Compounds: Definition and Modeling

Gianina Iordăchioaia

MGK, SFB 732, University of Stuttgart
17 July, 2015
Compounding

- A word formation process next to derivation and inflection
- **Inflection**: formation of new words by means of functional affixes: boy – boys; walked
- **Derivation**: formation of new words by means of contentful affixes: unhappy; reader
- **Compounding**: the formation of new words by putting together two (or more) existing words/roots: playground
Background

• Most literature focuses on N-N compounds

• Head (vs. non-head) of an (endocentric) compound:
  • The Right-hand Head Rule (RHHR; Williams 1981): the head of a compound word (in Germanic languages) is the right-hand member of that compound: e.g., apple pie
  • Determines the lexical category of the compound: e.g., blackboard$_N$: black$_A$ + board$_N$
  • Encodes the core meaning: a blackboard is a board
  • Carries inflectional morphemes: bookshops
  • The non-head modifies the head: drive$_V$ + way$_N$: driveway$_N$
Background

- **Exocentric** (vs. Endocentric) compounds have no head: e.g., must$_v$ + have$_v$ $>$ must-have$_n$; football (game)

- Phonology: a compound behaves like one phonological word, so it has **one primary stress**

- The primary stress of a compound in English is on the **leftmost element**:
  
  - cf. black 'board (phrase) vs. 'blackboard (compound)
  
  - green 'house vs. 'greenhouse
The Problem of Definition

- Bauer (2003: 40): 'the formation of a new lexeme by adjoining two or more lexemes'

- Marchand (1967): Expansion vs. Derivation (no Compounding!)
  - Expansion: the head is an independent morpheme:
    - prefixation: re-heat, out-run
    - compounding: steam-boat, color-blind
  - Derivation: the head is not independent: suffixation: read-er

=> The problem of a universally applicable definition of compounds
Two Main Issues

1. The elements that make up compounds are not words but stems or roots in some languages.

2. We cannot make a clean distinction compounds vs. phrases.
The Elements in Compounds

- Marchand (1960): 'when two or more words are combined into a morphological unit, we speak of a compound' (for English!)

- But in Slovak: *rychl-o-vlak* 'express train' (compound): *rychl*ₐ 'fast' has no inflection and there is a linking element 'o'
  - *rychly vlak* 'fast train': *rychly* is inflected to agree with the noun
  - cf. German: *Kleinwagen* 'supermini/subcompact'
    vs. *kleiner Wagen* 'small car'

=> English compounds seem to be made up of words because English has too little inflection.

- Bauer's definition in terms of *lexemes* covers words/stems/roots; a lexeme stands for one lexical item (dictionary entry)
Compounds vs. Phrases

• Bauer (2003): compounds are "new lexemes" made up of two or more lexemes;
• How do we know that some expression is a new lexeme?
• Some compounds are clearly lexicalized: cf. blackboard vs. black board
• What about tomato bowl referring to a bowl with tomatoes?
• What about: a floor-of-a-birdcage taste, a wouldn't-you-like-to-know-sneer, a ate-too-much headache?
Finding Criteria for Compoundhood

- **Spelling** is not a good criterion for English: different compounds are spelt in different ways and some have all three versions: e.g., *flowerpot*, *flower-pot*, *flower pot*.
  - cf. German: spelling in one word seems more reliable
  - Spelling in one word should be the result of other criteria that identify the compound and not the other way around

- More reliable criteria:
  - **Phonology**: stress pattern
  - **Syntactic** impenetrability, inseparability and unalterability
  - **Inflection**
Phonology: Stress

- English compounds bear stress on the left-hand constituent, whereas syntactic phrases carry a level stress or are stressed on the head (right-hand constituent).

- There may be individual variation or variation depending on context: e.g. Spencer (2003) distinction between compounds: 'toy factory vs. toy 'factory.

- There are various attempts to relate the presence of stress to the structure of the compound.

- Olsen (2000): all synthetic compounds (including a deverbal noun) have left-hand stress: e.g. 'truck driving, 'truck driver.
Phonology: Stress

• Giegerich (2004):
  • Attribute-head N+N constructions are phrases and have right-hand stress: e.g. steel 'bridge
  • Complement-head N+N constructions are compounds and have left-hand stress: e.g. 'battlefield, 'hand cream

• Plag (2006) shows experimentally that both types exhibit left-hand stress in new compounds

=> difficult to find a structural explanation for the variability of stress in English compounds
Phonology: Stress

- Semantic criteria are just as hard to argue for in support of the different stress patterns;
- Olsen (2000): right-stressed vs. left-stressed collocations:
  - Non-head indicating temporal/locational relations => right
  - e.g., summer 'dress, summer 'night, hotel 'kitchen
  - But see: 'restaurant kitchen, 'winter coat, 'summer school
- Conclusion: left-hand stress is often a mark for English compounds, but it is not either a necessary or sufficient condition to distinguish them from phrases
Syntactic Criteria

• Inseparability: black (shiny) board vs. shiny blackboard
  • Exception in coordination: e.g. wind and water mills

• Impossibility to modify the non-head: (*very) blackboard
  • Exception: Serious Fraud Office; instant noodle salad

• Inability to replace the head with 'one':
  I bought a black board and a green one.
  *I bought a blackboard and a green one.
  • Exception: He wanted a riding horse, as neither of the carriage ones would suffice (Bauer 1998)
Inflection and Linking Elements

- In languages that have nominal inflection, 2 possibilities:
  1. The **head** of a compound **bears inflection**, but its non-head doesn't;
  2. Non-head bears a **compound-specific inflection**
Inflection on the Head, not on Non-Head

• E.g.: apple cakes - *apples cake; doghouses - *dogshouse
• Exceptions (Selkirk 1982): overseas investor; parks commissioner; arms-conscious; programs coordinator;
• Selkirk: pragmatic function to indicate plural (vs. sing)
  - But: "dress manufacturer" despite production of more dresses;
  - "programmes list": "programme list" wouldn't be a list if it had only one programme.

=> a plural is possible but not necessary in a compound to denote plurality of the first stem!
Compound-Specific Inflection on Non-Head

• A linking element is a meaningless extension that occurs between the first and second elements of compounds.

• e.g. German: Stelle-n-anzeige 'job advertisement'
  
  English frozen forms: hunt-s-man, state-s-man

  Slovak: rychl-o-vlak 'express train'

• Linking elements indicate compoundhood for languages that have them, but they are rare in some languages;
Conclusions: Compounds vs. Phrases

• Maybe there is no (unitary) compounding process(?)

• Compoundhood is a relative notion: there are constructions that are more or less like compounds without a clear categorical distinction.

• Three problems and possible solutions:
  • Definition: compounding is a gradient, rather than categorical phenomenon, with prototypical examples and fuzzy edges;
  • Interpretation: compounds vs. idioms; determining the interpretation of compounds; prediction of interpretation
  • Components: what analysis fits compounds? What do compounds tell us about the architecture of grammar, the split between morphology and syntax and 'wordhood'?
Possible Analyses: Lieber (2009)

- 'Root'/primary compounds (no relation to verbal nexus):

  ![Diagram](image1)

  1. Dog, bed, rock, hard, hat

- Synthetic compounds (deverbal N head; argumental non-head):

  2. Cannot account for the argumental status of the non-head 'truck'

  3. Accounts for argumental relation, but predicts non-existent/productive *to truck-drive
Theoretical Approaches to Compounds

- 'Lexicalist' vs. 'Syntactic' approaches

- Lexicalist approaches: word formation processes (and morphological processes, in general) take place in the lexicon and syntax deals only with words

- Syntactic approaches: word formation obeys the same syntactic rules that phrase-level syntax follows; e.g. in Distributed Morphology (DM) there is no lexicon
Lexicalist Approach: Giegerich (2009)

- Stratification: ordered sequence of two or more domains.
- Stratum 1: root-based; output is listed and formally/semantically irregular (fraternal – fraternize – fraternity)
  - Bases are prone to stress shifts or other phonological distortions (solemn-solemnity; serene-serenity)
  - Stratum 1 words are morphologically like simple words
- Stratum 2: word-based; rule-driven morphology; productive
  - Blocking: cf. wept (stratum 1) – *weeped (stratum 2)
  - Ordering of suffixes: *homeless-ity (-ity: 1, -less: 2)
- Bracket Erasure Convention: at the end of a stratum brackets are erased; morph. complexity visible to morph/phon operations within a stratum and not above it: cf. damning vs. damnation
Lexicalist Approach: Compounds

- *Lexical Integrity Hypothesis*: syntactic processes cannot manipulate the morphological elements of words
  
e.g. watchmaker
  *watch skilled maker
  *a watchmaker and a clock one

- Where is compounding: stratum 1 or 2?

- Kiparsky (1982): 3 strata: stratum 1 for irregular inflection; stratum 2 for compounds; stratum 3 for regular inflection:
  
e.g. lice-infested vs. *rats-infested (cf. Rat-infested)
  BUT: drinks dispenser

- All identifiable morphemes are the realizations of terminal nodes of a hierarchical morpho-syntactic structure
- Abstract feature bundles are manipulated by syntactic operations to form an appropriate syntactic representation
- This syntax then splits in two subderivations: Logical Form (LF: gives a semantically interpretable object) and Phonetic Form (PF: gives a well formed phonological representation)
- Terminal nodes:
  - Feature bundles (subject to Vocabulary insertion; competition):
    - e.g. past tense T[past] realized as hit-Ø, lef-t, play-ed
  - Roots: non-grammatical, encyclopedic meaning, no category
  - Categorizing heads: n, v, a categorize roots (catØ; marriage)
DM Syntactic Approach: Compounds

- A fundamental difference between roots & other terminal nodes
- No hard distinction between inflectional and derivational terminal nodes.

(1) a. That student with short hair and this one with long hair sit together.
   b. ?*That student of chemistry and this one of physics sit together.
   c. She studies physics, and he studies chemistry.

- The internal argument comes with the root (not with v or n)

(2)
(4) truck-driver, truck-driving (*[the-truck]-driver, *trucks-driver)
(5) driver/driving of a/the truck/trucks
   • The complement noun combines/incorporates with the root before the latter is categorized by the nominal suffix in compounds (6a); in AS-nominals it is a DP like with the verb

(6) nP (compound) nP (AS-nominal)
   \[\begin{array}{c}
   \text{n} \\
   \text{-er/-ing} \\
   \sqrt{\text{DRIVE}} \\
   \text{n} \\
   \text{Ø} \\
   \end{array}\] nP (V-ExP)
   \[\begin{array}{c}
   \text{n} \\
   \text{√TRUCK} \\
   \text{Ø} \\
   \sqrt{\text{DRIVE}} \\
   \text{DP} \\
   \end{array}\]
Acknowledgments

- These slides are based on the introduction and various other chapters from Lieber & Stekauer (2009) 'The Oxford Handbook of Compounding', as presented in the course material for HS Advanced Morphology jointly taught with Artemis Alexiadou (IfLA, SS 2015)
References


