Rethinking Coercion as a Cognitive Phenomenon: Processing, Frequency, and Semantic Compatibility

Suzanne Kemmer
Rice University

Soyeon Yoon
Seoul National University/
Rice University

ICLC-12, Edmonton, June 2013
Coercion

- Resolution of semantic incompatibility between a construction and a lexical item occurring in it (Michaelis 2005)
  - Give me some pillow!
  - I sneezed the napkin off the table.
  - I’m liking it.
- A contextual reinterpretation that occurs when semantic specifications clash (Pustejovsky) –
  - An adjustment of specifications “repairs the mismatch” (de Swart)
  - “Special meaning effects” (de Swart)
Issues with coercion as typically conceived

- Changes in binary feature specifications (Michaelis 2005; de Swart 2000) are inconsistent with a frame-based, gradient semantics.

- It is not clear how the coercion ‘mechanism’ relates to online processing or other aspects of language use:
  -- Theories (Construction Grammar; formal grammars) are generally either silent or inexplicit about how processing relates (but see Traxler et al. 2002, Piñango et al. 1999, Piñango et al. 2006).
  --or, they explicitly divorce language structure from processing (e.g. Sign-based Construction Grammar).
Usage-based Model: Correlation of Four Dimensions

**General prediction** (Kemmer 2008 following from Langacker 1987, 1990, 2000 inter alia.)

There should be some correlation between:

- Semantic **compatibility** of a host construction with lexical item
- **Frequency** of use (distributional/behavioral correlate of cognitive entrenchment)
- **Processing** time
- **Acceptability** judgments


But an empirically-demonstrated overall correlation will support the usage-based model as described by Langacker: **dynamic, gradient, integrated**
Why?

Semantic compatibility;
Frequency;
Processing;
Acceptability

--Why and how should these relate?
In a dynamic usage-based language system:

- **Greatest semantic compatibility**: Maximal conventionality, minimal semantic extension; schemas and exemplars fit together in their specifications, no clash.

- **Frequency**: Constructions are schematizations over many exemplars; they derive exactly from repetition of exemplars that (therefore) best fit them. Highly frequent exemplars are analogical attractors for novel exemplars of less frequency and less compatibility – including coercions.
Interrelation, cont.

• **Processing:** Generally, cognitive mismatches should be harder to process. Specifically, in a cognitive competition model, ambivalence/difficulty of categorization should take more time. Also--a well-known property of cognition: the more frequent the experience, the easier (and therefore faster) it is to process.

• **Acceptability judgments:** Speakers like most what they have most heard before: schemas with their usual exemplars in prototypical relations. Minimal mismatches. (Boas 2011 shows relation of coercion, semantic compatibility, and variable acceptability)

All subject to incremental change over time; and construction and its conventional and productive uses developing as the individual’s language system matures.
Investigated for one construction in Yoon (2012)

English Ditransitive Construction [V NP1 NP2]

- *Sally gave John the book.*
- Constructional meaning: transfer of possession from an Agent to a Recipient

The criteria of semantic compatibility

- the number of participants in the prototypical event scene of the verb
- the possibility that the Patient is transferrable as a result of the action prototypically designated by the verb (e.g. *kill*)
## Semantic compatibility of verbs in Ditransitive Construction (DC)

<table>
<thead>
<tr>
<th>Semantic Compatibility (1 most, 5 least)</th>
<th>Semantic Type of Verb</th>
<th>Eg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>SemCom1</td>
<td>Inherent transfer</td>
<td>give, send</td>
</tr>
<tr>
<td>SemCom2</td>
<td>Potential transfer</td>
<td>cook, find</td>
</tr>
<tr>
<td>SemCom3</td>
<td>Prevented transfer</td>
<td>refuse, deny</td>
</tr>
<tr>
<td>SemCom4</td>
<td>Impossible transfer</td>
<td>cut, break</td>
</tr>
<tr>
<td>SemCom5</td>
<td>Events internal to the Agent</td>
<td>think, stay</td>
</tr>
</tbody>
</table>

Categories 1-3 based on Pinker (1989) and Goldberg (1995)
Verbs said to not occur with the DC (Goldberg 1995: 128)

• Verbs of fulfilling (X gives something to Y that Y deserves, needs, or is worthy of)
  
  present, donate, provide

• Verbs of continuous causation of accompanied motion in some manner
  
  pull, carry, push

• Verbs of manner of speaking
  
  shout, murmur, whisper

• Verbs of proposition and propositional attitude
  
  say, claim

• Verbs of choosing
  
  choose, pick
1. Frequency of verbs in DC

Method
• Collexeme Analysis (Stefanowitsch and Gries 2003)

Corpus
• BNC, spoken subcorpus - ca. 1,450,000 words
• # of DC exemplars: 1,374
• # of verbs used in the DC: 49
Result

- Verbs more compatible with the DC tend to be more frequently associated with the DC (higher collostructional rank).
- Verbs less compatible tend to occur less frequently in the DC or do not occur at all.

Table shows the relation:
## Compatibility and Collostructional Rank

<table>
<thead>
<tr>
<th>Compatibility</th>
<th>Average Collostruction rank</th>
<th># of verbs found</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>SemCom 1</td>
<td>17</td>
<td>25</td>
<td>give, send, tell</td>
</tr>
<tr>
<td>SemCom2</td>
<td>33</td>
<td>20</td>
<td>buy, make, find</td>
</tr>
<tr>
<td>SemCom3</td>
<td>29</td>
<td>2</td>
<td>refuse, deny</td>
</tr>
<tr>
<td>SemCom4</td>
<td>-</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>SemCom5</td>
<td>34</td>
<td>2</td>
<td>run, wish</td>
</tr>
</tbody>
</table>
Next chart shows relation of specific semantic classes of verbs (and their individual verbs) and collostructional rank.
2. Processing effort and acceptability judgments (DC)

- **Experiment Design**
  - **Stimuli**
    - 35 verbs selected from semantic compatibility categories and result of corpus analysis
    - 35 sentences where each verb was used as a main verb in the DC
      - (1) Eddie *told* Kim the news last month. (*tell* from SemCom1)
      - (2) Billy *found* Jane the ring six days ago. (*find* from SemCom2)
Task

- 27 participants read the sentences in a self-paced reading task.
- The time taken to read the second NP (underlined in (1) and (2)) was recorded.

Acceptability judgments

- After reading each sentence, the participants judged its naturalness on 7-point-scale.
<table>
<thead>
<tr>
<th><strong>Verb semantic class</strong> (from 1, most compatible, to 5, least compatible)</th>
<th><strong>Verb subclass</strong></th>
<th><strong>Selected verbs</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Verbs of inherent transfer</td>
<td>Inherently signifying giving Communication Instrument of communication Future having Sending Deictic</td>
<td><em>give</em>, <em>tell</em>, <em>fax</em>, <em>owe</em>, <em>promise</em>, <em>leave</em>, <em>allow</em> <em>send</em>, <em>bring</em></td>
</tr>
<tr>
<td>2. Verbs of possible transfer</td>
<td>Ballistic motion Creation Obtaining</td>
<td><em>throw</em>, <em>drop</em>, <em>create</em>, <em>cook</em>, <em>find</em>, <em>buy</em>, <em>rent</em> (<em>hire</em> in BE)</td>
</tr>
<tr>
<td>3. Verbs of refused transfer</td>
<td>Refusal</td>
<td><em>refuse</em>, <em>deny</em></td>
</tr>
<tr>
<td>4. Verbs of impossible/impaired transfer</td>
<td>Damaging</td>
<td><em>break</em>, <em>cut</em></td>
</tr>
<tr>
<td>5. Verbs of events internal to the Agent</td>
<td>Emotion/cognition/desire intransitive</td>
<td><em>think</em>, <em>want</em>, <em>wish</em>, <em>stay</em>, <em>sneeze</em></td>
</tr>
<tr>
<td>Verbs occurring only in the corpus (placed in 2\textsuperscript{nd} most compatible)</td>
<td>Location General causation</td>
<td><em>put</em>, <em>set</em>, <em>cause</em></td>
</tr>
<tr>
<td>Verbs that were expected not to occur in the DC (the least compatible)</td>
<td></td>
<td><em>present</em>, <em>donate</em>, <em>provide</em>, <em>push</em>, <em>whisper</em>, <em>say</em>, <em>choose</em></td>
</tr>
</tbody>
</table>
Significant trend

- If semantically less compatible, processed slower

- Judged as less acceptable
Semantic compatibility with processing time; with acceptability judgments

Average processing time of each semantic compatibility category
(Linear Trend: $t(26) = 3.02, p < .01$)

Average naturalness score of each semantic compatibility category
(Linear Trend: $t(26) = 30.29, p < .001$)
Excluding outliers (misclassified?) *put, set, and cause*

A more linear trend

Average processing time of each semantic compatibility category (excluding *put, set, and cause*)

Average naturalness score of each semantic compatibility category (excluding *put, set, and cause*)

\[
y = 21.968x + 766.94 \\
R^2 = 0.65421
\]

\[
y = 1.16x + 0.21 \\
R^2 = 0.91509
\]
All four aspects were significantly correlated with each other.

<table>
<thead>
<tr>
<th>(*p &lt; .01, **p &lt; .001)</th>
<th>SemCom</th>
<th>ColloRank</th>
<th>NatScore</th>
<th>ProcessingT</th>
</tr>
</thead>
<tbody>
<tr>
<td>SemCom</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ColloRank</td>
<td>.42**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NatScore</td>
<td>.54**</td>
<td>.41**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ProcessingT</td>
<td>.09*</td>
<td>.12 **</td>
<td>.13**</td>
<td></td>
</tr>
</tbody>
</table>
If semantic compatibility is gradable, will coercion be the same for all different degrees of semantic compatibility?

→ No, coercion is also gradable:

*Kelly sent* Ryan the card.
*Billy found* Jane the ring.
*Larry refused* Kim the lunch.
*Jean broke* David the bread.
*Ricky stayed* Sue the space.

Semantic compatibility correlates with:

- how often the resolved co-occurrences are used
- how difficult the resolution is to process
- how natural the speakers feel the co-occurrences are
Coercion is closely related with usage, specifically, processing.

- In comprehension, speaker requires different amounts of actual processing effort, depending on the amount of semantic incompatibility.
- Coercion can be thought of not as a theoretical mechanism in the “grammar”, separate from processing (and usage in general) but as part of an actual psychological process during language use: resolving semantic incompatibility online in usage events.
Directionality of coercion

Semantics of the target lexical item and the construction

Sometimes, the meaning of a lexical item overrides the constructional meaning.

→ challenges Override Principle (Michaelis 2005) claiming construction always coerces lexical item

- Larry refused Kim the lunch.
- Kevin caused Liz the fire.
Linguistic and extralinguistic context affect coercion

- Linguistic context
  Sometimes, coercion is easier with particular linguistic contexts – particularly V NP collocations (via activation of general or specific frames)
  - *Larry owed Jane the watch.* vs. *Larry owed Jane $10.*
  - *Kevin caused Liz the fire.* vs. *Kevin caused Liz trouble.*

- Extra-linguistic context
  - Speakers try to resolve the incompatibility by exploiting extra-linguistic context.
  - *David broke Jean the bread.*
  - *She squinted into the room.* (Kemmer 2008)
What is coercion, really?

What people call “coercion” is a subcase of dynamic semantic integration of constructional schemas, lexical schemas, their associated conventional frames, and contextual elements

--where the incompatibility is noticeable (there’s some violation of a generalization that works in prototypical cases)

--during syntactic/semantic composition of open-slot constructions with lexical items

(purely semantic composition/resolution as in colorless green ideas, has not been of much interest in modern Linguistics).
Why investigate coercions in particular?

Relevance:

• Coercions are relatively novel motivated usages that partly conform to an existing constructional schema. Thus they are relevant to syntax.

• We can closely observe the synchronic grammar and its processing at an interesting point: where conventionalization of a construction is intermediate, and it works with some classes of lexical items but not others.

• **Diachrony:** Emergence and change of constructions can be studied. As exemplars of a particular type become more entrenched, the construction changes its specifications (cf. Israel 1996).

• **Acquisition:** Can investigate learning of a construction and expansion to new lexical items/classes of lexical items.

• **Variation:** Can observe variation among and within speakers.
Coercion is a concept widely invoked to allow for/explain semantic mismatches and to argue for existence of constructions.

We conclude:

1. Since theoretical ideas rest on it, its nature should be more closely investigated.

2. Coercion is a gradient cognitive process reflected in variable processing time. It is not a unitary or “all or nothing” device or process.
Conclusions

3. Coercion is the set of syntactically relevant subtypes of the **dynamic process of semantic integration** of:

- conventional linguistic specifications
- frame-based knowledge not specific to language
- contextual elements

This general process occurs in language usage in general, not just in syntactic constructions noticed by linguists.
Phenomena given the name “coercion” are disparate, e.g.:

- “NP-coercion” – specifically mass construed as unit (Give me a beer) – is highly conventionalized in English

- A schema with semantics ‘conventionally unitized drink’ has entrenched exemplars with particular lexical items associated with particular frames; is compatible with count noun constructions (singular indef. article, pluralization); and can be licensed for non-conventionalized nouns (new drink names, masses not usually unitized etc.), in contexts activating the frames associated with the schema

- Under usage-based model, entrenched cases like a beer do not actually involve coercion. They are expected to be processed more quickly, show higher frequency, and have greater acceptability than found in cases of real incompatibility (genuine coercions)
4. Coercion can be investigated for specific constructions, but we need to take into account the degree of entrenchment of relevant constructional schemas, specific and general.

Doing so will provide:

- A more general and accurate description of coercion phenomena
- Stronger theoretical grounding
- Natural relation to acquisition, synchronic variation, and diachrony
References


Details on regressions.
Details: Regressions correlating the factors

Regression 1

\[ y_{\text{NatScore}} = 0.79x_{\text{SemCom}} + 0.03x_{\text{ColloRank}} + 0.001x_{\text{ProcessingT}} - 2.87 \]

\[ (p < .001) \quad (p < .001) \quad (p < .05) \]

- Semantically less compatible construction and verb
- Less frequently used together
- Processed with more effort
- Judged less acceptable
- More coercion

Regression 2

\[ y_{\text{ProcessingT}} = 7.79x_{\text{SemCom}} + 1.47x_{\text{ColloRank}} - 67.72 \]

\[ (p = .22) \quad (p < .01) \]
## Multiple Regression

<table>
<thead>
<tr>
<th>Step 1</th>
<th>Unstandardized coefficient</th>
<th>Standardized coefficient</th>
<th>p</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-34.70</td>
<td>.09</td>
<td>p &lt; .01</td>
<td>r² = .01 (p &lt; .01)</td>
</tr>
<tr>
<td>SemCom</td>
<td>15.37</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td>Unstandardized coefficient</td>
<td>Standardized coefficient</td>
<td>p</td>
<td>R²</td>
</tr>
<tr>
<td>Constant</td>
<td>-67.72</td>
<td>.04</td>
<td>p = .218</td>
<td>r² = .02 (p &lt; .001)</td>
</tr>
<tr>
<td>SemCom</td>
<td>7.79</td>
<td>.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ColloRank</td>
<td>1.47</td>
<td></td>
<td>p &lt; .01</td>
<td></td>
</tr>
</tbody>
</table>

Diagram:

- **ProcessingT**
  - **ColloRank**
    - **SemCom**