Towards a constructicon using patterns and frames

Florent Perek

Amanda Patten

University of Birmingham

UNIVERSITY OF BIRMINGHAM | COLLEGE OF ARTS AND LAW
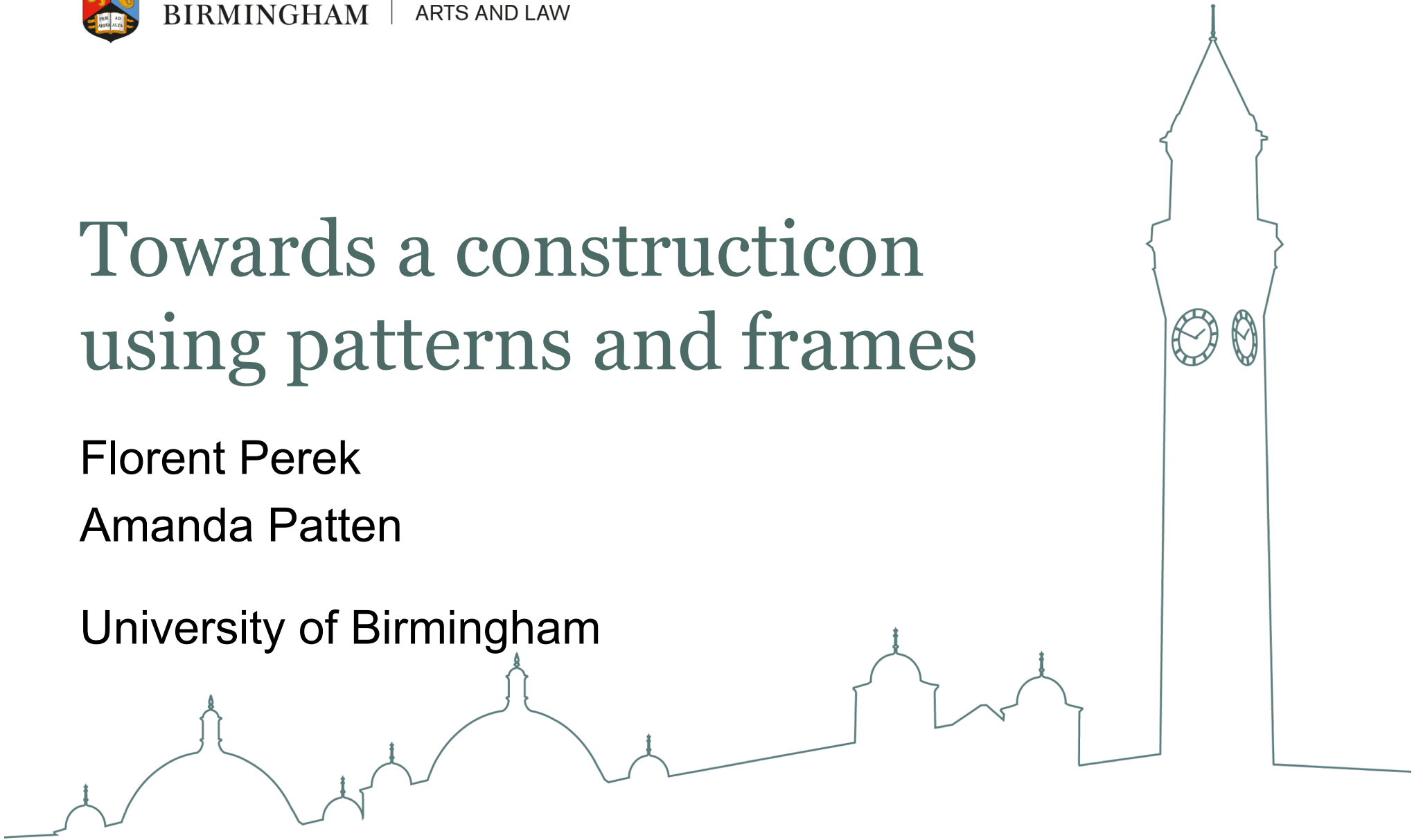
Overview

- Outline and first results of a new project

- Proposal: merge two corpus-based resources, the COBUILD grammar patterns and FrameNet

  - Automatic method and quantitative results

 - Two qualitative case studies

COBUILD

- Lexicographic project started in the 1980s by John Sinclair with Collins publishers in Birmingham

- Design dictionaries entirely from authentic corpus data

- One key insight in particular

  – A word is better described in terms of its typical uses

  – This notably includes the syntactic frames or "patterns" it can occur in
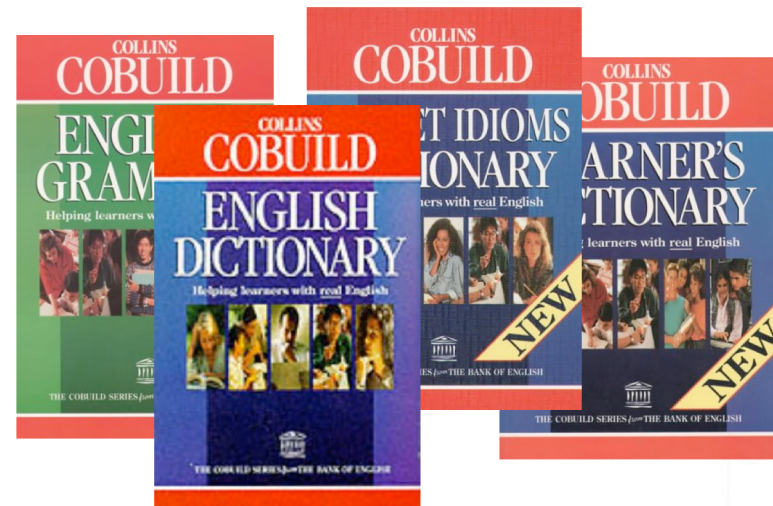
The COBUILD Grammar Patterns

- Proposals for compiling a pattern grammar of English (Francis 1993, Hunston & Francis 2000)

  → The COBUILD Grammar Patterns series

- List of all the patterns mentioned in the COBUILD entries

  - Volume 1: verbs (Francis et al. 1996)

  - Volume 2: nouns and adjectives (Francis et al. 1998)

- List all lexical items attested in these patterns
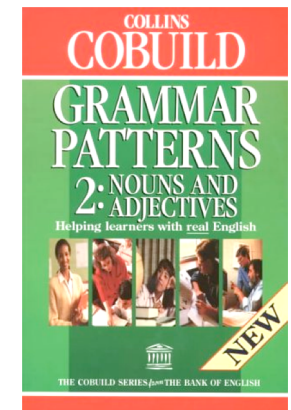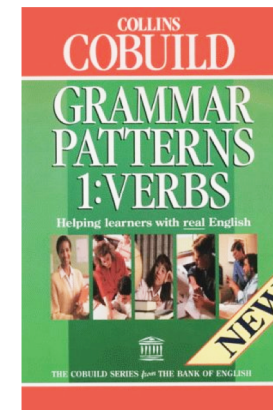
Francis, G. (1993). A corpus-driven approach to grammar – principles, methods and examples. In Baker, M., Francis, G. & Tognini-Bonelli, E. (eds). *Text and Technology: in Honour of John Sinclair*. Amsterdam: Benjamins, pp. 137–156.
Francis, G., Hunston, S. & Manning, E. (1996). *Collins COBUILD Grammar Patterns 1: Verbs*. London: HarperCollins.
Francis, G., Hunston, S. & Manning, E. (1998). *Collins COBUILD Grammar Patterns 2: Nouns and Adjectives*. London: HarperCollins.
Hunston, S. & Francis, G. (2000). Pattern Grammar: A corpus-driven approach to the lexical grammar of English. Amsterdam: Benjamins.

The COBUILD Grammar Patterns

- ☐ 124 patterns for lexical verbs in Francis et al. (1996)

- ☐ Simple notation: V n, V that, V with n, V n to n, …

- ☐ 10,522 verbs listed under the patterns

- ☐ In each pattern, the verbs are grouped into meaning groups (816 in total, avg. 6.6 groups per pattern)

(figures calculated from the XML version provided by HarperCollins)
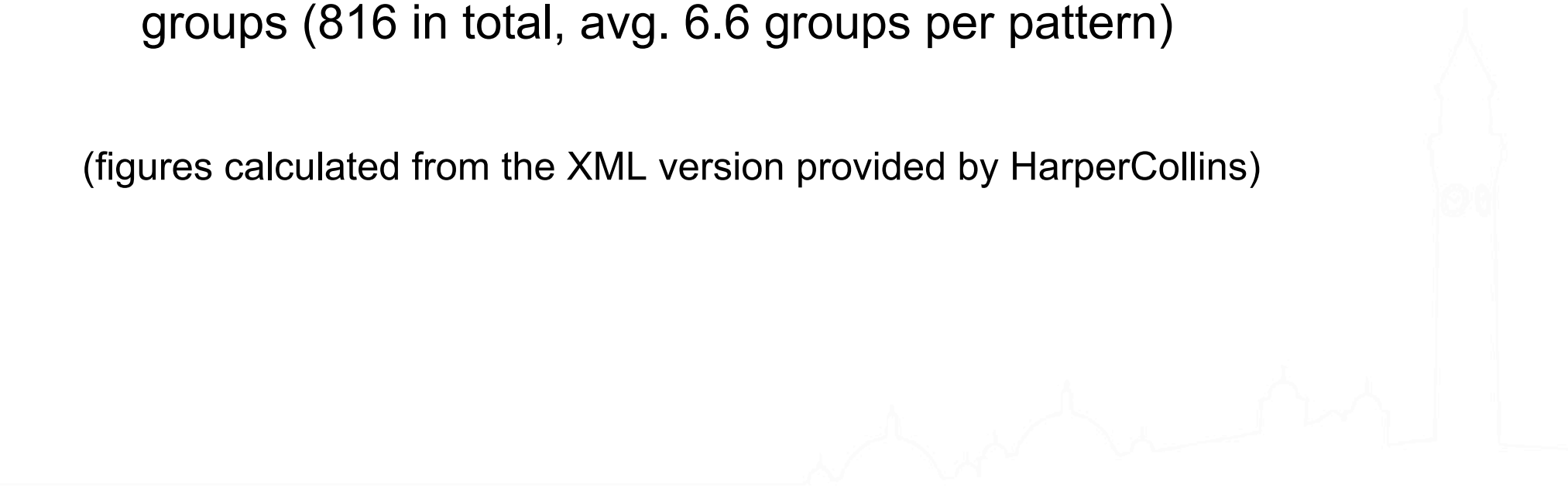
The COBUILD grammar patterns

Example: **V n *of* n**

☐ Verb followed by NP and *of*-PP

☐ Three meaning groups

- The 'rob' and 'free' group: … **cure** *her of a disease,* … **robbed** *them of their watches* (24 verbs)

- The 'inform' group: … **assured** *us of their help* (11 verbs)

- The 'acquit' and 'convict' group: … **clear** *him of attempting to murder,* … **suspected** *him of perjury* (5 verbs)

- 11 other verbs

FrameNet

https://framenet.icsi.berkeley.edu

- Aims to describe the lexicon of English in terms of **semantic frames**

- Frames describe basic scenarios or situations that underlie word meanings

- Contain actors and props, called **frame elements** (FEs)
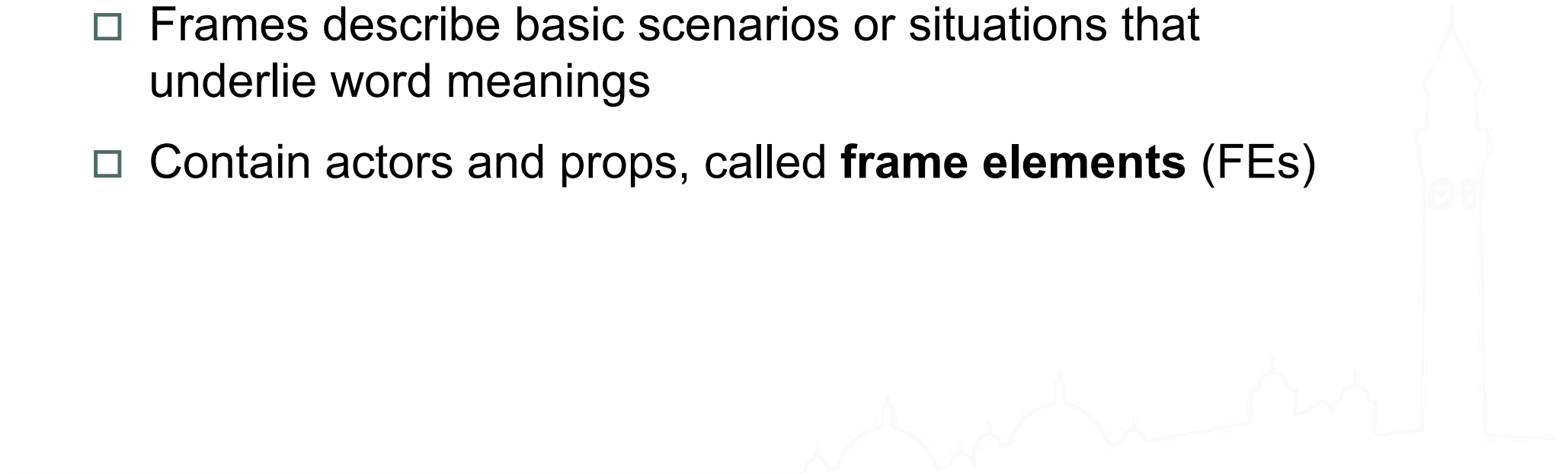
Giving

Definition:

A Donor transfers a Theme from a Donor to a Recipient. This frame includes only actions that are initiated by the Donor (the one that starts out owning the Theme). Sentences (even metaphorical ones) must meet the following entailments: the Donor first has possession of the Theme. Following the transfer the Donor no longer has the Theme and the Recipient does.

> Barney GAVE the beer to Moe.

> $300 was ENDOWED to the university to build a new performing arts building.

FEs:

Core:

Donor [Donor]	The person that begins in possession of the Theme and causes it to be in the possession of the Recipient .
Recipient [Rec]	The entity that ends up in possession of the Theme.
Theme [Thm] **Semantic Type:** Physical_object	The object that changes ownership.

Non-Core:

|	
Circumstances [cir]	The Circumstances are the conditions under which the Theme is given. I GIVE my services free of charge.
Depictive [dep]	A description of the Donor, Recipient, or Theme given independently of the giving event per se.
Explanation [Exp]	The Explanation for which the Donor gives the Theme to the Recipient.

FrameNet

- A word can belong to more than one frame

- Frame + lemma = Lexical Unit (LU)

- Frame elements (FEs) can be realized with the LUs

  - Core FE: obligatorily present in all uses of the frame, may be realized as major clause elements (subject, object etc.)

  - Non-core FE: peripheral and typically optional information (often adverbials and modifiers)

- A frame is <u>not</u> a definition; rather, a higher level of lexicographic description
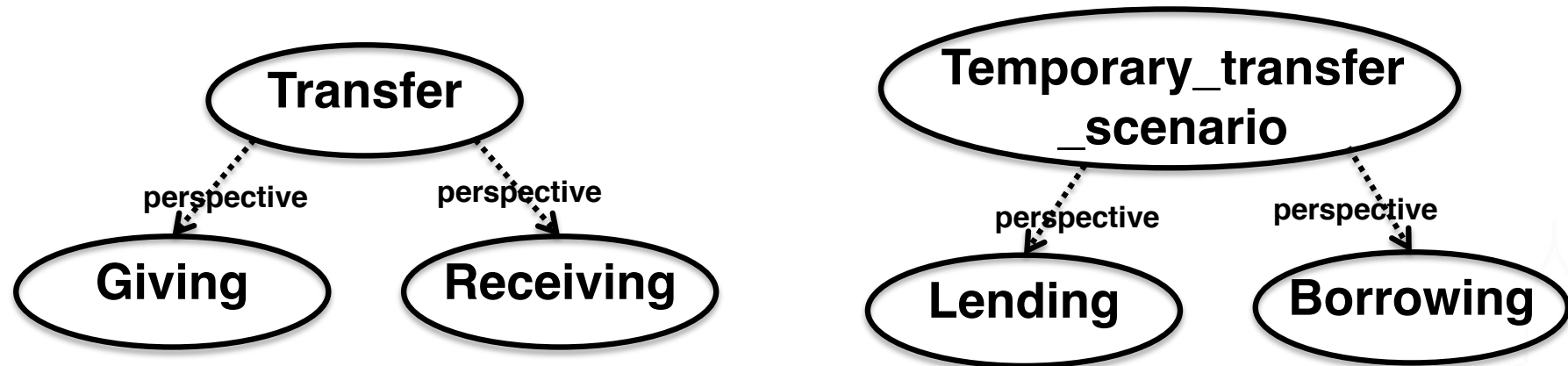
Frame-to-frame relations

- FrameNet also describes how each frame is related to other frames in the database

- **Inheritance:** relates frames in a taxonomy
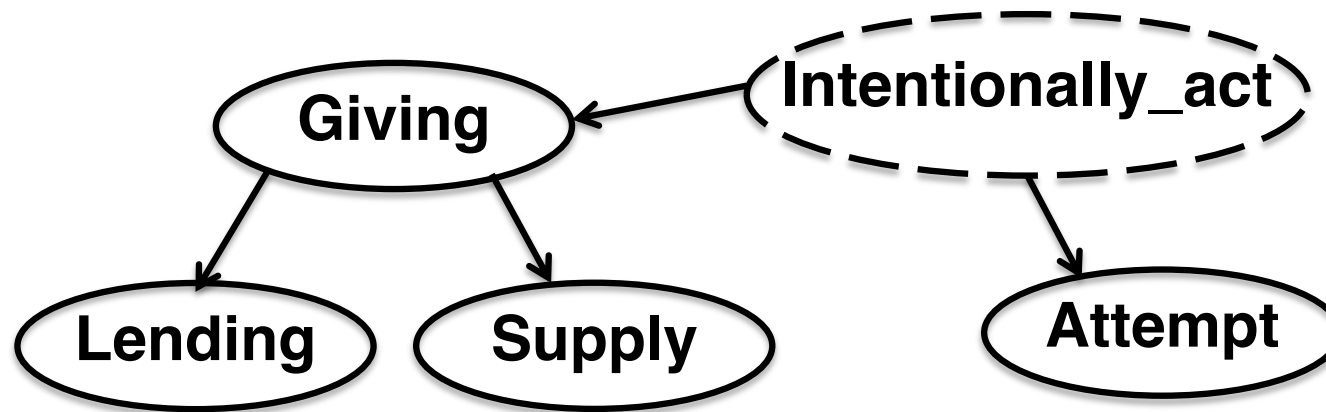
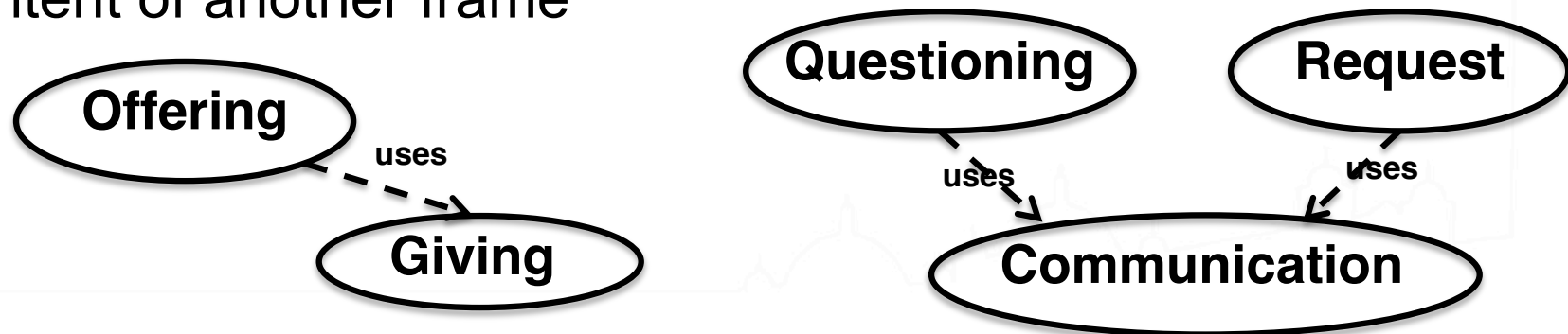- "Intentionally_act" = non-lexical frame: frame with no LUs

Frame-to-frame relations

☐ **Perspective**: construes an event from a certain perspective, in particular one of the FEs'

☐ **Use:** the content of a frame is required to understand the content of another frame

FrameNet

☐ Corpus data is used to discover and document frames

☐ The database contains selected corpus examples with a description of how frame elements are realized

☐ Makes it possible to extract argument realization information of LUs

Lexical Entry

give.v

Frame: Giving

Definition:

COD: freely transfer the possession of; cause to receive or have.

Frame Elements and Their Syntactic Realizations

The Frame Elements for this word sense are (with realizations):

Frame Element	Number Annotated	Realization(s)
Donor	(52)	CNI.-- (12) DNI.-- (2) NP.Ext (37) PP[by].Dep (1)
Manner	(3)	AVP.Dep (3)
Purpose	(4)	VPto.Dep (4)
Recipient	(52)	PP[to].Dep (16) DNI.-- (5) INI.-- (2) NP.Ext (4)

[X] Katy and Jamie got ready very quickly and Mum GAVE each of them two wee spoons .
[X] They wrapped it up and GAVE it to her , and it did have a head like a baby .
[X] I 'm just going to GIVE her some milk . "
[X] I GIVE him coffee .
[X] Once they stopped the drugs they were GIVING me , my hair started to come back .

COBUILD vs. FrameNet

COBUILD

Focus on lexicogrammar

> What patterns are there?

> What words can be used in them?

Meaning is secondary

> Ad hoc meaning groups in each pattern

> No systematic pairing with meaning

FrameNet

Focus on meaning

> What frames are there?

> What words evoke them?

Lexicogrammatical information = addendum

> Added through examples

> No systematic inventory, by word or across words

COBUILD vs. FrameNet

□ Complementary resources

□ Proposal: match the verbs in the COBUILD patterns entries to FrameNet lexical units

□ Potential to turn the patterns into a constructicon: inventory of form-meaning pairs (Goldberg 1995)

– Form = pattern

– Meaning = generalization over frames used in the pattern

– More than one possible construction for the same pattern

Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. University of Chicago Press.

Method

☐ Automatic procedure using the XML version of FrameNet and the COBUILD patterns (provided by HarperCollins)

☐ Every verb listed in each pattern is looked up in FrameNet

  – If found, this returns one or more LUs

  – For each lexical unit, the annotated examples are consulted (if any)

  – If the valency realization of the frame elements matches the pattern, the LU is mapped onto the COBUILD entry

 – NB: only core frame elements are considered

Method

- Phrasal verbs were ignored

- Some patterns could not be matched to FrameNet

  - Patterns with 'dummy' *it*

    e.g., V *it* adj that

  - Missing grammatical distinctions in FrameNet

    e.g., V n-pl   (NP number not coded in FrameNet)

- 78 patterns matched to FrameNet
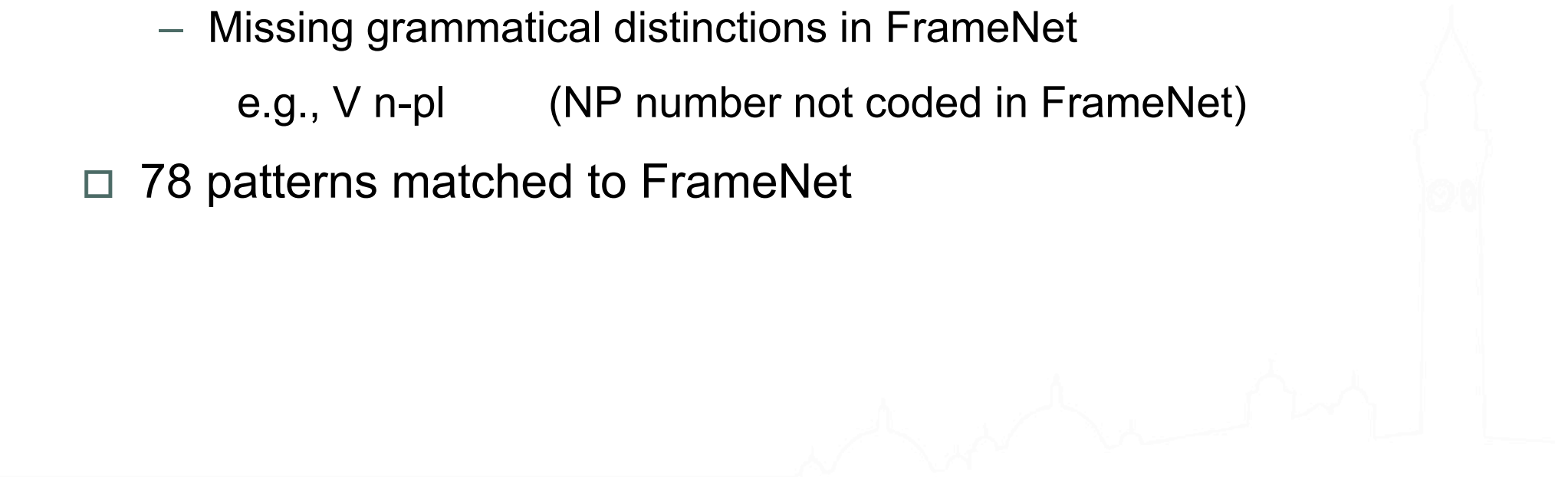
Results

Only **40.5%** of the entries in the COBUILD verb patterns matched to at least one LU in FrameNet (3063 out of 7572)
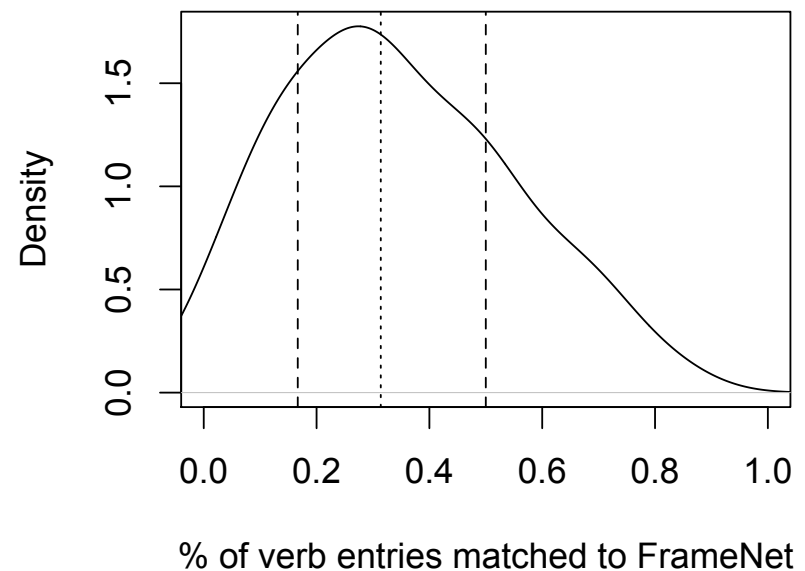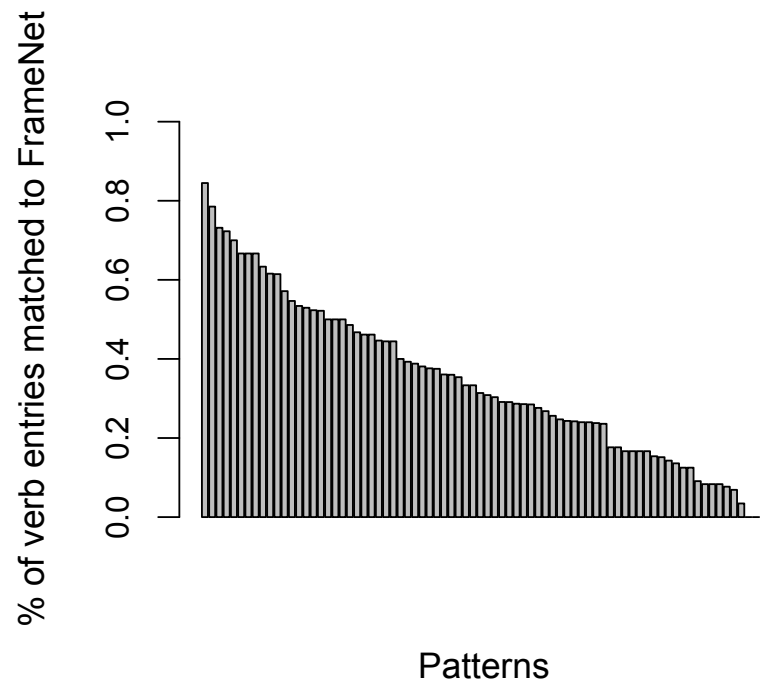

Only about 25% patterns have 50% or more matches

50% have between 17 and 50% matches

25% have less than 17% matches

Results

☐ Still insufficient coverage in FrameNet

☐ Problems with non-core frame elements

  – E.g., Addressee for Communication, Explanation for Death

  – Prevents these frames from being matched to "V n to n" and "V of" (for instance)

☐ Annotation errors and inconsistencies
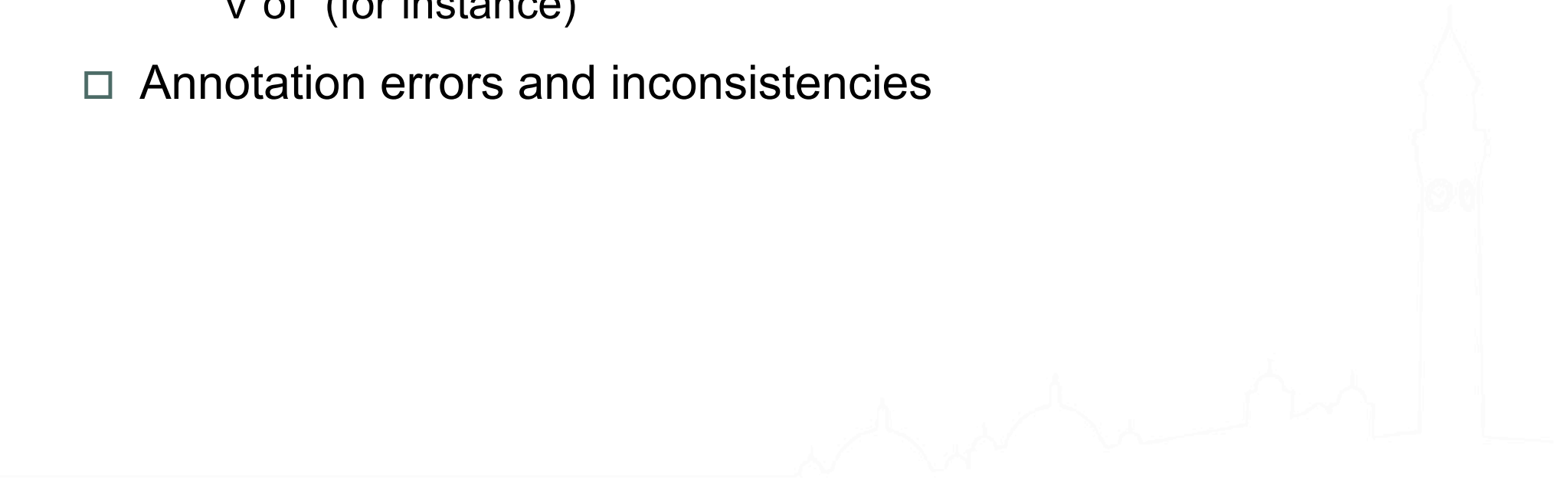
Two case studies

- Matching the patterns to FrameNet will necessitate a lot of manual intervention

- Yet this would create a useful new resource

- Two case studies:

    - From patterns to frames: what frames do we get when we look at a particular pattern? How are they related?

    - From frames to patterns: what verbs evoke a particular frame and in what patterns can they be used?

From patterns to frames

- ☐ Example: "V that"

- ☐ 255 verbs (w/o phrasal verbs)

- ☐ 10 meaning groups, for instance:
  - – The 'say' group: *claim, complain, insist, report, say,* …
  - – The 'think' group: *assume, know, think, understand,* …
  - – The 'show' group: *confirm, demonstrate, reveal, show,* …

- ☐ 62% were matched to at least one lexical unit

- ☐ Further annotation work was carried out to provide a better picture

A tight network: the 'say', 'add', and 'scream' groups (172 LUs)
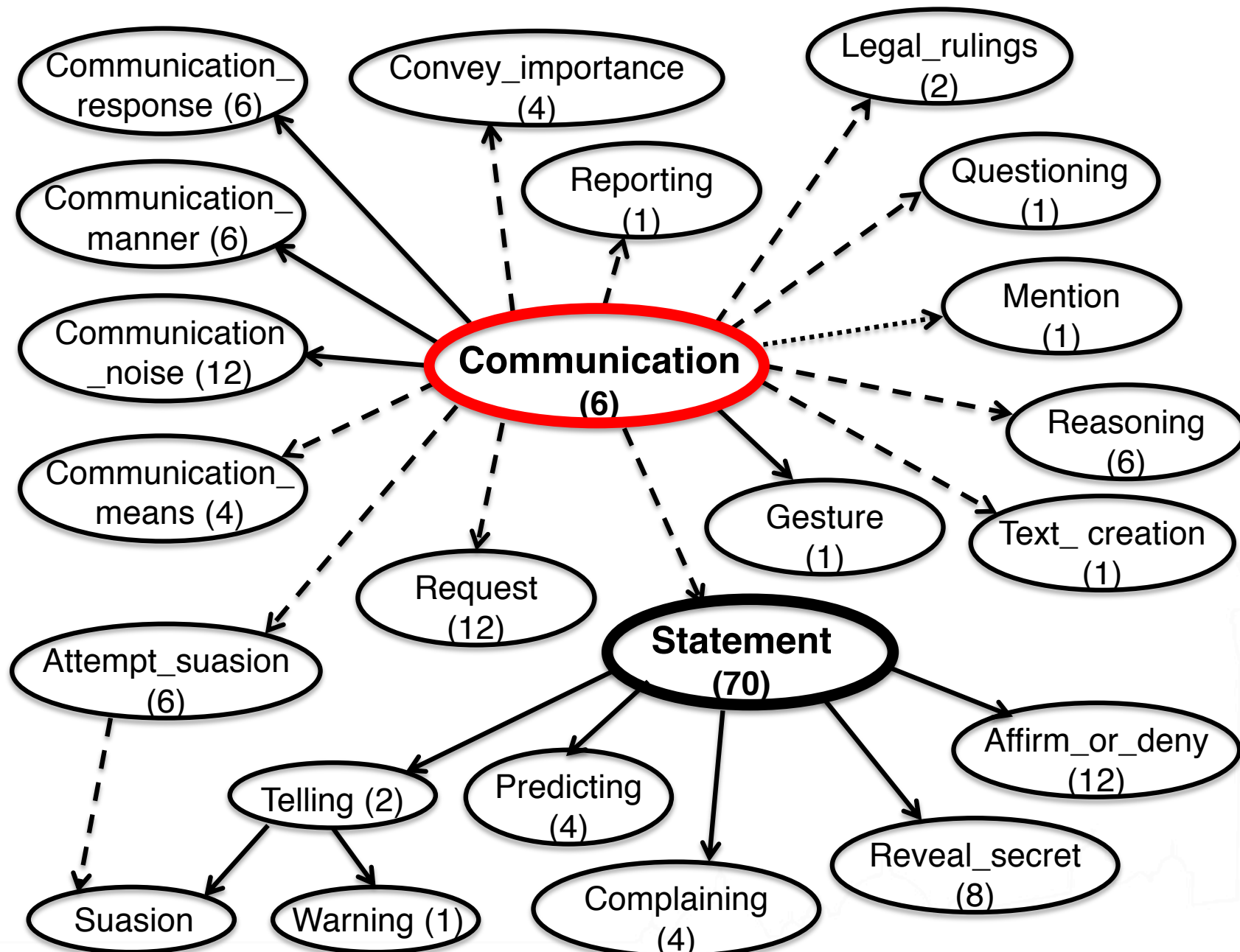
The "V that" Communication construction

- ☐ Communication frame
  - – The one frame that unifies all lexical units
  - – Can be seen as the 'schema' shared by all uses

- ☐ More about different uses of communication than different forms: make a statement, a request, persuading, etc.

- ☐ Statement frame (verbal communication to make a claim)
  - – The most typical use: 70 LUs (101 with subframes)
 - – Can be seen as prototype, or 'core' constructional meaning

A looser network: the 'think', 'discover', and 'love' and 'hate groups (110 LUs)
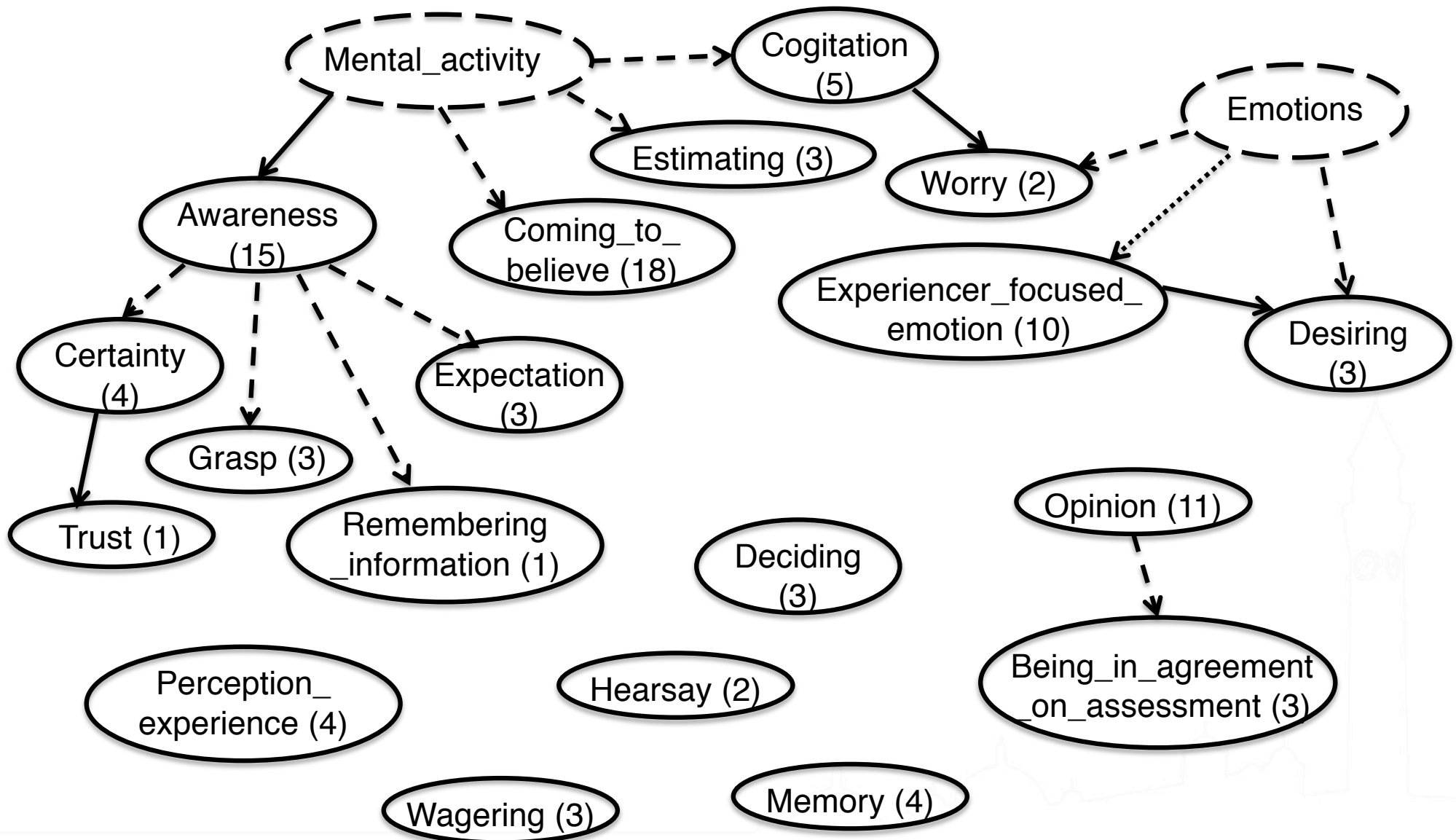
The "V that" Mental_activity & Emotions construction(s)

- ☐ Two partially overlapping networks centered on Mental_activity and Emotions

- ☐ A lot of orphans: Deciding, Memory, Opinion, …

- ☐ Highlights frame relations that are not recorded in FN

- ☐ Awareness (*know*), Opinion (*believe*), Experiencer_focused_emotion (*fear*), and Coming_to_believe (*realize*) are among the

- ☐ Cluster of related constructions rather than single generalization
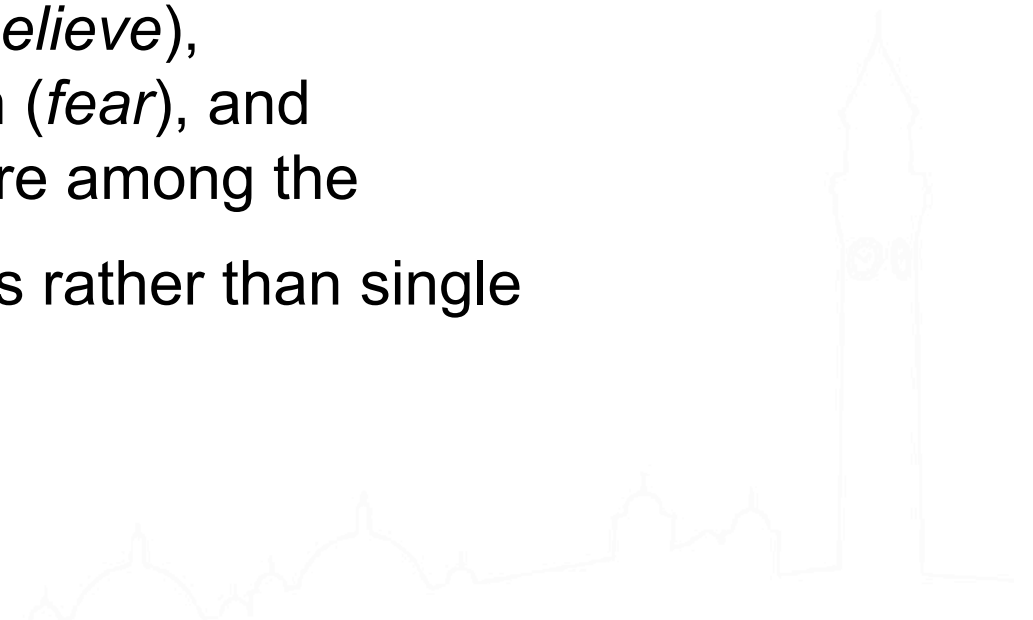
From frames to patterns

- We can also use FrameNet + COBUILD to compile lexico-grammatical information from the perspective of meaning

- Example: the **Evidence** frame

  "The Support, a phenomenon or fact, lends support to a claim or proposed course of action, the Proposition"

  "Proposition: This is a belief, claim, or proposed course of action to which the Support lends validity"

  "Support: Support is a fact that lends epistemic support to a claim, or that provides a reason for a course of action"

- Highly relevant to academic writing

- What verbs and patterns can be used to express it?

From frames to patterns

V n

Support (n) *confirm* Proposition (n)
indicate
prove
reveal
rule out
show
suggest
support
tell

From frames to patterns

V that

Support (n)   *attest*   Proposition (that)
confirm
demonstrate
indicate
mean
prove
reveal
show
suggest
testify
verify

From frames to patterns

V wh & V wh-to-inf

Support (n)   *illustrate*   Proposition (wh)
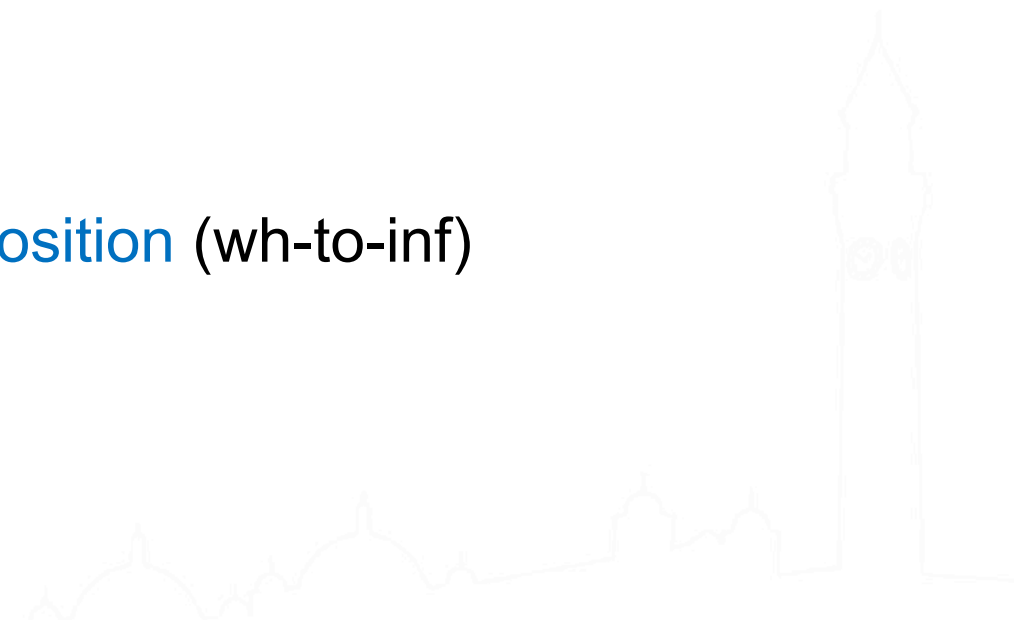   *indicate*
                           prove
                           reveal
                           show
                           suggest

Support (n)   *illustrate*   Proposition (wh-to-inf)
                           indicate
                           reveal
                           show
                           suggest

From frames to patterns

V to n

Support (n)   *attest*   Proposition (to n)
             testify

V for n

Support (n)   *argue*   Proposition (for n)

**V against n**

Support (n)   *argue*   Proposition (against n)

V in favour of n

Support (n)   *argue*   Proposition (in favour of n)

Summary

☐ The COBUILD Grammar Patterns and FrameNet can benefit a lot from each other

☐ A lot of manual processing still necessary to merge the two resources

☐ Frames can be used to turn patterns into constructions

☐ Many applications for language teaching: constructions as teaching tools, course material & course book design, etc.
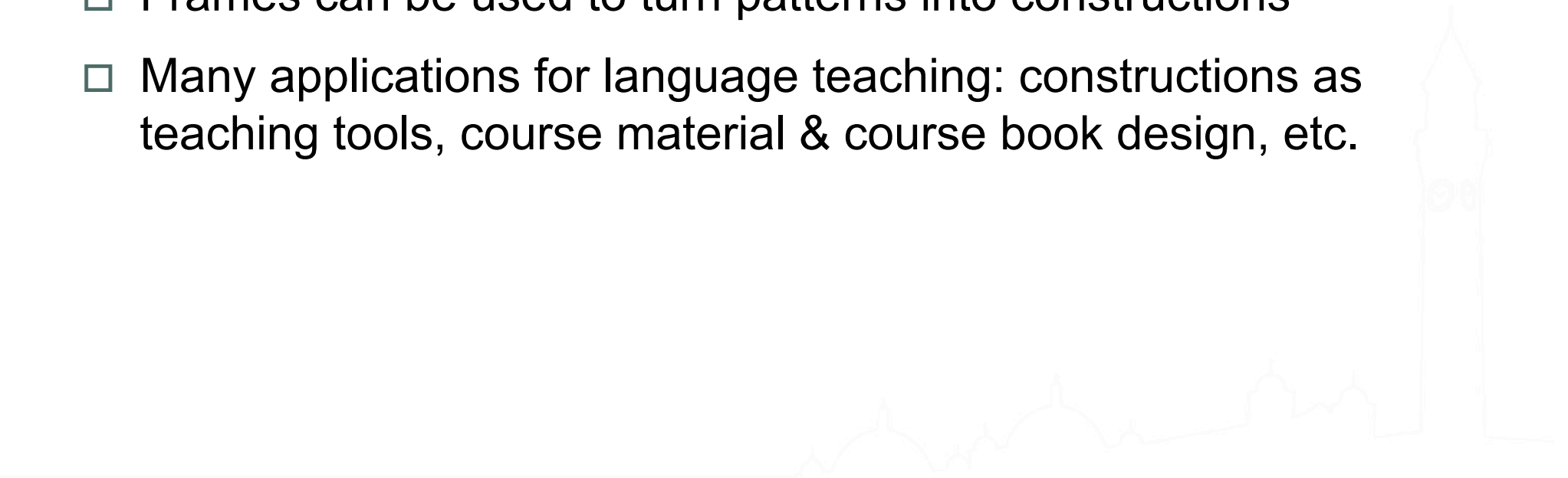
UNIVERSITY OF BIRMINGHAM | COLLEGE OF ARTS AND LAW

Thanks for your attention!

f.b.perek@bham.ac.uk

www.fperek.net