Computational Semantics

Introduction

Shu-Kai Hsieh

September 24, 2021

Graduate Institute of Linguistics

National Taiwan University

Table of contents

Introduction

Course logistics

- Course website (`http:https://loperntu.github.io/comsem/`)
- Check the updated schedule constantly
- Active participation: lectures, reading and coding assignments, (seminar discussion), term paper

Computational Semantics: What and Why

- Semantics is the study of meaning (communicated through language).
- Linguistic description is an attempt to reflect a speaker's linguistic knowledge, the semanticist is committed to describing **semantic knowledge**.
- Computational semantics and AI core (NLP, NLU).

- **自然語言處理** | Natural Language Processing **vs** **計算語言學** | Computational Linguistics
- **自然語言理解** | Natural Language Understanding **vs** **計算語意學** | Computational Semantics

Meaning and Computation

Linguistics-oriented : Using computational models to gain a better understanding of how language works.

Engineering-oriented : Using computational models to build language technology/applications.

Computational Semantics

- formally (and scientifically) describe the meanings of human/natural language.
- reveal the mechanisms of (meaning) understanding (incl. how meanings are learned, processed, stored, transferred, etc.).

The nature of meaning

取決於你從什麼角度與興趣問這個問題

- Ontologically, what meaning IS/NOT ?
- Technically, operationalize the concept of Natural Language **Understanding** via its Evaluation Benchmark).

我們要如何理解「語意理解」？

Form-Meaning Pairs

先要了解意義是怎麼「產生」的？

- 符號層次：signs (**Pierce**: icon, idenx, symbols) and signal
- 文字與語言層次：script (*alphabet, character, kana*) | morpheme, lemma, word, phrases, sentences, text, discourse, .. (emoticon?)

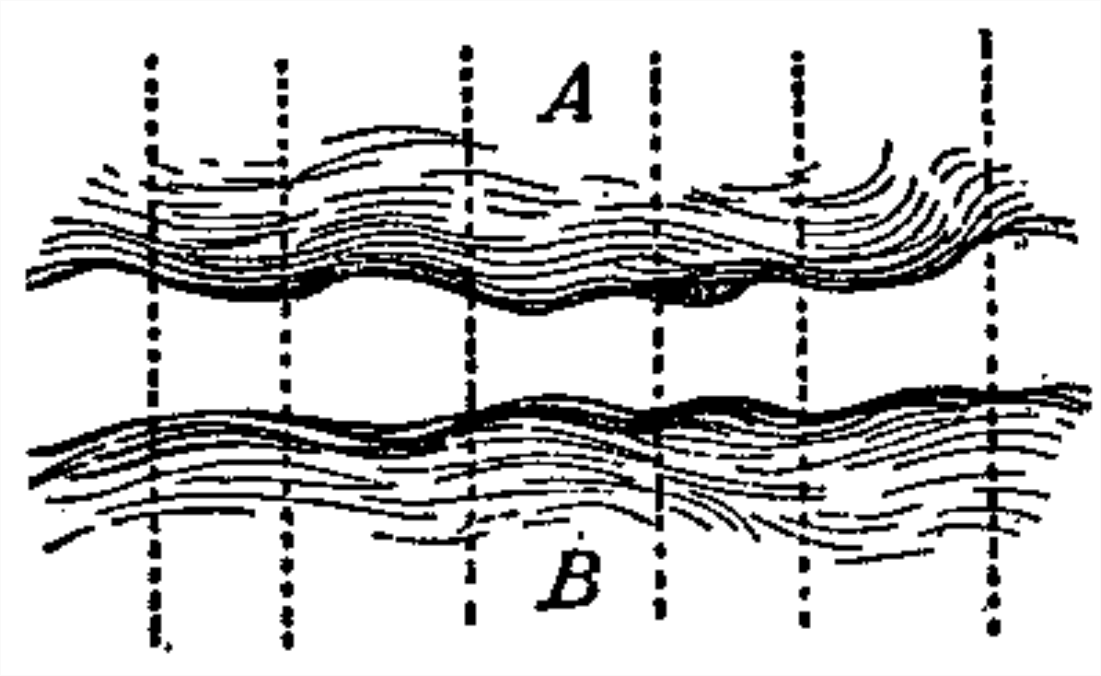
Not all languages have a word for 'word'

Not all languages have a word corresponding to English 'word': Warlpiri, again, makes no distinction between 'word', 'utterance', 'language' and 'story', all of which are translated by the noun *yimi*. In Cup'ik (Yup'ik, Central Alaska), the word for 'word' also means 'sayings, message' and 'Bible' (Woodbury 2002: 81). Dhegihan (Siouan, North America) has a single word, *íe*, referring to words, sentences and messages (Rankin *et al.* 2002).

source: Riemer

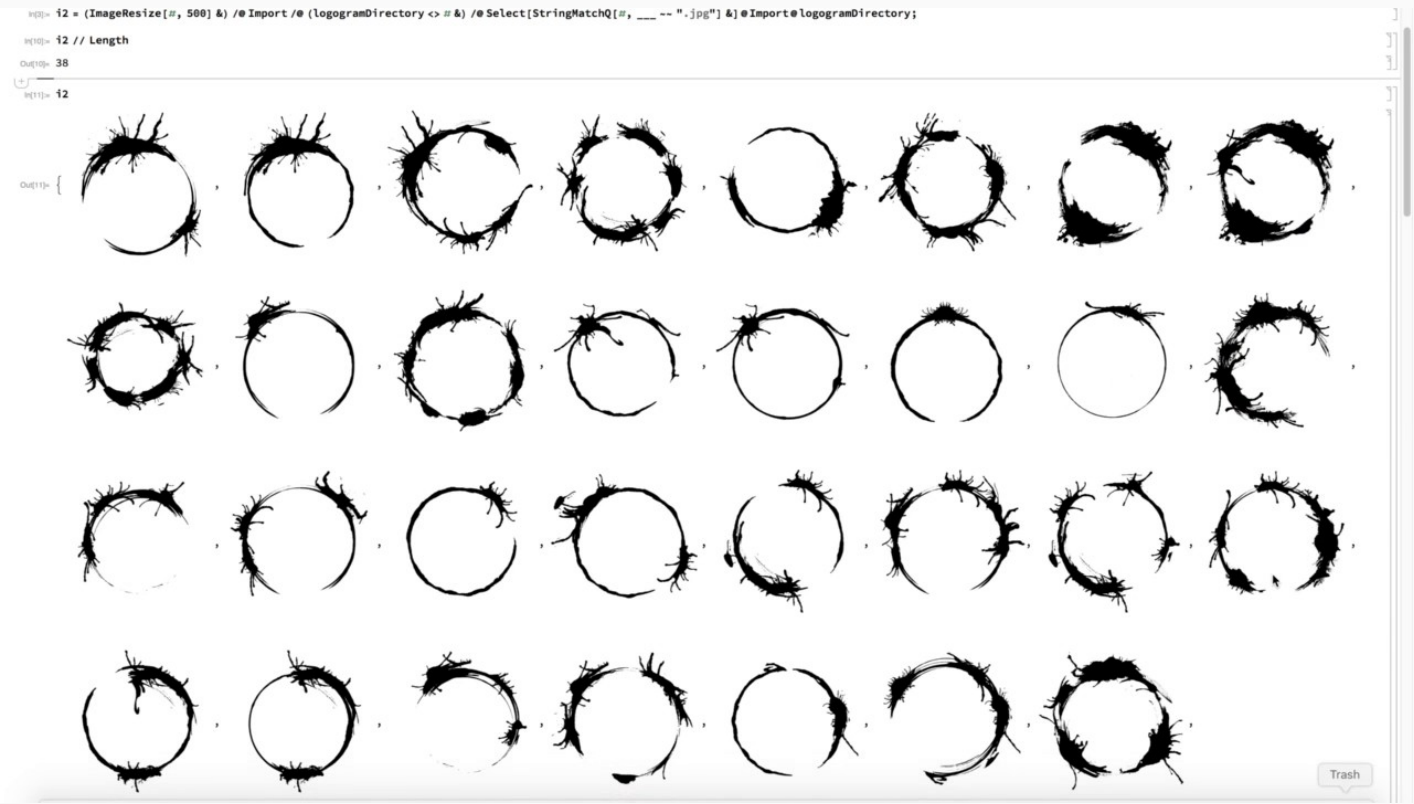
Saussure

可以想像一下形式 (A) 與意義 (B) 都是流動的配對



Arrival: heptapod logograms

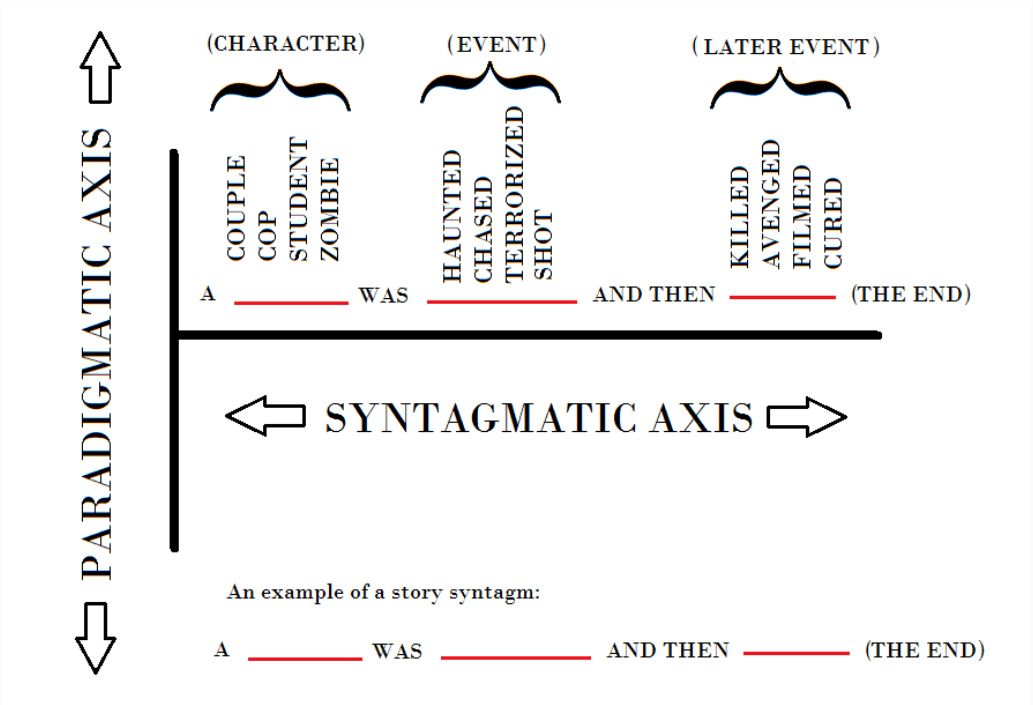
離題一下：有時候形式可以是圖形的對比



Semantic phenomena

Meaning relations

- conceptually,Paradigmatic and Syntagmatic relations
- logically, entailment, contradiction, paraphrase



Polysemy and Prototype Effect

多義、歧義、原初意義

Connotations

隱藏含義

a subtle aspect of meaning

> CONNOTATIONS *are shifting and idiosynchractic associations which a word may have for some speakers but not for others. e.g., 'feminist'.*

Anaphora / co-reference resolution Task

The task of resolving what a pronoun, or a noun phrase refers to.

- John took two trips around Taiwan, *they* were both wonderful.
- The cat caught the mouse because *it* was slow/quick.
- The physician hired the secretary because *he/she* was overwhelmed with clients.

This leads us to the involvement of Linguistic knowledge, Common-sense knowledge (and bias)

Emotion also plays a role

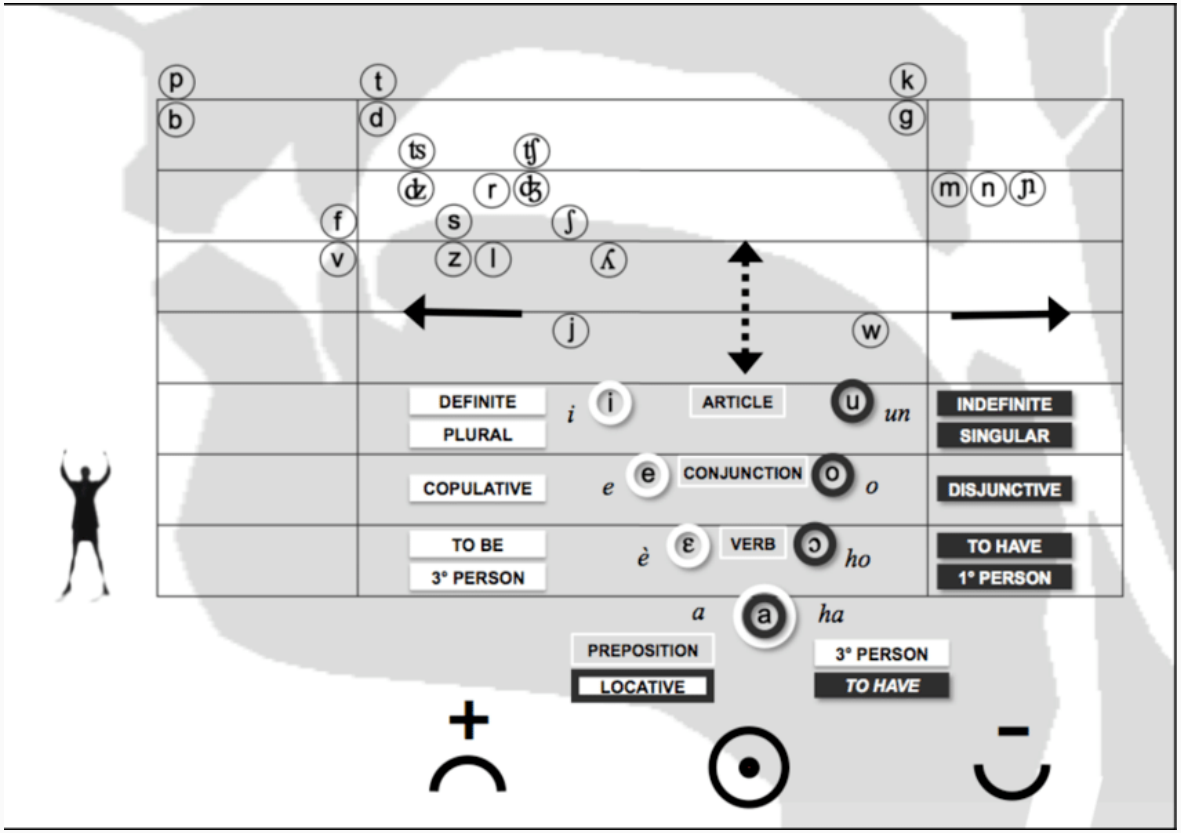


Figure 1: Iconic and imagic diagram of the vowel monophonemes.[2]

Linguistic approaches

語言如何複雜

單位、多義、組合、關係、脈絡、變遷、個體性 | How to cope robustly with the richness, variety of meaning in language with its ambiguities and underspecification.

Emergent Subfields

- Computational semantics

Emergent Subfields

- Computational semantics
- Semantics of social media

Emergent Subfields

- Computational semantics
- Semantics of social media
- Historical semantics

Emergent Subfields

- Computational semantics
- Semantics of social media
- Historical semantics
- Neuro semantics

Emergent Subfields

- Computational semantics
- Semantics of social media
- Historical semantics
- Neuro semantics
- Multimodal / cross-cultural / multi-brain semantics

Linguistic Approaches to Meaning

兩大傳統：

- 形式邏輯 **Truth-conditional theories**
- **認知概念 Conceptualist theories**

計算語意學作為語言學的分支

- 傳統的語意學下的計算語意學都以邏輯運算為核心（可以參考 Patrick Blackburn and Johan Bos 的書與**課程**

語意分析：求解還是求了解

- 求功能最適化 *engineering* question and/or 求真相大白 *scientific* question (兩者不衝突，但人生精力有限 XD)

思考練習
世界語言類型學 (linguistic typology) 和 **英漢機器翻譯**之間有關係嗎？

(Computational) Representation

Semantics and Computational Semantics

Semantics in computational terms | 計算方法驗證理論模型與產出應用

Key challenges

- 單位意義 formalization of lexical meaning and related world knowledge
- 組合意義 compositional analysis of sentence meaning
- 語境意義 modeling context and conversation to connect semantics and pragmatics

([3])

- 邏輯與形式語意 | logical metalanguage of formal semantics

Symbolic logic

符號邏輯與實在論

符號邏輯語意，與形上的實在 reality 有個鴻溝。

A random walk through NLP

What does the achievement reveal?

Coenen et al: Visualizing and Measuring the Geometry of BERT, arXiv https://arxiv.org/abs/1906.02715

Representation issues in semantic theory

這裡可以看出核心的技術關鍵在於**語意表徵** | The key technical idea bridging computer science and semantics is that of representation.

- *Any* semantic theory needs some way to state the meaning of an expression, called **semantic representation**.[1], and the metalanguage of semantic representation is called SEMANTIC METALANGUAGE (e.g., higher-order intensional logic (Montague 1974, Dowty, Wall & Peters 1981), set theory (Barwise & Cooper 1981), discourse representation theory (Kamp & Reyle 1993), dynamic semantics (Groenendijk & Stokhof 1990, Muskens 1996))

表徵是一種系統性的符碼結構

A representation is a symbolic structure in a system that can be understood to encode specified information—to carry meaning—because of mechanical processes that produce the representation as a function of appropriate states of the world and use the representation to guide the system's behavior in appropriate ways.

Issues

Semantic primitives/primes

- *ars combinatoria* (UNIVERSAL CHARACTERISTIC)
- M. Bierwisch, J.Katz, R. Jackendoff, A. Wierzbicka

Issues

Universality and Anglocentrism

Ethnocentrism [1]

The term refers to the distortions that can arise when the concepts, values, or practices of people of one culture are described through the prism of concepts from an alien culture (the culture of the investigators)

2

[2] will be discussed in UNIT 7

Issues

Discreteness

- vagueness and subjectivity
- FAMILY RESEMBLANCE (Wittgenstein)

Issues

Linguistic, Encyclopedic/Common sense knowledge

- impossible to draw line between linguistic knowledge and other knowledge?
-

Issues

Semantics vs Pragmatics

- 語詞意義 (lexical meanings)、文法知識 (grammatical knowledge) 與世界知識 (常識) (world knowledge) 是本質不可分的（但權宜可分）。
- 語意是在不同的語境溝通中突現 (emerged)、學習與交換的。
- 完整的語意表徵需要包含語境、溝通者的目的與意圖 (goals and intentions of interlocutors in communicative exchanges)

語用無塵室練習：

Analysis and Applications

Semantic applications

計算語意學不等同於機器的自然語言理解 (NLU)，也關心人類的

- 語言 (語意) 與情緒、主張、態度 | 輿情、社會對話
- 語言習得與學習、治療復健
- 跨域、跨語言、多模語意應用

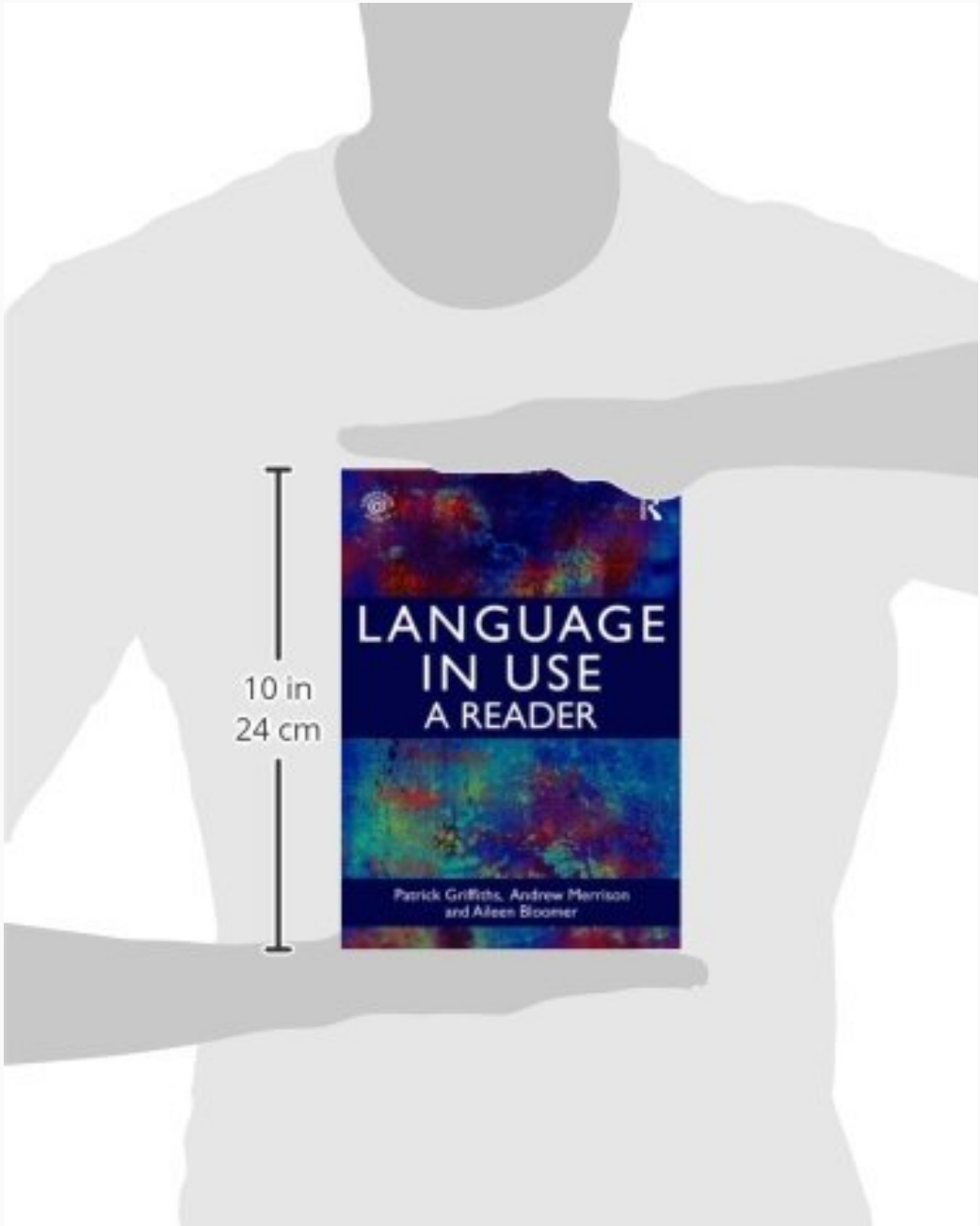
離題分享

mindsets that promote long-term (computational) linguistic learning

Academic tenacity

- Search mindset to develop knowledge
- Analytics-aware to connect society

慢慢體會這個概念：Language in use



Plan for the semester

- Vector Semantics [2]
- Graph Semantics
- Multimodal Semantics and Other applications

[2] Required math will be gently touched when encountered

References i

C. Goddard.

***Semantic analysis: A practical introduction*.**

Oxford University Press, 2011.

L. Nobile.

Words in the mirror.

Semblance and Signification, 10:101, 2011.

M. Stone.

Semantics and computational semantics.