On explanation in syntax: Evidence for abstract syntactic structures

Jason Merchant
Department of Linguistics, University of Chicago
June 2013
Types of explanation in linguistics:

(1) Socio-historical
(2) Functional
(3) Formal
Socio-historical:

(4) why is the plural of ‘foot’ = ‘feet’ and not ‘foots’?

(5) PIE *ped- (cf. Latin pes, pedis; Greek pous, podos) > ProtoGmc *fot+i *foti > *föti > *föt > OE *feet, Modern Eng feet

(cf. doom/deem; old/elder; brother/brethren; fall/fell; book/*beech>books)
Functional:

(6)  a. Why is there no nasalized glottal stop?
     b. Can’t close the glottis and still have airflow through the nose

(7)  a. Universal: [m] → [n]
     b. dental/alveolar place of articulation is less marked than bilabial
Formal:

(8) **syntax**

a. \( S \rightarrow NP \ VP \)

b. \( NP[sg] \rightarrow (Det) (AP^* \ N[sg] \ (PP^*) \ (CP^*) \)

(9) **phonology**

a. \( dV[-hi,\ -rd] \rightarrow V[\alpha_{back}] / V[\alpha_{back}]C^*\)

b. Turkish: karpuz + dV = karpuzda; ev + dV = evde

(10) **semantics**

a. \[ more/-er \] = \( \lambda d \lambda g_{<d,et>} \lambda x \{ max\{d'|g(d')(x)\} > d \} \)
One central discovery from syntactic studies of human languages: the properties of sentences cannot be modeled without abstract structures; mere strings of words (WYSIWYG or ‘surfacist’ approaches) do not suffice.

Today’s goal:
Review some of the evidence for this; present new evidence for phonologically null syntactic structures; argue against naïve (and sophisticated) surfacism
Outline

1. The claim: Sentences have abstract structure
2. The evidence: Structural ambiguities
3. The evidence: Ellipsis
   - Ellipsis and agreement
   - Ellipsis and case
   - Ellipsis and preposition-stranding
   - Ellipsis and priming
4. Conclusion: Sentences have abstract structure
Some definitions

(11) A grammar $G$ consists of a pair of a set of lexical elements $L$ and a set of operations $O$:

$$G = \langle L, O \rangle$$

(12) A derivation on a numeration $D_N$ is a pair:

$$D_N = \langle N, < PM_1, ..., PM_n > \rangle,$$ where

- a. $N$, called the Numeration, is a nonempty set of lexical elements drawn from $L$ and a possibly empty set $S$ of phrase markers $PM$ (each of which is itself the result of a separate convergent or semi-convergent derivation), and

- b. $< PM_1, ..., PM_n >$ is an ordered n-tuple of phrase markers $PM$. 
Some definitions

(13) A grammar $G$ consists of a pair of a set of lexical elements $L$ and a set of operations $O$:
$$G = \langle L, O \rangle$$

(14) A derivation on a numeration $D_N$ is a pair:
$$D_N = \langle N, \langle PM_1, \ldots, PM_n \rangle \rangle,$$ where

a. $N$, called the Numeration, is a nonempty set of lexical elements drawn from $L$ and a possibly empty set $S$ of phrase markers $PM$ (each of which is itself the result of a separate convergent or semi-convergent derivation), and

b. $\langle PM_1, \ldots, PM_n \rangle$ is an ordered n-tuple of phrase markers $PM$. 
Some definitions

(15) A derivation $D_N$ is said to be convergent (or to converge) iff
    a. $PM_n$ contains no unvalued (:__) features
    b. $PM_n$ contains no unchecked strong (*) features
    c. $PM_n$ contains no unchecked selectional features
    d. All elements in the Numeration have been Merged
    e. For each adjacent pair of phrase markers $< PM_k, PM_{k+1} >$ in $D_N$, there is an operation $\omega \in O$ such that $\omega$ applied to $PM_k$ yields $PM_{k+1}$.

(16) A derivation $D_N$ is semi-convergent iff it satisfies conditions b-e of (15).
Some definitions

(17) A phrase $P$ (including a sentence) is \textit{well-formed} iff there is at least one convergent derivation for $P$.

(18) The Minimalist Program, in essence = $\text{min}|O|$ (Minimize the number of operations in O).
Some definitions

(19) A phrase $P$ (including a sentence) is well-formed iff there is at least one convergent derivation for $P$.

(20) The Minimalist Program, in essence $= min|O|$
(Minimize the number of operations in $O$).
A lexical item $LI$ has the following feature structure, with categorial, inflectional (or morphological), and selectional feature arrays:

$$LI = \begin{bmatrix}
\text{CAT} [...] \\
\text{INFL} [...] \\
\text{SEL} [...]
\end{bmatrix}$$
Some examples:

(21) \textit{dog} \begin{align*}
\left[ & \begin{array}{c}
\text{CAT} \ [N, \phi : 3sm] \\
\text{INFL} \ [\text{Case:__}] \\
\text{SEL} \ [ ] \\
\end{array} \right] \\
\end{align*}

(22) \textit{see} \begin{align*}
\left[ & \begin{array}{c}
\text{CAT} \ [V, -aux] \\
\text{INFL} \ [ ] \\
\text{SEL} \ [D] \\
\end{array} \right] \\
\end{align*}
Some examples:

(23)  \[ v_{trans} \]

\[
\begin{bmatrix}
\text{CAT} & [v, -aux] \\
\phi : \_
\end{bmatrix}
\begin{bmatrix}
\text{INFL} & V^* \\
\text{Infl:}\_
\end{bmatrix}
\begin{bmatrix}
\text{SEL} & < V, D > \\
\text{Case:ACC}
\end{bmatrix}
\]

(24)  \[ T_{pres} \]

\[
\begin{bmatrix}
\text{CAT} & [T, +fin, Tns:pres] \\
\phi : \_
\end{bmatrix}
\begin{bmatrix}
\text{INFL} & \text{Case:NOM}
\end{bmatrix}
\begin{bmatrix}
\text{SEL} & v
\end{bmatrix}
\]
Operations

(25) **Merge**($\alpha, \beta$):
For any syntactic objects $\alpha, \beta$, where $\alpha$ bears an unchecked selectional feature $F$, and $\beta$ bears a matching categorial feature $F'$, call $\alpha$ the head and

a. let $\alpha = \{ \gamma, \{ \alpha, \beta \} \}$
call $\gamma$ the label (or projection) and

b. let $F$ be checked (written $<F>$), and

c. let $\gamma = \alpha \cap \mathcal{I}$, where $\mathcal{I}$ is the set of all unchecked non-inflectional features
(In other words, all category features project, all unchecked selectional features project, and no inflectional features project. Inflectional features are therefore found only on heads, never on projections.)
Operations

(26) **Adjoin**($\alpha$, $\beta$):
For any syntactic objects $\alpha$, $\beta$, where neither $\alpha$ nor $\beta$ has any unchecked selectional feature, call $\alpha$ the host, and
a. let $\alpha = \{ \gamma, \{ \alpha, \beta \} \}$
call $\gamma$ the label (or projection) and
b. let $\gamma = \alpha$
Operations

(27) **Agree**(X,Y; F) (read: ‘X triggers agreement on Y with respect to F’ or ‘Y agrees with X in F’ or ‘X controls agreement on target Y for F’)

For any syntactic objects X and Y, where X bears a feature F with value Val(F) and Y bears a matching unvalued inflectional feature F':__ (that is, Val(F') = ∅), and either X c-commands Y or Y c-commands X,

a. let Val(F') = Val(F)
Operations

(28) \textbf{Move}_{phrase}(Y, X) \text{ (read: ‘Y moves to specXP’)}
If X is a projection with a feature F, Y a maximal projection with a matching feature F, and X contains Y, and F is strong (marked F*) on X or Y or both, then
a. let X = \{X, \{Y, X\}\} and
b. let all occurrences of F* = F<*> , and
c. let Y = <Y>
Three **major** syntactic phenomena have been factored out of the above definitions and must be added to the system to make it account for word order and other important syntactic facts:

(29) Linearization (an algorithm or principle to determine the linear order of any two sister nodes)

(30) Locality of application (Relativized Minimality)

(31) The spellout of complex heads by the Morphology

(Also, Merge must be made sensitive to the order of the selectional features, if there is more than one; either we define different kinds of features, e.g., COMP and SPEC features, and declare that COMP features must be checked (trigger Merge) before SPEC features, or, equivalently, the definition of Merge must be altered to be sensitive to an ordered n-tuple (list) of selectional features.)
(32) The big truck’s on the street

(33)

Phonological structure:
(A representation of the sounds and their groupings)

Prosodic structure
Syllabic structure
Segmental structure

• This is still an over simplification, because each segment itself consists of a complex bundle of features/articulatory gestures:
  /s/ = [+consonantal, -vocalic, -sonorant, -nasal, +continuant, -voiced, +coronal, +anterior]
Two hypotheses about syntax:
Hypothesis 1: Syntax as beads on a string
- The syntactic component of (mental) grammar consists of a set of successor functions:

For each word \( w \), the grammar associates with \( w \) the output of a function \( S, S(w) \), which is the set of possible successor words to \( w \) (plus some mechanism for choosing an element from that set)

(34) Example: \( S(\text{big}) = \{\text{truck}, \text{tree}, \text{nun}, \text{trouble}, \text{aardvark}, \text{lie}, \ldots\} \)
Two hypotheses about syntax:

Hypothesis 1: Syntax as beads on a string

- The syntactic component of (mental) grammar consists of a set of successor functions:

  For each word \( w \), the grammar associates with \( w \) a the output of a function \( S, S(w) \), which is the set of possible successor words to \( w \) (plus some mechanism for choosing an element from that set)

(35) Example: \( S(\text{big}) = \{\text{truck, tree, nun, trouble, aardvark, lie, \ldots}\} \)
Two hypotheses about syntax:
Hypothesis 1: Syntax as **beads on a string**
- The syntactic component of (mental) grammar consists of a set of successor functions:

\[ \text{For each word } w, \text{ the grammar associates with } w \text{ a the output of a function } S, S(w), \text{ which is the set of possible successor words to } w \text{ (plus some mechanism for choosing an element from that set)} \]

(36) Example: \( S(\text{big}) = \{ \text{truck, tree, nun, trouble, aardvark, lie, . . .} \} \)
Hypothesis 1: Syntax as **beads on a string**
This is equivalent to a kind of **finite state automaton**:
This approach works well for some applications:

**Google n-gram** corpus (books.google.com/ngrams)
This approach works well for some applications:

**Google n-gram** corpus (books.google.com/ngrams)
The most extremist power any political leader can assert is the power to target his own citizens for execution without any charges or due process, far from any battlefield. The Obama administration has not only asserted exactly that power in theory, but has exercised it in practice.

But who would believe this is all there is to it? Well:

How hierarchical is language use?
Stefan L. Frank¹,*, Rens Bod² and Morten H. Christansen³

¹Department of Cognitive, Perceptual and Brain Sciences, University College London, 26 Bedford Way, London WC1H 0AP, UK
²Institute for Logic, Language and Information, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands
³Department of Psychology, Cornell University, 228 Uris Hall, Ithaca, NY 14853-7601, USA

It is generally assumed that hierarchical phrase structure plays a central role in human language. However, considerations of simplicity and evolutionary continuity suggest that hierarchical structure should not be invoked too hastily. Indeed, recent neurophysiological, behavioural and computational studies show that sequential sentence structure has considerable explanatory power and that hierarchical processing is often not involved. In this paper, we review evidence from the recent literature supporting the hypothesis that sequential structure may be fundamental to the comprehension, production and acquisition of human language. Moreover, we provide a preliminary sketch outlining a non-hierarchical model of language use and discuss its implications and testable predictions. If linguistic phenomena can be explained by sequential rather than hierarchical structure, this will have considerable impact in a wide range of fields, such as linguistics, ethology, cognitive neuroscience, psychology and computer science.

Keywords: language structure; language evolution; cognitive neuroscience; psycholinguistics; computational linguistics
Problem 1 for Hypothesis 1:
Repeated elements:
(1) The guy who said **he was great** wouldn’t listen to Abby.
(2) The guy who said **he was great** wouldn’t listen to anyone who didn’t think **he was great**.
(3) *The guy who said **he was great** wouldn’t listen to anyone who didn’t think **he was great** wouldn’t listen to Abby.*
Problem 1 for Hypothesis 1:
Repeated elements:
(1) The guy who said he was great wouldn’t listen to Abby.
(2) The guy who said he was great wouldn’t listen to anyone who didn’t think he was great.
(3) *The guy who said he was great wouldn’t listen to anyone who didn’t think he was great wouldn’t listen to Abby.
Problem 1 for Hypothesis 1:
Repeated elements:
(1) The guy who said he was great wouldn’t listen to Abby.
(2) The guy who said he was great wouldn’t listen to anyone who didn’t think he was great.
(3) *The guy who said he was great wouldn’t listen to anyone who didn’t think he was great wouldn’t listen to Abby.
Problem 2 for Hypothesis 1:
Structural ambiguity...
Words with multiple meanings give rise to sentences with multiple meanings:

*John has a* bat. *Abby went to the* bank.

- **Lexical ambiguity**
But this is not the source of ambiguity in some sentences:
(37) Susan saw the man with a telescope.
   a. ◊ = Susan saw the man who was carrying a telescope.
   b. ◊ = Susan looked through the telescope and saw the man.

(38) The man in the chair with a broken leg is Joe.
   a. ◊ = The man’s leg is broken.
   b. ◊ = The chair’s leg is broken.

(39) The man with a broken leg in the chair is Joe.
   a. ◊ = The man’s leg is broken.
   b. ⊳ The chair’s leg is broken.
   (◊ = The leg is in the chair.)
(40) Susan saw the man with a telescope.
   a. ◊ = Susan saw the man who was carrying a telescope.
   b. ◊ = Susan looked through the telescope and saw the man.

(41) The man in the chair with a broken leg is Joe.
   a. ◊ = The man’s leg is broken.
   b. ◊ = The chair’s leg is broken.

(42) The man with a broken leg in the chair is Joe.
   a. ◊ = The man’s leg is broken.
   b. \(\neq\) The chair’s leg is broken.
   (◊ = The leg is in the chair.)
(43) Susan saw the man with a telescope.
   a. ◊ = Susan saw the man who was carrying a telescope.
   b. ◊ = Susan looked through the telescope and saw the man.

(44) The man in the chair with a broken leg is Joe.
   a. ◊ = The man’s leg is broken.
   b. ◊ = The chair’s leg is broken.

(45) The man with a broken leg in the chair is Joe.
   a. ◊ = The man’s leg is broken.
   b. ≠ The chair’s leg is broken.
      (◊ = The leg is in the chair.)
Hypothesis 2: Syntax involves abstract structures

- The syntactic component of (mental) grammar consists of a finite set of phrase (or constituent) structure rules (‘grouping laws’) for combining words into phrases, and such phrases into larger phrases.

For each phrase $P$, the grammar has a phrase structure rule $PSR_P$ that determines what kind of categories or phrases occur in $P$ and determines their relative position in $P$.

(46) The big truck’s on the street.

(47) PS-‘rules’ (equations):
   a. $S = NP$ VP
   b. $NP = Art A N$
   c. $VP = V$ PP
   d. $PP = P$ NP
Hypothesis 2: Syntax involves **abstract structures**

- The syntactic component of (mental) grammar consists of a finite set of **phrase (or constituent) structure rules** (‘grouping laws’) for combining words into phrases, and such phrases into larger phrases

*For each phrase $P$, the grammar has a phrase structure rule $PSR_P$ that determines what kind of categories or phrases occur in $P$ and determines their relative position in $P*  

(48) The big truck’s on the street.

(49) PS-‘rules’ (equations):
   a. $S = NP \ VP$
   b. $NP = Art \ A \ N$
   c. $VP = V \ PP$
   d. $PP = P \ NP$
The claim

The evidence: Structural ambiguities

The evidence: Ellipsis

Conclusion

S

NP

VP

Art

A

N

V

PP

the

big

truck

's

P

NP

on

Art

N

the

street
Two structures for *Susan saw the man with a telescope*:
Two structures for *Susan saw the man with a telescope*:
Moving the boxed elements can lead to disambiguation:

(50) With a telescope, Susan saw the man.
(51) The man, Susan saw with a telescope.
(52) The man with a telescope, Susan saw.
Moving the boxed elements can lead to disambiguation:

(53) With a telescope, Susan saw the man.
(54) The man, Susan saw with a telescope.
(55) The man with a telescope, Susan saw.
(56) Competent women and men got jobs.

(57) a. [Competent women] and [men] got jobs. (‘unfair world’)
b. Competent [women and men] got jobs. (‘fair world’)

(58) a. NP = A NP
b. NP = N
c. $\alpha = \alpha \text{ Conj } \alpha$
d. Conj = \{and, \ldots\}

- Structural ambiguity
(59) Competent women and men got jobs.

(60) a. [Competent women] and [men] got jobs. (‘unfair world’)
b. Competent [women and men] got jobs. (‘fair world’)

(61) a. NP = A NP
b. NP = N
c. $\alpha = \alpha \text{ Conj } \alpha$
d. Conj = \{and, \ldots\}

Structural ambiguity
(62) Competent women and men got jobs.

(63) a. [Competent women] and [men] got jobs. (‘unfair world’)
    b. Competent [women and men] got jobs. (‘fair world’)

(64) a. NP = A NP
    b. NP = N
    c. \( \alpha = \alpha \text{ Conj } \alpha \)
    d. Conj = \{and, \ldots\}

- Structural ambiguity

this crime covers anyone who intentionally accesses a federal computer without authorization, and by means of one or more instances of such conduct alters, damages, or destroys information

*this crime covers anyone who intentionally accesses a federal computer without authorization, and by means of one or more instances of such conduct alters, damages, or destroys information*

1 intentionally [access ... and by means of ...] (defendant)

2 [ intentionally access...] and [ by means of ... ] (plaintiff)
Conclusion:
Abstract structure (phrase structure) can explain certain types of ambiguity.
A different kind of abstract structure: phrase structures that have no phonology. Strings of words that are not sentences can have sentential meaning:

(65) Bill should collect butterflies. Jill should, too.

(66) Bill should collect butterflies. Jill should collect butterflies, too.

How can *Jill should* mean *Jill should collect butterflies*?
Analysis 1: Deletion
Full sentence structure, but part of the sentence is unpronounced.

The missing words are not really missing.
Analysis 2: WYSIWYG structure

The missing words are really missing.

S
   NP   Aux
     Jill  should

Context fills in the missing parts of the meaning.
If the Deletion analysis is correct, elliptical material has abstract structure, but no pronunciation.

(67) Four domains of evidence:
   a. Agreement
   b. Case
   c. Preposition-stranding
   d. Syntactic priming
Subject-verb agreement is a syntactic phenomenon; agreement is not (always) about meaning:

(68) Beth’s wedding was in Bond Chapel, and Rachel’s wedding was in Rockefeller Chapel.

(69) Beth’s nuptials were in Bond Chapel, and Rachel’s nuptials were in Rockefeller Chapel.

(70) *Beth’s wedding was in Bond Chapel, and Rachel’s wedding were in Rockefeller Chapel.

(71) *Beth’s nuptials were in Bond Chapel, and Rachel’s nuptials was in Rockefeller Chapel.
Subject-verb agreement is a syntactic phenomenon; agreement is not (always) about meaning:

(72) Beth’s wedding was in Bond Chapel, and Rachel’s wedding was in Rockefeller Chapel.

(73) Beth’s nuptials were in Bond Chapel, and Rachel’s nuptials were in Rockefeller Chapel.

(74) *Beth’s wedding was in Bond Chapel, and Rachel’s wedding were in Rockefeller Chapel.

(75) *Beth’s nuptials were in Bond Chapel, and Rachel’s nuptials was in Rockefeller Chapel.
Noun ellipsis preserves the syntactic properties of agreement:

(76) Beth’s wedding was in Bond Chapel, and Rachel’s was in Rockefeller Chapel.

(77) Beth’s nuptials were in Bond Chapel, and Rachel’s were in Rockefeller Chapel.

(78) *Beth’s wedding was in Bond Chapel, and Rachel’s were in Rockefeller Chapel.

(79) *Beth’s nuptials were in Bond Chapel, and Rachel’s was in Rockefeller Chapel.
Noun ellipsis preserves the syntactic properties of agreement:

(80) Beth’s wedding was in Bond Chapel, and Rachel’s was in Rockefeller Chapel.

(81) Beth’s nuptials were in Bond Chapel, and Rachel’s were in Rockefeller Chapel.

(82) *Beth’s wedding was in Bond Chapel, and Rachel’s were in Rockefeller Chapel.

(83) *Beth’s nuptials were in Bond Chapel, and Rachel’s was in Rockefeller Chapel.
Agreement is sensitive to abstract structure (the unpronounced head N, =nuptials):

```
S
  /\  /
/   /NP VP
/  /   /
/ /    /
/ /     /
/ Possessor N V PP
/ Rachel’s were in Rockefeller Chapel
```
Postverbal agreement in *there*-sentences:

*On Saturday, there was a wedding in Bond Chapel.*
*On Saturday, there were two weddings in Bond Chapel.*

*On Saturday, there were a wedding in Bond Chapel.*
*On Saturday, there was two weddings in Bond Chapel.*
Postverbal agreement in *there*-sentences:

*On Saturday, there was a wedding in Bond Chapel.*

*On Saturday, there were two weddings in Bond Chapel.*

*On Saturday, there were a wedding in Bond Chapel.*

*On Saturday, there was two weddings in Bond Chapel.*
Agreement is sensitive to the elided noun phrase:

*On Saturday, there was a wedding in Bond Chapel, but on Sunday, there wasn’t.*

*On Saturday, there were two weddings in Bond Chapel, but on Sunday, there weren’t.*
Agreement is sensitive to the elided noun phrase:

*On Saturday, there was a wedding in Bond Chapel, but on Sunday, there wasn’t.*

*On Saturday, there were two weddings in Bond Chapel, but on Sunday, there weren’t.*
**Ellipsis and case:** English pronouns have case...

- **Subject**
  
  \( \{I/*Me\} \text{ saw John.} \)

- **Object**
  
  \( \text{John saw } \{me/*I\}. \)

...but other nouns do not: there’s only one \textit{John}
Ellipsis and case: English pronouns have case...

- Subject
  
  \{I/*Me\} saw John.

- Object
  
  John saw \{me/*I\}.

...but other nouns do not: there’s only one John
A Basque sentence:

- Peru-k Jone ikusi zuen.
  - *Peru-ERG Jone-ABS seen had*
  - ‘Peru saw Jone.’

All nouns have case:

- **Subject is ergative:** -k (ergative)
- **Object is absolutive:** no marking

- Jone-k Peru ikusi zuen.
  - *Jone-ERG Peru-ABS seen had*
  - ‘Jone saw Peru.’
A Basque sentence:

- Peru-k Jone ikusi zuen.
  *Peru-ERG Jone-ABS seen had*
  ‘Peru saw Jone.’

All nouns have case:

- **Subject** is **ergative**: -k (ergative)
- **Object** is **absolutive**: no marking

- Jone-k Peru ikusi zuen.
  *Jone-ERG Peru-ABS seen had*
  ‘Jone saw Peru.’
Case is about structure:

```
S
  NP     VP
    |    |   |
    ergative    NP   V
      absolutive
```

Jone-k Peru ikusi zuen.

*Jone-ERG Peru-ABS seen had*

‘Jone saw Peru.’
It’s not about meaning:

Jone-\textit{k} \textbf{hitz-\textit{egin}} zuen.
\textit{Jone-ERG talked had}

‘Jone talked.’

\[
\begin{tikzpicture}
  \node {S}
  \child {node {NP}}
  \child {node {VP}}
  \child {node {ergative}}
  \child {node {NP}}
  \child {node {V}}
  \child {node {hitz-}}
  \child {node {egin}}
\end{tikzpicture}
\]
It’s not about meaning:

Jone-<b>k</b> <b>hitz-egin</b> zuen.  
*Jone-ERG talked had*  
‘Jone talked.’

\[
\text{S} \\
\text{NP} \quad \text{VP} \\
\text{ergative} \quad \text{NP} \quad \text{V} \\
\text{hitz-} \quad \text{-egin}
\]

\[
\text{S} \\
\text{NP} \quad \text{VP} \\
\text{absolutive} \quad \text{V} \\
\text{mintzatu}
\]

Jone <b>mintzatu</b> zen.  
*Jone-ABS talked was*  
‘Jone talked.’
Ellipsis in wh-questions (sluicing):

*I know John saw someone, but I don’t know who.*

= 

*I know John saw someone, but I don’t know who *John* saw.*
Ellipsis in wh-questions (sluicing):

*I know John saw someone, but I don’t know who.*

=  

*I know John saw someone, but I don’t know who John saw.*

Deletion analysis:

WYSIWYG analysis:
Wh-question ellipsis (sluicing) in Basque:

- Badakit norbaite-k **hitz-egin** zuela, baina ez dakit nor-k.  
  *I.know someone-ERG talked had but not I.know who-ERG*
  
  ‘I know that someone talked, but I don’t know who.’

- Badakit norbait **mintzatu** zela, baina ez dakit nor.  
  *I.know someone-ABS talked was but not I.know who-ABS*
  
  ‘I know that someone talked, but I don’t know who.’
The case of the subject is determined by the deleted verb:

*hitz-egin*: ergative

```
S
  NP    VP
    nor-k NP    V
      hitz-    -egin
```

*mintzatu*: absolutive

```
S
  NP    VP
    nor NP   V
      mintzatu
```
The case of the subject is determined by the deleted verb:

$hitz$-$egin$: ergative

$mintzatu$: absolutive
In WYSIWYG analysis, the structure is the same in both cases:

```
  S
   |   NP
    |   who
```

- The verb is not part of the structure, so there’s no way to assign the right case to the subject.
Likewise in German:

(84) Anke hat jemandem gedroht, aber ich weiss nicht, Anke has someone.dat threatened but I know not
wem/*wen sie gedroht hat.
who.dat she threatened has
‘Anke threatened someone, but I don’t know who she threatened.’

(85) Anke hat jemanden gelobt, aber ich weiss nicht, Anke has someone.acc praised but I know not
wen/*wem sie gelobt hat.
who.acc she praised has
‘Anke praised someone, but I don’t know who she praised.’
Sluicing in German:

(86) Anke hat jemandem gedroht, aber ich weiss nicht,  
_Anke has someone.dat threatened but I know not_  
*wem/*wen.  
*who.dat/who.acc  
‘Anke threatened someone, but I don’t know who.’

(87) Anke hat jemanden gelobt, aber ich weiss nicht,  
_Anke has someone.acc praised but I know not_  
*wem/*wen.  
*who.acc/who.dat  
‘Anke praised someone, but I don’t know who.’
Important point: Other anaphoric devices (e.g., pronouns) do not agree in case with their antecedents (though they may agree in person, number, and gender):

(88) Anke hat jemandem gedroht, aber ich weiss nicht, ob er reagiert hat.
Anke has someone.dat threatened but I know not whether he reacted has
‘Anke threatened someone, but I don’t know whether he reacted.’

(89) Anke hat jemanden gelobt, aber ich weiss nicht, ob er reagiert hat.
Anke has someone.acc praised but I know not whether he reacted has
‘Anke praised someone, but I don’t know whether he reacted.’
Case assignment is a lexical, idiosyncratic—that is, syntactic—matter for many predicates in many languages: e.g. in German, you can’t predict on the basis of meaning that helfen ‘to help’ assigns the dative, while assistieren ‘to assist’ assigns the accusative.

Explaining this effect is straightforward if the ellipsis contains an instance of the case-assigning verb.
Case assignment is a lexical, idiosyncratic—that is, syntactic—matter for many predicates in many languages: e.g. in German, you can’t predict on the basis of meaning that helfen ‘to help’ assigns the dative, while assistieren ‘to assist’ assigns the accusative.

Explaining this effect is straightforward if the ellipsis contains an instance of the case-assigning verb.
Code-switching: switching from one language system to another, typically within a single sentence or utterance:

(90) Juan amenazó a alguien, aber ich weiss nicht, wem
Juan threatened someone.acc but I know not who.dat
Juan gedroht hat.
he threatened has

(91) Juan amenazó a alguien, aber ich weiss nicht, wen
Juan threatened someone.acc but I know not who.acc
Juan amenazó.
Juan threatened

‘Juan threatened someone, but I don’t know who Juan threatened.’
Gonzalez and Ramos (2012): Tested speakers’ ratings for sluiced, Spanish, and German continuations:

Test sentences:

(92) Juan amenazó a alguien, aber ich weiss nicht, wem.
Juan threatened someone.acc but I know not who.dat

(93) Juan amenazó a alguien, aber ich weiss nicht, wen.
Juan threatened someone.acc but I know not who.acc

‘Juan threatened someone, but I don’t know who.’
Results:

Table 1. Verbs that assign accusative in Spanish (ratings on a 1-5 Likert scale, M=mean, SD=standard deviation)

<table>
<thead>
<tr>
<th>NOM</th>
<th>NOM M</th>
<th>NOM SD</th>
<th>ACC M</th>
<th>ACC SD</th>
<th>DAT M</th>
<th>DAT SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sluiced</td>
<td>1.38</td>
<td>0.58</td>
<td>4.00</td>
<td>1.29</td>
<td>2.08</td>
<td>1.21</td>
</tr>
<tr>
<td>Spanish</td>
<td>1.21</td>
<td>0.66</td>
<td>4.00</td>
<td>1.25</td>
<td>2.17</td>
<td>1.43</td>
</tr>
<tr>
<td>German</td>
<td>1.13</td>
<td>0.34</td>
<td>1.71</td>
<td>0.81</td>
<td>5.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>
The unpronounced Spanish verb (here: amenazar) assigns the accusative case to the wh-phrase in German (so (95) is predicted); the equivalent German verb is not possible (so (94) is ruled out).
**Preposition-stranding:**

Some languages allow for prepositions to be ‘stranded’: separated from their objects when those objects must appear somewhere other than adjacent to the preposition:

(96)  

a. English: Who was he talking with?  
b. Swedish: Vem har Peter talat med?  
c. Norwegian: Hvem har Per snakket med?  
d. Danish: Hvem har Peter snakket med?  
e. Icelandic: Hvern hefur Pétur talað við?
Others (most others) don’t:

(97)  
a. Greek: * Pjon milise me?  
b. Russian: * Kem ona govorila s?  
c. Slovene: * Kom je govorila Anna s?  
d. Bulgarian: * Koj e govorila Anna s?  
e. Persian: * Ki ali ba harf mi-zad?  
f. German: * Wem hat sie mit gesprochen?  
g. Yiddish * Vemen hot zi mit geredt?  
h. Hebrew: * Mi Dani katav le?
P-stranding is a constraint on the application of the rule that maps one phrase marker to another:
Prediction:

*If abstract syntax underlies elliptical questions, then this language-particular constraint should be in effect in such questions as well.*

*(If there is no syntax inside an ellipsis site, there need be no correlation between nonelliptical forms of wh-phrases and ones that appear in elliptical constructions.)*
(98) a. Peter was talking with someone, but I don’t know (with) who.

b. Peter har talat med någon; jag vet inte (med) vem.
   *Peter has talked with someone I know not with who*

c. Per har snakket med noen, men jeg vet ikke (med) hvem.
   *Per has talked with someone but I know not with who*

d. Peter har snakket med en eller anden, men jeg ved ikke
   *Peter has talked with one or another but I know not with who*

e. Pétur hefur talað við einhvern en ég veit ekki (við) hvern.
   *Peter has spoken with someone but I know not with who*
(99)  a. I Anna milise me kapjon, alla dhe ksero *(me) pjon.
   the Anna spoke with someone but not I. know with who

b. Anna e govorila s njakoj, no ne znam *(s) kjo.
   Anna AUX spoken with someone but not I. know with who

c. Anna je govorila z nekom, ampak ne vem *(s) kom.
   Anna aux spoken with someone but not I. know with who

d. Anja govorila s kem-to, no ne znaju *(s) kem.
   Anja spoke with someone, but not I. know with who

e. Ali ba kasi hard mi-zad, ama ne-mi-dan-am *(ba) ki.
   Ali with someone talk PROG-hit.3sg but not-PROG-know-I with who

f. Anna hat mit jemandem gesprochen, aber ich weiss nicht, *(mit) wem.
   Anna has with someone spoken but I know not with who

g. Zi hot mit emetsn geredt, ober ikh veys nit *(mit) vemen.
   she has with someone spoken but I know not with who

h. Dani katav le-mishehu, aval ani lo yode’a *(le-)mi.
   Dani wrote to-someone, but I not know to-who
**Conclusion:** the local syntactic constraints on interrogative phrases hold even when there is no pronounced syntax.
Evidence from **priming**

- Humans can be primed to use recently encountered syntactic structures
- Example: NP NP vs NP PP in ditransitives

\begin{align*}
(100) \quad & \text{Ralph sang } [_{NP \text{ Sheila}}] [_{NP \text{ a song}}]. \quad \text{ (NP NP)} \\
(101) \quad & \text{Ralph sang } [_{NP \text{ a song}}] [_{PP \text{ to Sheila}}]. \quad \text{ (NP PP)}
\end{align*}
**Hypothesis**: Syntactic ellipsis will give rise to syntactic priming effects.

<table>
<thead>
<tr>
<th>Prime Sentence 1</th>
<th>Prime Sentence 2</th>
<th>Continuation Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP NP Primes First Ralph sang Sheila a song,</td>
<td>and then Marcus did.</td>
<td>(Ellipsis)</td>
</tr>
<tr>
<td></td>
<td>and then Marcus sang her one.</td>
<td>(Nonelliptical)</td>
</tr>
<tr>
<td></td>
<td>and then Marcus groaned.</td>
<td>(Neutral control)</td>
</tr>
<tr>
<td>NP PP Primes First Ralph sang a song to Sheila,</td>
<td>and then Marcus did.</td>
<td>(Ellipsis)</td>
</tr>
<tr>
<td></td>
<td>and then Marcus sang one to her.</td>
<td>(Nonelliptical)</td>
</tr>
<tr>
<td></td>
<td>and then Marcus groaned.</td>
<td>(Neutral control)</td>
</tr>
</tbody>
</table>

- 18 6-condition items, 36 fillers
**Experiment**: Expose speakers (N=82) to priming sentences, then ask them to describe a picture, thereby producing a sentence containing a different ditransitive verb. (joint work with Ming Xiang and Julian Grove)
The claim

The evidence: Structural ambiguities

The evidence: Ellipsis

Conclusion
**Conclusion**: The properties of sentences cannot be modeled solely by treating them as strings of words. We need ‘abstract’ structures:

- There is no succor in surfacism.
**Conclusion:** The properties of sentences cannot be modeled solely by treating them as strings of words. We need ‘abstract’ structures:

- There is no succor in surfacism.

Thank you!