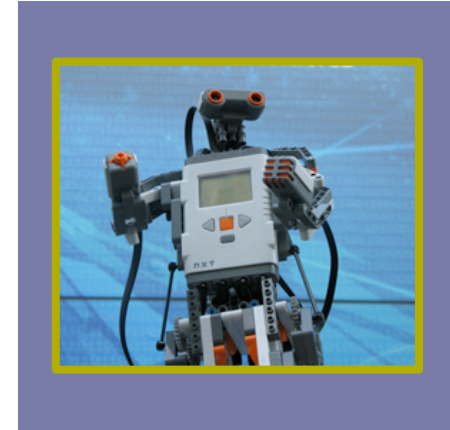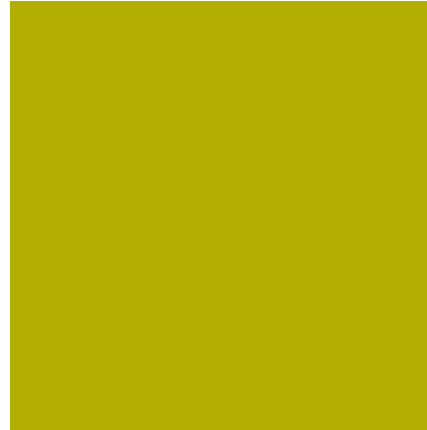# COMP40020
Human Language Technologies

Colourless green ideas sleep furiously

February 2019

Prof. Julie Berndsen
School of Computer Science

Julie.Berndsen@ucd.ie

+ HLT5

Contents:

- The domain of syntax, semantics and pragmatics.
- Structure, meaning and ambiguity
- Definitions and examples

Aim:

- To give a brief introduction to the basic concepts of syntax and syntactic analysis, and to the some of concepts of semantics and pragmatics. This is preparation for HLT 6, 8 and 9.

Intuitions about Sentences

- He got up late this morning.

- *Late got this morning up he.

- *John and Jane is playing in the garden.

- Colourless green ideas sleep furiously.

- The rat the cat the dog chased killed ate the malt.

- Anyone who feels that if so-many more students whom we haven't actually admitted are sitting in on the course than ones we have that the room had to be changed, then probably auditors will have to be excluded, is likely to agree that the curriculum needs.

See: Karlson (2007): https://www.jstor.org/stable/40057996

Syntax = Sentence Syntax

- Every sentence is a sequence of words but not every sequence of words is a sentence.

- "Sentences are by definition **grammatical** (i.e. grammatically well-formed) ... they may be regarded as well-formed **word-strings** (i.e. sequences of word-forms)..." (Lyons, 1981:104)

→ Grammaticality has nothing to do with the content and thus well-formed sentences need not be true.

+ Some Examples

- **All** the students might have been sitting together.
- The students **all** might have been sitting together.
- The students might **all** have been sitting together.
- The students might have **all** been sitting together.
- The students might have been **all** sitting together.
- The students might have been sitting **all** together.
- *The students might have been sitting together **all**.
- *The **all** students might have been sitting together.

- I saw the baby with the telescope

→ ambiguity

- This is an exam you must not fail to miss.

→ garden path

+ Function vs. Syntactic Category

Function:

E.g.

The child found the puppy.

→ The noun phrase **the child** functions as a subject and the noun phrase **the puppy** functions as an object.

Syntactic Category:

- Determiner (Det) the (→ also Art)
- Noun (N) cat
- Verb (V) sat
- Adjective (Adj) lazy
- Adverb   (Adv) slowly
- Preposition (P or Prep) on

- Noun Phrase (NP) the cat
- Verb Phrase (VP) reads the book
- Prepositional Phrase (PP) on the mat

+ Language Types: Function

Six possible language types:

- SVO
  - e.g. English, French, Swahili, Hausa, Thai

- VSO
  - e.g. Tagalog, Irish, Classical Arabic, Biblical Hebrew

- SOV
  - e.g. Turkish, Japanese, Persian, Georgian, Eskimo

- OVS
  - e.g. Apalai (Brazil), Barasano (Columbia), Panare (Venezula)

- OSV
  - e.g. Apurina and Xavante (Brazil)

- VOS
  - e.g. Cakchiquel (Guatemala), Huave (Mexico)

Fromkin & Rodman (2013)

UCD DUBLIN

+ Phrase Structure Trees: Categories

- Phrase structure trees show:
  - the linear structure of the words in the sentence
  - the groupings of words into syntactic categories
  - the hierarchical structure of syntactic categories

- Constituency:
  - Groupings of words behave as single constituents (i.e. appear in similar syntactic environments)
    - Ten students from UCD
    - A level 4 module such Human Language Technologies
    - *towards eat
    - *the is

+ Phrase Structure Trees

```
                                S
                               / \
                                  VP
                                 /|\
                                / | \
                               /  |  \
                              /   |   PP
                             /    |   / \
                            /     |  /   \
                          NP     NP     NP
                         /  \   /  \   /  \
                        /    \ /    \ /    \
                      Det    N V  Det N Prep Det N
```
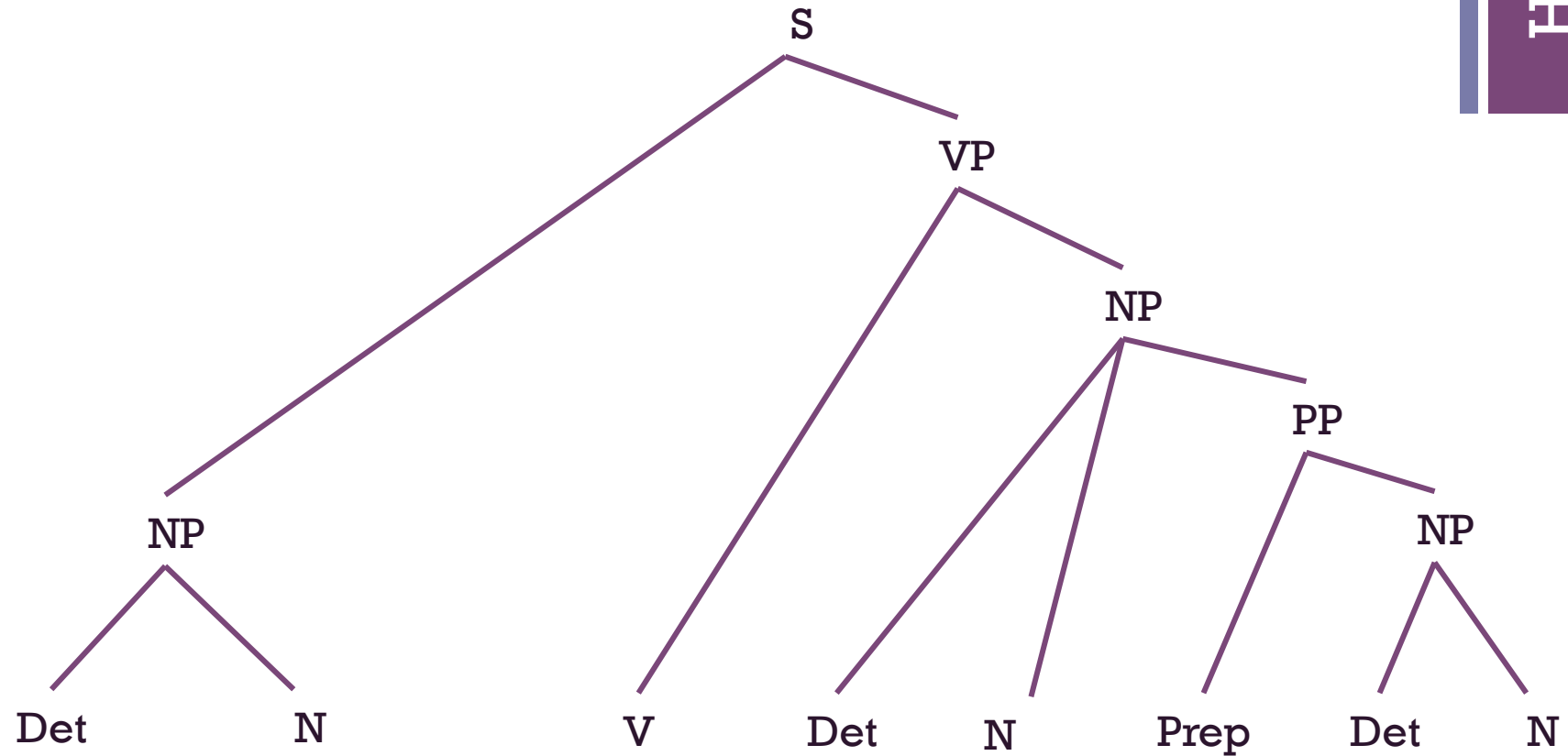
Det   N   V   Det   N   Prep   Det   N
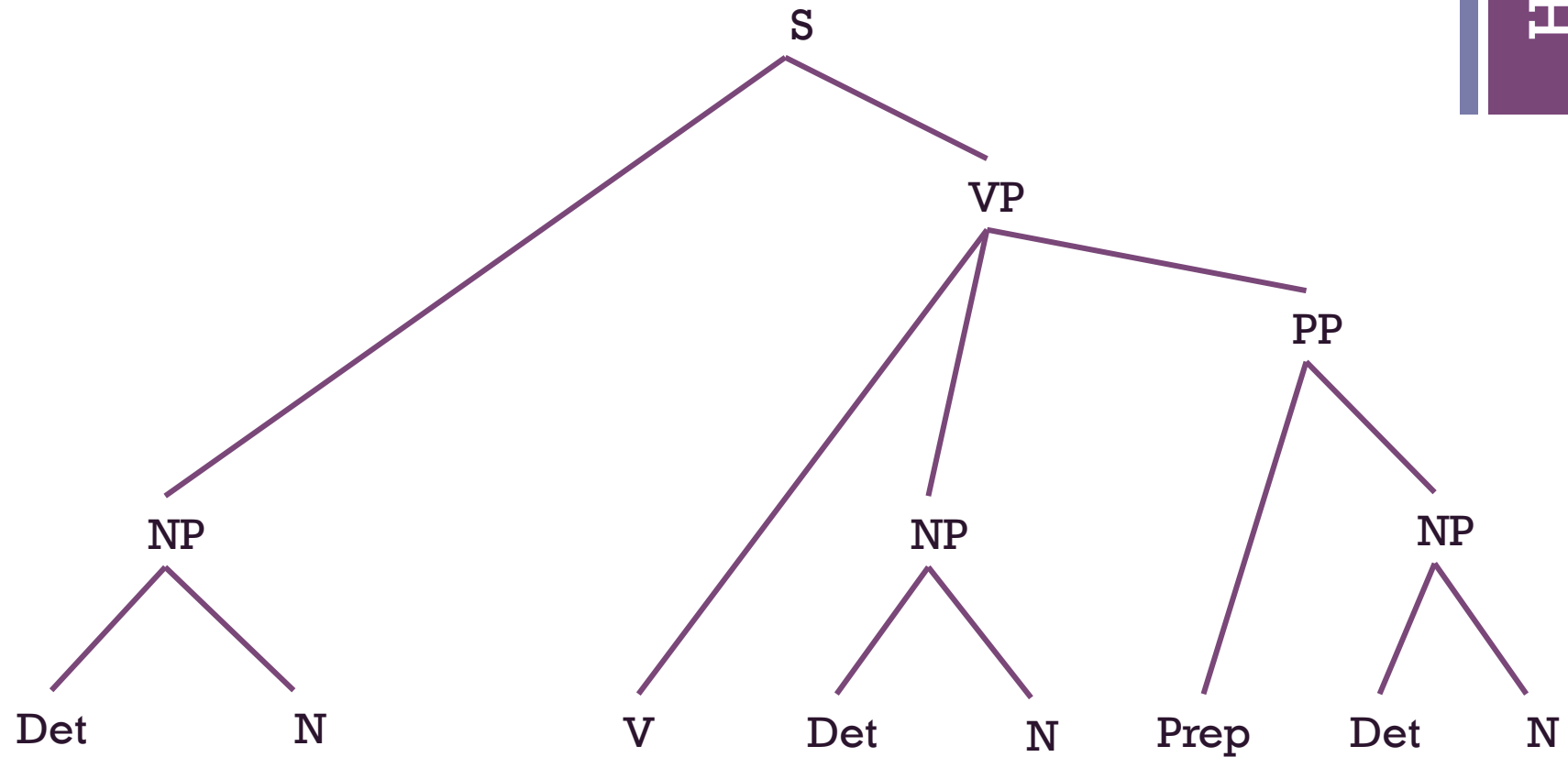
[[The gangster] [hit [the man] [with the gun]]]

+ Syntactic Ambiguity

[[The gangster] [hit [the man [with the gun]]]]

+ Phrase Structure Rules

The two alternatives can be produced by the following context-free (CF) phrase structure rules.

- S → NP VP
- NP → Det N
- NP → Det N PP
- VP → V NP
- VP → V NP PP
- PP → Prep NP

- Det → the
- N → gangster
- N → man
- N → gun
- V → hit
- Prep → with

What other sentences can be produced (or recognised) by the these phrase structure rules?

+ Phrase Structure Rules

- It is possible to generate other sentences with these phrase structure rules
  - the gangster hit the man.
  - the gangster hit the gun.
  - the gangster hit the gangster.
  - the man hit the gangster.
  - the man hit the gun.
  - the man hit the man.
  - the gun hit the man.
  - the gun hit the gangster.
  - the gun hit the gun. ......

+ Context Free Grammars

- How can we represent "agreement" in CF-phrase structure rules?

  - S → NPsg VPsg

  - S → NPpl VPpl

  - NP → NPsg

  - NP → NPpl

  - NPsg → Det Nsg

  - NPpl → Det Npl

  - VPsg → Vsg NP

  - VPpl → Vpl NP

  - Det → the

  - Nsg → author

  - Npl → men

  - Nsg → novel

  - Npl → novels

  - Vsg → writes

  - Vpl → write

Problem is the explosion in the size of the grammar – even for very simple sentences!

+Categories vs Tags

Sentence from: http://stp.lingfil.uu.se/~nivre/docs/05133.pdf

Economic	news	had	little	effect	on	financial	markets	.
Adj	N	V	Adj	N	Prep	Adj	N	.
JJ	NN	VBD	JJ	NN	IN	JJ	NNS	PU

Penn Treebank Part of Speech Tags: https://www.ling.upenn.edu/courses/
Fall_2003/ling001/penn_treebank_pos.html

+ Constituent Structure

From: http://stp.lingfil.uu.se/~nivre/docs/05133.pdf
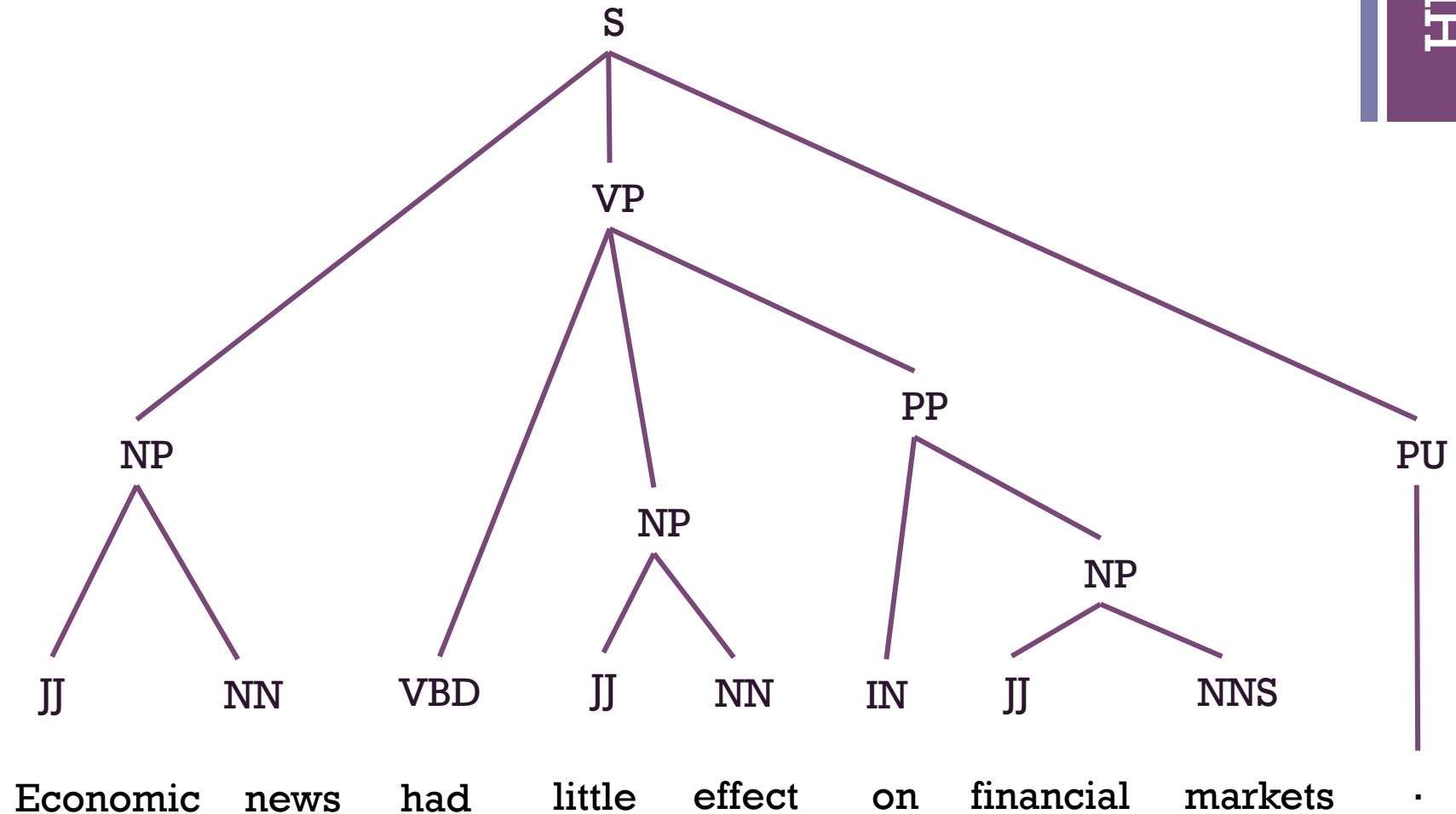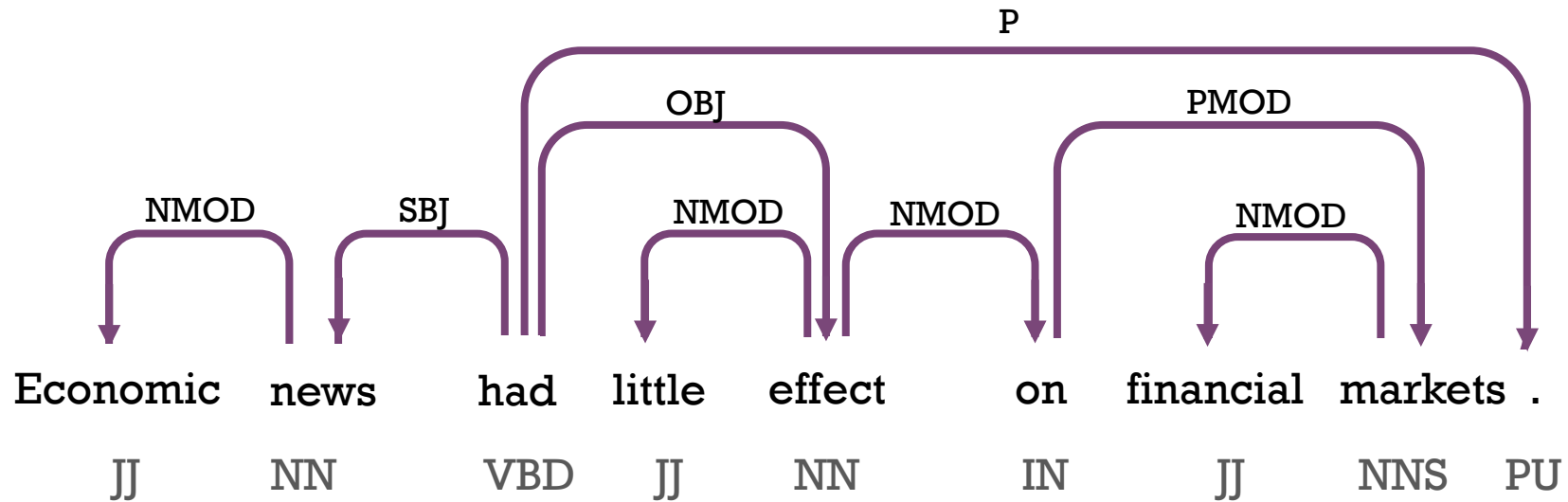
⁺Other Approaches to Syntax

- Transformational Grammar

- Immediate Dominance & Linear Precedence (ID & LP)

- Generalised Phrase Structure Grammar (GPSG)

- Lexical Functional Grammar (LFG)

- Government and Binding Theory (GB Theory)

- Head-driven Phrase Structure Grammar (HPSG)

- Tree Adjoining Grammar (TAG)

+ "Spoken" Language Syntax

- Utterance vs. well-formed sentence

- Noise vs. Disfluencies

filled pause

How much does a single   [uh]   return ticket into Dublin cost?

reparandum   editing phase   repair
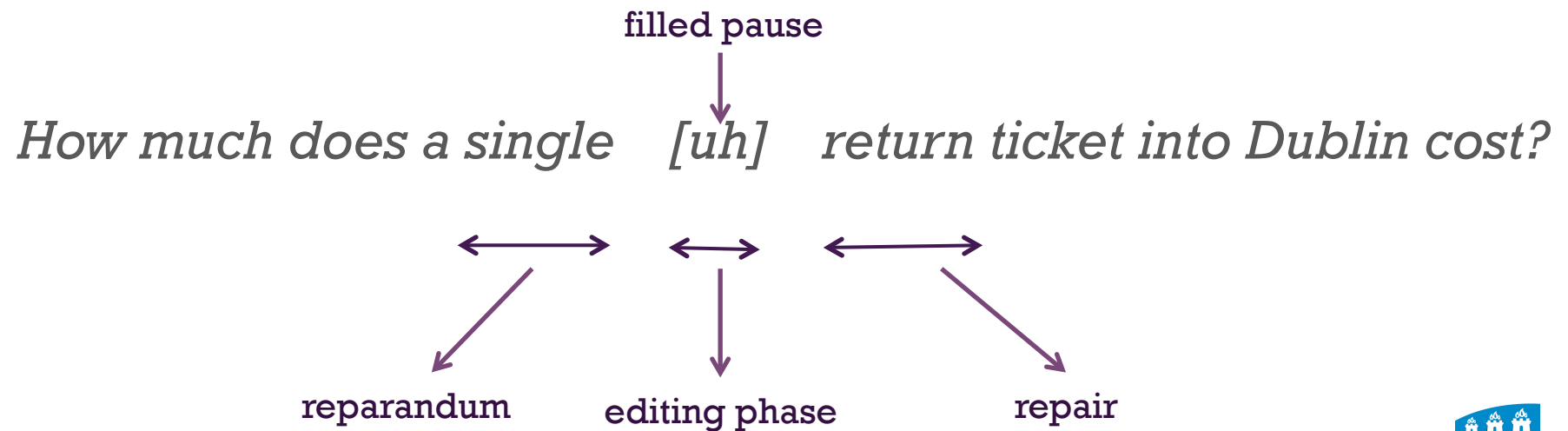
+ Semantics

- Some varieties of meaning (Akmajian et al, 2017)

```
                          ┌──────────┐
                          │ Meaning  │
                          └──────────┘
                  ┌──────────────┴──────────────┐
          ┌──────────────┐              ┌──────────────┐
          │  Linguistic  │              │   Speaker    │
          │   Meaning    │              │   Meaning    │
          └──────────────┘              └──────────────┘
          ┌───────┴───────┐             ┌──────┴────────┐
   ┌──────────────┐ ┌──────────────┐ ┌──────────┐ ┌──────────────┐
   │   Language   │ │   Idiolect   │ │ Literal  │ │  Nonliteral  │
   │   Meaning    │ │   Meaning    │ └──────────┘ └──────────────┘
   └──────────────┘ └──────────────┘
   ┌──────────────┐
   │   Dialect    │
   │   Meaning    │
   └──────────────┘
  ┌──────┴───────┐
┌──────────┐ ┌──────────┐        ┌──────────┐ ┌──────────┐
│ Regional │ │  Social  │        │  Irony   │ │ Metaphor │
└──────────┘ └──────────┘        └──────────┘ └──────────┘
                                       ┌──────────┐
                                       │ Sarcasm  │
                                       └──────────┘
```
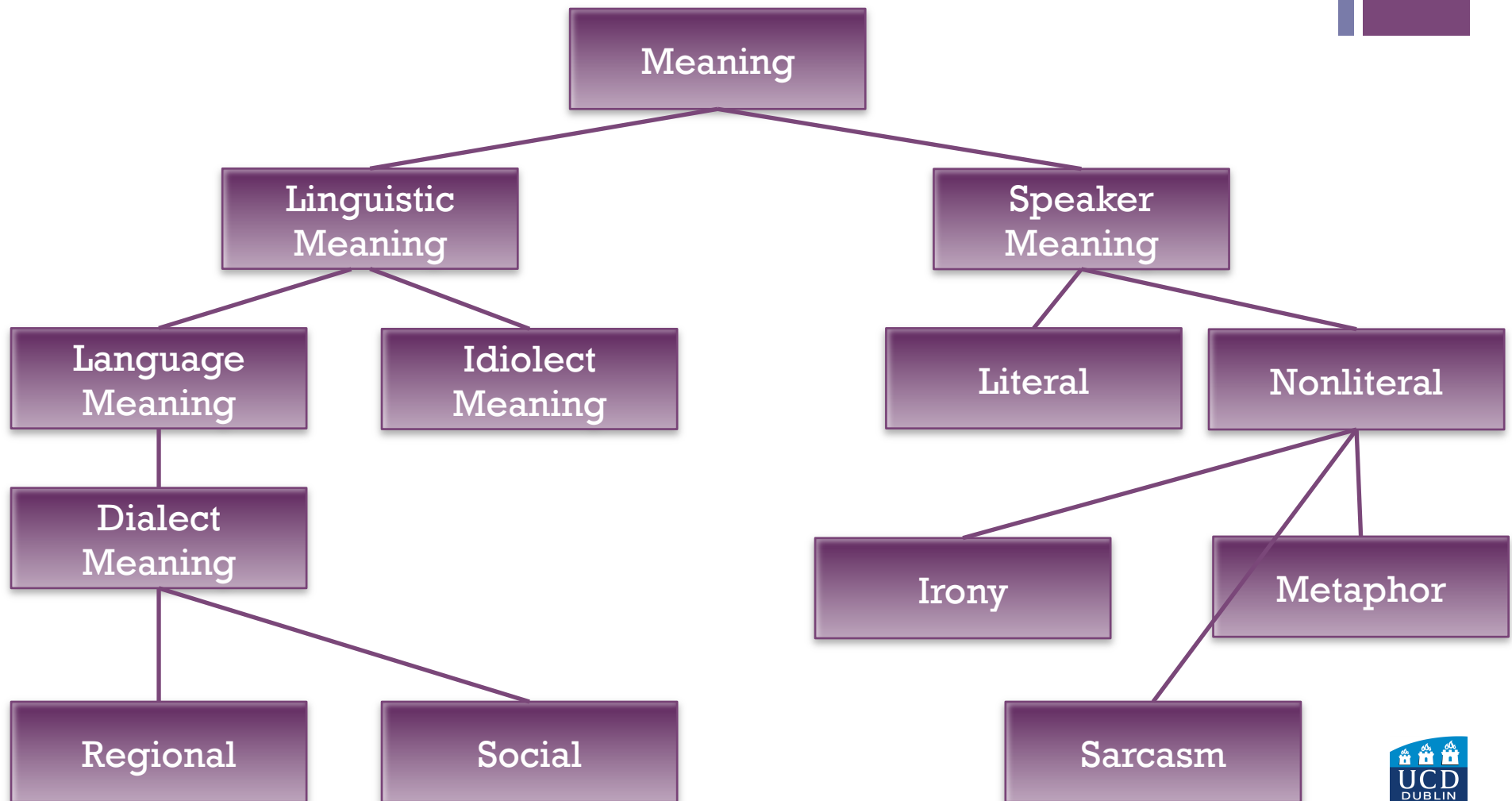
+ Other Types of Ambiguity

- This is the last **leg**.

- You should have seen the **bull** we got from the pope.

- **He** is a Nobel prize winner?

- **He** thinks that Homer is sick.

- If **he** is sick, Homer will stay at home.

+ # Some Semantic Properties

■ **Homonyms**: lead/lead, tail/tale

   → homo<u>graphs</u> vs homo<u>phones</u>

■ **Polysemy**: leg, mouth, branch.

■ **Synonymy**: target/goal, shallow/superficial, think/believe.

■ **Antonymy**: false/true, expensive/cheap, parent/offspring, beautiful/ugly

   → **complementary, gradable, relational opposites**.

+ Features and Redundancy Rules

	woman	father	mare	girl
human	+	+	-	+
female	+	-	+	+
young	-	-	-	+
equine	-	-	+	-
animate	+	+	+	+

Features and Redundancy Rules

	woman	father	mare	girl
human	+	+	-	+
female	+	-	+	+
young	-	-	-	+
equine	-	-	+	-
animate	+	+	+	+

[+ human] $\Rightarrow$ [+ animate]

Features and Redundancy Rules

	woman	father	mare	girl
human	+	+	-	+
female	+	-	+	+
young	-	-	-	+
equine	-	-	+	-
animate			+	

[+ human] ⇒ [+ animate]

+ Features and Redundancy Rules

	woman	father	mare	girl
human	+	+	-	+
female	+	-	+	+
young	-	-	-	+
equine	-	-	+	-
animate			+	

[+ human] ⇒ [+ animate]

[+ equine] ⇒ [- human]

Features and Redundancy Rules

	woman	father	mare	girl
human	+	+		+
female	+	-	+	+
young	-	-	-	+
equine	-	-	+	-
animate			+	

[+ human] $\Rightarrow$ [+ animate]

[+ equine] $\Rightarrow$ [- human]

+ Features and Redundancy Rules

	woman	father	mare	girl
human	+	+		+
female	+	-	+	+
young	-	-	-	+
equine	-	-	+	-
animate			+	

[+ human] $\Rightarrow$ [+ animate]

[+ equine] $\Rightarrow$ [- human]

Note that from this subset we do not know that
[- human] $\Rightarrow$ [+ equine]
does not hold

UCD
DUBLIN

+ Thematic Roles

Of the verb:

- Agent
  - e.g. **the boy** hit the toy with a hammer in the garden

- Theme
  - e.g. the boy hit **the toy** with a hammer in the garden

- Instrument
  - e.g. the boy hit the toy with **a hammer** in the garden

- Location
  - e.g. the boy hit the toy with a hammer in **the garden**

- Source
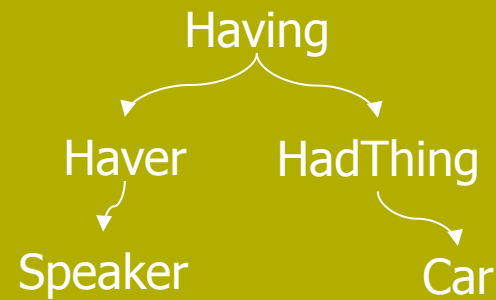 - e.g. the boy lifted the toy from **the table**

Ways of Representing Meaning

I have a car

First Order Predicate Calculus

- $\exists x,y$ Having(x) $\wedge$ Haver(Speaker, x) $\wedge$ HadThing(y,x) $\wedge$ Car(y)

Semantic Networks

Having

Haver   HadThing

Speaker   Car

Conceptual Dependency

Car

$\Uparrow$ POSS-BY

Speaker

Frame-based Representation

Having:

    Haver: Speaker

      HadThing: Car

+ Pragmatics

- Pragmatics looks at the way in which context contributes to meaning
  - Did you know that he is a Nobel prize winner?
  - We can eat the pizza now.
  - It's over there
  - You're fired!

- Language in Use
  - Interaction
  - Deixis
  - Maxims of Conversation (Grice, 1975)
  - Speech Acts (Austin (1962), Searle (1975),…)

+ Discourse

- "Knowing" a language enables us to combine sentences together to express complex thoughts and ideas.

- Such larger linguistic units are called discourse → concerned with the meaning relations among sentences.

- Pronouns, Articles
  - It seems that the man loves the woman. Many people think he loves *her*.
  - Whenever I see you, I think of *her*.

+ Functions of Language

- Transactional (Expression of Content)
  - message oriented
  - informative detail correct: 'The maternity unit is on the fourth floor.'
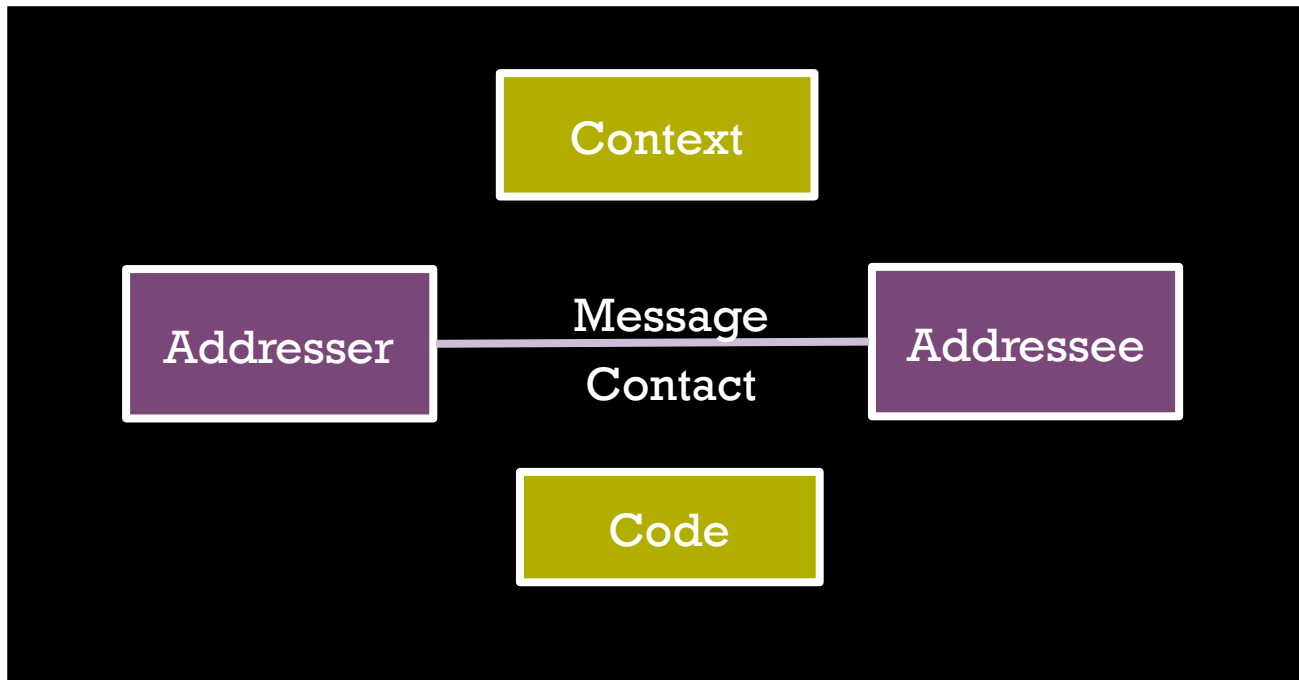
+ Functions of Language

- Interactional (Expression of Relations/Attitudes)
  - phatic use of language
  - contributions to conversation: weather: 'My goodness, it's cold!'