

ELBA 2024

El silencio de la sintaxis:

Una introducción al fenómeno de la elipsis

Clase 2. 24 de julio de 2024

1. Review of last class

Ellipsis

Phenomenon in which certain linguistic material is **omitted** but it's still **understood**.

- The mapping between form and meaning breaks down → *there is meaning without form*
- Not all that's missing is ellipsis.**
- Types of ellipsis → categorized based on what's silenced/omitted/elided

I. Nominal ellipsis: a noun is missing.

(1) La tesis de Chomsky es mejor que la tesis de Halle.

II. Verbal/Predicate ellipsis: a predicate or a verbal phrase is missing.

(2) Jason has taught syntax, and Karlos has ~~taught syntax~~ too.

III. TP ellipsis: an entire sentence, except for one word/phrase, is missing.

(3) Conocí a alguien en la fiesta, pero no recuerdo a ~~quién conocí en la fiesta~~.

Terminology

(4) Ana comió algo ayer pero no sé qué (comió ayer).

antecedent remnant ellipsis site

correlate

source

- remnant* = what is not omitted
- ellipsis site* = what is omitted
- source* = the sentence formed by the remnant and the ellipsis site
- antecedent* = the sentence that provides the meaning for the ellipsis site
- correlate* = the element in the antecedent that corresponds to the remnant in the source

- Most of the research on ellipsis (from a syntactic perspective) has focused on:
 - Structure question:** In ellipsis, is there syntactic structure that is unpronounced?
 - Identity question:** the elided material and its linguistic antecedent?
 - Licensing question:** What syntactic configurations allow for ellipsis?
- Surface vs. Deep Anaphora** (Hankamer & Sag 1976) (syntactically vs. pragmatically controlled anaphora)
 - A non-exhaustive list of tests and diagnostics:

	surface anaphora	deep anaphora
pragmatic control	*	ok
missing antecedents	ok	*
syntactic parallelism	required	optional
extraction	ok	*
agreement	required	optional
inverse scope	ok	*

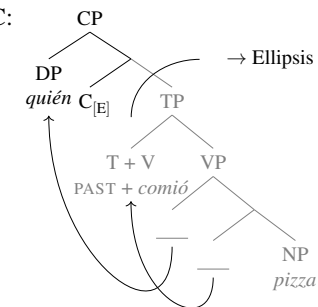
Note: No single test should be used in isolation or considered definitive proof.

Working Hypothesis: In ellipsis, there is a full-fledged syntactic structure that is not pronounced.

• What triggers ellipsis? What determines the size of the ellipsis site?

An [E] feature, that can optionally appear on certain heads (Merchant 2001). This feature gives the instruction to omit/not-pronounce its complement. The size of the ellipsis site is determined by which head bears the [E]-feature.

E.g., in TP-ellipsis, [E] is on C:



• What are the constraints on what can and cannot be elided?

Elliptical sentences have a linguistic antecedent. The ellipsis site has to be 'similar' to its antecedent. If the two structures are not similar enough, ellipsis is not licensed.

2. The identity condition(s) on ellipsis

- Ellipsis is **anaphoric** and depends on its **linguistic context** to get its meaning.
- The attested meanings for ellipses are usually connected to the **linguistic antecedent**.
→ What is the relation between an elliptical structure and its antecedent?

There are broadly three kinds of answers to these questions:

- Semantic/Weak Identity: the relation between the ellipsis site and its antecedent involves a kind of **identity of meaning**.
- Syntactic/Strict Identity: the relation between the ellipsis site and its antecedent involves a kind of **identity of structure**.
- Hybrid/Mixed Identity: a bit of both.

Assumptions

- Elliptical sentences are the result of ‘silencing’ certain parts of a syntactic structure.
- If a non-elliptical sentence is grammatical, its elliptical counterpart is also grammatical; if a non-elliptical sentence is ungrammatical, its elliptical counterpart is also ungrammatical.
→ Any deviations from (ii) should be explained.

Rule of Thumb: Always compare elliptical sentences with their non-elliptical counterparts.

3. Evidence for syntactic/strict identity

→ Evidence will come from contexts where just similarity in meaning is not enough—i.e., structural isomorphism is required:

i. Voice mismatches (Merchant 2013):

- Active sentences mean (truth-conditionally) the same than their passive counterparts:
- However, an active ellipsis site cannot have a passive antecedent, and a passive ellipsis site cannot have an active antecedent:

ii. Auxiliaries & Tense:

- Emily played beautifully at the recital and Ana will, too.
E. played beautifully at the recital and A. will (~~play~~ beautifully at the recital), too.
- *Emily was beautiful at the recital and Ana will, too.
Emily was beautiful at the recital and Ana will (~~be~~ beautiful at the recital), too.

4. Evidence for semantic/weak identity

→ Evidence will come from contexts where strict identity is not met—i.e., where we find structural mismatches between the ellipsis site and its antecedent:

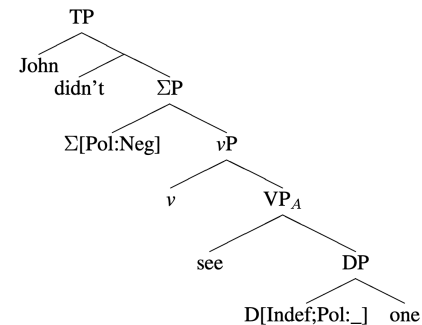
i. Negative polarity items

- While it can be argued that the semantics of *someone* and *anyone* are identical—both corresponding to an existentially bound variable—it is less clear if their syntax is as well.

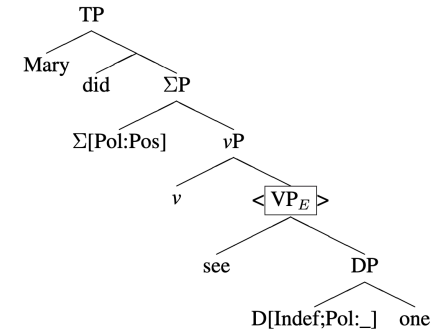
- Juan no vio a nadie, pero María sí.
 - Juan no vio a nadie, pero María sí (vio a alguien).
 - *Juan no vio a nadie, pero María sí (vio a nadie).
- Juan vio a alguien, pero María no.
 - ≠Juan vio a alguien, pero María no (vio a alguien).
 - Juan vio a alguien, pero María no (vio a nadie).

- Merchant (2013): the examples above are compatible with a syntactic identity condition.
 - Giannakidou (2000, 2007): polarity items have a syntactic feature [Pol:_] which is valued under Agree with a c-commanding ‘licensor’ such as negation:

(9) a. Antecedent:



b. Ellipsis-site



(10) Lexical Insertion rules

- [Cat[D, Indef]; Infl[Pol:Neg]] ↦ any
- [Cat[D, Indef]; Infl[Pol:Pos]] ↦ some (sm)/a

→ certain expressions have varying morphological realizations, depending on their syntactic environment; which morphology is realized is determined by agreement with a valuer

ii. Vehicle Change

- ‘Vehicle change’ (Fiengo and May, 1994): pronoun/name-equivalences under ellipsis:

- (11) a. They arrested Alex_i, even though he_i thought they wouldn’t.
 b. *They arrested Alex_i, even though he_i thought they wouldn’t ~~arrest Alex_i~~.
 c. They arrested Alex_i, even though he_i thought they wouldn’t ~~arrest him_i~~.

– If the VP-ellipsis site were completely identical to its antecedent, this example would violate Principle C of the Binding Theory just as the non-elliptical does.

– The fact that it is grammatical that the proper name has ‘transformed’ into the pronoun *him*, thus avoiding the binding violation.

- While it can be argued that the denotation of *Alex* and *he* is identical under the relevant assignment function, syntactically there is substantial difference between these two DPs.
- Murphy & Müller (2022): ellipsis applies in a successive-cyclic fashion and that the absence of Principle C effects follows from the fact that the R-expression is no longer syntactically accessible at the relevant point in the derivation.

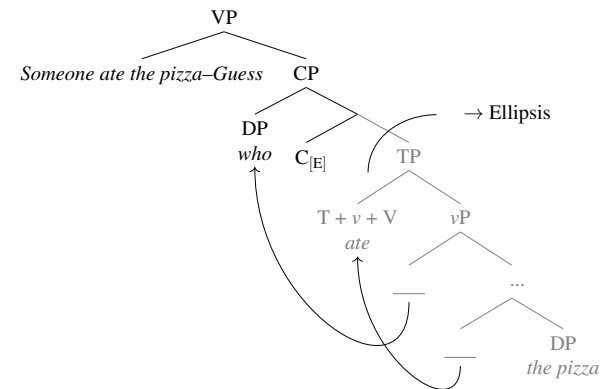
5. Case study: Sluicing

Sluicing

Sluicing is a type of TP-ellipsis where an entire wh-question is missing, to the exclusion of a wh-word/phrase. (Ross 1969)

- (12) Someone left...
 a. ...Guess who left!
 b. ...Guess who!
- (13) Yesterday someone gave me a red book in the library after hitting me in the head...
 a. ...Guess who gave me a red book in the library after hitting me in the head yesterday!
 b. ...Guess who!
- (14) Alguien se fue...
 a. ...¿Adiviná quién se fue!
 b. ...¿Adiviná quién!
- (15) Alguien me dio un libro rojo en la plaza después de golpear me en la cabeza...
 a. ...¿Adiviná quién me dio un libro rojo en la plaza después de golpear me en la cabeza!
 b. ...¿Adiviná quién!

- The *Ross-Merchant* approach:



- Assumptions of the *Ross-Merchant* approach:

1. there is structure inside the ellipsis site
2. the structure inside the ellipsis site goes silent/unpronounced
3. elliptical sentences ‘behave’ like their non-elliptical counterparts (e.g., wh-movement, case marking, pronoun binding, etc.)
4. the elliptical sentence ‘resembles’ its linguistic antecedent (i.e., the identity question)

→ To what extent can the elliptical sentence differ from its linguistic antecedent? What are the possible and impossible mismatches between the antecedent and the ellipsis site, and what does that tell us about the identity condition that licenses ellipsis?

→ Two hypotheses:

- (16) [Someone left]_{Antecedent} Guess [who ____]!
 a. Guess [who left]_{Source} → *isomorphic source*
 b. Guess [who it was]_{Source} → *non-isomorphic (copular) source*
- (17) [Alguien se fue]_{Antecedent} ¿Adiviná [quién ____]!
 a. Adiviná [quién se fue]_{Source} → *isomorphic source*
 b. Adiviná [quién es/era]_{Source} → *non-isomorphic (copular) source*

How do we decide between these two alternatives? Does elliptical sentences pattern with isomorphic sources or with copular sources?

→ We look for the effects of Structure X (i.e., either isomorphic or copular) in non-elliptical contexts and we compare that to their elliptical counterparts.

5.1 Case (mis)matches

Case is a purely *grammatical* marking, with no meaning consequences.

- German: *schmeicheln* ‘flatter’ assigns DAT; *loben* ‘praise’ assigns ACC:

- (18) a. Er will jemandem schmeicheln.
 he wants someone.DAT flatter
 ‘He wants to flatter someone.’
 b. Er will jemanden loben.
 he wants someone.ACC praise
 ‘He wants to praise someone.’

- In elliptical contexts, the case of the remnant must ‘match’ the case of its correlate:

- (19) a. Er will jemandem schmeicheln, aber sie wissen nicht, { *wer |
 he wants someone.DAT flatter but they know not who.NOM
 *wen | wem }.
 who.ACC who.DAT
 ‘He wants to flatter someone but they don’t know who.’
 b. Er will jemanden loben, aber sie wissen nicht, { *wer |
 he wants someone.ACC praise but they know not who.NOM
 wen | *wem }.
 who.ACC who.DAT
 ‘He wants to praise someone but they don’t know who.’

- Similarly, Spanish human/animate objects bear DOM:

- (20) a. Tom vio *(a) Iván.
 b. Tom vio *(a) la película.

- In elliptical contexts, human wh-phrases must bear DOM, as in non-elliptical contexts:

- (21) a. Tom vio a alguien, pero no recuerdo *(a) quién.
 b. Tom vio algo, pero no recuerdo *(a) qué.

→ What predictions do the different sources make?

- A non-isomorphic source predicts DOM-less remnants to be grammatical (22b)...

- (22) a. *Tom vio a alguien pero no recuerdo a quién es (la persona que vio).
 b. Tom vio a alguien pero no recuerdo quién es (la persona (a la) que vio).

... contrary to fact (21a).

- Similarly, in Greek, copular sources predict the remnant to bear NOM case, contrary to fact:

- (23) a. I astinomia anekrine enan apo tous Kiprious prota, ala dhen ksero
 the police interrogated one.ACC from the Cypriots first but not I.know
 { *pjos | pjon } anekrine i astinomia.
 who.NOM who.ACC interrogated the police
 ‘The police first interrogated one of the Cypriots, but I don’t know who.’
 b. I astinomia anekrine enan apo tous Kiprious prota, ala dhen ksero
 the police interrogated one.ACC from the Cypriots first but not I.know
 { pjos | *pjon } itan.
 who.NOM who.ACC it.was
 ‘The police interrogated on of the Cypriots first, but I don’t know who (it was).’

Form-Identity Generalization I: Case-matching

The remnant of sluicing must bear the case that its correlate bears.
 (adapted from Merchant 2001, p.91)

→ Should we just reduce this to case *matching*? No.

- *Sprouting*: nothing to ‘match’ with—e.g., German → *helfen* ‘help’ assigns DAT:

- (24) Er meinte, er hätte geholfen, aber wir wüßten nicht, { wem | *wen }.
 he thought he had helped but we knew nor who.DAT who.ACC
 ‘He claims he helped, but we wouldn’t be able to say who.’

→ Case-matching effects are easily explained if the case assigner is in the elided structure.

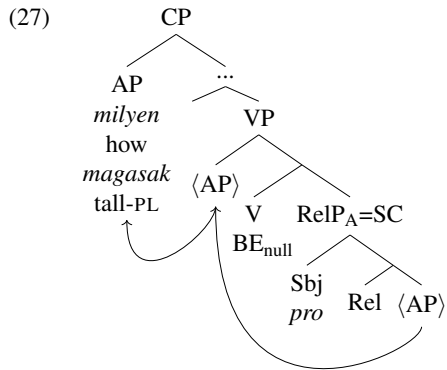
Case mismatches

- There are instances of apparent case mismatches:

- (25) Hungarian
 a. Mari ismer néhány magas lány-t, de nem tudom milyen magas-ak-at.
 Mary knows some tall girls-ACC, but not know.I how tall-PL-ACC
 ‘Mary knows some tall girls, but I don’t know how tall.’
 b. Mari ismer néhány magas lány-t, de nem tudom milyen magas-ak.
 Mary knows some tall girl-ACC, but not know.I how tall-pl
 ‘Mary knows some tall girls, but I don’t know how tall.’

- Ronai & Stigliano (2021): apparent exceptions to the Form-Identity Generalization I in Hungarian are not elliptical, but they arise from copular sentences with null copular verbs.

- (26) M. ismer néhány magas lány-t, de nem tudom milyen magas-ak BE_{null} pro.
 M. knows some tall girl-ACC, but not know.I how tall-PL
 ‘Mary knows some tall girls, but I don’t know how tall.’



- (32) a. I Anna milise me kapjon, alla dhe ksero **me** pjon.
 the Anna talked with someone but not I.know with who
 ‘Anna talked with someone but I don’t know with who.’
 b. *I Anna milise me kapjon, alla dhe ksero pjon.
 the Anna talked with someone but not I.know who
 Intended: ‘Anna talked with someone but I don’t know with who.’

Form-Identity Generalization II: P-Stranding

A language *L* will allow preposition stranding under sluicing iff
L allows preposition stranding under regular wh-movement.
 (adapted from Merchant 2001, p.92)

- Maybe there are true cases of case-mismatches in ellipsis like Mongolian (see Vicente 2015):

- (28) Bat hennege**n-d** ene nom-ig ug-sun, gevch bi {hen-ig |
 Bat.NOM someone-DAT this book-ACC give-PERF but I who-ACC
 *hen-d} 'n med-eh-gui.
 who-DAT POSS know-INF-NEG
 ‘Bat gave someone this book, but I don’t know who.’

5.2 P-Omission

- In English P-stranding in regular wh-questions is possible and P-stranding in elliptical contexts (e.g., sluicing) is also possible:

- (29) a. I don’t know who was he talking **with**.
 b. I don’t know **with** whom he was talking.
 (30) a. Ben was talking with someone, but I don’t know **with** who(m).
 b. Ben was talking with someone, but I don’t know who.

→ This is expected if elliptical sentences arise from non-elliptical sentences.

- In Greek P-stranding in regular wh-questions is impossible and P-stranding in elliptical contexts is also impossible:

- (31) a. *Pjon milise me?
 who she.talked with
 Intended: ‘Who did she talk with?’
 b. Me pjon milise?
 with who she.talked
 ‘Who did she talk with?’ (lit. ‘With who did she talk?’)

- According to Merchant (2001):
 - Languages that allow P-stranding in regular wh-questions and in sluicing: Frisian, Swedish, Norwegian, Danish, Icelandic.
 - Languages that do not allow P-stranding in regular wh-questions nor in sluicing: Greek, German, Dutch, Yiddish, Russian, Czech, Bulgarian, Serbo-Croatian, Slovene, Persian, Catalan, French, Italian, Hebrew, Moroccan Arabic, Basque, a.o.
- BUT, there are also languages that do not allow P-stranding in regular wh-questions but do allow it in sluicing: Spanish, Brazilian Portuguese, Emirati Arabic, Hijazi Arabic, Iraqi Arabic, Polish, a.o.

Spanish → apparent counterexample to Merchant’s generalization:

- (33) a. *No sé quién habló con.
 b. No sé con quién habló.
 (34) a. Sara habló con alguien pero no sé **con** quién.
 b. Sara habló con alguien pero no sé quién.

- How do we explain (34b)?
 - Option 1: the source of (34b) is the ungrammatical sentence (33a), that gets ‘repaired’.
 - Option 2: the source of (34b) is not a wh-question, but a copular source.
 - Option 3:

Let's explore Option 2 (Rodrigues et al 2009)...

- (35) a. Sara habló con alguien pero no sé quién.
b. Sara habló con alguien pero no sé quién es (la persona con la que habló).

- **Hypothesis:** the ellipsis site can either contain a structure that is isomorphic to the linguistic antecedent, or a copular source:

- (36) Someone ate my pizza but I don't know who.
a. Someone ate my pizza but I don't know who ~~<ate my pizza>~~.
b. Someone ate my pizza but I don't know who ~~<it was>~~.

Problems?

Testing copular sources (van Craenenbroeck 2010)

- Evidence 1: Adjuncts and implicit arguments in Sprouting
→ copular sentences are ungrammatical, but elliptical sentences are grammatical:

- (37) a. He fixed the car, but I don't know how.
b. *He fixed the car, but I don't know how it was.
(38) a. Mario arregló el auto, pero no tengo idea cómo.
b. *Mario arregló el auto, pero no tengo idea cómo fue.

- Evidence 2: Exhaustive readings
→ copular questions entail exhaustivity, wh-questions do not; hence copular questions (but not wh-questions) are incompatible with *mention-some* modifiers:

- (39) Copular questions → require an exhaustive list of all the people the addressee saw:
a. Who is it that you saw?
b. ¿Quién es la persona a la que viste?
(40) Regular wh-question → no exhaustivity requirement:
a. Who did you see?
b. ¿A quién viste?

– Then, an interrogative cleft cannot be combined with a modifier such as *for example*, since it means that the answer to the question need not be exhaustive.

- (41) A: Deberías hacer algo para solucionar tus problemas financieros.
B: ¿Qué, por ejemplo?
a. ¿Qué debería hacer para solucionar mis problemas financieros, por ejemplo?
b. *¿Qué es lo que debería hacer para solucionar mis problemas, por ejemplo?

- (42) A: You should talk to somebody in the legal department for help with that.
B: Who, for example?
a. Who should I talk to, for example?
b. *Who is it, for example?

- Another example → *else* modification

- (43) Ana estaba en la plaza, pero no sé quién más.
a. Ana estaba en la plaza, pero no sé quién más estaba en la plaza.
b. *Ana estaba en la plaza, pero no sé quién más era (la persona que estaba...)
(44) Ana was there, but I don't know who else.
a. Ana was there, but I don't know who else was there.
b. *Ana was there, but I don't know who else it was.

→ Back to exceptions to the P-Stranding Generalization:

- (45) Sara habló con alguien, pero no sé quién.

Using the tests above, determine whether they can arise from copular sources or not.

- Further evidence against copular sources: Polish is a non-P-stranding language that allows P-stranding in sluicing BUT the remnant must have the case assigned by the missing preposition (rather than INST, the case of clefts pivots):

- (46) Adam dostaje prezenty od kongós, ale nie wiem {kogo | *kim}
Adam gets presents from someone.GEN but not know who.GEN who.INST
'Adam gets presents from someone, but I don't know who.'

→ The preposition seems to be present even if it's not pronounced

- Some questions:

Q1 What does the P-Stranding Generalization actually convey?

Literal: A language L will allow *preposition stranding* under sluicing iff L allows preposition stranding under regular wh-movement.
→ i.e., whatever mechanisms/constraints/restrictions a language has in non-elliptical sentences, we'll also find them in elliptical sentences

Revised: A language L will allow *preposition omission* under sluicing iff L allows preposition stranding under regular wh-movement.
→ empirically false and requires an explanation for the counterexamples

Q2 How do we explain the counter examples to the P-Stranding Generalization?

→ An answer to Q2 (Stigliano 2021):

Ellipsis type	P-Omission	Movement of the correlate in the antecedent	Copular source
Sluicing	✓	✗	✓
Fragment answers	✗	✓	✓
Contrast sluicing	✓	✗	✗
	✗	✓	✓/✗
Split questions	✗	✓	✓
Fragments	✓	✗	✓
	✗	✓	✓
Stripping	✓	✗	✗
Pseudostripping	✗	✗	✓

(47) The *P-(re)position Omission Generalization* for Spanish:
P-Omission in ellipsis in Spanish is only allowed when the remnant's correlate in the antecedent does not move.

– Basic contrast:

(48) Ana habló con alguien pero no sé (con) quién.

(49) A: ¿Con quién habló Ana? B: *(Con) Pedro.

6. Sprouting and P-Omission

• Sprouting: a type of sluicing, where the remnant doesn't have an overt correlate:

(50) Ben found a diamond but he didn't tell me where.

(51) Ben encontró un diamante, pero no me dijo dónde.

• A preposition cannot be omitted in ellipsis if the remnant is sprouted:

(52) a. They're jealous, but it's unclear of who.

b. *They're jealous, but it's unclear who(m).

(53) a. Paula está celosa, pero no sé quién.

b. *Paula está celosa, pero no sé de quién.

Why is P-Omission in ellipsis only possible when there is a PP in the antecedent?

• Option 1 – Only Semantics (Merchant 2001): there should be mutual entailment between the antecedent and the source (ellipsis site + remnant)

What would be the difference between the sentences with and without the preposition in terms of entailment? **None.**

(54) Who is he jealous of? ≡ Of who is he jealous?

• Option 2 – Lexicon/Syntax (Chung 2005): every lexical item in the source that ends up (only) in the ellipsis site must be identical to an item in the numeration of the antecedent.

→ In other words: except for the moved interrogative phrase, the lexical items in the ellipsis site must be a subset of the lexical items from which the antecedent is constructed.

P-Stranding in Sprouting:

(55) Licit derivation:

[Antecedent They're jealous], but it's unclear [of who [Ellipsis site they're jealous ___]].

a. {they, are, jealous}^{Ellipsis site} ⊆ {they, are, jealous}^{Antecedent}

(56) Illicit derivation:

[Antecedent They're jealous], but it's unclear [who [Ellipsis site they're jealous of ___]].

a. {they, are, jealous, of}^{Ellipsis site} ⊈ {they, are, jealous}^{Antecedent}

P-Stranding in Regular Sluicing:

(57) Licit derivation:

[Antec. They're jealous of someone], but it's unclear [of who [Ellipsis site they're jealous ___]].

a. {they, are, jealous}^{Ellipsis site} ⊆ {they, are, jealous, of, someone}^{Antecedent}

(58) Licit derivation:

[Antec. They're jealous of someone], but it's unclear [who [Ellipsis site they're jealous of ___]].

a. {they, are, jealous, of}^{Ellipsis site} ⊆ {they, are, jealous, of, someone}^{Antecedent}

→ Does this work? Why/Why not? What evidence do we need to look at?

A proposal:

• Ellipsis is triggered by [E] and licensed by [†]:

(59) Assignment of [†]: Assign [†] to every head *h* in the complement of a head $z_{[E]}$.

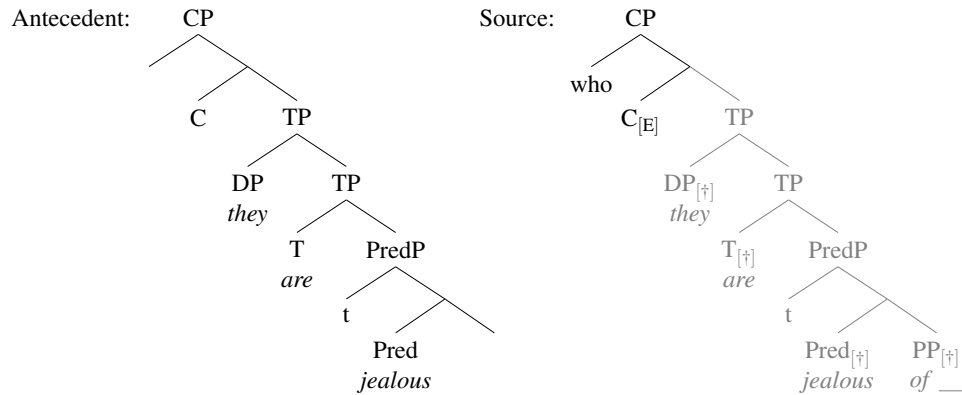
(60) Phonology of [†]: A head containing the feature [†] is not pronounced.

(61) Identity Condition: A head $h_{[†]}$ is licensed iff *h* has an identical correlate *h'* in *A*, where *A* is the antecedent.

(62) Identity: A morpheme α is identical to another morpheme β if and only if α and β match all their semantic and syntactic features.

- Let's see how this works:

(63) [_{Antecedent} They're jealous], but it's unclear [who [_{Ellipsis site} they're jealous of ___]].



→ the preposition *of* in the ellipsis site doesn't find an identical correlate in the Antecedent:

(64) Identity reference set: {⟨DP_E, DP_A⟩, ⟨T_E, T_A⟩, ⟨jealous_E, jealous_A⟩, ⟨of_E, ∅_A⟩}

7. Ellipsis & Islands

- We expect that the grammaticality of an elliptical sentence will match the grammaticality of its non-elliptical counterpart, but...
 - ...we discussed examples where the elliptical sentence was ungrammatical and its non-elliptical counterpart was grammatical
 - there was something wrong in the elliptical 'operation' (e.g., it was illicit).
 - ...there are also cases where an ungrammatical non-elliptical sentence becomes grammatical in the context of ellipsis.

- Islands: movement dependencies are subject to certain restrictions:

(65) *[Which Romance language]_i did they hire someone who speaks ____i?

(66) *¿[Cuál lengua romance]_i contrataron a alguien que habla ____i?

- However, it seems that some of those restrictions disappear in contexts of ellipsis:

(67) a. They hired someone who speaks a Romance language, but I don't know which.
 b. Contrataron a alguien que habla una lengua romance, pero no recuerdo cuál.

→ Think of an island and see if it can be 'repaired' by ellipsis.

→ How do we explain 'island repair by ellipsis'? Any thoughts?

- Option 1: back to copular sources :(

(68) They hired someone who speaks a Romance language, ...
 ...but I don't know which Romance language it is.

(69) Contrataron a alguien que habla una lengua romance, ...
 ...pero no recuerdo qué lengua romance es.

- Option 2: short sources

(70) They hired someone who speaks a Romance language, ...
 ...but I don't know which Romance language {they_{sg}/he/she} speak.

(71) Contrataron a alguien que habla una lengua romance, ...
 ...pero no recuerdo qué lengua romance habla.

- Pros—We seem to find short sources elsewhere:

(72) I remember meeting him, but I don't remember when.

a. but I don't remember when I met him.

b. ≠but I don't remember when I remember meeting him.

- Cons—Some cases of 'repaired islands' don't have a possible short source:

(73) She bought a big car, but I don't know how big.

a. ...but I don't know how big it is.

(74) Quieren contratar a alguien que hable una lengua romance, pero no sé cuál.

a. ≠pero no sé cuál lengua romance habla.

b. *pero no sé cuál lengua romance hable.