Constructionalization

How new constructions are created
Constructions are repeated and conventionalized sequences of morphemes or words which bear a particular semantic, functional or probabilistic relation to one another, and in addition, specify at least one sequential position that can be filled by more than one word or morpheme (Fillmore et al, Goldberg 1995, Croft 2001, Langacker 1995, Tomasello 2003, Bybee 2001)
What is a construction?

Form / meaning pairing (a linguistic sign)

Goldberg:
Morphemes
Words
Prefabs
Morpheme sequences
Word sequences
Clause sequences
What is a construction?

Entirely fixed

Partially fixed +

having open slots filled by categories of items.
Most constructions involve specific words or morphemes

Specific grammatical morphemes:

Poss: [NP] of [NP]; [NP]’s [N]

page of the book; the boy’s mother

Passive: [NP] BE + VERB + participle (by NP)

Small closed classes:

NP → Det + N

Det → the, a, some, this, that, these, those

These constructions arise by grammaticalization.
Constructions have open slots

Some are highly schematic

\[ \text{NP} + \text{going to} + \text{VP} \] (non-finite form)

English subject construction

\[ [\text{NP}]_{\text{SUBJECT}} + [\text{VP}] \]

More specific constructions for individual Vs.
Types of schematicity

Structure of the schematic slots in constructions

1. Semantically defined; centered around a prototype
2. Purely schematic

Type frequency counts the number of distinct lexemes that occur in the slot.
productivity

How likely is the construction to apply to a new lexeme?

English irregular verbs: many classes, fewer than 200 verbs all together.

English regular construction:

\[ [[\text{VERB}] + /d/, /t/, /\dd/]_{\text{PAST}} \]

Applies to thousands of verbs.
Types of schematicity

Spanish verbs of becoming:

*quedarse* ‘to remain + reflexive’

several semantic foci: alone, lacking

still, quiet

surprised
Types of schematicity

Spanish verbs of becoming:

hacerse ‘make + reflexive’
  bueno ‘good’
  rico ‘rich’
  presente ‘present’
  viejo ‘old’
  realista ‘realistic’
  responsable ‘responsible’

Highly schematic semantically. Lower type frequency.
Constructionalization

Sources of new constructions
1. Discourse > grammar (Givon)
2. An instance of an existing construction (this includes grammaticalization)
3. Expansion of a prefab or idiom
Existing constructions $>$ new constructions

This leads to a lot of regularity within the syntax.

**SVO:**

$[\text{VERB}]_{\text{MAIN}} + [\text{VERB} + (\text{NP...})] > [\text{AUX}] + [\text{VERB} ...]$  
The AUX will be preposed.

**SOV:**

$[\text{VERB} + (\text{NP...})] + [\text{VERB}]_{\text{MAIN}} > [\text{VERB}...] + [\text{AUX}]$
Expansion of a prefab

Prefab? A conventionalized word sequence.

idioms: *pull strings*

highly entrenched phrases: *good morning*

less frequent sequences but conventional: *best friend, mixed message, prominent role*
Spanish quedarse solo (Wilson 2009)

E el conde quando vio que de otra manera no podia ser sino como queria el comun delos romeros no quiso ay quedarse solo & fazia lo mejor & cogio sus tiendas & fue se empos delos otros.

‘And when the count saw that there could be no other way than the common wishes of the pilgrims to Rome wanted it, (he) didn’t want to be left alone and did his best and gathered his tents and went after the others.’ (Gran conquista de Ultramar, anon., 13th c.; Davies 2005)
Figure 1: The *solo* clusters in the 1200’s.
Figure 2, 1800s: The *solo* clusters.
The structure of schematic slots

There is often a central member of the category. (Bybee and Eddington 2006)

A high frequency item with general meaning. (Goldberg, Casenheiser and Sethuraman 2004)
Item-based analogy

solo ‘alone’  viuda ‘widow’
abandonado ‘abandoned’

sin padre ‘without a father’
sin jefe ‘without a boss’
sin hijo ‘without a son’

sin pluma ‘without a pen’
sin armas ‘without weapons’
Evidence that the most frequent member is central

• The larger, more productive categories have a semantically central exemplar that is of high frequency.

• Family resemblance structure uses the most frequent adjective

petrificado - inmóvil - atrapado
‘petrified’ ‘immobile’ ‘trapped’
How does high type frequency affect productivity?

From the production point of view:
1. Lots of types mean that there are lots of models available for item-based analogy.
2. High type frequency makes the construction’s representation stronger and easier to access.

From the decoding point of view:
1. Rare occurrences (hapax legomena) have to be parsed to be interpreted and the parsing strengthens the construction (Baayen 1993).
What makes a construction emerge?


Constructionalization is the creation of a new form – meaning pairing.

‘It is accompanied by changes in degree of schematicity, productivity and compositionality.’
Constructionalization

It’s a construction if its meaning is non-compositional.

Smirnova, Diewald, Heine and others note that the gradual change in pragmatic and semantic content occurs much as in grammaticalization.

Inferences occurring in the ‘bridging’ context lead to uses in the ‘switch’ context.
Constructions

• Constructions can carry specific pragmatic implications while still maintaining at least a surface resemblance to the construction from which they arose.
Diner:
Waiter, what’s this fly doing in my soup?

Waiter:
Why, madam, I believe that’s the backstroke.

(From Kay and Fillmore 1994)
What’s X doing Y?

C. Johnson points out that there is an inherent ambiguity in certain locative expressions in that they can apply to an activity or a participant in that activity:

What are you doing with that knife
  = ‘why do you have that knife?’
or the literal meaning
  = ‘what are you doing with it?’
Early examples of ‘WXDY?’

(9) \textit{c1430} Syr Tryam. 431
What do ye here, madam? Fro whens come ye?

(10) 1470-85 \textcolor{blue}{MALORY} Arthur XIV. v,
What dost thow here? He ansered I doo neyther good nor grete ylle.
Incongruity / disapproval

(11) 1656 COWLEY Misc., Swallow 3

Foolish Prater, what do 'st thou
So early at my window do
With thy tuneless Serenade?
1835 Aristophanes' *Clouds* in Blackw. Mag. XXXVIII. 520

*Str.* Pray who's that in the basket hung up in the air?.. Do tell me, I pray, what you're doing up there.

*Soc.* Aerobating, sun-musing, pacing air.
Updating of form shows

1. The construction was not frequent enough nor autonomous enough to maintain the old form with simple *do*. (cf. *How do you do?*)

2. It was still closely related to its mother construction, the Wh-question.
Implications

• The incongruity implication becomes conventionalized with this set of exemplars of the Wh-question because this implication commonly occurs with it.

What’s that box doing up there?
Incongruity implication

• In order for this construction to acquire special implications, the hearer/learner must record in memory the implications of incongruity from the very first exposure.

• Exemplar representations allow this record of detail about contexts of use and implications.
Constructions

- Research on constructions has focused on the fact that many constructions have idiosyncrasies of form, meaning or pragmatic value. Ex:
  - *What’s that box doing up there?*

Requires a non-literal interpretation of *doing* and implies the situation is incongruous.
Why do constructions have meaning?

• Constructions develop (both diachronically and in acquisition) by the domain-general process of chunking.
• Chunks are repeated sequences of behavior or experience that come to be stored in memory as units (Newell 1990, Ellis 1996).
As Newell 1990 put it:

A chunk is a unit of memory organization, formed by bringing together a set of already formed chunks in memory and welding them together into a larger unit. Chunking implies the ability to build up such structures recursively, thus leading to a hierarchical organization of memory. Chunking appears to be a ubiquitous feature of human memory. (p. 7)
Chunks

- Chunks are storage and access units.
- Chunks are found at many different levels in language. Prefabs, formulaic sequences and constructions are chunks.
- Chunks can have schematic slots and can be embedded in one another.
Cross-modal association

- The assignment of meaning to linguistic form is the result of the domain-general process of cross-modal association (Ellis 1996:110).

- James ‘Law of Contiguity 1890:
  ‘Objects once experienced together tend to become associated in the imagination, so that when any one of them is thought of, the others are likely to be thought of also, in the same order of sequence or coexistence as before.’
Cross-modal association

• Ellis goes on to say:

‘The implicit, automatic pattern-detection processes that occur within these modalities of representation entail that any such cross-modal associations typically occur between the highest chunked levels of activated nodes.’ (1996:110)
Playing fetch
Assigning meaning to chunks

• Constructions acquire meaning because of the human tendency to assign meaning at the chunk level.

• To make sense of our environment we have to identify patterns, sequential or otherwise, and once identified these patterns form categories in cognition which can then be associated with other categories.

• The meaning of the chunk can override the meanings of the parts.
Constructions

• The assignment of meaning to common sequences of words, such as occur in constructions, is driven by this process.

• When a sequence of words is commonly associated with a certain context, the understanding associated with that context becomes associated with that group of words.
Constructional meaning

- Thus Kay and Fillmore write about the *What’s X doing Y?* construction:
  - ‘While the WXDY construction may have had its origin in conversational implicatures—through situations in which an individual A is clearly up to no good and B asks what A is doing—the semantic of incongruity is now conventionally associated with the special morphosyntax of WXDY constructs (emphasis in original).’ (1999:5)
Assigning meaning at the highest chunked level

Leads to

- loss of compositionality
- loss / change of constituent structure

The construction determines constituent structure, that is, the parts of the sequence that belong together.
Recurrent sequences become units
Recurrent sequences become units

A
  B
  |
  C
  |
  D

semantic-pragmatic context
Recurrent sequences become units

\[
\begin{array}{c}
\text{A} \\
\text{B} \\
\text{C} \\
\text{X}
\end{array}
\]

Semantic-pragmatic context
Recurrent sequences become units

semantic-pragmatic context
Recurrent sequences become units

Z B C E

semantic-pragmatic context
Recurrent sequences become units

```
U    B    C    D
```

```
{semantic-pragmatic context}
```
Recurrent sequences become units

V B C F

semantic-pragmatic context
Recurrent sequences become units

\[
\begin{align*}
&Z & B & C & G \\
&\text{semantic-pragmatic context}
\end{align*}
\]
Recurrent sequences become units

semantic-pragmatic context
Constituent structure

Do complex prepositions have internal constituent structure?

```
PP
  /  
NP  PP
  /   
NP  NP
 /    /
in  spite of the king
```
Constituent structure

Are *in spite of*, *in front of*, *on top of*, etc. constituents themselves?
Degrees of constituency

• We expect to find mixed indicators of constituency since analyzability is lost gradually.

• Method: check corpus data to find out how speakers are really using this phrase.

• Along with other factors, examine semantics as evidence of reanalysis, and as a mechanism for change.
Constituency tests

I. Fronting (Seppänen et al. 1994)

(1) *Of what obstacles did he say he would do it in spite?

In spite of fails this test. (No evidence that there is an internal constituent boundary).
Constituency tests

II. Coordination

(2) *In spite of* your objections and *of the point raised by Dr Andersson*, we feel confident that we can proceed with the project.

Note the repeated *of* in this (constructed) example. Such usage would suggest internal structure where *of* still heads a PP.
Corpus data: Coordination

In COCA (385 million words), there are 3 instances that match the coordination in (2):

(3) it's one of the easiest rip-offs to execute, in spite of the growing array of removable stereos and of car alarms that address you in the imperative. (COCA 1993)
Corpus data: Coordination

However, in such cases it is far more common to use *in spite of* as an uninterrupted unit.

There are 43 corpus examples of the following type:

(4) **In spite of** motorbikes, **in spite of** karaoke music, **in spite of** the stink of gasoline fumes that seeps into each kitchen. (2005)
Corpus data: Coordination

When *in spite of* is conjoined with other complex preposition sequences, 7 examples are of this type:

(5) The prime minister remains unable to reap the credit for economic success, which is perceived to have occurred *in spite*, not *because*, *of* his policies...(1995)

(6) ...a lesson in how Congress makes politically expedient decisions *at the expense (or in spite)* of the constitutional implications of their actions (2002)
Corpus data: Coordination

In contrast, there are 35 examples of the following types:

(7) Do it in spite of, or more accurately because of, how difficult it is to do. (2005)

(8) ...in this allegedly anti-American country Sarkozy would be elected (as early as the spring of 2007) either because of or in spite of the public perception that he is somehow “American.” (2005)
Corpus data: Coordination

Conjoined with simple prepositions:

(9) Scorsese's strongest works are fictions of formation, in which a religious conviction comes with or in spite of a vocation. (1991)

(10) Commitment is healthiest when it is not without doubt, but in spite of doubt. (1991)
Constituency tests

III. Interpolation (Seppänen et al. 1994)

(11) The morning air was clear and clean, in spite, one might add, of the traffic and crowds. (constructed example)
Corpus data: Interpolation

One written example, a quote from 1877:

(12) The religionists of our time are occupying about the same ground occupied by heretics and infidels of one hundred years ago. The church has advanced in spite, as it were, of itself.

Compare with ~6200 corpus instances of in spite of that occur without interruption.
Syntactic tests

• We find minor indications that *of* is analyzable as a separate preposition, but a lot of evidence that it is part of the chunk *in spite of*.

• This is not surprising if constituency is due to analyzability and is gradient; it is quite surprising if constituent structure is discrete and uniquely determinable.
Gradient constituency

• The other vexing property of complex prepositions is that they all seem to demonstrate differing degrees of constituency (Quirk et al. 1985).
• This is another argument for the gradient constituency point of view.
• Each complex preposition can change at its own rate.
Additional usage evidence

\[ \text{in spite} \xrightarrow{\text{TP: 99.5\%}} \text{of} \]

\begin{itemize}
  \item Frequency: 6254
  \item Frequency: 6224
\end{itemize}
Competition between old and new constructions
Resistance to change

• Exemplars of morpho-syntactic constructions, like morphologically complex words, are resistant to change if they are highly frequent.

• Just as irregular verbs that are of high frequency resist regularization, so constructions with irregular properties resist change in the particular exemplars that are of high frequency.
Not-negation and no-negation

• Tottie 1991

(13) He did not see anything. (not-negation)
(14) He saw nothing. (no-negation)
Examples

(15) by the time they got to summer there was no more work to do

(16) the Fellowship had no funds

(17) as a nation we are not doing well enough. This is no new discovery
older and newer constructions

Old English had negative concord

Ic wyrce þa tacnu þe næfre nan man ne geseah ær
I do the miracles that never no man not saw before
on nanum lande
in no land.
‘I will do miracles that no man has ever seen before in any land.’

(Haeberli 1991: 58; Exod, 34.10)
*not* developed from *ne + wiht* ‘nothing’

Because *wiht* was a noun, it followed the verb

(3) he ne  edstont nawt as foles  doð II   ah...  
he not stops  not  as fools do   but...

(4) my wife rose nott
Examples

(15) *there was no* more work to do =  
    there wasn’t any more work to do

(16) the Fellowship *had no* funds =  
    the Fellowship did not have any funds

(17) This *is no* new discovery =  
    This is not a new discovery
Table 2: Proportion of *no*-negation (Tottie 1991)

<table>
<thead>
<tr>
<th></th>
<th>Spoken</th>
<th></th>
<th>Written</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Existent <em>be</em></td>
<td>34/38</td>
<td>89%</td>
<td>96/98</td>
<td>98%</td>
</tr>
<tr>
<td>Stative <em>have</em></td>
<td>18/28</td>
<td>64%</td>
<td>41/42</td>
<td>98%</td>
</tr>
<tr>
<td>Copular <em>be</em></td>
<td>12/20</td>
<td>60%</td>
<td>26/47</td>
<td>55%</td>
</tr>
<tr>
<td>Lexical verbs</td>
<td>20/76</td>
<td>26%</td>
<td>67/104</td>
<td>64%</td>
</tr>
</tbody>
</table>
Some frequent lexical verbs use *no*-negation

(18) no, Marilyn *does no* teaching I imagine she’s a research assistant

(19) I’ve *done nothing* except you know bring up a family since I left school

(20) I *know nothing* about his first wife
Prefabs used in writing

(21) the ballads *make no mention* of the trapping of rabbits

(22) *Make no mistake* about it, the divisions are very serious

(23) the split in the Conservative Party over Africa *gives me no joy*
Particular instances of constructions resist change

• Even after a construction has lost its productivity, specific exemplars of the construction may live on because they had accrued strength through repetition and so continue to be used.

• Evidence for exemplar representation of instances of constructions.
Parataxis to syntaxis
From discourse to syntax (Givón)

The development of relative clauses

Loosely adjoined clauses that share a referent become more consolidated into a mono-clausal construction.
Bambara (Bird 1968)

Chained (paratactic) configuration:

n ye ce min ye, o ye muru san.
I PAST man REL see, he PAST knife buy
‘The man that I saw, he bought the knife.’
(Hist.: ‘I saw that man, he bought the knife.’)

With REL-clause:

n ye ce min ye [ø] ye muru san.
I PAST man REL see [ø] PAST knife buy
‘The man that I saw bought the knife.’
(Hist.: ‘I saw that man, [he] bought the knife.’)

*min* is also used as a demonstrative (pronoun)
Givón (2015:211-212)

a. **Clause-chaining source** (Bambara, Hittite):
   You saw the man, **that/he** is my friend

b. **Finite non-RRC source** (German):
   The man – **that** (one) you saw – is my friend

c. **Nominalize non-RRCc source** (Ute):
   The man – of your [0] seeing – is my friend

d. **WH-complement source** (English):
   (i) **Paratactic source**: I don’t know. **Who** did it?
   (ii) **Syntactic V-complement**: I don’t know **who** did it.
   (iii) **Syntactic REL-clause**: I don’t know the man **who** did it.
Multiple source constructions converge on a similar construction-type.
Motivations

- **Communicative motivation.** The underlying communicative goal is that of using an event/state clause to identify a referent, thus a co-referent.

- **Available paratactic source.** The move from parataxis to syntaxis is a major component of the change from pre-grammar to grammar (Givón, 1979: ch. 5). Parataxis is the hallmark of the various pre-grammatical modes of communication: pidgins, Broca’s aphasia, and early child language (Givón, 2009).
Mechanisms

‘grammar acquires the characteristics of skilled performance’ (Givón 1979 On understanding grammar, chapter 5).

1. Complex hierarchic structure
2. Automated, speeded up production
3. Rigid rules