About pulling strings,??spilling beans, and *kicking buckets

Sascha Bargmann

in collaboration with Gert Webelhuth, Manfred Sailer, and Christopher Goetze

University of Frankfurt and Goettingen

Goettingen, January 2014

Some Facts About Idioms

Idioms ...

- are extremely pervasive in natural languages: English is estimated to contain at least 25,000

- are highly metaphorical/figurative/graphic.

- also have a literal meaning.

- are the playing field (i.e. the most creative area) of language.

- may but need not have a literal translation or equivalent in another language.

- are called proverbs if they make up an entire sentence and give advice.

What is an Idiom?

A Definition:

"An idiom is an expression larger than a word whose meaning cannot be systematically derived from meanings that the parts have when used independently of each other."

(Huddleston and Pullum 2002:273)

Prototypical Properties of Idioms:

- phrasal: multi-word expression
- idiomatic: non-literal and holistic meaning
- lexically fixed: none of the words can be replaced
- syntactically fixed: the idiom parts cannot be separated

An Example of a Prototypical Idiom

kick the bucket (= die):

- phrasal, i.e. multi-word expression: yes

- idiomatic, i.e. non-literal and holistic meaning: yes

- lexically fixed, i.e. none of the words can be replaced: yes:
  *kick the container [1]
  *throw the bucket
  *kick a bucket

- syntactically fixed, the idiom parts cannot be separated: yes:
  *The bucket has been kicked (by Tom).
  *The bucket appeared to have been kicked (by Tom).
  *The bucket, Tom has kicked.

[1] The asterisk here indicates that the idiomatic interpretation is unavailable.

Idioms in Principles & Parameters

Nunberg et al. (1994):

Assumptions:

- Idioms are inserted *en bloc* at Deep Structure.
- Transformations always and only apply at Deep Structure.

Predictions:

1. All idioms have a regular syntactic structure.

2. Idioms can either have:
   - a canonical form only or
   - a canonical and a non-canonical (= transformed) form,
   - but never a non-canonical (= transformed) form only.

3. Only the idiom as a whole has a meaning,
   whereas parts of idioms do not have a meaning.

Evidence Against Prediction 1

There are idioms that do not have a regular syntactic structure:

(1) a.   *trip the light fantastic* (= dance lightly)

      b.   *kingdom come* (= paradise)

      c.   *easy come, easy go* (= What you get easily, you lose easily.)

(Chafe 1968; Nunberg et al. 1994)

Evidence Against Prediction 2

There are idioms that only occur in non-canonical form:

(2) a. Passive only: *fit to be tied* (= angry and agitated)

b. Wh-moved only: *What the hell!* (= Why not!)

c. Yes/No-question only: *Is the Pope catholic?* (= Of course!)

d. Imperative only: *Break a leg!* (= Good luck!)

e. Tough-movement only: *to be hard to take* (= to be difficult to accept)

(Nunberg et al. 1994; Wasow et al. 1983)

Evidence Against Prediction 3

The parts of some idioms appear to carry meaning:

(3)   Internal Modification (Ernst, 1981):

    a.   Katz and I had by then become good friends, having long before buried the old hatchet (L. Melamed, *Escape to the Future*)

    b.   My girls should've buried the damn hatchet when they were in their prime. (www; expressive modifier)

    c.   Pat pulled some important strings. (= Pat used some important connections.)

(4)   Determiner variation:

    a.   Pat finally kicked the/*a bucket last night.

    b.   Pat pulled some/a lot of strings to get you the job.

☞   The existence of internal modification and determiner variation is strong evidence that idiom parts can be meaningful.

Interim Conclusion

The existence of:

- idioms with obligatory non-canonical sentence forms

- idioms with irregular syntactic structure

- idioms that allow for internal modification

shows that the en bloc insertion theory of idioms is incomplete at best.

Two Types of Idioms

Wasow et al. (1983) and Nunberg et al. (1994) distinguish
two types of idioms:

- **I**diomatic **Ph**rases (**IPh**s):
  kick the bucket, *saw logs* (sleep/snore), *trip the light fantastic*
  - ▶ Idiom parts cannot occur in positions/constructions that require content.
  - → less/no syntactic flexibility

- **I**diomatically **C**ombining **E**xpressions (**ICE**s):
  spill the beans, *bury the hatchet*, *keep tabs on somebody*
  - ▶ Idiom parts can occur in positions/constructions that require content.
  - → syntactic flexibility

  Nunberg et al. (1994) see a strong connection between semantic
  decomposability and syntactic flexibility.

Tests to Identify ICEs

If an idiom part can occur in a position/construction that has meaning, the idiom is semantically decomposable, i.e. an ICE.

- Internal modification by an adjective or relative clause:

  (5) Pat pulled some important strings.

  (6) Internal modification by a relative clause:
       Pat pulled the strings that got Chris the job
       *Pat kicked the bucket that nobody expected.

- Determiner change:

  (7) We've made some headway this year.

  (8) * Alex was sawing many logs last night.

Tests to Identify ICEs

- Fronting:

  (9)   The strings Pat has pulled.

  (10) * The bucket Pat has kicked.

- Pronominalization:

  (11)   Eventually they spilled the beans, but they didn't spill them deliberately.

  (12)   Kim's family pulled some strings on her behalf, but they weren't enough to get her the job. (Nunberg et al., 1994)

  (13) * Pat kicked the bucket and Chris kicked it too.

  (14) * Pat tripped the light fantastic but Alex didn't want to trip it.

Caveat

Decomposability is taken as a purely semantic notion!
It is **not** to be confused with:

- Transparency of the expression as a whole:

    - *saw logs* (transparent, non-decomposable)

    - *spill the beans* (non-transparent, decomposable)

    - *shoot the breeze* (non-transparent, non-decomposable)

- Paraphrasability:
  kick the bucket = *end one's life* (non-decomposable)

Our Analysis

Just like Horn (2003), we assume that there are not only 2, but 3 types of idioms. We distinguish the following types:

1. Syntactically frozen idioms (= IPhs) like *kick the bucket*

2. Mobile Idioms (= ICEs)
   a. Semantically and syntactically connected idioms like *spill the beans*
   b. Semantically connected idioms like *pull strings*

Our General Analytic Strategy

1. A syntactically frozen (i.e. *non*-decomposable) idiom is listed as *one single lexical entry* (i.e. as a completely fixed tree) in the *phrasal* lexicon.

2. A mobile (i.e. decomposable) idiom is composed of *separate lexical entries* that are listed in the *word* or *phrasal* lexicon and syntactically combined in the normal way.

   ▶ Each of these separate lexical entries requires that the linguistic representation of the clause or discourse containing their part of the idiom also contain a property that is only licensed if the other parts of the idiom are present in the structure. This property may be syntactic or semantic.

Type 1: Syntactically Frozen Idioms

Fraser (1970), p. 32:

(15) a. Your friend's [$_{VP}$ kicking the bucket] caused great concern.
     b. * Your friend's kicking of the bucket caused great concern.

(16) * The bucket was kicked.

Schenk (1995), p. 254:

(17) a. * The bucket Pete kicked.
     b. * It was the bucket that Pete kicked.
     c. * The bucket John kicked was astonishing.
     d. * Which bucket did John kick?

Wasow et al. (1980), Nunberg et al. (1994), Jackendoff (1995):

Syntactically invariable idiomatic expressions are analyzed in terms of surface(!) phrasal lexical entries: [$_{VP}$ kick+INFL the bucket]

Types 2a and 2b: Mobile Idioms

- The pieces of mobile idioms must be connected in the larger context containing them.

- Present-day grammatical theory contains a number of concepts that connect linguistic expressions through syntactic, semantic, or psycholinguistic relations that do not require the connected expressions to be linearly adjacent or to even occur within the same clause/sentence:

  1. Chains (A and A')
  2. Binding (e.g. variable binding, long distance reflexives)
  3. Scope (e.g. negative polarity items)
  4. Salience in context (pronominal anaphora, ellipsis)

Strategy: try to capture the different degrees of mobility of the pieces of idioms by imposing different syntactic and/or semantic connectedness conditions on them within the larger linguistic context containing them.

The LF-Account of Mobile Idioms

Two Central Assumptions:

1. Each part of a mobile idiom makes a unique contribution to the Logical Form (LF) of the larger linguistic context.

2. A part of an idiom may require the unique LF-contribution of the other parts of the idiom to be present in the LF of the larger linguistic context containing it.

Assumptions about LF:

- LF is a level of representation.
- The LF of a sign (word, phrase, sentence, . . .) is an expression of some semantic representation language.
- The LF is part of the representation of the sign, i.e. there can be mutual constraints on the syntactic form and LF of a sign.
- Discourse: We assume a DRT-like architecture in which
  - ▶ A semantic representation of the preceding discourse is available.
  - ▶ The LF of the current sentence is still set apart from that of the preceding discourse.
  - ▶ However, anaphoric relations have already been resolved.

Two Case Studies

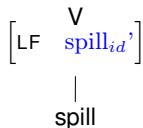Case Study 1: *spill the beans*

Case Study 2: *pull strings*

Case Study 1: *spill the beans*

Descriptive generalizations covering the empirical claims in the literature:

1. An NP of exactly the form *the beans* must be present.
2. The NP can undergo A-movement but not A'-movement.
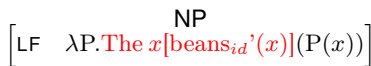3. Ellipsis of the verb is possible.
4. The NP can be pronominalized.

spill the beans – a Semantically and <u>Syntactically</u> Connected Idiom
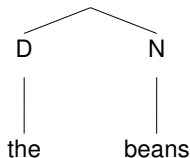
Lexical Entry	**Constraint**

$\begin{bmatrix} \text{LF} & \begin{matrix} \text{V} \\ \text{spill}_{id}\text{'} \\ | \\ \text{spill} \end{matrix} \end{bmatrix}$

In LF, the verb's idiomatic argument is specified by a term of the form $\text{The } x[\text{beans}_{id}\text{'}(x)]$ (possibly after anaphor resolution).
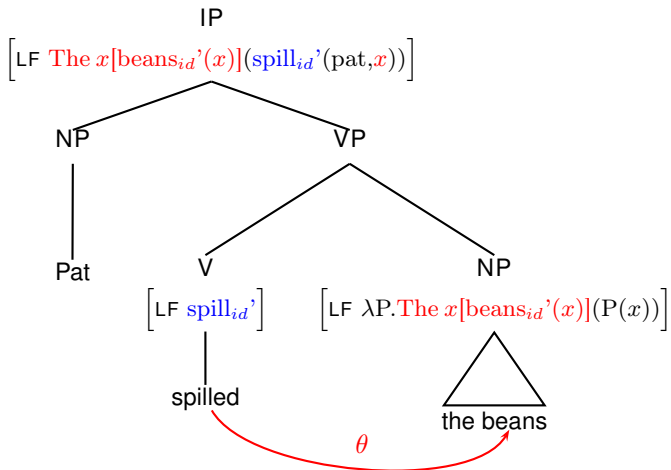
$\begin{bmatrix} \text{LF} & \lambda\text{P.The } x[\text{beans}_{id}\text{'}(x)](\text{P}(x)) \end{bmatrix}$

$\begin{matrix} & \text{NP} & \\ & \diagup \diagdown & \\ \text{D} & & \text{N} \\ | & & | \\ \text{the} & & \text{beans} \end{matrix}$

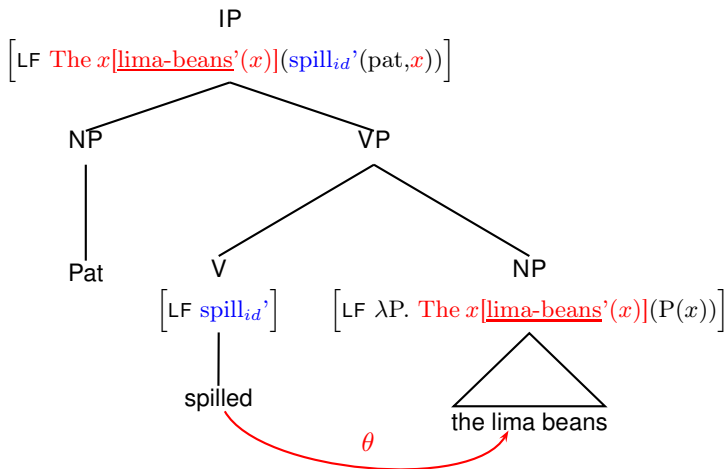The NP <u>heads an A-chain</u> theta-marked by a verb with the LF $\text{spill}_{id}\text{'}$.

☞ Crucial assumption: The LF-constants in these lexical entries are contributed solely by these lexical entries themselves or pronouns/ellipsis sites licensed by them.

Pat spilled the beans.

*Pat spilled the lima beans.

IP
$$\left[\text{LF The } x[\underline{\text{lima-beans'}}(x)](\text{spill}_{id}\text{'}(\text{pat},x))\right]$$

NP — VP

Pat

V
$$\left[\text{LF spill}_{id}\text{'}\right]$$

NP
$$\left[\text{LF } \lambda\text{P. The } x[\underline{\text{lima-beans'}}(x)](\text{P}(x))\right]$$
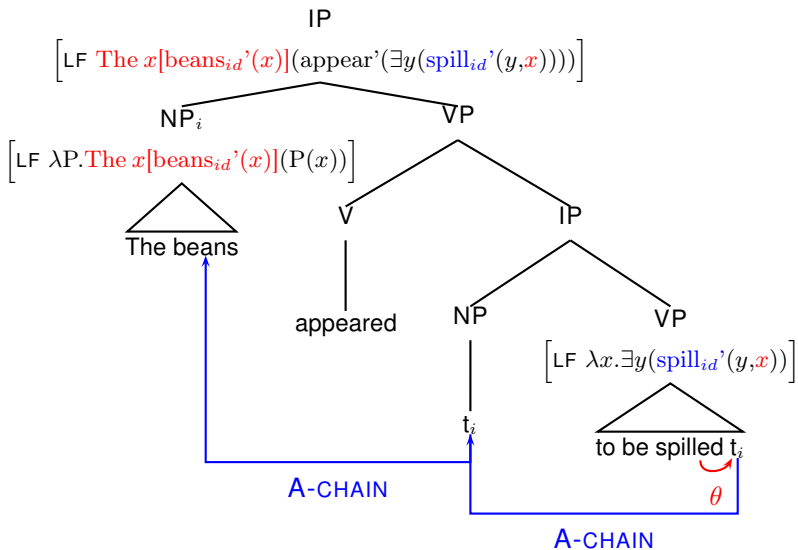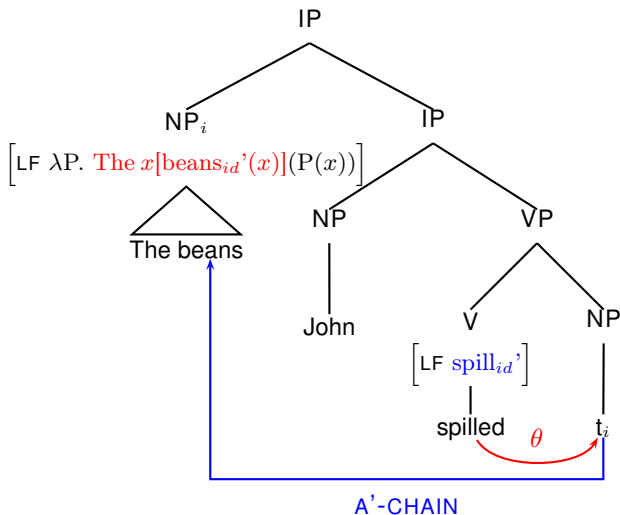
spilled

the lima beans

θ

☞ [NP the beans] is not present: hence the verb's idiomatic argument fails to be specified by The $x[\underline{\text{beans}_{id}}\text{'}(x)]$ in LF.

The beans appeared to be spilled.

The beans John spilled.

☞ [$_{NP}$ the beans] does not head an A-chain θ-marked by a verb with LF: spill_{id}'.

Which beans did Harry spill?
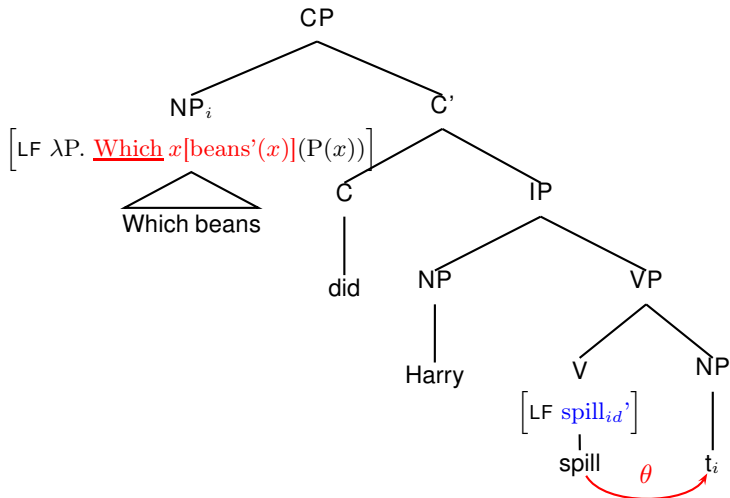
☞ [NP the beans] is not present: hence the verb's idiomatic argument fails to be specified by $\underline{\text{The}}\, x[\text{beans}_{id}\text{'}(x)]$ in LF.

*The beans that Joe spilled caused us a lot of trouble.

The Partee structure:

☞ [NP the beans] is not present: hence the verb's idiomatic argument fails to be specified by The x[beans$_{id}$'(x)] in LF.

The beans that Joe spilled caused us a lot of trouble.
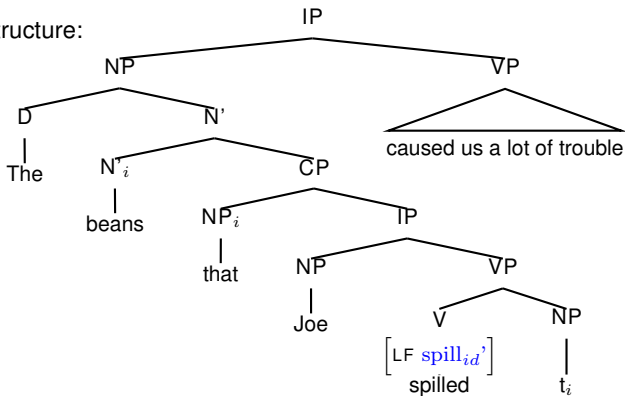
The Bach-Cooper structure:
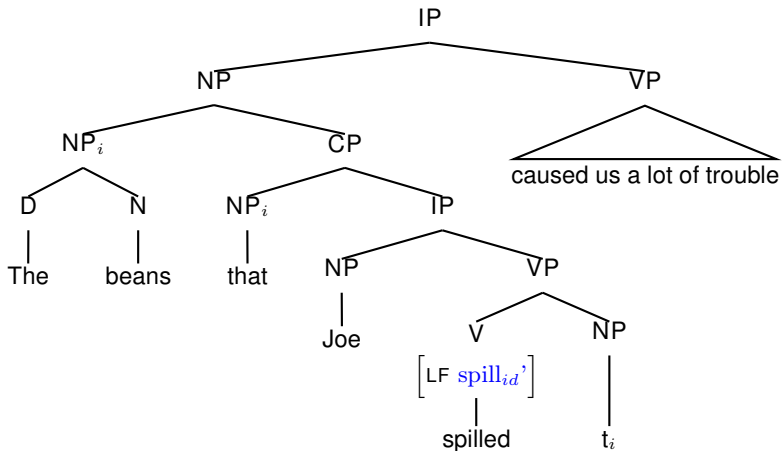
☞   [$_{NP}$ the beans] is present: but it fails to head an A-chain θ-marked by a verb with LF: spill_{id}'.

Assumptions on Pronouns

- LF account: A pronoun is interpreted as a definite NP whose restrictor is identical with that of its antecedent.

  (18)   a.   Preceding discourse: [A woman]$_i$ entered the room.
             Current sentence: She$_i$ whistled.

         b.   Preceding discourse: $\exists x[\text{woman'}(x)](\text{enter-room'}(x))$
             Current sentence: The $x[\text{woman'}(x)](\text{whistle'}(x))$

We worried that Pat might spill the beans, but it was Chris who finally spilled them.

We worried that Pat might spill [the beans]$_k$, but it was Chris

CP

$\left[\text{LF } \lambda y.\text{The } x[\text{beans}_{id}\text{'}(x)](\text{spill}_{id}\text{'}(y,x))] \right]$

NP

who$_i$

IP

NP

t$_i$

VP

V

$\left[\text{LF spill}_{id}\text{'} \right]$

spilled

NP$_k$

$\left[\text{LF } \lambda \text{P.The } x[\text{beans}_{id}\text{'}(x)](\text{P}(x)) \right]$

them
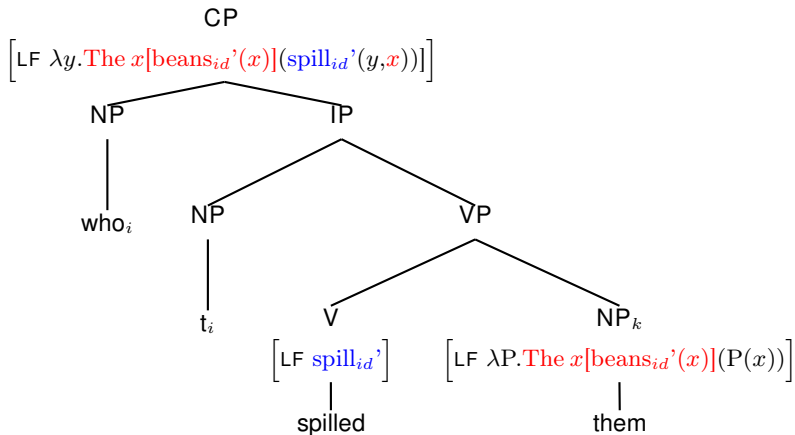
I was worried that the beans might be spilled, but they weren't.

I was worried that [the beans]$_i$ might be [spilled]$_k$,

LF: worried'(speaker',The x[beans$_{id}$'(x)]($\exists z$[spill$_{id}$'(z,x)]))

but

IP

$\Big[$LF The x[beans$_{id}$'(x)]($\neg \exists z$[spill$_{id}$'(z,x)])$\Big]$

NP$_i$

$\Big[$LF λP.The x[beans$_{id}$'(x)](P(x))$\Big]$

they

I'

I

$\Big[$LF λP$\lambda u.\neg$(P(u))$\Big]$

weren't

VP$_k$

$\Big[$LF $\lambda y.\exists z$[spill$_{id}$'(z,y)]$\Big]$

e
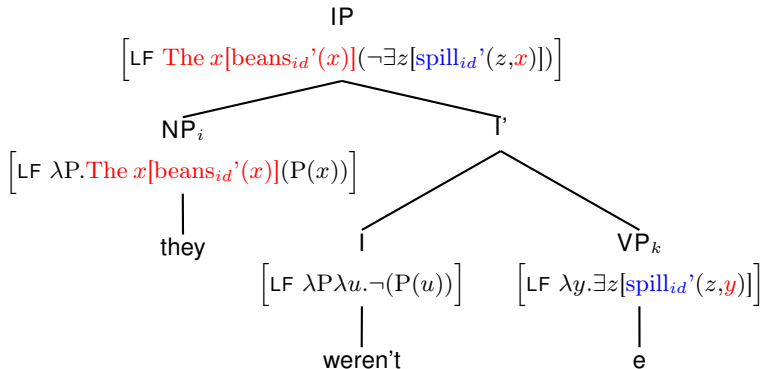
I was worried that the beans might be spilled, but they weren't.

- Neither idiomatic *spill* nor idiomatic *the beans* is present. Therefore, the constrains on them cannot be violated; they are vacuously fulfilled.

- For the interpretation of the sentence to be idiomatic, idiomatic *spill* and idiomatic *the beans* in the context sentence must have meanings that support anaphoric links, so that the pronoun *they* in the current sentence can refer to *the-beans$_{id}$'* and the ellipsis site can refer to *spill$_{id}$'*. This condition is fulfilled.

Case Study 2: *pull strings*

pull strings is more flexible than *spill the beans* in several respects:

1. The determiner of the NP headed by *strings* is not frozen.
2. The NP can undergo both A- and A'-movement.
3. *Strings* can occur in a main clause without *pull*, if another occurrence of *strings* did cooccur with *pull* in the preceding discourse.

pull strings

Illustration of the last claim:

Wasow et al. (1980), p. 93f

The authors claim that *strings* can occur without *pull* if the whole idiom has been introduced in the preceding discourse:

(19)   Pat and Chris graduated from law school together with roughly equal records. Pat's uncle is a state senator, and he pulled strings to get Pat a clerkship with a state supreme court justice. Chris, in contrast, didn't have access to any strings, and ended up hanging out a shingle.

pull strings – a Semantically Connected Idiom
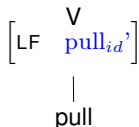
Lexical entries	Constraints
$\begin{bmatrix} & \text{V} \\ \text{LF} & \text{pull}_{id}\text{'} \end{bmatrix}$ $\vert$ pull	In LF, the verb's idiomatic argument is specified by a quantifier Qx that is restricted by $\text{strings}_{id}\text{'}(x)$ (possibly after anaphor resolution).

$\begin{bmatrix} & \text{N} \\ \text{LF} & \text{strings}_{id}\text{'} \end{bmatrix}$
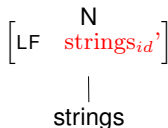 	
 $\vert$
 strings
	$\text{strings}_{id}\text{'}(x)$ restricts a quantifier Qx and either
	
	    a Qx binds the idiomatic argument of $\text{pull}_{id}\text{'}$ (possibly after anaphor resolution)
	
	      or
	
	    b $\text{strings}_{id}\text{'}$ is salient in the present discourse.

☞ Crucial assumption: The LF-constants in these lexical entries are contributed solely by these lexical entries themselves or pronouns/ellipsis sites licensed by them.

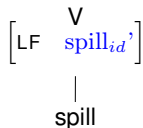# *spill the beans* – a Semantically and Syntactically Connected Idiom

Lexical Entry	Constraint

$$\left[\text{LF} \quad \begin{array}{c} \text{V} \\ \text{spill}_{id}\text{'} \end{array}\right]$$
$$|$$
$$\text{spill}$$

In LF, the verb's idiomatic argument is specified by a term of the form $\text{The } x[\text{beans}_{id}\text{'}(x)]$ (possibly after anaphor resolution).

$$\left[\text{LF} \quad \begin{array}{c} \text{NP} \\ \lambda\text{P.The } x[\text{beans}_{id}\text{'}(x)](\text{P}(x)) \end{array}\right]$$

The NP <u>heads an A-chain</u> theta-marked by a verb with the LF $\text{spill}_{id}\text{'}$.

```
      NP
     /  \
    D    N
    |    |
   the  beans
```
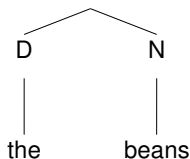
☞ Crucial assumption: The LF-constants in these lexical entries are contributed solely by these lexical entries themselves or pronouns/ellipsis sites licensed by them.
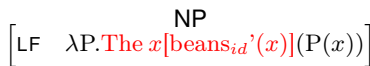
Kim's family pulled some strings on her behalf.

Tree diagram:

IP
$[\text{LF } \exists x[\text{strings}_{id}\text{'}(x)](\text{pull}_{id}\text{'}(\text{kim-family'},x))]$

- NP: Kim's family
- VP
  - VP
    - V: $[\text{LF } \text{pull}_{id}\text{'}(y,x)]$ — pulled
    - NP: $[\text{LF } \lambda P.\exists x[\text{strings}_{id}\text{'}(x)](P(x))]$
      - D: some
      - N: strings
  - PP: on her behalf
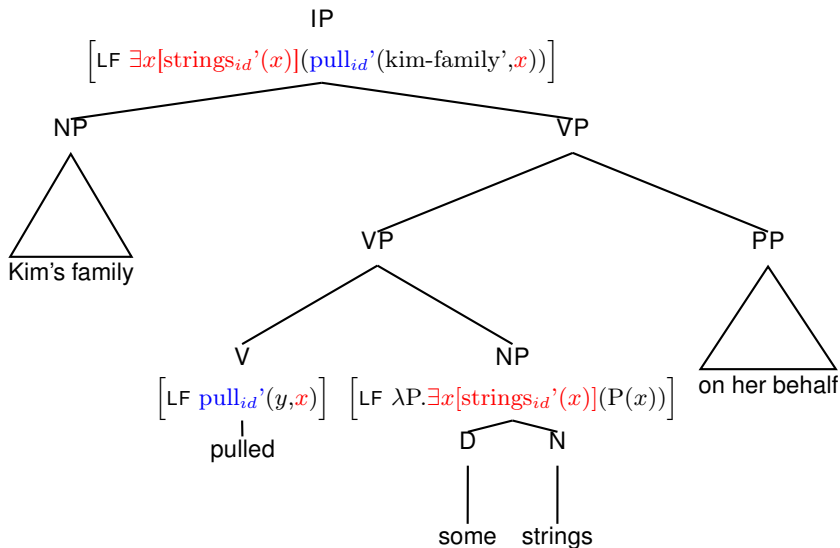
Strings seem to be pulled every time he applies for a promotion.

IP

$$\Big[\text{LF} \ldots \text{seem'}(\exists x[\text{strings}_{id}\text{'}(x)](\exists y(\text{pull}_{id}\text{'}(y,x))))\Big]$$

```
          IP
         /  \
      NP_i    VP
       |     /  \
    Strings V    IP
            |   /  \
          seem NP    VP
               |    /  \
              t_i  VP    NP
                   /\    /    \
      to be pulled t_i  every time he applies for a promotion
```
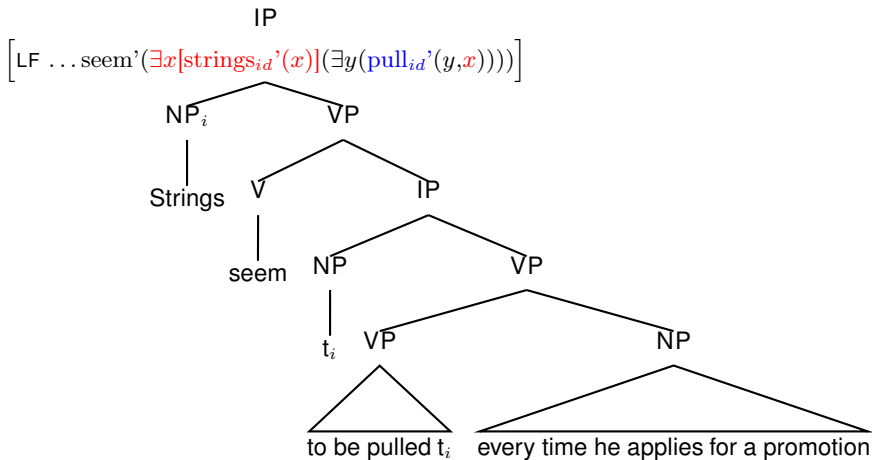
[*The strings that Pat* <u>*pulled*</u>] *helped Chris get the job.*

IP

$\Big[$ LF The $x[\text{strings}_{id}\text{'}(x)$ & $\text{pull}_{id}\text{'}(\text{pat'},x)](\text{help'}(x,\text{chris'},\text{get-the-promotion'})) \Big]$

NP

VP

D

N'

The

N'$_i$

CP

strings

helped Chris get the promotion

NP$_i$

IP

that

NP

VP

Pat

V

NP

pulled

t$_i$
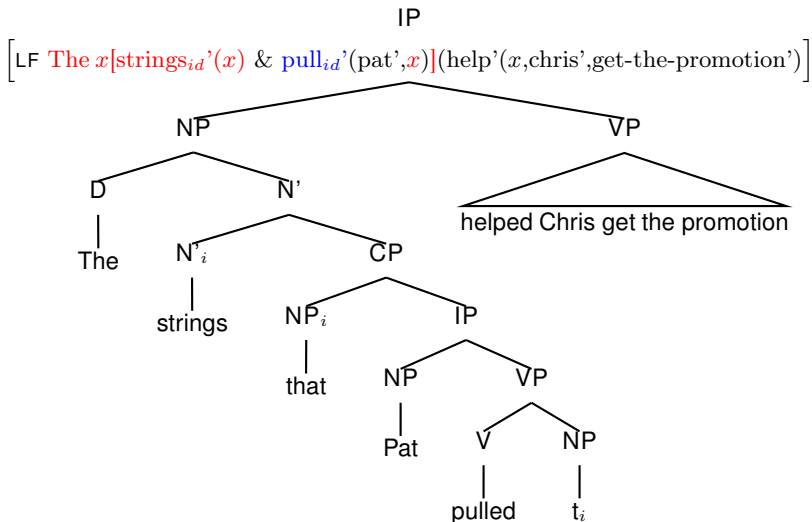
Pat _pulled_ [_strings that Chris had no access to_].

IP

$\Big[$LF $\exists x[\text{strings}_{id}\text{'}(x)$ & $\neg(\text{access'(chris',}x))](\text{pull}_{id}\text{'(pat,}x))\Big]$

- NP — Pat
- VP
  - V — pulled
  - NP
    - N'
      - N' — strings
      - CP
        - NP_i — that
        - IP
          - NP — Chris
          - VP — had no access to
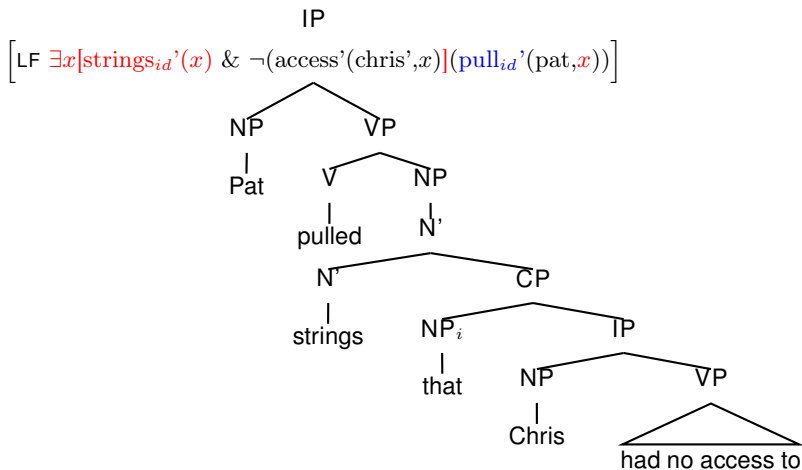
How many strings did Chris pull to get the promotion?

$$\left[\text{LF how-many } x[\text{strings}_{id}\text{'}(x)](\text{pull}_{id}\text{'}(\text{chris'},x) \ \& \ \text{purpose'}(\text{get'}(\text{chris'},\text{promotion'}))) \right]$$

CP

NP — C'

How many strings

C — IP

did

IP — CP

NP — VP

Chris

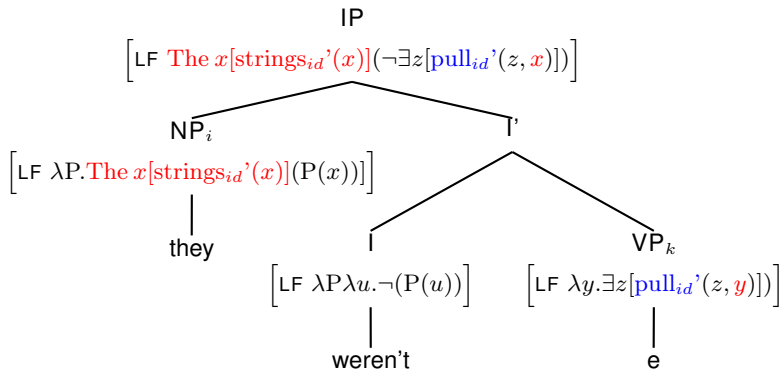to get the promotion

V — NP

pull — t_i
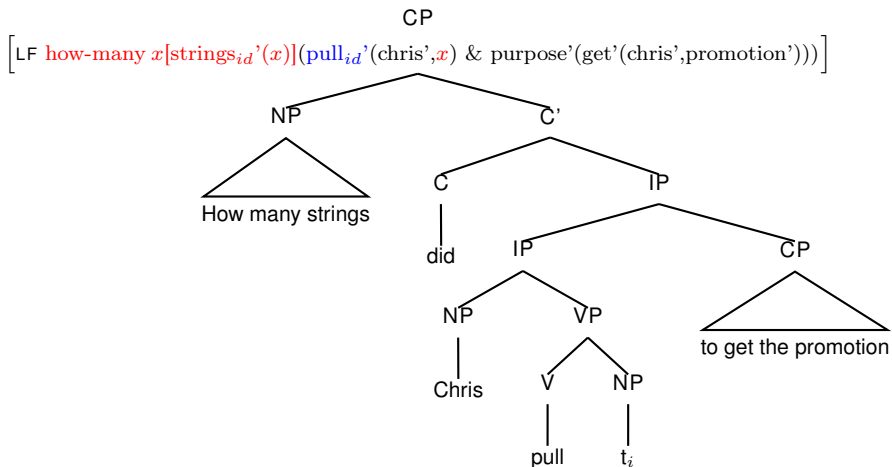
I was worried that strings might be pulled, but they weren't.

I was worried that [strings]$_i$ might be [pulled]$_k$,

LF: $\text{might}'(\exists x[\text{strings}_{id}'(x)](\exists y(\text{pull}_{id}'(y, x))))$

but

IP

$\Big[\text{LF The } x[\text{strings}_{id}\text{'}(x)](\neg\exists z[\text{pull}_{id}\text{'}(z, x)])\Big]$

NP$_i$

$\Big[\text{LF } \lambda P.\text{The } x[\text{strings}_{id}\text{'}(x)](P(x))]\Big]$

they

I'

I

$\Big[\text{LF } \lambda P\lambda u.\neg(P(u))\Big]$

weren't

VP$_k$

$\Big[\text{LF } \lambda y.\exists z[\text{pull}_{id}\text{'}(z, y)])\Big]$
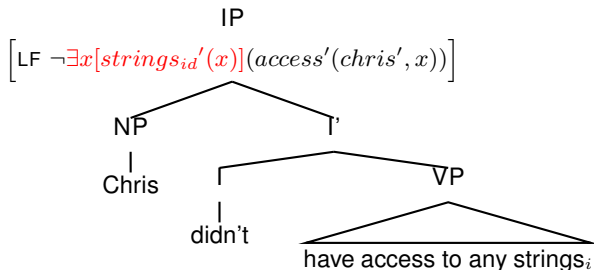
e

☞ Remember that a pronoun is interpreted as a definite NP whose restrictor is identical with that of its antecedent.

Chris didn't have access to any strings.

Pat and Chris graduated from law school together with roughly equal records. Pat's uncle is a state senator, and he pulled strings$_i$ to get Pat a clerkship with a state supreme court justice.

Context LF: $\dots \text{strings}_{id}' \dots$

$$\text{IP} \quad \left[\text{LF} \ \neg \exists x [strings_{id}'(x)](access'(chris', x)) \right]$$

NP — Chris

I'

didn't

VP

have access to any strings$_i$

Summary and Conclusion

1. Idioms differ from each other in their degree of syntactic flexibility.
2. The en bloc insertion theory of idioms is thus insufficient, as it needs to be predicted which set of surface forms a given idiom can occur in.
3. Syntactically frozen idioms are treated as entries in the phrasal lexicon (which, however, must specify surface entries.)
4. Syntactically mobile idioms (ICEs) differ from each other in how their parts are linguistically connected.
5. I have presented an account in which:
    1. The parts of mobile idioms each have their own lexical entry.
    2. These lexical entries make reference to some syntactic or semantic property of the other parts of the idiom.
6. This approach predicts a hierarchy of idioms in terms of the syntactic mobility of their parts:

phrasal lexical entry < syntactically connected < semantically connected.

References I

Chafe, Wallace (1968). Idiomaticity as an Anomaly in the Chomskyan Paradigm. *Foundations of Language 4*, 109–127.

Ernst, Thomas (1981). Grist for the Linguistic Mill: Idioms and 'Extra' Adjectives. *Journal of Linguistic Research 1*, 51–68.

Nunberg, Geoffrey, Sag, Ivan A., and Wasow, Thomas (1994). Idioms. *Language 70*, 491–538.

Wasow, Thomas, Sag, Ivan A., and Nunberg, Geoffrey (1983). Idioms: An Interim Report. In S. Hattori and K. Inoue (Eds.), *Proceedings of the XIIIth International Congress of Linguistics*, pp. 102–115.