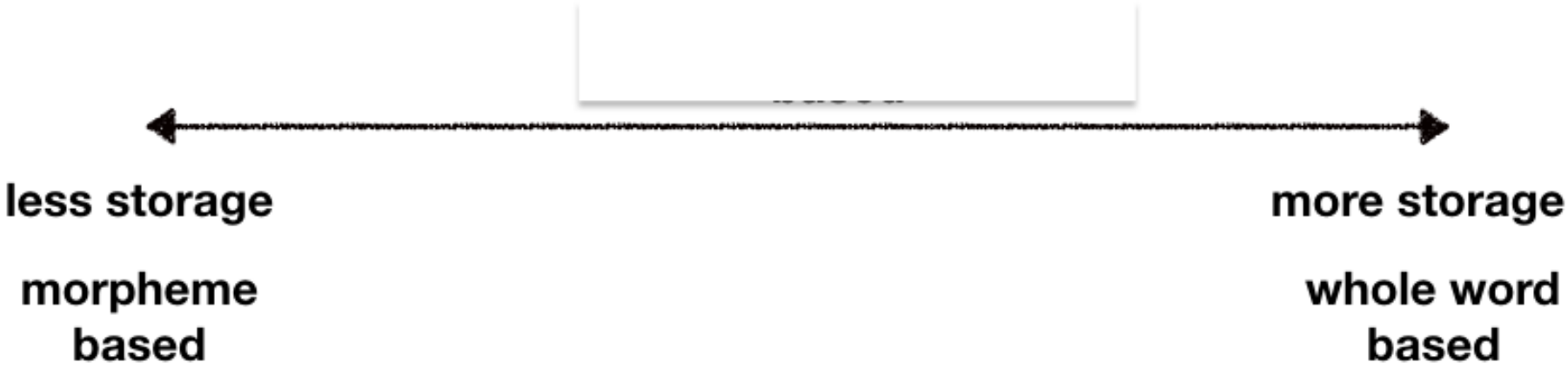
Dual route/Situations

HUL 243

23rd Sep 2024

Two approaches differ based on storage



Whole-word lexicon

- In the whole-word lexicon approach, there is an idea that storage is large, unlimited
- Morphemes are not the basic units that are stored, but larger chunks (words themselves or even larger chunks) are also possible
- The whole word approach has far less restrictiveness and much more redundancy
- Emphasis on system-external explanations, such as language use

Past tense aquisition in children

- Hoeffner et al, (1993) re-examined the child language acquisition data on past tense inflections
- Children also produced unmarked verb forms such as *I play-∅ with the baby rabbits yesterday*
- The past tense was not produced in an obligatory context
- Therefore, the over-regularization was just a part of a more general change
- Children went from unmarked *play* to *played* or *holded*
- Eventually, they switched to *hold-holded* but this was based on a *gradual change* (not a sudden switch)

Emphasis on experience

- System-external motivation for learning process
- Based on an idea of maximal storage, storage is cheap!

Joan Bybee's 'The Emergent Lexicon'

"I am going to argue that if such a lexicon does exist it is because it is **emergent** from the storage of linguistic experience, which is of a very different nature than the traditional conception of a lexicon would suggest. "

Storage/Memory

- We store a very large number of details: visual, auditory, olfactory etc.
- In a similar way, we are sensitive to linguistic experience from an early age
- Fairly powerful pattern recognition capability, even with very little exposure

Saffran, Aslin and Newport (1996)

Infants exposed to only two minutes of an artificial language were able to infer word boundaries in test scenarios, i.e. able to distinguish 'words' they had heard- and 'non-words' they hadn't during training.

Storage/Memory

- Not only can we extract patterns, but we can also recall them (after long periods of time)
- E.g. recalling voices, words, sentences, phrases even after brief exposure

Standing (1973)

Demonstrated the power of visual memory:- participants saw 10,000 pictures (people on vacation) over a course of a week, and were able to recall quite reliably whether they had seen a given picture before or not.

- When we have unlimited storage, we don't need to minimize the size of the lexicon
- We store the monolithic surface forms (without attending to their individual structure) e.g *yellow brick road* or *ferris wheel*

Patterns

- Does this mean there is no room for generalization/productivity/creativity?
- **Analogy**: form-meaning correspondences are used to produce some generalization
- If certain patterns occur often enough during linguistic experience, we can predict how other forms might behave
- E.g `wugs` $\rightarrow$ *bugs*, *mugs*, *rugs*, *hugs*

Over-regularization

- How to account for phenomena like over-regularization ?
- Analogy and experience together can explain the child language data of *holded* and *heared*
  1. Initially, no past tense forms are produced, even in obligatory contexts
  2. Then, via analogistic pressure, there is a form-meaning correspondence between regular past *-ed* and irregular verbs *hold*, *hear*
  3. Gradual increase in input over time competes/does not agree with the analogical forms
  4. Finally, there is recovery from over-generalization
- More than one cue for learning is possible, importance of experience

Strict word-form lexicon

- If a notion of a word is the largest possible 'chunk' of meaning, then it works with languages like English, German which have fewer affixes
- However – it still does not explain how speakers learn agglutinative language structures, where each word consists of many morphemes

(1) oku-ya-ma-yabil-ir-im
read-POT-NEG-POT-AOR-1SG
'I might not be able to read'

- There also exist issues with assuming that words have **no** internal structure (there seems evidence that they do)
- Further a 'whole-word' lexicon would be a lot less elegant and modular than a morpheme based lexicon

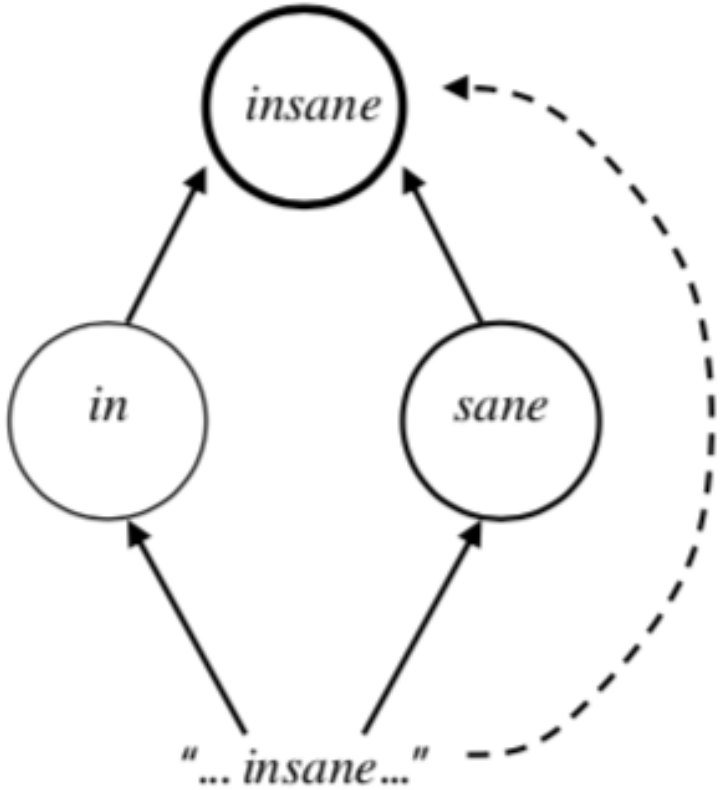


Figure: Schema of the dual-route model, Hay (2001)

Moderate word-form lexicon

- *Insane* is more common than the root *sane*, more frequently used
- Hence *insane* has better memory strength and is retrieved via direct route
- On the other hand, if *insanely* needs to be retrieved, it is also less frequent than *insane*– does it get decomposed?
- Entirely possible that in this model, some partial decomposition takes place e.g. *insane+ ly* rather than *in+sane+ly*
- The 'word form' that is memorized does not need to correspond to a minimal, analytical form

Situations

Jack and Jill went up the hill,
To fetch a pail of water.
Jack fell down and broke his crown,
And Jill came tumbling after

- When we hear this, we get the intuition that there are some discrete events
- What are these events ?
- Which linguistic structures express them ?
- In the end, these are connected into a coherent whole
- How are different kinds of events encoded in language?

Situation types

- Verbs can be described in terms of **situation** types
- Situation type describes the events encoded in the semantics of a language
- This is most commonly encoded in *verbs*

Situation types: states

- For example, some verbs describe **states**:-
  - Robert loves pizza
  - Mary knows the way to the LHC
- In such examples, we don't know much about the internal structure of the state: it just holds for some point
- The beginning or end of this situation is not known (e.g. Even if I have *Mary loved pizza* – we don't know when she stopped doing so
- Verbs like *be, have, remain, know, love* are stative verbs