Terminology

Lesson Four

More on recognizing terms

Ferns are a very ancient family of plants

Early fern fossils predate the beginning of the Mesozoic era, 360 million years ago. They are older than land animals and far older than the dinosaurs. They were thriving on Earth for two hundred million years before the flowering plants evolved. Most ferns are leafy plants that grow in moist areas under forest canopy. They are "vascular plants" with well-developed internal vein structures that promote the flow of water and nutrients. Unlike the other vascular plants, the flowering plants and conifers, where the adult plant grows immediately from the seed, ferns reproduce from spores and an intermediate plant stage called a gametophyte.

Fern reproduction

- *Flowering plants (and conifers) reproduce when pollen from a male flower - carried by wind, insect or other vector - fertilises the female flower. Many flowering plants, of course, include both male and female parts in the same flower. The male pollen cell carries half the genetic material of the adult plant and fuses its genetic material with that of the female cell, which carries the other half. The complete, fertilised cell grows into the seed, which, when ripe and when it finds itself in suitable soil and moisture, is capable of producing a complete adult plant. Higher plants have a very robust propagation system: the pollen from the male flower is very hardy, and the female flower nurtures the seed until it is ready to grow. The seeds themselves are often very durable, able to wait for long periods in adverse conditions before they grow. So the higher vascular plants have evolved to occupy nearly every niche on the land surface of the earth.*

- *Ferns do it differently. They have a more complicated method that depends on there being liquid water for the process to complete. As a result, they can only reproduce where there is sufficient moisture: the reproduction process itself requires moisture.*

How do the criteria help?

- A reminder of the criteria previously mentioned:

  - semantic : the term refers to a single concept
  - quantitative : frequency in similar sorts of texts
  - taxonomic : part of a hierarchical classification
  - synonymic : synonym of a recognized term
  - neological: name for a new concept in the field
  - typographical: use of inverted commas, italics, etc.

Terms refer to a single concept

- In a new field, it is not always easy to identify concepts (we need terms for this).

- It may help to identify the concept (i.e. a class) from the referent:
  - *male pollen cell* : this refers to one (sub)class of cells
  - *fertilised cell*: this is not a subclass of cells but a stage (single concept) in the life cycle.

quantitative

- The extracts we have are far too short to use statistics **but**

  - certain words uncommon in ordinary English (see monitor corpus)

    are frequent here:

    - *fern*
    - *plant*
    - *seed*
    - *moist*

taxonomic

- Part of an established system of classification

    - ***Mesozoic era*** is one geological era of the Phanerozoic eon
    - It is itself further divided into periods.

    This is a pure taxonomy.

Part of a hierarchy

- – Without being part of a strict taxonomy, concepts can be ordered into a hierarchy. This is an element of termhood.

- – We learn that there are three classes of *vascular plants*: *ferns*, *conifers* and *flowering plants*.

- – These three hyponyms are thus terms, and so is the corresponding hyperonym, *vascular plant*.

synonymic

– Synonymy is a difficult concept to handle. One reason for terminology work is to keep tabs on synonyms, and eliminate them when they cause confusion.

– But if we find a synonym of a word we already have identified as a term, then this is also a term.

– This **may** be the case for

- *Higher plant* and *higher vascular plant* (reduced form)

neological

- Very often in an cutting-edge subject field, new concepts are developed all the time and the words to express them will also be novel.

- Note the three main classes of neology
  - New combinations of morphemes
    - Free morphemes : *vascular plant* (no longer new!)
    - Bound morphemes : *fertilize* (no longer new!); *termhood*
    - Combining forms *: meso-zo-ic*
  - New meanings to existing words
    - Cf. *material* in *genetic material*
  - Borrowings

typographical

- 'vascular plants' in inverted commas, so the author is indicating to the reader that it is a term.

- Authors do this in other ways as well:
    - **called a** *gametophyte*

    Can you think of any other similar textual markers?

Other criteria

- Artificiality
  - Terms are words made up to express specialised knowledge. When this knowledge is new, it needs to be expressed in words. Sometimes specialists use systematic naming systems (taxonomies) or use Greek or Latin
    - *gametophyte; Mesozoic era…*

    So the more artificial a word looks, the more likely it is to be a term.

Artificiality (2)

- The contrary is not necessarily true:

  - An ordinary word can also be a term
    - Process of terminologization
      - Cf *water* in chemistry, physics

      - *Flowering plant* is a term, because it is part of a hierarchy

Requiring special learning

- Since terms are the names given to concepts in specialized fields, it follows that it takes special learning to understand them.
  - Thus *vascular plant* seems more a term than *flowering plant.*
  - But even words of everyday language may require special learning to be correctly understood:
    - cf *genetic* ***material.***

The meaning of the whole is different (more?) than that of its parts

- This is another semantic criterion
  - We can say that *genetic material* has a different meaning from that of its parts (at least for *genetic*).

Appearing in a
knowledge rich context

- « Knowledge rich context, indicating at least one item of domain knowledge that could be useful for conceptual analysis. In other words, the context should indicate at least one conceptual characteristic, whether it be an attribute or a relation »

  – Meyer, Ingrid (2001), « Extracting Knowledge-rich Contexts for Terminography: A Conceptual and Methodological Framework » in Bourigault et al. *Recent Advances in Computational Terminology*. Benjamin, Amsterdam/Philadelphia, p. 281

Part of a conceptual structure

- This is probably the most basic criterion, but one that is difficult to determine, when one is not familiar with the field.
- It is the case for hierarchies
- It is the case for parts of a whole
  - *male pollen cell* – part of the flower
- It is the case for stages of a cycle
  - *adult plant* (what is the co-hyponym?)
  - *gametophyte* (intermediate plant stage)
          and many other terms in plant life cycle

Supplementary criteria

- Marie-Claude L'Homme (2004 : 64-68) suggests four criteria for determining what a term is.

1. A lexical unit whose meaning is linked to a specialized field

    (This is equivalent to our previous indices)

Criterion 2

- Predicative lexical units (verbs, adjectives) can be terms if their actants are terms.

  - **Example :** *Ce système n'est pas conçu pour adresser plus de 1024k de mémoire.*

  The actants are *système (informatique)* and *mémoire*. The verb *adresser* thus has term status in the field of information technology.

Proviso for Criterion 2

- The predicative unit thus identified must only have a specialised meaning when used with the actants which are terms.

- If it has the same meaning as it does with non-specialised actants, then it is not a term.
  - *La carte principale comporte des circuits intégrés…*

Criterion 3

- If an already identified term is a member of a morphological family, other members of the family will probably also be terms:
  - *compilateur* is identified as a term
  - Therefore *compilable*, *compiler, recompiler*… will also be terms.

  - If *fertilisation* is a term…

Criterion 4

- Any lexical unit in a paradigmatic relationship with an already identified term is also a term:
  - If *interface* is a term, then *menu* and *window* will be too,

    as these are part of the interface (meronymic relationship).

And what are (probably) **not** terms

- Vague words
  - *Male and female **parts***

- Errors in bracketing

  adult fuse (fuse here is a verb!)

Why *gamtophyte* is a term

- it is a word that is not heard in ordinary conversation
- it would be statistically infrequent in a general language corpus (eg. newspapers)
- it would be generally frequent in a specialised corpus (eg. Articles on botany)
- therefore it is typical of a subject field or domain
- it requires special learning to be understood
- it is "artificial" in as much as it was made up by botanists from Greek roots

And why *vascular plant* is a term

Hyponyms of *vascular plant*:
- *fern*
- *flowering plant*
- *conifer*.
- This sort of series is called a **paradigm**

  – (We should look for *non-vascular plants*)

Concept structuring

- First establish hierarchical relationships
  - This entity is a member of this class
    - « isa » relationship
- Secondly establish meronymic relationships
  - This entity is part of this whole
    - Or this whole has a <entity> (« hasa » relationship)

Ordering relationships

- Thirdly establish other types of relationships and label them clearly

- Fourthly label relations which you are not sure of so that you can come back and decide when you have more information.

Terminology assignment

- An in-depth analysis of five terms from corpus linguistics
  - Aims:
    - explain how you identified the terms
    - Explain how you treated them
      - In particular how you defined them
    - Comment on the difficulties you experienced
      - Possibly including term candidates you eventually rejected

Identifying terms

- How do you indentify terms from non-terms in your text?

- How do you determine what is part of a term and what is not?

definition

- What sort of definitions do you use?
- How do you justify this choice?
- What modifications are you obliged to make to achieve a satisfactory definition?

- How do you validate your definitions ?

Concept structure

- ISO 704 recommends situating the term in its conceptual structure.

- Have you been able to visualize the relevant conceptual structure, for example in the form of a tree diagram?

Morphological treatment

- Is the morphology of the terms different from that of everyday language
  - Eg *pollen* is non-count in ordinary English, but count in aerobiology
- Corresponding derivatives
  - French *plaquette: plaquettaire*
    - *tissu – tissulaire*

- Do you find synonyms in complementary distribution? *kidney, renal, nephr* ?

Syntax and context

- Cooccurences

- Syntactic constraints?

To be handed in

- On 17th December – on paper.