Background
○○
○○○○○○○○○○

Questions and goals
○
○○○
○○

Concept posession
○○○○○
○○○○○○○
○○○○○○

Morphology
○○○○
○○○○

Linking
○○○○○
○○○

Conclusion

How, what for and since when

does

word meaning

influence

syntactic composition?

Matthijs Westera (superv. M. Marelj)

March 17, 2011

Outline

Background

Questions and goals

Concept posession

Morphology

Linking

Conclusion

Outline

How words are choosy

(1) Semantics:
   a. John/#his binoculars saw the mountain
   b. Mary drank soy milk/#a veggie burger.

(2) Linking:
   a. Bill threw the ball.
   b. *The ball threw Bill. (with Bill as Agent)
   (Pesetsky 1995 p.3)

(3) Verb-frame alternations:
   a. John defrosted the ice. (causative)
   b. The ice defrosted. (inchoative)
   c. John saw the mountain.
   d. *The mountain saw.

Methodological aim

Align linking and verb frame alternations to (aspects of) lexical semantics.

- But why should this be so?

  ... we will have the best theory of the lexicon we can hope for: children learn pairings of sound and meaning, UG [Universal Grammar] does the rest. (Pesetsky 1995 p.4)

The Theta System

The Lexicon-Syntax interface has three main components (Reinhart 2002):

- Verbs assign *theta roles*, composed of causal (c) and mental (m) involvement.
- A *linking procedure*, that ensures e.g. that +c and +m arguments become Subjects.
- An *active lexicon* with operations, that modify a verb's set of roles.

Theta roles (some examples)

(4)   a.   *break*([+c],[-c-m])

       b.   *walk*([+c+m])

       c.   *see*([+m],[-c-m])

       d.   *befriend*([+c+m],[-c+m])

       e.   *buzz*([-c-m])

Linking procedure

(5)   Marking: given a verb-entry with several arguments:

    a. Mark minus roles with '2'

    b. Mark plus roles with '1'

(6)   Merging:

    a. The '2' argument becomes Object

    b. The '1' argument becomes Subject

    c. Otherwise, the argument becomes Subject

Marking (some examples)

(7)   a. $break([+c]_1, [-c-m]_2)$

       b. $walk([+c+m])$

       c. $see([+m]_1, [-c-m]_2)$

       d. $convince([+c+m]_1, [-c+m])$

       e. $buzz([-c-m])$

Background   Questions and goals   Concept posession   Morphology   Linking   Conclusion
oo   o   ooooo   oooo   ooooo
oooo●oooooo ooo   ooooooo   oooo   ooo
             oo   oooooo

Transitive derivation

(8) a. Max saw him
     b. Base entry: *see*([+m],[-c-m])
     c. Marking: *see*([+m]$_1$,[-c-m]$_2$)
     d. Merging: *Max* ([+m]$_1$) becomes Subject, *him* ([-c-m]$_2$) becomes Object

Unergative derivation

(9)   a. The bell buzzed
       b. Base entry: *buzz*([-c-m])
       c. Marking: *buzz*([-c-m])
       d. Merging: *the bell* ([-c-m]) becomes Subject
       (Marelj 2004, p. 26)

The Lexicon Uniformity Hypothesis

How come there's a *break*([+c],[-c-m]) and a *break*([-c-m])?

  (10)   **Lexicon Uniformity Hypothesis**
        Each verb-concept corresponds to one lexical entry with one
        thematic structure. The various thematic forms of a given
        verb are derived by lexicon operations from one thematic
        structure. (Reinhart 2000, p. 20)

Active lexicon

(11) **Passivisation** existentially closes an argument:

    a. John$_{[+c]}$ hit Mary$_{[-c-m]}$ /
       Mary$_{[-c-m]}$ was hit (with a rock)

(12) **Decausativisation** removes a [+c] role (causative/inchoative alternation):

    a. John$_{[+c]}$ broke the window$_{[-c-m]}$ /
       The window$_{[-c-m]}$ broke _ (*with a rock)

(13) **Reflexivisation** and **reciprocalisation** bundle two roles:

    a. John$_{[+c+m]}$ washed himself$_{[-c-m]}$ / John$_{[+c+m]\&[-c-m]}$ washed

    b. John$_{[+c+m]}$ fought Bill$_{[+c]}$ / John and Bill$_{[+c+m]\&[+c]}$ fought.

Why decausativisation?

Some inchoative verbs behave differently from unergative verbs:

(14)   The broken vase.

(15)   *The seen mountain.

But there is no meaning component that distinguishes such inchoatives from the class of unergatives.

Unaccusative derivation

(16) a. The vase$_i$ broke ___$_i$

b. Base entry: *break*([+c],[-c-m])

c. Marking: *break*[+c]$_1$,[-c-m]$_2$)

d. Decausativisation: *break*([-c-m]$_2$)
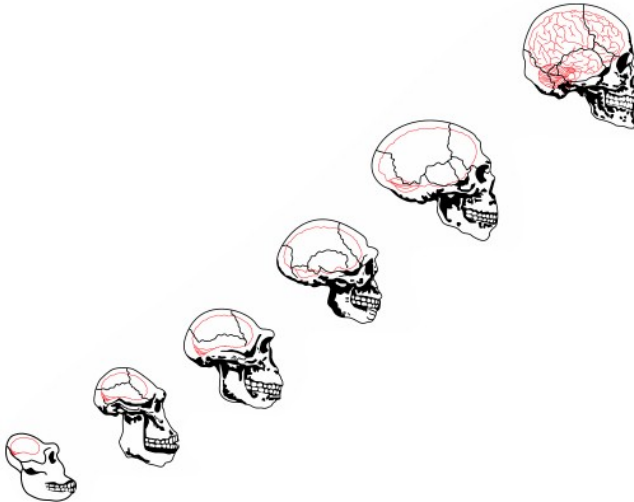
e. Merging: *the vase* ([-c-m]$_2$) becomes Object

(Marelj 2004, p. 26)

Outline

Is the Theta System *real*?

What's in the lexicon

The Theta System makes two claims regarding the lexicon:

- Only one thematic form of each verb is stored in the lexicon (LUH).

- No linking instructions of a lexical entry are stored in the lexicon.
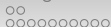
Wait... what?

- Why would our mind have evolved to *ignore* certain linguistic data?

- We need separate storage anyway if we want to explain the acquisition of exceptions/idioms.

- The same operations/linking procedure would have to be invoked over and over again.

- What about connectionist representations (storage = program)?

- Problematic both if concepts are innate and if concepts are acquired.

Lexicon vs. vocabulary

The Theta System postulates separate storage of form and meaning:

- **Lexicon**: The storage of (innate, universal) concepts.
- **Vocabulary**: The storage of (acquired) morpho-phonological forms.

Problems:

- One could not tell an *active lexicon* from an *active vocabulary*.
- What concepts do the vocabulary labels of inchoatives link to (e.g. *die*)?

Background   Questions and goals   Concept posession   Morphology   Linking   Conclusion
oo   o   ooooo   oooo   ooooo
ooooooooooo   ooo   oooooooo   oooo   ooo
                    ●o   oooooo

Goal

I will increase the Theta System's explanatory power by revisiting some of its core concepts. To start with, I assume that:

- All thematic forms of a verb are stored in the lexicon;
- Lexical entries are stored with their linking instructions;

Background   Questions and goals   Concept posession   Morphology   Linking   Conclusion
oo   o   ooooo   oooo   ooooo
ooooooooooo   ooo   ooooooo   oooo
                            o●   oooooo

Approach

Main idea: verb frame alternations can be studied independently in three domains:

- **Concepts**: Why do we have some verb concepts and not others? E.g. why do all causatives have an inchoative alternate?

- **Morphology**: Why are different thematic forms of a verb concept often morphologically similar? E.g. why are causative and inchoative BREAK both 'break'?

- **Linking**: Why do different thematic forms of a verb behave in syntax the way they do? E.g. why do some inchoatives derive unaccusatively?

Outline

Background
○○
○○○○○○○○○○○

Questions and goals
○
○○○
○○

Concept posession
●○○○○
○○○○○○○
○○○○○○

Morphology
○○○○
○○○○

Linking
○○○○○
○○○

Conclusion

Concept nativism

Concept nativism   ▸ Skip

Word learning is a tremendously fast and accurate process.

> *This can only mean that the concepts are already available, with all or much of their intricacy and structure predetermined, and the child's task is to assign labels to concepts, as might be done with very simple evidence. (Chomsky 1987, p. 29)*

Wait... what?

- How can 'doorknob' or 'computer' be innate?
  - Well, given the architecture of our brains, some concepts just 'happen to come naturally' to us...
- Chomsky's argument only holds if concepts indeed are indeed intricate/structured (cf. Fodor).
- Given the enormous number of concepts, is assigning labels to concepts really easier?

Concept nativism and the Theta System

- Unergative verbs are stored in the lexicon, while inchoative verbs are not (*decausativisation*).

- Hence somehow, unergative verbs come naturally to us while inchoative verbs do not.

- The reasons for this must be semantic.

- *However*, there is no categorical semantic difference between unergative and inchoative verbs, or between their causative alternates.

- *Therefore*, the Theta System is incompatible with concept nativism.

Concept *non-nativism* and the Theta System

- Unergative verbs are stored in the lexicon, while inchoative verbs are not (*decausativisation*).

- Hence somehow, unergative verbs are acquired while inchoative verbs are not.

- The reasons for this must be semantic.

- *However*, there is no categorical semantic difference between unergative and inchoative verbs, or between their causative alternates.

- *Therefore*, the Theta System is incompatible with concept acquisition, too.

Two solutions

- Either concepts are innate; then
  - *All* thematic forms are innate;
  - The *appearance* of verb frame alternations is due to the (non)-availability of certain vocabulary labels in the language.
- Or concepts are acquired; then
  - *All* thematic forms are acquired;
  - Verb frame alternations are due to the (non)-availability of certain concepts in the language.

I explore both in my thesis, with equivalent results.

Concept availability and concept usefulness

- Why do we have some verb concepts and not others?
- Simple answer: some concepts are useful to have, while others are not.

Usefulness of reflexive and reciprocal verbs

Concept usefulness immediately explains why:

- Typical reflexives are 'grooming' verbs such as *wash*, *shave*, *equip* and *arm* (Reinhart and Siloni 2005).
- Typical reciprocals are verbs of competition (e.g. *fight*), joint action (e.g. *communicate*), connecting (e.g. *combine*) and dividing (e.g. *separate*) (Haspelmath 2007).

But how about *productive* verb frame alternations like causative-inchoative?

Patterns in concept usefulness

Core idea: productive verb frame alternations reflect *patterns* in the usefulness of concepts.

- Assumed pattern: Inchoative verbs are always more useful than their causative alternates.

- *Therefore*, if a language has a causative verb, it also has a corresponding inchoative alternate, but not vice versa.

Background
OO
OOOOOOOOOO

Questions and goals
O
OOO
OO

Concept posession
OOOOO
OOO●OOO
OOOOOO

Morphology
OOOO
OOOO

Linking
OOOOO
OOO

Conclusion

Criteria for concept usefulness

Eleanor Rosch (1978):

- *Cue validity*: how predictive a category's attributes are of the concept (e.g. WINGS for BIRD);

- *Inclusiveness*: how widely applicable a category is (e.g. ARMCHAIR vs. CHAIR vs. FURNITURE.

Background
○○
○○○○○○○○○○

Questions and goals
○
○○○
○○

Concept posession
○○○○○
○○○○○●○○
○○○○○○

Morphology
○○○○○
○○○○

Linking
○○○○○
○○○

Conclusion

The causative-inchoative alternation

- If you remove a [+c] role from a verb concept, then:
  - Cue validity increases because the [+c] role is a lousy cue.
  - Inclusiveness remains the same.
- Hence, every inchoative verb is more useful than its causative alternate.
- Hence, if a language has a causative verb, it has its inchoative alternate (not vice versa).
- Learners may discover this generalisation and apply it productively.

(Research idea: How learners discover such generalisations is a different question. Pinker's (1989) account is unsatisfying/dated.)

Frozen forms

- In some languages 'derived' concepts exist without their causative alternates (e.g. *vanish* in English). More specifically:

  - Inchoatives/reflexives/reciprocals may exist without their causative alternate.
  - Passives never exist without their causative alternate.

- Horvath and Siloni (2005): This is due to an 'inert' [+c] role that must be removed before entering syntax.

Wait... what?

Concept usefulness and frozen forms

Concept usefulness to the rescue:

- There is no 'derivation' going on, so underived 'derived' forms are in fact expected.
- Inchoatives/reflexives/reciprocals may be more useful than their causative alternates.
- Passives are never more useful than their causative alternates.

Background
○○
○○○○○○○○○○

Questions and goals
○
○○○
○○

Concept posession
○○○○○
○○○○○○○
●○○○○○

Morphology
○○○○
○○○○

Linking
○○○○○
○○○

Conclusion

The Lexicon-Syntax Parameter

(17) Operations may apply in the lexicon or in the syntax.
     (Reinhart and Siloni 2003)

   a. Lexicon setting: Hebrew, Dutch, English, Russian,
      Hungarian

   b. Syntax setting: Romance, German, Serbo-Croatian,
      Czech, Greek

(18) Many diagnostics, e.g.:

   a. She dresses slowly because she is an elegant dresser

   b. Jean est un excellent habilleur/maquilleur
      Jean is an excellent dresser/make-up-er (of others)

Concept usefulness and the Lexicon-Syntax Parameter

- Reinhart and Siloni (2003): There's an innate parameter.
- My idea: concept usefulness depends also on the availability of an alternative (e.g. syntactic) way of expression.
  - If a language has an item for absorbing a thematic role, then (most) operations can happen in the syntax.
  - Romance (syntax) languages have a SE/SI clitic that does just that (Reinhart and Siloni, 2005)

(Research idea: Compare sentence processing in lexicon and syntax languages to isolate the computational cost of thematic role absorption.)

Examples: Italian clitic

(19) a. Giovanni e Maria si sono abbracciati.
         Giovanni and Maria SI are hugged
         'Giovanni and Maria hugged each other'

   b. Giovanni si preoccupa di questo.
         Giovanni SI worries of this
         'Giobanni worries about this'

   c. Questi vestiti si lavano facilmente.
         These suits SI wash easily
         'These suits wash easily'

   d. Si mangiano le mele.
         SI eat the apples
         'The apples are (being) eaten'

   e. . . .

(Reinhart and Siloni 2005, pp.2-3)

Background
oo
oooooooooo

Questions and goals
o
ooo
oo

Concept posession
ooooo
ooooooo
ooo●oo

Morphology
oooo
oooo

Linking
ooooo
ooo

Conclusion

The instability of the Lexicon-Syntax Parameter

- Dutch (lexicon) and German (syntax) are closely related.
- Hence, the Lexicon-Syntax Parameter is expected to be unstable.
- The Theta System has nothing to say about this.

Grammaticalisation

My explanation: clitics appear/disappear due to grammaticalisation.

- Grammaticalisation: The gradual demotion of content words, to syntactic markers, to morphology, to zero.
  - English *going to* (movement + goal) → *gonna* (future tense).
  - words for BODY → reflexive markers → reciprocal/middle markers. (Heine and Kuteva, 2002)

(Research idea: Cross-linguistic study. A possible in-between language may be Greek. See Dimitriadis on Greek reciprocals.)

Background   Questions and goals   Concept posession   Morphology   Linking   Conclusion
oo   o   ooooo   oooo   ooooo   ooo
oooooooooo   ooo   ooooooo   oooo
                    oo   oooooo●

Interim summary

- We have a concept (or: a concept has a label) if it is useful.
- Verb frame alternations reflect patterns in concept usefulness (cue validity, inclusiveness).
- Concept usefulness is influenced also by e.g. the availability of clitics (Lexicon-Syntax Parameter).
- Grammaticalisation may switch the Lexicon-Syntax Parameter.

Background   Questions and goals   Concept posession   **Morphology**   Linking   Conclusion
oo   o   ooooo   oooo   ooooo   ooo
ooooooooooo   ooo   ooooooo   oooo   ooooo
                    oo   ooooooo

Outline

The (ir)relevance of morphology

- The Theta System: morphology/vocabulary is acquired and hence irrelevant for the study of the language faculty.
  - Language is primarily for thought, not communication. Thought operates on the concepts directly. (Chomsky)

Background   Questions and goals   Concept posession   **Morphology**   Linking   Conclusion
oo   o   ooooo   ●ooo   ooooo   o
ooooooooooo   ooo   ooooooo   oooo   ooo
                  oo   oooooo

The (ir)relevance of morphology

- The Theta System: morphology/vocabulary is acquired and hence irrelevant for the study of the language faculty.
  - Language is primarily for thought, not communication. Thought operates on the concepts directly. (Chomsky)
- But I propose to consider it, because:
  - It mediates our knowledge of the lexicon.
  - If irrelevant, then why make explicit that "derivational morphology does not tell us anything about the direction of derivation"?
  - My explanation of morphological regularities will also account for linking regularities.
  - What if Chomsky is wrong?

Regularisation of the form-meaning mapping

- Language learners have to generalise (acquisition bottleneck).
- *Iterated* learning gradually regularises a form-meaning mapping (e.g. K. Smith 2003).
- Morphological markers may occur, giving the impression of 'derivational morphology'.

Morphology examples (1)

(20)   Marking on the causative:

    a. Georgian (Haspelmath 1993):
*du*γ*-s* ('cook-intr') / *a-du*γ*-ebs* ('cook-tr')

    b. Khalka Mongolian (Piñon 2001):
ongoj-x ('open-intr') / *ongoi-lg-ox* ('open-tr')

    c. Japanese (Hasegawa 2001):
kawak-u ('dry-intr') / *kawak-as-u* ('dry-tr')
ugok-u ('move-intr') / *ugok-as-u* ('move-tr')
tob-u ('fly-intr') / *tob-as-u* ('fly-tr')

Morphology examples (2)

(21)   Marking on the inchoative:
   a. Russian (Haspelmath 1993):
      katat'-sja ('roll-intr') / *katat'* ('roll-tr')
   b. Polish (Piñon 2001):
      złamać-się ('break-intr') / złamać ('break-tr')
   c. Serbo-Croat:
      otvoriti se ('open-intr') / *otvoriti* ('open-tr')
   d. French:
      se casser ('break-intr') / *casser* ('break-tr')

Background
○○
○○○○○○○○○○

Questions and goals
○
○○○
○○

Concept posession
○○○○○
○○○○○○○
○○○○○○

Morphology
○○○○
●○○○

Linking
○○○○○
○○○

Conclusion

'Derivational' morphology?

- Interesting observation:
  - Morphology is consistently on the 'derived' entry for e.g. reflexives and reciprocals.
  - Morphology is variable for the causative-inchoative alternation.

- Theta System: Morphology is uninformative of direction of derivation.

- My explanation: There is no derivation going on. But how about reflexives/reciprocals?

Background
○○
○○○○○○○○○○

Questions and goals
○
○○○
○○

Concept posession
○○○○○
○○○○○○○
○○○○○○

Morphology
○○○○
○●○○

Linking
○○○○○
○○○

Conclusion

A syntactic explanation

Decausativisation, unlike reflexivisation/reciprocalisation, can never happen *in the syntax*, despite the Lexicon-Syntax Parameter.

(22) **The Lexicon Interface Guideline**:
      The syntactic component cannot change theta grids:
      Elimination and modification of a theta role as well as
      addition of a role to the theta grid are illicit in syntax.
      (Siloni 2002, p.2)

Morphological markers through grammaticalisation

- Variable morphological markers emerge through iterated learning, in no particular order.
- For reflexivisation/reciprocalisation etc. in syntax languages, grammaticalisation yields apparently 'derivational' morphology, that persists in lexicon languages.
- ... but not for the causative-inchoative alternation.

(Research idea: Test and quantify this hypothesis by computer simulation (e.g. K. Smith 2003))

Background   Questions and goals   Concept posession   Morphology   Linking   Conclusion
oo   o   ooooo   oooo   ooooo
oooooooooooo   ooo   oooooooo   ooo●   ooo
                  oo   oooooo

Interim summary

- Verb frame alternations reflect world structure (concept
  usefulness).
- Morphology aligns with verb frame alternations due to iterated
  learning (regularisation of form-meaning mapping).
- The Lexicon Interface Guideline predicts when 'derivational'
  morphology is properly derivational (i.e. consistent).
- This all seems quite relevant.

Outline

Background   Questions and goals   Concept posession   Morphology   Linking   Conclusion
oo   o   ooooo   oooo   ●oooo   ooo
ooooooooooo ooo   ooooooo   oooo
             oo   oooooo

An acquired linking procedure

Main idea: the linking procedure, like productive operations, is the result of iterated learning. *However*:

- If we have language primarily for thought (Chomsky), some linking mechanism must be innate, so how could it be acquired, too?

- My answer: a deterministic linking procedure is needed only for communication, not for thought.

Linking procedure

(23)   Marking: given a verb-entry with several arguments:
   a.   Mark minus roles with '2'
   b.   Mark plus roles with '1'

(24)   Merging:
   a.   The '2' argument becomes Object
   b.   The '1' argument becomes Subject
   c.   Otherwise, the argument becomes Subject

Transitive derivation (repeated)

(25) a. Max saw him
      b. Base entry: $see([+m],[-c\text{-}m])$
      c. Marking: $see_{ACC}([+m]_1,[-c\text{-}m]_2)$
      d. Merging: *Max* ($[+m]_1$) becomes Subject, *him* ($[-c\text{-}m]_2$)
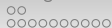   becomes Object

Unergative derivation (repeated)

(26)   a. The bell buzzed
       b. Base entry: *buzz*([-c-m])
       c. Marking: *buzz*([-c-m])
       d. Merging: *the bell* ([-c-m]) becomes Subject
       (Marelj 2004, p. 26)

Unaccusative derivation (repeated)

(27)   a.   The vase$_i$ broke ___$_i$

b.   Base entry: $break([+c],[-c-m])$

c.   Marking: $break_{ACC}([+c]_1,[-c-m]_2)$

d.   Decausativisation operation: $break([-c-m]_2)$

e.   Merging: *the vase* ($[-c-m]_2$) becomes Object

(Marelj 2004, p. 26)

Linking regularities through iterated learning

- Main idea: Just like the form-meaning mapping is regularised, so the syntax-meaning mapping is regularised.
- To understand how, we must understand the evolutionary pressures for and against regularisation.

Background   Questions and goals   Concept posession   Morphology   Linking   Conclusion
oo   o   ooooo   oooo   ooooo
ooooooooooo   ooo   oooooooo   oooo   o●o
                  oo   oooooo

Evolutionary pressures on linking

Ordered constraints:
Equality (total) > Economy > Similarity (c and m)

- **Equality**: Identical roles link identically
  - Breakee of causative *break* and Breakee of inchoative *break*.
- **Similarity**: Similar roles link identically:
  - Breaker and Thrower share their +c feature.
- **Economy**: Roles of intransitive verbs become Subject.
  - Glowee becomes Subject despite its sharing the -c feature with Breakee.

(Research idea: Compare synchronic with diachronic optimisation;
optimality theory (Sorace 2004) vs. language evolution)

Interim summary

- Verb frame alternations reflect world structure (concept usefulness).
- Both morphology and linking regularities align with verb frame alternations due to iterated learning (regularisation of form-meaning mapping).
- Linking regularities are driven by similarity and economy.

Outline

Conclusion

Consider where we started:

- A problematic view of the lexicon.
- Innate concepts, innate operations ('derivational' morphology unaccounted for), innate linking procedure, innate Lexicon-Syntax Parameter (fluidity unaccounted for), 'inert' roles.
- Not even a clue regarding their origins.

And where we ended:

- An unproblematic view of the lexicon.
- Only an innate tendency to generalise.
- Origins explained in terms of world structure (usefulness) and iterated learning.

The new view is still compatible with Chomsky's 'language for thought' or generative linguistics in general.

Methodological aim (revisited)

Why align linking and verb frame alternations to (aspects of)
lexical semantics?

> ... we will have the best theory of the lexicon we can hope
> for: children learn pairings of sound and meaning, UG
> [Universal Grammar] does the rest. (Pesetsky 1995 p.4)

It seems that we should (and safely can) replace 'Universal
Grammar' by 'language change', at least in this domain.