Polysemy and Syntax

Stephen Wechsler, The University of Texas at Austin
Course time: Monday/Thursday 9:00-10:50 AM

Class handouts for Part I

1. Words and senses
   1.1. Homonymy, polysemy, and generality
   1.2. Tests for homonymy vs. polysemy vs. generality
   1.3. Facets
   1.4. A zeugma test for prepositions
   1.5. Refining the tests
2. Generative theories of polysemy
   2.1. Pragmatic roots of polysemy
   2.2. Cross-linguistic variation in sense extension rules
   2.3. Lexical licenses
   2.4. Coercion
   2.5. Constructional polysemy and sense extensions
   2.6. Type presupposition accommodation
3. Sense indeterminacy as a route to polysemy
   3.1. Inquiry-resistant polysemy
   3.2. Gavagai, repurposed for a theory of polysemy
   3.3. Putting the brakes on generalization
   3.4. Semantic change
   3.5. Semantic change in modal verbs
   3.6. The sound emission puzzle
4. Causative alternations
   4.1. Why are causative alternations so common?
   4.2. Causative ~ Inchoative ~ Resultative Alternations
   4.3. Spatiotemporal overlap and direct causation
   4.4. Which verbs alternate and why
1. Words and senses

1.1. Homonymy, polysemy, and generality

Traditionally, two types of lexical ambiguity are distinguished: **homonymy** refers to cases in which two words “accidentally” have the same phonological form (e.g., *bank* ‘financial institution’ versus *bank* ‘side of a river’), whereas **polysemy** refers to the phenomenon that one and the same word acquires different, though obviously related, meanings, often with respect to particular contexts. Consider the following examples of **homonymy** (cf. Pustejovsky (1995): p. 27):

(1) a. Mary walked along the **bank** of the river.
   b. Harbor**Bank** is the richest **bank** in the city.

(2) a. The judge asked the defendant to approach the **bar**.
   b. The defendant was in the pub at the **bar**.
   c. He bought a **bar** of soap.

And contrast this with the following cases of meaning variation, which illustrate **polysemy**:

(3) a. The **bank** raised its interest rates yesterday.
   b. The store is next to the newly constructed **bank**.
   c. The **bank** appeared first in Italy in the Renaissance.

Words are said to be **polysemous**: each word form has a range of meanings that are related, whether closely or distantly. Polysemy is ubiquitous, ‘the rule rather than the exception’ (Cruse 1986, 50).

**Homonyms** accidentally take the same phonological shape but are unrelated in meaning, such as *light* in weight versus *light* in color. (*Homophones* have the same sound (break ~ brake); *homographs* have the same spelling (bow of a ship ~ bow and arrow); and *homonyms* have the same sound and spelling (the dogs bark ~ the bark of the tree)).

**Polysemy** involves meaning variation, such as *bank* as ‘financial institution’ (3a) versus *bank* as ‘physical building housing a financial institution’ (3b) (Pustejovsky 1995). This example of polysemy differs from homonymy in two respects.

1. The financial institution and the building housing it are clearly related, while in contrast there is no apparent relation between financial institutions and riversides (but see below).
2. The polysemy relation connecting the two uses of *bank* in (3) is systematic.

A parallel polysemy can be found across virtually all English words and phrases referring to buildings that house institutions, including even proper names such as *Austin City Hall*, *The University of Texas*, *The Performing Arts Center*, and so on:
(4) a. The University of Texas raised its tuition rates.
b. The University of Texas is located several blocks north of the state capitol building.

This alternation (institution $X \sim$ building that houses $X$) is an instance of systematic (or regular) polysemy, since the relation is regular within the language.

Idiosyncratic polysemy: Individual words are often extended to new uses that bear some semantic relation to the old ones. If these extensions catch on for a single word but fail to generalize to semantically related words of the language then they remain as isolated instances of idiosyncratic polysemy.

\[ \text{cup} \ '\text{part of a bra that covers a breast'.} \]

(n.b. A hat covers a head but it is not called a cup)

With idiosyncratic polysemy one of the two criteria distinguishing polysemy from homonymy has been lost: idiosyncratic polysemy is not regular. The connection between senses may still be motivated, in that one can explain it after the fact, while not being predictable in the sense that a general rule applies. This leaves only the criterion of semantic similarity. But if the connection between uses becomes opaque over time, either due to semantic drift or to changes in the extralinguistic world, then such cases of polysemy can grade off into homonymy.

ex. bank illustrates homonymy, namely ‘riverside’ and ‘financial institution’; but they have a common historical origin in a form denoting a “shelf, natural or artificial, of earth, rock, sand, or wood.” (Oxford English Dictionary, OED). The bank of a river is such a shelf. Regarding the ‘financial institution’ sense, the OED notes that “The original meaning ‘shelf, bench’ … was extended in Italian to that of ‘tradesman’s stall, counter, money-changer’s table, … whence ‘money-shop, bank,’ a use of the word which passed, with the trade of banking, from Italy into other countries.”

This connection between the two senses of bank is unknown to many contemporary speakers of English, so for them this is a case of homonymy, not polysemy. But since polysemy can gradually evolve into homonymy, the line between the two categories is fuzzy. These systematic aspects make polysemy an important field of study of synchronic and generative linguistics.

Distinct from both polysemy and homonymy is generality, where a word is simply general in its application: a sweater can refer to either a red sweater or a green sweater; a teacher can be male or female. The word teacher refers to male teachers and female teachers but it lacks discrete senses for ‘male teacher’ and ‘female teacher’. (The term vagueness is sometimes used instead of generality, but we will reserve vagueness for the problem of fuzzy boundaries around senses.)

Summarizing: generality does not involve discrete senses; when there are discrete senses then it is polysemy if the senses are related and possibly systematic; and homonymy if they are neither related nor systematic.
1.2. Tests for homonymy vs. polysemy vs. generality

(i) identity test for discreteness (homonymy/polysemy vs. generality)

(Zwicky and Sadock 1975)

sweater is **general** between ‘blue sweater’ and ‘red sweater’:

(5) Mary is wearing a sweater. Susan is wearing one too.
    Mary bought a sweater. Susan did too.
    (can combine different colors)

bank has discrete senses ‘financial institution’ and ‘riverbank’:

(6) Mary saw a bank. Susan saw one too.
    Mary saw a bank. So did Susan.
    (can’t combine financial institution with riverbank.)

(ii) independent truth condition test for discreteness.

Truth conditions can depend on which of the two senses is intended.

(7) Mary is wearing a light coat. (No she isn’t—it’s a rather heavy coat.)
    (If Mary’s coat is light in color but heavy in weight, this is true under one sense but false
     under the other.)

(8) Mary saw a teacher. (#No she didn’t—she saw a man.)

(iii) the zeugma test for relatedness of senses: Use a single word token in two different ways in a
sentence. Combining accidental homophones leads to an intuition of zeugma; combining closely
related senses does not.

(9) I can walk to a bank with low interest rates.

Zeugma is a metalinguistic trope that intentionally exploits polysemy, often for humorous effect,
as in these examples (10a and b from Cruse 1986:13):

(10) a. Arthur and his driving licence **expired** last Thursday.
    b. He was **wearing** a scarf, a pair of boots, and a look of considerable embarrassment.
    c. I heard a Californian student in Heidelberg say, in one of his calmest moods, that
        he would rather decline two drinks than one German adjective. (from Mark Twain, ‘The
        Awful German Language’)
    d. The Mad Hatter’s riddle: "Why is a raven like a writing desk"? Answer: "Because
        Poe wrote **on** both". (From Martin Gardner, the Annotated Alice, an answer attributed to
        Sam Loyd)
The symbol \( \text{z} \) means that the sentence is judged to be ‘zeugmatic’. The traditional term for this figure of speech is *zeugma*, or more accurately, *syllepsis*. *Zeugma* originally referred more generally to cases in which a word is shared between clauses, regardless of whether it has different senses in each context, while *syllepsis* specifically refers to those cases of *zeugma* in which the word appears in construction with two clauses ‘while properly applying to or agreeing with only one of them… or applying to them in different senses (e.g. literal and metaphorical).’ (OED entry for *syllepsis*, emphasis added)

Cruse 1986:72: example of *mouth*:
(11) ?The poisoned chocolate entered the Contessa’s mouth at the same instant that the yacht entered that of the river.

Cruse 1986:72 *sense spectrum*: We can get from *John’s mouth* to *the mouth of the river* without encountering zeugmatic incompatibility, by combining ‘contiguous local senses’, i.e. contiguous on a spectrum:

(12) a. John keeps opening and shutting his mouth like a fish.
    b. This parasite attaches itself to the mouths of fishes, sea-squirts, etc.
    c. The mouth of the sea-squirt resembles that of a bottle.
    d. The mouth of the cave resembles that of a bottle.
    e. The mouth of the enormous cave was also that of the underground river.

There are *degrees of antagonism* between senses, a continuum from *homophony* (full ambiguity) to *polysemy* (related senses) to *generality* (vagueness, not distinct senses at all).

1.3. **Facets**

*book* as [TOME]: The book weighs four pounds / has a red cover / etc.
*book* as [TEXT]: John has written a book, but it has not yet appeared in print.

According to Cruse (1995:36ff), two *facets* of a single sense have these properties:

(a) They are non-antagonistic.

(13) a. Mary is reading a book.
    b. The book is difficult both to read and to carry around. (NOT zeugmatic)

(b) They are autonomous.
i. one form can be used in explicit contrast to the other, cp. *the X itself / the real X*
A word can refer to an entity possessing properties of only one of the facets. E.g. a blank book is still a book in the [TOME] (but not [TEXT]) sense. Books on tape or e-books could be books in the [TEXT] (but not [TOME]) sense.

(c) They give rise to identity constraints (?) and the possibility of independent sets of truth conditions.

(d) Not ambiguous, but may give rise to ambiguous phrases.

- a long book[TEXT] e.g. too long to read
- a long book[TOME] e.g. too long to fit on the shelf
- a new book[TEXT] e.g. written recently
- a new book[TOME] e.g. purchased recently

(e) each facet may be independently involved in lexical relations.

hyponyms of different facets of book:

book[TOME] < hardback
book[TEXT] < novel

1.4. A zeugma test for prepositions

Prepositions can be shown to have semantic content event when they are selected by verbs to mark specific arguments. The verbs search and examine take a ‘desiderative’ PP[for], representing the thing which is desired; these verbs can be coordinated (15a) (from Wechsler 1995) Similarly, praise and condemn take a for-PP representing the cause or reason for the praise or condemnation; these can be coordinated (15b). But they cannot be mixed, as shown in (15c).
(15)  a. They will search and examine George for the letter.
      b. They will either praise or condemn George for the letter.
      c. *They will either condemn or search George for the letter.

(16)  a. Somehow we must safely pile or cram the weapons into trucks.
      b. Somehow we must safely transform or convert the weapons into trucks.
      c. *Somehow we must safely pile or transform the weapons into trucks.

(17)  a. Tommy’s mother sometimes tapes or pins a note to him.
      b. Tommy’s mother sometimes mails or faxes a note to him.
      c. *Tommy’s mother sometimes mails or tapes it to him.

Desiderative and benefactive for:

(18)  a. ??Mikael longed and danced for[des?/ben?] the Princess of Muu.
      b. Mikael longed for[des] and danced for[ben] the Princess of Muu.

1.5. Refining the tests

Identity tests are best done with indefinite rather than definite expressions. For example consider a clear case of non-distinct senses, teacher as ‘male teacher’ versus ‘female teacher’. In (19a) we find that the putative senses ‘male teacher’ vs. ‘female teacher’ fails the identity test: the two teachers can be of different sexes, as expected. But in (19b) with a definite they must, of course, be of the same sex, since the teacher Mary likes is the same individual that Sue likes.

(19)  a. Mary likes a teacher; so does Sue.
      b. Mary likes the teacher; so does Sue.

Thus the diagnostic in (19a) give the right result while the one in (19b) gives the wrong result.

(20)  Mary was wearing a light coat; so was Sue.

In (19a) and (20) the existential quantifier introduced by the indefinite article is reconstructed in the second sentence. For (20) this gives these two meanings depending on the sense of light:

(21)  \[\text{[wearing a light coat]} = \lambda y \exists x [\text{light-color}^*(x), \text{coat}^*(x), \text{wearing}^*(y,x)]\]
      \[\text{[wearing a light coat]} = \lambda y \exists x [\text{light-weight}^*(x), \text{coat}^*(x), \text{wearing}^*(y,x)]\]

Indefinites allow us to abstract away from the referent and look at the predicate (here, light’) as applied across two different referents. Definites evoke a discourse referent, so whatever is said about that referent in the first sentence carries over to the second, thus clouding the issue.
Pronouns pose a problem as well.

(22) They took the door off its hinges and went through it.

Pronouns can often refer to something salient in the discourse context, even if there is no linguistic antecedent (Heim and Kratzer 1998, 240).

(23) a. I met a couple that I think you know.
    — Oh, is she Norwegian?

b. Aunt Doris called.
    — Oh, how are they?

One-anaphora and other forms of sense anaphora (*so, do so, etc.*) are better tests, because these forms picks out the concept, not the discourse referent.

Conclusion:
\[ \begin{align*}
\text{concept-anaphora:} & \text{ indefinites, ellipsis, one-anaphora} \quad \Rightarrow \text{ good identity tests} \\
\text{referent-anaphora:} & \text{ pronouns, definite NPs} \quad \Rightarrow \text{ bad identity tests}
\end{align*} \]
2. Generative theories of polysemy

meaning transfer:

\[chicken\text{ (animal)} \rightarrow chicken\text{ (meat)}\]

Views of the origins of polysemy:

- pragmatic processes (reference transfer)
- derivational (word-formation) rules
- sense indeterminacy: utterance contexts are often compatible with multiple senses

2.1. Pragmatic roots of polysemy

Systematic polysemy is thought to be rooted in the pragmatic phenomenon of reference transfer (Nunberg 1979, 1995).

\(\text{(24) Reference transfer.}\) In the right utterance context speakers can use any predicate \(P\) to refer to an entity \(x\), even if \(P\) does not apply to \(x\) directly, but rather to an entity \(y\) that is related to \(x\). The speaker uses \(x\) to mean \(y\), where \(y\) bears a systematic relation to \(x\); for example \(x\) may be a part of \(y\).

\(\text{(25a) contains the demonstrative this, where the demonstratum is a key but the referent is a car.}\)

In (25b) it is not the speaker but the speaker’s car that is located out back (examples 25-27 are from Nunberg 1995, p. 110-111):

\(\text{(25) a. Parking valet: This (displaying a key) is parked out back.}\)
\(\text{b. I am parked out back.}\)

The subject this in (25a) actually refers to the car and not the key, as shown by the following.

\(\text{(26) a. This is parked out back and may not start.}\)
\(\text{b. ???This fits only the left front door and is parked out back.}\)

But the subject I in (25b) refers to the speaker, not the car:

\(\text{(27) a. I am parked out back and have been waiting for 15 minutes.}\)
\(\text{b. *I am parked out back and may not start.}\)

Both examples in (25) involve a meaning transfer in the sense of (24): in the first case it involves the subject NP, while in the second it involves the predicate \(parked out back\):
(28) a. (25a): *this* (key) ⇒ ‘the car that this key fits’
    b. (25b): *be parked out back* ⇒ ‘be the driver of a car that is parked out back’

Note agreement: *I am/*is parked out back. When NP (instead of VP) transfers, agreement goes the other way:

(29) The hash browns at table 20 wants/*want* his/*their* check.

Here the verb and pronoun show singular, not plural, agreement, suggesting a transfer from *the hash browns at table 20* (which is plural) to ‘the person who ordered the hash browns at table 20’ (which is singular).

The property contributed to the subject by the new predicate must be obvious or ‘noteworthy’, i.e. a useful classification in the context of utterance. A painter is more likely to say *I’m in the Whitney Museum* than *??I’m in the second crate from the right*, because when a painting goes to a museum the artist acquires a noteworthy property, but not so for the crate (Nunberg 1995, 113-4). It’s not clear that it is always the context of utterance that matters so much as the context of the described situation, as shown by contrasts in a past tense narrative:

(30) a. I was out back. / I was idling. / I was leaking oil.
    b. #I was for sale. /#I was brand new at the time.

The predicates in (a) allow a shift to predication on the car driven by me: ‘I [drove a car that] was parked out back’, and so on. It is not clear why (b) do not allow such a shift, but the constraint, whatever it is, seems to apply to the past situation and not to the utterance context.

For Nunberg such meaning shift is essentially a pragmatic phenomenon; these shifts are not instances of lexical polysemy. But when a usage becomes less context-dependent and more useful, that is, where the relation between x and y is obvious in many or most contexts for a given word, the result is systematic polysemy.

(31) Examples of systematic polysemy (from Nunberg 1995)
    a. transmissions for cars: *4 speed, automatic*, etc.
    b. texts for inscriptions: a *Webster’s Third*, a *Guide Bleu*, etc.
    c. painters for works: a *Picasso*, a *Derain*, etc.
    d. containers for volumes of stuff: *She drank two glasses*, etc.
    e. writer for oeuvre: *fifty pages of Wordsworth*
    f. place for inhabitants: *Indianapolis voted for the referendum*
    g. tree for wood: *The table is make of oak.*

Even these meaning shifts fall under a pragmatic theory for Nunberg, but he also claims that they can become ‘idiomatic’ or conventionalized, which thus allow for shifts that are less and less context-dependent.
2.2. **Cross-linguistic variation in sense extension rules**

Importantly, the sense extension conventions vary from language to language:

A Russian (Apresjan 1974, 18) and English rule:

\[
\text{type of vessel} \rightarrow \text{the quantity of substance that the vessel is capable of containing}
\]

(32) a. John hauled three wheelbarrows from the shed. => there are 3 wheelbarrows
b. John hauled three wheelbarrows of bricks from the shed. \(\neq\) there are 3 wheelbarrows

Russian (but not English) rule: bodily organ → its disease (Apresjan 1974, 24).

(33) У нее почки.

by 3SG.F.GEN kidney.NOM.PL
‘She has (a disease of the) kidneys.’

(34) English rule: object → stuff derived from the object
a. There was too much apple in the cake.
b. ??We fried the chicken in safflower/olive/corn.
c. ??I enjoyed a glass of orange.

(35) French rule: names of fruits → brandies made from them

\[\text{une prune} \rightarrow \text{a prune}, \text{un poire} \rightarrow \text{a pear}\]

(Nunberg 1995, 118)

2.3. **Lexical licenses**

Nunberg and Zaenen (1992) argue that polysemy rules differ from more traditional cases of derivational or inflectional rules. Many polysemy rules are conventional but not tied to a particular lexical item or lexical type. For example, in

(36) The table is make of oak.

A rule takes us from a count noun for a tree of a certain genus (colloquially called ‘species’) to a mass noun for the wood derived from such a tree:

\[
\text{an oak} \Rightarrow \text{too much oak}
\text{a birch} \Rightarrow \text{made of birch}
\]

Nunberg and Zaenen’s (1992) *lexical licenses*:

- can be exploited to produce new lexical items
- distinct from *lexical rules*; not subject to arbitrary syntactic or morphological conditions
depend on background beliefs, not on strictly grammatical knowledge of words
‘conventions of use’ that determine not what is grammatical but what is appropriate or in accordance with ordinary linguistic practice

Universality and variation:
The principles or mechanisms permitting extended uses are based on general schemas of conceptual organization; hence the same patterns recur across many languages.
What is conventionalized, and hence varies across languages, are not the mechanisms of transfer themselves but rather the semantic restrictions on the inputs or outputs of such mechanisms.

Example:
A very general rule of ‘universal grinding’ (Pelletier and Schubert 1989) takes us from a thing to the stuff derived from the thing
Lexical licenses restrict the application of this rule in language-specific ways.

English rule of grinding: applies to names of an animal to yield its meat or hide (chicken, rabbit), or to trees to yield the wood of the tree; grinding in West Greenlandic Eskimo applies only to trees, not to animals (Nunberg and Zaenen 1992, 394, citing Jerrold Sadock p.c.).

(37) ‘beverage portioning’: mass noun ⇒ count noun

| a. I drank too much beer last night. (mass) |
| b. Would you like a beer? (count) |

Applies at the lexical level: all the mass/count properties change in lockstep. For example, English count nouns, but not mass nouns: (a) require a specifier when singular; (b) can appear with a(n) in singular; (c) have plural forms; (d) cannot appear with too much; (e) can be antecedents for one-anaphora:

(38) a. Q: What do you want?  
   A: Evidence! / *Clue!  
 b. a clue / *an evidence  
 c. clues / *evidences  
 d. *too much clue / too much evidence  
 e. John discovered the clue; Mary wants to discover one too.  
    *John discovered the evidence; Mary wants to discover one too.

beverage mass ⇒ portion count noun:
a beer; two beers; John liked the beer, so Mary wants one now.

Only for beverages: a water must refer to a drink: cp. *You’re dripping wet, and so now there’s a water on the floor.  
two waters = two drinks, not two puddles.
one-anaphora works for drinks but not for puddles: John ordered the water; Mary wants one too; but *John saw the water that dripped on the floor; Mary saw one too.
**Conclusion:** the English portioning rule applies to the words. If instead portioning were a direct consequence of the local syntactic and semantic environment, then there would have to be multiple portioning rules, each with the same semantic effect and the same restriction to beverages: a beverage portioning rule for the article *a(n)*, another for plural morphology, another for *one*-anaphora, and so on.

2.4. **Coercion**

The word *fast* seems to apply differently to different nouns:

(39) a fast typist/motorway/car/driver

Putejovsky (1995) analyses such modification as *selective binding* of an event variable within the decomposition of the noun. In addition to argument structure, a word has a *qualia structure* with 4 subtypes:

(40) Qualia structure (Putejovsky 1995)

a. constitutive: the relation between an object and its constituents, like material, weight, or parts and component elements.

b. formal: that which distinguishes the object within a larger domain, like orientation, magnitude, shape, dimensionality, color, position

c. telic: the purpose and function of the object, in particular
   -- the purpose that an agent has in performing an act,
   -- the built-in function or aim which specifies certain activities

d. agentive roles: factors involved in the origin or bringing about of an object referring to the creator of an artefact or a cause.

The function of a typist is to type, so the TELIC qualia of the word *typist* contains a typing event:

(41)

```
PHON: typist
ARGSTR: [ ARG1 x:human ]
QUALIA: [ FORMAL x
          TELIC type(e,x) ]
```

An modifier like *fast* selects the event variable e as its argument.

(42) coercion: similar to the better-known semantic phenomenon of type shifting. Like type-shifting, coercion is triggered by the linguistic context: a functor seeks an argument of a particular semantic type, but if it combines with a constituent of the wrong type, then coercion rules allow that type to be shifted or coerced to the right one.

Aspectual verbs *begin, start, finish, stop, end*, etc. select an event-denoting complement. But they can combine with certain entity-denoting DPs to yield an event reading.
Salient readings of these sentences involve the *telic* quale of the object noun. Such sentences also have readings involving the object’s *agent* quale. The event variables are specified in the lexical entry of the nouns:

\[
\begin{array}{|c|c|}
\hline
\text{PHON: novel} & \\
\text{QUALIA:} & \\
\text{CONST} & \text{narrative(x)} \\
\text{FORMAL} & \text{book(x)} \\
\text{TELIC} & \text{reading(e1,x)} \\
\text{AGENT} & \text{writing(e2,x)} \\
\hline
\end{array}
\]

On Pustejovsky’s analysis, the verb *begin* effectively coerces its complement into an event denotation, so that *a beer* in the context of a sentence like (43a) means ‘drinking a beer’; *a novel* in such a context means either ‘reading a novel’ or ‘writing a novel’.

2.5. **Constructional polysemy and sense extensions**

Copestake and Briscoe (1995) distinguish between *constructional polysemy* and *sense extensions*:

*constructional polysemy*: sense differences determined by syntactic/semantic context

\[(45)\]  a fast typist/motorway/car/driver

* e.g. a fast typist = ‘a typist who can type fast’ (default interpretation). But ‘in the context of a race between typists and accountants, for example, a fast typist might be one who can run, ski or ride a motorbike quickly’ (C&B 1995, p. 33)

C&B’s analysis (cp. Pustejovsky’s coercion account): *specialization* of an underspecified lexical item, which takes place through syntactic combination. Words like *fast* semantically compose with an appropriate meaning component of a complement or other element in their context. E.g. *fast* requires an event argument; a typist or a motorway is not an event, so it combines instead with the *typing event* found in the QUALIA structure of the word *typist*, with the *driving event* in the QUALIA structure of the word *motorway*, and so on.

*broadening*:
\[(46)\]  a. a cloud default: *water vapor*
  b. a cloud of mosquitos overridden by *mosquitos*

this appears to be a non-monotonic change (info is not just added but changed)
C&B analyze broadening using a default specification, which is overridden in (46b).
sense extensions

• genuine cases of systematic polysemy, where a class of words productively alternates between systematically related senses
• susceptible to blocking by synonymous terms
• ‘formally identical to other rules of conversion and derivational morphology’ (C&B p. 18-9)

(47) ‘grinding’: count noun => mass noun
   a. Bugs is eating a carrot. (count)
   b. There’s too much carrot in this cake. (mass)

(48) ‘portioning’: mass noun => count noun
   a. I drank too much beer last night. (mass)
   b. Would you like a beer? (count)

C&B’s analysis: The sense extensions are generated with productive lexical rules, e.g. a ‘grinding rule’ taking a count noun as input and returning a mass noun, a ‘portioning rule’ applying in the opposite direction, and so on.
As with uncontroversial word-formation rules, we find blocking by existing words:

(49) a. Sam ate lamb / chicken / rabbit / etc.
    b. Sam ate ?pig / pork.
    c. Sam ate ?cow / beef.

Negative attitude conveyed by violations (due to a taboo against killing?):

(50) There were five thousand extremely loud people on the floor eager to tear into roast cow with both hands and wash it down with bourbon whiskey. (Tom Wolfe, The Right Stuff).

C&B discuss an apparent prediction: that constructional polysemy should allow co-predication while sense extension should not allow it. In the following case of constructional polysemy—

(51) a fast and intelligent typist

—the respective adjectives pick out the typing event and the individual, respectively, yielding:

(52) [x][fast(e) & type(e,x) & intelligent(x) & typist(x)]

But copredication should be harder with cases of sense extension, since the word as it appears in a sentence has either one sense or the other, depending on whether the lexical rule has applied or not. Words for regular publications like newspaper and magazine can refer to the published object (a copy of the newspaper) or to the publishing institution. Can predicates appropriate to these two senses copredicate of a single token of the word? The facts are unclear; (53a) seems worse than (53b):
(53)  a. *That newspaper is owned by a trust and is covered with coffee.  
b. The newspaper has been attacked by the opposition and publicly burned by  
demonstrators.

C&B argue for a lexical component to sense extension: that it is not purely in the pragmatics:

1. Lack of context yields a default interpretation.  (p. 42)

(54)  Sam enjoyed the rabbit.  uninformative context => meat-grinding rule applies  
‘enjoyed eating rabbit meat’

(55)  There was rabbit splattered all over the road.  rich context => overrides default

But is it really about the word rabbit?

(56)  Sam enjoyed it / that / etc.  
A lot of it was splattered all over the road.

2. Sense extensions can affect agreement.  Nunberg’s (1995) reference transfer, where the hash  
browns is shifted to refer to the person who ordered it:

(57)  The hash browns at table six wants / *want the check.

(58)  Pannkak-or är gott.  (Swedish) 
pancake-PL be.PRES good.NT.SG  
‘Situations involving pancakes are good.’  (e.g. ‘Eating pancakes is good.’)

(59)  Steroids is big business.

3. The lexical rule feeds (other) type-shifting rules.  (a); the ham sandwich, shifted to refer to the  
person who ordered it, can be converted to a predicate nominal (b):

(60)  a. The ham sandwich at table three wants his check.  
b. I am the ham sandwich.

Overgeneration?  C&B note that one lexical rule can feed another.  E.g.

(61)  We ordered three lambs (…and a poached salmon).  (‘three portions of lamb’)

lamb(count; animal) =animal-grinding-LR=> lamb(mass)  
lamb(mass) =portioning-LR=> lamb(portion)

But this also overgenerates.  C&B give the following example (p. 56).  There is a rule taking  
animal-denoting nouns and deriving nouns denoting humans with some properties of the  
corresponding animal (pig, lamb, etc.).  Applied to humans, a pig is a slovenly person, a lamb is
meek, and so on. So (62a) can have interpretation (62c or d). But it can’t be interpreted as (62e), even though the grammar would generate this:

(62)  
   a. John saw some lambs.  
   b. John saw some animals.  
   c. John saw some meek humans (people with lamb-like properties).  
   d. John saw some portions of lamb meat. (see above)  
   e. John saw some portions of substance derived from meek humans.  

(But would interpretation (62e) really be generated? As long as the input to the animal-grinding rule specifies a (non-human) animal then it could not apply to human-denoting lamb.)  

Some words allow both constructional polysemy and sense extension, e.g. container words like reel:

(63)  
   a. *film reel: a type of container (for film)  
   b. constructional polysemy: reel of film: the contents of the container; reel inherits semantic content from the PP complement (p. 30)  
   c. sense extension: The mystery is only resolved in the final reel.  

No PP; must be a lexical rule

2.6. Type presupposition accommodation  

Different aspectual verbs license different coercions (Asher 2011, 90):

(64)  
   a. Mary finished/stopped eating the apple.  
   b. Mary finished the apple. (non-motion; default: ‘finished eating’)  
   c. Mary stopped the apple. (motion only)  

(65)  
   a. *The apple finished  
   b. The apple stopped.  

• The meaning-shifting is in the verb rather than the noun.  
• Hence Asher rejects a pan-pragmatic view; coercions are ‘sensitive to actual lexical items’ (Asher 2011, 90).  
• As with Pustejovsky, coercion is a consequence of mismatches of semantic type.  
• But that sensitivity to actual lexical items affects presuppositions rather than truth conditional meaning (unlike Pustejovsky (1993, 1995) and Copestake and Briscoe 1995).  
• Coercion is a special case of the more general phenomenon of presupposition accommodation.  

prsupposition accommodation: Often presupposed content, if not already salient, can be accommodated: the phrase Sylvain’s son is almost three years old presupposes that Sylvain has a son, but that information is easily accommodated and added to the new updated context. Presuppositions are introduced via the semantic types that a predicate specifies for its arguments
3. Sense indeterminacy as a route to polysemy

*The Gavagai Hypothesis:* The royal road to polysemy is sense indeterminacy.

Summary:
- One source of polysemy is *sense indeterminacy:* multiple word senses are consistent with a given reference context (cp. Quine’s (1960) *gavagai problem*).
- As a consequence, *narrowing* and *broadening* of meaning are common semantic changes (Bréal 1900, among many others).
- This leads to *autohyponymy* as long as the earlier sense does not fall out of use.

3.1. Inquiry-resistant polysemy

Even in a fully specified discourse context, a word is typically consistent with multiple senses.

Erk et al. (2012, 513) asked annotators to pick WordNet definitions for words in a corpus context, allowing them to pick multiple senses and rate them for applicability. Annotators picked multiple senses, with high inter-annotator agreement. Example:

(66) This can be justified thermodynamically in this case, and this will be done in a separate paper which is being prepared.

Highest ratings shown for *paper*:
- (i) a material made of cellulose pulp
- (ii) an essay (especially one written as an assignment) (3, 3, 5)
- (iii) a daily or weekly publication on folded sheets; contains news and articles and advertisements
- (iv) a medium for written communication (5, 3, 1)
- (v) a scholarly article describing the results of observations or stating hypotheses (5, 5, 5)
- (vi) a business firm that publishes newspapers
- (vii) the physical object that is the product of a newspaper publisher

-- Language is inherently unstable, since we lack direct access to other people’s thoughts.
-- Think of word interpretation in context as a probability distribution over senses.

3.2. Gavagai, repurposed for a theory of polysemy

Quine’s (1960) tale of *gavagai:* A linguist has arrived at a village to do fieldwork on a language that he knows nothing about yet. A native says ‘Gavagai!’ just as a rabbit runs by.

The question: What does *gavagai* mean? Its meaning cannot be determined.

To illustrate referential indeterminacy, Quine cites many examples of concepts compatible with the utterance setting: ‘rabbit’; ‘animal’; ‘dinner!’; ‘rabbit parts’; ‘time slices of rabbit’; etc.

Interestingly, natural language words are often polysemous in just these ways.
‘dinner!’
languages where ‘rice’ means ‘meal’:
Japanese: gohan desu yo lit. ‘It’s cooked rice’ means ‘dinner is ready’
Cantonese: sihk faahn lit. ‘eat rice’ means ‘have at it’, even if meal contains no rice
Many Asian languages: ‘Have you eaten rice?’ is a greeting

‘rabbit parts’
We are having rabbit for dinner. (‘rabbit meat’)  
She was wearing rabbit. (‘rabbit fur’) 

time slices of rabbit
child as relational: Adult children of alcoholics
child as time-slice: When I was a child, I loved to play in the woods.

Correlation hypothesis: The likelihood that two senses A and B are senses of a single word is proportional to the strength of the correlation between A and B across spatiotemporal coordinates.

Nouns are used for reference to things. Things are related by taxonomic relations between natural kinds: rice is food; a rabbit is a quadruped.

Two strongly correlated descriptions:
(67)   a. Look! (pointing to location l at time t) There goes a rabbit!  
       b. Look! (pointing to location l at time t) There goes a quadruped!

Hence autohyponymy is common: cow (‘bovine’ ~ ‘female bovine’); finger (five digits of the hand; four digits opposed to the thumb).

Verbs are used for descriptions of situations (= events and states). Like things, situation types are related by taxonomies (walking is going). But they are also related by laws of nature: situations supporting killing(x,y) also support dying(y); situations supporting causing-to-break(x,y) also support breaking(y).

Two strongly correlated descriptions:
(68)   a. Look! (pointing to location l at time t) Someone is breaking a stick!  
       b. Look! (pointing to location l at time t) A stick is breaking!

Hence causative alternations are common.

3.3. Putting the brakes on generalization
Bayesian theories of lexical acquisition:
(Xu and Tenenbaum 2007) (Nematzadeh, Grant, and Stevenson 2015)
3.4. Semantic change

Research questions.
1. What sorts of changes in word meaning over time are attested (and unattested) in the world’s languages?
2. How can we explain the patterns of attested semantic change?

How word meanings change. Semantic change from meaning A to B involves a transitional phase of polysemy where the form has both meanings.

Traugott and Dasher’s (2002) model:

Stage I. Form L encodes meaning A

\[ L \rightarrow A \]

intermediate process leading from Stage I to II:
- new utterance-token meaning: speaker exploits inferences innovatively
- differential weighting of inferences: relevance, cultural factors, etc.
- utterance-type meaning: some inferences become new coded meanings

Stage II. Form L is polysemous

\[ L \rightarrow A + B \]

(Stage III. Form L loses original meaning: \( L \rightarrow B \). This is rather rare.)

Modified model: Speakers/addressees are never entirely certain of the senses of words.

Stage I. For a given word in context, the speaker/addressee has a weighted probability over different related senses, all of them compatible with the context:
- 84% sense1; 10% sense2; 6% sense3

Stage II. Given the speakers’ lack of definitive evidence for selecting any given sense, the system is inherently unstable. The coefficients shift. Now sense2 has become a more serious possibility:
- 55% sense1; 35% sense2; 10% sense3

Different patterns of synchronic polysemy correspond to different diachronic pathways of semantic change.

Example of semantic change: the verb click
Stage I. Coded meaning
[click] → ‘make a thin, dry, hard sound’

   Click here! (In the context of computing.)

sense1: ‘Place the cursor here and do something that causes a thin, dry, hard sound’
sense2: ‘Press one of the buttons on a mouse and release instantaneously or hold down while performing another action.’

Stage II. New coded meaning

Likelihood of sense2 increases until polysemy develops.

[click] → ‘make a thin, dry, hard sound’ + ‘press mouse button to initiate’

3.5. Semantic change in modal verbs

(69) a. They must be married, I demand it.  deontic: obligation
     b. They must be married, I am sure of it.  epistemic: probability

(70) a. You may marry him; I will allow it.
     b. You may marry him, for all I know.

Across languages, the deontic meaning precedes the epistemic:

DEONTIC < EPISTEMIC

Deontic modality (Gk. deon ‘what is binding’): social obligation or compulsion
Epistemic modality (Gk. episteme ‘knowledge’): probability

More examples:

(71) Japanese beki ‘must’: obligation < probability

(72) a. I promise to do my best.  obligation
     b. This promises to be a great class.  probability

Why this order? Traugott and Dasher attribute it to a general tendency to move towards greater subjectivity. They assume that epistemic modals express subjective attitudes.

Subjectification (pp. 30ff): The development of less truth-conditional and less referential meanings.
Alternative explanation: For a given modal operator, almost all situations supporting the deontic reading also support the epistemic one. E.g. if John is permitted to leave, then John might leave; if John is obligated to leave, then John will certainly leave, as long as he fulfills his obligation. So addressees hearing a modal in a context supporting a deontic interpretation are uncertain as to whether the epistemic meaning is intended instead. But the converse is not the case: many situations supporting the epistemic reading fail to support the deontic one, either because the subject is non-human, or else because the proposition is not subject to obligation.

(73) a. It must be raining. #It is obligated to rain.  
b. Your father must have blue eyes. #Your father is obligated to have blue eyes.

Stages of English:
OE  Old English   450-1150
ME  Middle English 1150-1500
EMDe Early Modern English 1500-1770
MdE Modern English 1770-1970
PDE Present Day English 1970-present

The history of English must

Stage I. ability, permission

Indo-European *med- ‘take appropriate measure, be fitting/mete’  
(cf. medical, modal, modify, commodity, …)

Early Germanic to OE mot- ‘be able, be permitted’

participant-internal ability:
(33) Wilt þu, gif þu most,11 wesan usser her aldorma, 
will you if you able:are be-INF our army leader
leodum  lareow?
people-DAT teacher
“Are you willing, if you are able, to be the leader of the army, the teacher of the people?”

(8th century, Genesis, 2482)

participant-external ability (= permission):
(34) Ic hit þe þonne gehate þæt þu on Heorote most12
I it you then promise that you in Heorot will:be:able
sorleas  swefan.
anxiety-free sleep
“|promise you that you will be able to sleep free from anxiety in Heorot.”

(8th century, Beowulf, 1671 [Visser: 1969: 1791])
Stage II. obligation/deontic

OE and Early ME saw a shift from permission to obligation (both still deontic).

(37) a. Hit is halig restendeg; ne most ṣa styrigan þine
   it is holy rest-day NEG may/can/must thou move thy
   beddinge.
   bed
   “This is a holy rest-day; you may/must not move your bed.”
   (c. 1000 æCHom II, 42 [Goossens 1987b: 33])

   b. we motions eow seegan eowre sawle þearfe, licige
   we must you-DAT tell your soul-GEN need please
   eow ne licige eow.
   you-DAT not please you-DAT
   “we must tell you about your soul’s need, whether it please you or not.”
   (c. 1000 æCHom I, 17 (App) 182.240 [Goossens 1987b: 32])

How did this change happen? Perhaps from the negative contexts, such as ‘you may not…’

Originally participant-external obligation, from an external authority. Later, in ME, some participant-internal uses appeared:

(39) I moste han of the perys that I se, Or I moot dye.
    “I must have some of the pears that I see, or I will die.”
    (1395 Chaucer, CT, Merchant, p. 167, l. 2331 [MED moten 3])

Also, there came to be cases where the obliged individual is inanimate:

(41) nota þæt every centre mot ben also smal as a needle &
    note that every center must be as small as a needle and
    in every equant mot be a silk thread.
    in every equant must be a silk thread
    “Note that every center must be as small as a needle and there must be a silk
    thread in every equant.”
    (c. 1392 Equatorie of the Planets, p. 26)

Stage III: epistemic

Some epistemic uses appeared in OE, especially with impersonal constructions (‘one can…’, and in conditional clauses:

(42) & raðe æfter þam, gif hit mot gewiderian, mederan settan.
    and quickly after that if it may be fine madder plant
    “and quickly after that, if the weather may be fine, [one can] plant madder.”
    (950–1050 LawGer 12, 454 [Denison 1993: 300])

This example seems ambiguous between deontic and epistemic:

(43) Ealle we motion sweltun.
    all we must die
    “We must all die.”
    (?8th century Exodus, 12.33 [Warner 1993: 162])

But the context suggests it was epistemic: the Egyptians are telling the Jews that if they don’t leave Egypt, we will all necessarily die.
‘must’ started appearing with ‘necessarily’: mot nede ‘must necessarily’. If the authority requiring an action is God, the law, or logic, then the deontic necessity entails epistemic necessity. Obligation to act in the future entails a (present) modal necessity, as in this example:

(45) Ah heo mot nede beien, þe mon þe ibunden bit. but he must necessarily submit the man that bound is “But he who is bound ought necessarily to submit/necessarily submits.” (c. 1225 (?1200) Lay Brut, 1051 [MED moten 2c])

By the middle ME, unambiguous epistemic uses appeared:

(46) a. He moste kunne muchel of art, he must know much of art (cunning)

sat þu woldest geve þerof þarte, that you would yield thereof part

“He must know much of art since you are willing to give part of it.” (c 1300 (?1250) Floris (Cmb), 521 [MED moten 4])

b. I have wel concluded that blisfulnesse and God ben the sovereyn good; for whiche it mote nedes be that sovereyne blisfulnesse is sovereyn devynite.

“I have properly deduced that blissfulness and God are the supreme good; therefore it must necessarily be that supreme blissfulness is supreme divinity.” (c. 1389 Chaucer, Boece p. 432, l. 124)

c. For yf that schrewedenesse makith wrecches, than mot he nedes ben moost wrecchide that lengest is a schrewe.

“For if depravity makes men wretched, then he must necessarily be most wretched that is wicked longest.” (c. 1380 Chaucer, Boece, p. 447, l. 47)

The history of Chinese de

Chinese de has this history:

‘obtain’ ⇒ ability ⇒ (deontic) permission ⇒ (epistemic) possibility

Modern Mandarin has a main verb de ‘obtain’:

(70) De le jiangxuejin.

get PERF scholarship

“X got a scholarship.” (Chao 1968: 741)

And an auxiliary verb de ‘be able, be possible, be permitted’:
Late Archaic Chinese had both:

\[ (73) \quad \begin{array}{ll}
\text{er} & \quad \text{de} \quad \text{tianxia}.\\
\text{and} & \quad \text{obtain} \quad \text{world} \\
\end{array} \]

“and have the kingdom.”

\[ (300 \text{ BC Mengzi, Gongxun Chou [Sun 1996: 112]} \]

Many examples are ambiguous between permission and epistemic readings.

\[ (75) \quad \begin{array}{ll}
\text{wu} & \quad \text{de} \quad \text{you} \quad \text{qi} \quad \text{yi} \quad \text{yi} \quad \text{man} \quad \text{qi} \quad \text{er}.\\
\text{how} & \quad \text{allowed/possible}\text{?} \quad \text{possess} \quad \text{this one} \quad \text{with} \quad \text{despire} \quad \text{these two} \\
\end{array} \]

“How is it allowed/can it be/is it possible that the person who possesses one of these (properties) despises the one who possesses the other two (properties)?”

\[ (300 \text{ BC Mengzi, Gongxun Chou [Sun 1996: 113]} \]

Conclusion

- Historically the deontic meaning precedes the epistemic, across languages.
- The reason for this direction of meaning change may be that contexts supporting the deontic are very likely to support the epistemic, while the converse is not nearly so likely.

3.6. The sound emission puzzle

Motion phrases: V must CAUSE motion

\[ (74) \quad \begin{array}{ll}
a. & \quad \text{We walked into the room.} \\
\text{walking} & \quad \text{CAUSES motion into the room} \\
b. & \quad \text{*We sang into the room.} \\
\text{singing} & \quad \text{does not CAUSE motion into the room} \\
\end{array} \]

Result phrases: V must CAUSE result
(75)  a. We pounded the metal flat.
    pounding the metal CAUSES the metal to become flat
b. *We watched the metal flat.
    watching the metal does not CAUSE it to become flat

A puzzle: sound emission verbs appear in contexts that otherwise require causation:

(76)  a. The truck rumbled down the street.
b. The wooden-legged man clumped into the room.
c. The train screeched into the station.
d. The fly buzzed out the window.
e. The door squeaked open.
f. As we began our drive to Nairobi, a tire hissed flat.

The rumbling sound, clumping sound, etc. does not CAUSE motion; yet these are OK
Why?

English sound emission verbs are subject to this polysemy rule:

(77)  to make sound \( x \)  
        \( \Rightarrow \) to perform activity that causes sound \( x \)

(78)  Examples: crack, smack, clash, crash, click, ...

\textit{crack, V} orig. To make a dry sharp sound \textbf{in breaking}, to break with this characteristic sound; hence, in branch I, mainly or exclusively of the sound; in II, of the act of breaking.
(here and below, from \textit{The Oxford English Dictionary})

I. Referring mainly to sound emission. (from c1000)
   • To make a sharp noise in the act of breaking, or as in breaking; to make a sharp or explosive noise (said of thunder or a cannon, a rifle, a whip, etc.)

1495. Comyn salt cracketh and sperkleth in fyre.
   • To shoot (with firearms), fire.
   • To cause (anything) to make a sharp noise.
   • To strike with a sharp noise; to slap, smack, box.
   • To utter, pronounce, or tell aloud, briskly, or with éclat (survives in crack a joke).

II. Referring mainly to the breaking indicated by the sound. (from c1300 AD)
   • To break anything hard with a sudden sharp report; now chiefly of things hollow, a skull, a nut, etc.

1553. To cracke the nutte, he must take the payne.
   • (from fig. to crack a nut) To puzzle out, make out, solve, discuss. (from 1640 on)
• To break or crush (corn, etc.) into small pieces. U.S.
• To break without complete separation or displacement of parts…

\textit{smack, V} To open or separate (the lips) in such a way as to \textbf{produce a sharp sound}; to do this in connection with eating or drinking, esp. as a sign of keen relish or anticipation.

1557. \textit{Not smackynge thy lyppes, As comonly do hogges.}

\textbf{To strike} (a person, part of the body, etc.) with the open hand or with something having a flat surface; to slap.

1835. \textit{Mrs. A. smacks Mrs. B.'s child for ‘making faces’.}

\textit{clash, V} To make … \textbf{the loud sound} of collision made by a heavy stroke or blow…

\textit{trans.}, where the object refers to the sound: 1667. \textit{And fierce with grasped arms Clash'd on their sounding shields the din of war. (Milton, Paradise Lost)}

\textbf{To come into violent and noisy collision} (incl. without the notion of noise). To come into, or engage in, conflict.

1820. \textit{It is not possible that the learned professors and the reading public should clash.}

\textit{click, V} To make a thin, dry, hard \textbf{sound.}

1682. \textit{The little clicking sound of the Dead-Watch.}

To \textbf{strike} with a click; to cause (anything) to make such a noise.

To press (one of the buttons on a mouse) and release instantaneously or hold down while performing another action; to activate (a program function) or select (a particular item) in this way, having first positioned the cursor on the appropriate part of the computer screen. (OED draft additions March 2001. \textit{Computing.})

2000. \textit{PC World}. Click an entry and drag it to the Insert menu..

\textit{to click through}: to click on a hyperlink using a mouse or similar device in order to access a specific file, web page, etc.; spec. to access an advertiser’s web site by clicking on an advertisement on another web page. Chiefly with \textit{to}. (OED draft additions October 2009. \textit{Computing.})

2009. \textit{Wired}. I clicked through to the user's photostream and determined it was the woman I had seen earlier.
An English polysemy pattern

L → ‘Make a type x sound’

> 

L → ‘Make a type x sound’ +
   ‘Do a particular activity that causes a type x sound’

**Result complements.** When a particular noisy activity has other effects besides the sound, then the verb denoting that noisy activity takes result complements encoding those other effects:

(79)

a. As we began our drive to Nairobi, a tire hissed flat.
b. The truck rumbled down the street.
c. The wooden-legged man clumped into the room.
d. The train screeched into the station.
e. The fly buzzed out the window.
f. The door squeaked open.

(exs. b-f from Goldberg 1995)

hiss → ‘make a hissing sound’ +
       ‘leak air, causing a hissing sound’

Causal relations:

leaking  hissing (sound)

flattening

• Leaking causes hissing, so leaking and hissing event types are correlated, across tire situations.
• The leaking and hissing events have a large spatiotemporal overlap.
• So hiss has come to mean ‘leak air, causing a hissing sound’; it is contextual polysemy, with this new sense manifest only in the context of a resultative complement.

(80) A tire hissed flat.

hiss → ‘make a hissing sound’ +
       ‘leak air, causing a hissing sound’ / ___ result

Q: Is it a generalized invited inference or a new coded meaning?

(81) The fly buzzed out the window.
[buzz] → ‘to make a buzzing sound’ +
‘to fly, causing a buzzing sound’ / __ path-of-motion

(82) Sense anaphora zeugma test:
a. One fly buzzed into the kitchen, and another did so into the dining room. (‘to fly, causing a buzzing sound’)
b. One trapped, immobile fly buzzed under the paper blinds, and another did so under a cup. (‘to make a buzzing sound’)
c. ??[Z] A trapped, immobile fly buzzed under the paper blinds, while a free one did so into the room.
   If c is non-zeugmatic ⇒ GIIN
   If c is zeugmatic ⇒ new coded meaning (polysemy).

(Levin and Rappaport Hovav 1995, 189ff) note the following condition: ‘the sound must be emitted as a necessary concomitant of the motion.’ (L&RH 1995, p. 191) Their examples 24b and 25a:

(83) a. At that moment, a flatbed truck bearing a load of steel rumbled through the gate.
b. *He yelled down the street.

Under the Correlation Hypothesis, this condition is explained. The common cause behind the sound and motion means that one is a necessary concomitant of the other. (83b) is ruled out because yelling and movement don’t have a common cause.

Comparison with constructional proposals.
The verb of sound emission is monosemous, but:
• the construction around V is polysemous (Iwata 2008, 1074); or
• the sound emission involves a special semantic relations between V and construction (Goldberg 1995, 2006).

Iwata’s (2008) proposal: the construction is polysemous.

(84) The door swung open.

   Syn: [... V ... [AP open/shut]...]
   ||
   Sem: “As a direct result of and concurrently with an internal motion described by V, the state of being open/shut obtains”

(Iwata 2008, 1074)
(85) The door squeaked open.

<table>
<thead>
<tr>
<th>Syn:</th>
<th>[... V ... [AP open/shut]...]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sem:</td>
<td>“As a direct result of and concurrently with a sound-emitting event described by V, the state of being open/shut obtains”</td>
</tr>
</tbody>
</table>

(Iwata 2008, 1082)

Iwata’s proposal: The *open/close* construction is polysemous; the construction in (85) ‘metonymically motivated’ by the one in (84). (Iwata 2008, 1083)

Problem: No mechanism for metonymic change.
Stage I. The door swung open. [ V - open ] construction
The mouse squeaked. squeak (sound)

Stage II. ?? (How does the new construction arise?)


(86) Let $e_v$ be the event type designated by the construction and $e_v$ the event type designated by the verb.
   a. $e_v$ may be a subtype of $e_c$
   b. $e_v$ may designate the **means** of $e_c$ (very common)
   c. $e_v$ may designate the **result** of $e_c$ (special relation used with sound emission verbs)
   d. $e_v$ may designate a precondition of $e_c$
   e. To a very limited extent, $e_v$ may designate the manner of $e_c$, the means of identifying $e_c$, or the intended result of $e_c$ (Goldberg 1995, 65)

Intransitive motion construction: X MOVE Y (place)
*X clumped* is the result: *The wooden-legged man clumped into the room.*

‘the clumping sound is the result of the man’s moving’ (Goldberg 1995, 62).
But is it?

*The wooden-legged man clumped into the room.*

walking — clumping sound

moving

A wooden-legged man walking in place still makes a clumping sound.
*cum hoc ergo propter hoc* fallacy?
N.b.: A fly buzzes because its wings vibrate. Motion into the room does not cause the buzzing sound!
• Can be fixed: “e_v and e_c may have a common cause”
• Suggests that correlation (and not causation per se) is the key
• Correlation is sufficient for inducing sense indeterminacy

Goldberg (1995, 65): ‘... why should these relations [in (86)] be privileged? ... if we consider certain verbs’ inherent semantics to bear a metonymic relationship to the semantics of the construction, we may find a partial explanation. The semantics associated with the construction defines a semantic frame, and the verb must inherently designate a particular salient aspect of that frame.’

Critique: It does not follow that two grammatical formatives must bear a metonymic relation, in order to be combined. In general they don’t. So an explanation for the relations in (86) is still needed. Perhaps the Correlation Hypothesis provides that explanation.

Conclusions

• Words develop new senses over time.
• Correlations, such as those introduced by causal relations, drive sense innovation, hence polysemy.
• New senses lead to new complement patterns, hence to certain diathesis alternations.
• Certain action types (a fly flying) cause both a characteristic sound (buzzing), and motion. Hence the three descriptions (flying, buzzing, moving) correlate across situations. Polysemy has developed between the sound (buzzing) and the causing action (flying).
• Proposals to move polysemy to constructions, or to the relation between the word and the construction, fail to provide any mechanism for innovation.
4. Causative alternations

4.1. Why are causative alternations so common?

The causative alternation is illustrated here:

(87) a. The string broke\textsubscript{intr}.
    b. Mary broke\textsubscript{trans} the string.
    (= ‘Mary \textbf{caused} the string to break\textsubscript{intr}.’)

The causative alternation is very common across languages. In contrast, labile transitivity
alternations in which the transitive variant of the verb has a meaning like (88b) are unattested, as
far as I know. Unlike \textit{causing}, relations like \textit{wanting}, \textit{expecting}, \textit{considering}, \textit{observing}, and so
on, do not modulate verb meanings.

(88) a. The string wugged\textsubscript{intr}.
    b. Mary wugged\textsubscript{trans} the string.
    (= ‘Mary \textbf{wanted} the string to wug\textsubscript{intr}.’)

(Some \textit{non-labile} alternations of this latter kind, e.g. where the transitive verb bears an affix
meaning ‘want’, are attested, as in Yupik-Inuit; cp. Woodbury 2005.)

The prevalence of (labile) causative alternations is captured— but not explained— by
positing a \textbf{CAUSE} operator in the lexical decomposition ontology (Dowty 1979) or ‘little \textit{v}’ to
introduce causal roles in the syntax. A typical conclusion of a survey of causative alternations
and related phenomena:

‘The conclusion: notions of causation and becoming are central to the semantic
architecture of language, part of grammar like tense and modality.’ (Chierchia and
McConnell-Ginet 1990, p. 438)

Causative-Inchoative alternation

• causative alternation is very common in language.
• Levin (1993:28-29) lists over 300 alternating verbs of English.
• found in a great many, perhaps all, languages of the world.
• Nedjalkov’s (1969) typological study covered 60 languages of diverse type, and

Q: Why is causation so common in systematic polysemy relations? (To put it differently: Why
is the \textbf{CAUSE} operator silent?)

\text{break\textsubscript{intr} $\sim$ break\textsubscript{trans} (‘\textbf{CAUSE} to break\textsubscript{intr}’)}

A: Summary of the proposed answer:

• One source of polysemy is \textit{sense indeterminacy}: multiple word senses are consistent with
  a given reference context (cp. Quine’s 1960 \textit{gavagai problem}).
• As a consequence, \textit{narrowing} and \textit{broadening} of meaning are common semantic changes
  (Bréal 1900, among many others).
• This leads to *autohyponymy* as long as the earlier sense does not fall out of use.
• The causative alternation is such a case: all reference contexts supporting (87b) also support (87a). Causation entails correlation. (In contrast, contexts supporting (88b) ‘Mary wanted the string to V’ need not support (88a) ‘The string V’ed.’)
• In sum, the prevalence of causative alternations arises from emergent regularities in the relation between word meaning and the world of described situations.

Note that this explanation does not rely on any particular assumptions about cognition. Notions like iconicity, concept complexity, salience, imageability, processing strategies, short term memory, and so on play no role in the account. Strictly speaking, it does not even require that causation be part of our ontology of concepts.

4.2. Causative ~ Inchoative ~ Resultative Alternations

The *causative alternation* results from optionality of a causer or agent argument:

(89) a. **causative:** Jonas dried the socks.
    b. **inchoative:** The socks dried.
    c. **result state:** The socks are dry.

Decomposition analysis (Lakoff (1965), (Dowty 1979)

(90) a. Joan opened the door.  CAUSE(j, BECOME(open’(d)))
    b. The door opened.  BECOME(open’(d))
    c. The door is open.  open’(d)

Haspelemath 1993: typological generalizations over causative-inchoative
Koontz-Garboden 2005: typological generalizations over inchoative-result state

(91) Varieties of morphological marking for causative alternations.
(Haspelemath (1993), following Nedjalkov 1969).

a. **causative:** inchoative is basic, causative is derived

ex. Georgian  \textit{duγ-s}  ‘cook (intransitive)’
\textit{a-duγ-ebσ}  ‘cook (transitive)’

b. **anticausative:** causative is basic, inchoative is derived

ex. Russian  \textit{katat’-sja}  ‘roll (intr.)’
\textit{katat’}  ‘roll (tr.)’
c. non-directed: neither is derived from the other; subtypes:

(i) equipollent: both derived from the same stem with different morphology:

ex. Japanese  
\textit{atum-aru} ‘gather (intr.)’  
\textit{atum-eru} ‘gather (tr.)’

(ii) suppletive: different roots. ex. English \textit{kill/die}.

(iii) labile: same form for both. ex. other English verbs described here

English causative alternations are non-directed: mainly \textit{labile}, with some possibly \textit{suppletive} pairs like \textit{kill/die}, \textit{put/go}, \textit{bring/come} (see below).

4.3. Spatiotemporal overlap and direct causation

(92) Spatiotemporal overlap condition: The situations described by the two senses of a word should overlap in time and location. The greater the size of the overlap, the stronger the tendency towards polysemy.

More precisely: For any given situation \(s_1\) describable by sense1, there should be a situation \(s_2\) describable by sense2, such that \(s_1\) and \(s_2\) overlap in time and location.

Prediction: Given an event \(s_1\) described by (93a), there is an event \(s_2\) described by (93b), and \(s_1\) and \(s_2\) overlap in time and location.

(93)  
\begin{enumerate}
  \item a. Sara is opening the door.
  \item b. The door is opening.
\end{enumerate}

A much-noted generalization: in comparison to periphrastic causatives, lexical causatives favor physical and temporal contiguity (cp. ‘direct causation’):

(94)  
\begin{enumerate}
  \item a. Sara opened the door. \textit{lexical causative}
  \item b. Sara caused the door to open. \textit{periphrastic causative}
\end{enumerate}

“...lexical causatives apply to cases of causation via direct or physical contact...” (Pinker, 1989, p. 48)

“the causer must physically manipulate the causee...[with lexical causatives]” (Shibatani, 1976, p. 31)

“The unacceptability of ...[the lexical causative shown above]...is due to the violation of the assumption that there is \underline{direct physical contact}” (Gergely & Bever, 1986, p. 230)

“...it would appear that ‘kill’ and innumerable other causative verbs require a specific reference to \underline{physical contact} between some two (or three) objects” (Wierzbicka, 1975, p. 495)
A variant: lexical causatives favor a direct relation between the causer (the verb’s subject) and causee (the verb’s direct object), i.e. without intervening causes.

<table>
<thead>
<tr>
<th>Causal chain</th>
<th>Linguistic expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sara caused the door to open. Sara opened the door.</td>
<td></td>
</tr>
<tr>
<td>Sara caused the door to open. *Sara opened the door.</td>
<td></td>
</tr>
</tbody>
</table>

(Wolff 2003, 2)

(95) a. Sara is opening the door.
     b. The door is opening.

(Wolff 2003, 11)

(96) a. The blue marble moved the green marble. (lexical causative)
     b. The blue marble made the green marble move. (periphrastic causative)

Woff 2003. Procedure: Participants watch an animation, and then “choose the sentence that best describes” the occurrence. Result:
Discussion. Under the Correlation Hypothesis, the situations described by *The blue marble moved the green marble* and *The green marble moved* should overlap in time and location. They do overlap in the unmediated chain condition; they do not overlap in the mediated chain condition. Hence the Correlation Hypothesis predicts the reported results.

Sentient causes ameliorate the effect of being mediated:

Initial causer is human; same procedure. Result (left): Participants were more willing to describe mediated chains with lexical causatives when the causer was sentient than when it was not.

Another experiment: Participants are exposed to intentional and unintentional mediated causal chains. Result (right):

(Wolff 2003, 25)

No obvious explanation for these results under the Correlation Hypothesis.
4.4. Which verbs alternate and why

The question of distribution can be split into two questions:

1. First, why do many intransitives (laugh, cry, smile, appear, and many others, for English) refuse to causativize?
2. Second, why do many transitives (cut, shelve, write, eat, and many others, for English) refuse to decausativize?

Addressing question 1:

Correlation theory analysis: Events of laughing, crying, smiling, talking, and so on, do not often overlap spatiotemporally with actions that cause them. (An exception: tickling!) The conditions necessary for the development of a word polysemous between ‘laugh’ and ‘cause-to-laugh’ are not met.

cp. *internal vs external causation* (Smith 1970):

- In the inchoative sentences ‘the activity or change can be said to occur without an external agent’ (*external agent*: external to the VP, i.e. an agent that is not also the affected theme),
- if the action is such that ‘external control of the change can be assumed by an agent’ then the causative variant is also possible (Smith 1970:101). [modify this to read: ‘direct, spatio-temporally contiguous external control…’]
- a verb alternates if the action it describes can occur with or without a cause that is external to the theme.

Addressing question 2:

Agent-Manner. Several scholars have observed that a causative verb specifying that the agent participates in the event in a specific way generally does not allow an inchoative variant (Levin 1993, (Guerssel, Hale et al. 1985) Haspelmath 1993, Hale and Keyser 1993, 1997, Kiparsky 1997).

(97) Agent-Manner. An inchoative verb must lack a specialized external agent-oriented meaning component.

Example: to cut something requires that the agent use a sharp instrument such as a knife or scissors, but tear places no such condition on the manner in which the agent participates in the event. Hence tear is a good candidate for omission of the agent, while cut is not (examples from Haspelmath 1993:93):

(98)  
a. The girl tore her pants. / The pants tore.  
b. The tailor cut the cloth. / *The cloth cut.

\[\text{cut} < \text{agent[\text{manner1}}, \text{theme}] > \quad (\text{manner1} = \text{use a sharp instrument})\]
\[\text{tear} < \text{agent}, \text{theme}] > \]

The agent-manner component can be optional, disappearing in the inchoative. 
smear has two specialized meaning components, involving agent and theme:

\[\text{smear} = \text{‘to spread or daub with a sticky, greasy, or dirty substance’} \ (\text{OED}).\]
**smear**<agent[manner1], theme[manner2], location>

manner1 = spread or daub (optional)  
manner2 = a sticky, greasy, or dirty substance (obligatory)

Many supposedly disallowed inchoative verbs improve dramatically under interpretations in which the agent-manner specification disappears or shifts from the agent to another argument. For example, (Hale and Keyser 1993) illustrate the purported Agent-Manner generalization with the following contrast (their examples 40-41):

(99) *splash, drip, dribble, pour, squirt, ...*  
a. The pigs splashed mud on the wall.  
b. Mud splashed on the wall.

(100) We smeared mud on the wall. causative: spread or daub

(101) inchoative: not spread/daub  
a. Maggie feeds it to him. He takes a bite of it, and it smears on his nose a bit.  
b. Bell said the problem stems from the plastic plate. After paint gets on it, it smears on the abutting wall.  
c. In that case, re-align your efforts into (a) making the sunblock last longer without reapplication ... and (b) making it smear on less goopily.  
d. Mud is part of life on the farm— it smears on jeans and sticks under fingernails.

Moreover, this account predicts that the inchoative should be difficult if neither agent nor theme is expressed. The verb *smear* is a locative alternation verb, but only one alternant can be decausativized:

(102)  
a. She smeared sunscreen on his nose.  
b. She smeared his nose with sunscreen.  
c. The sunscreen smeared on his nose.  
d. *His nose smeared with sunscreen.

(Web searches for *it smeared on* produced many hits; a search for *it smeared with* turned up no results). The operative condition on *smear*: any manner-rich argument must be expressed as a direct argument (subject or object).

**Conclusion.**

- According to the Correlation Hypothesis, the correlations introduced by causal relations drive sense innovation, hence causative / inchoative polysemy.
- This explanation does not depend upon any special assumptions about the conceptualization of causation.
References


