Identity with double ellipsis

Richard Stockwell

*School of Communication and Media, Ulster University, York Street, Belfast, BT15 1ED, United Kingdom.

Corresponding author(s). E-mail(s): r.stockwell@ulster.ac.uk;

Abstract

On the view that ellipsis sites contain full – just unpronounced – syntactic structure, it follows that two ellipsis sites should be able to mutually license one another for identity. This paper supports the structural view of ellipsis by exploring this prediction in relation to a series of otherwise recalcitrant counterexamples. With 'double ellipsis', two ellipsis sites can satisfy identity for one another without formally relating to overt antecedent material. This analysis allows structural identity to be maintained in the face of apparent mismatches, including argument structure in sluicing, Dahl's many clauses puzzle and elliptical answers to certain questions. It means that ellipsis sites must contain syntactic structure to evaluate for identity, and can do so even where no overt antecedent material is available.

Keywords: ellipsis, identity, voice mismatch, sluicing, verb phrase ellipsis, Dahl's puzzle, missing antecedents

1 Introduction

'Ellipsis' provides meaning without form. In (1), we understand in (a) that the question is about what John bought and in (b) that Mary also bought a book without having to repeat ourselves:

(1) a. John bought something, but I don't know what.
b. John bought a book, and Mary did too.

There are, very broadly, two views of how ellipses get their meaning in the absence of form. On one view, an ellipsis has its own, fully structured syntax that is permitted to go unpronounced based on a relationship of identity with antecedent linguistic material (Hankamer 1971, Sag 1976, Hankamer and Sag 1976, Merchant 2001). The

silent structure <thus indicated> in (2) then gives rise to the unspoken meaning of the elliptical clause by being interpreted in the usual compositional way:

(2) a. John bought something, but I don't know what <he bought t_{what}>.
b. John bought a book, and Mary did too <buy a book>.

On another view, ellipsis is a species of anaphora, akin to pronouns. Ellipsis sites consist of silent *pro*-forms (3) that depend on an antecedent for their meaning via the usual mechanisms of anaphora (Wasow 1972, Williams 1977, Hardt 1993, Lobeck 1995) or LF-copying (Fiengo and May 1994, Chung et al 1995), or even consist of nothing at all (Ginzburg and Sag 2000, Culicover and Jackendoff 2005, Jacobson 2019):

(3) a. John bought something, but I don't know what *pro*.
b. John bought a book, and Mary did too *pro*.

On the structural view, nothing prevents antecedents of ellipsis from themselves being elliptical. Jarvis (2022) presents the example of sluicing in (4) whose antecedent itself contains an elided verb phrase:

(4) A: Would you ever donate your hair?
B: Well, I probably would <donate my hair> to some organisations, yeah.
I don't really know which ones <I would donate my hair to t>, though.

In (4), the initial instance of verb phrase ellipsis in B has an antecedent in A's question. More extremely, however, on the structural view there is nothing in principle to prevent an antecedent for ellipsis from itself being entirely elliptical. And moreover, that ellipsis should in turn be able to serve as the antecedent for ellipsis of its antecedent, with the result that the two ellipses are mutually licensing as the antecedents for each other.

This paper will survey such cases of 'double ellipsis'. Two ellipsis sites can be mutually licensing in that each provides structured linguistic material for the other to satisfy identity. Such configurations will circumvent identity constraints that apply to single ellipsis, allowing for apparent mismatches with antecedent material by evaluating identity entirely between the two ellipses. That is, where a lone ellipsis is bad for violating identity, adding a second ellipsis will make both good, since the two ellipses then satisfy identity for one another.

We will proceed through a series of case studies starting from baselines where a single ellipsis fails, as schematically in (5). In these circumstances, ellipsis fails due to a lack of identity ($*\neq$) with antecedent material:

(5)
Overt material
$*\neq$
bad <ellipsis>

Yet keeping all else equal, the very same ellipsis becomes good after adding a second, intermediate ellipsis, as schematically in (6). With double ellipsis, identity ($=$) can now be established between the two ellipses. In other words, the ellipsis sites are mutually licensing with respect to identity:

(6)

Overt material
≠
intermediate <ellipsis>
=
previously bad <ellipsis> becomes good

This analysis is available only on the structural view of ellipsis. For the ellipsis sites to be mutual licensing, there must be structure inside them to evaluate for identity.

Several apparent counterexamples to identity can be accommodated under the structural view by taking this perspective on double ellipsis. We begin in section 2 by demonstrating the ability of double ellipsis to circumvent voice and other argument mismatches in sluicing. Section 3 shows that double ellipsis is not an all-powerful amelioratory device, but is crucially limited by parsing considerations and the need to establish structural identity between the two ellipses with reasonably available lexical resources. Sections 4 and 5 downsize to verb phrase ellipsis, applying the present perspective to resolve Dahl's many clauses puzzle and to make some progress on understanding elliptical answers to adjunct and alternative questions.

Even if formally the two ellipsis sites mutually license one another in double ellipsis, there needs to be some clue as to their meaning.[1] Usually, this clue will come from overt antecedent material. But section 6 will argue that it can also come from a non-linguistic scene, allowing double ellipsis to succeed entirely without an overt antecedent. Section 7 concludes the paper.

2 Argument structure

We begin with a case study in argument structure and clausal ellipsis. Sluicing requires structural identity in voice (Merchant 2013). The voice matches in (7) are grammatical, whereas the mismatches in (8) are not:

(7) a. Someone saved Alex, but we don't know who <t_{who} saved Alex>.
active = active

b. Alex was saved, but we don't know by whom <Alex was saved>.
passive = passive

(8) a. * Someone saved Alex, but we don't know by whom <Alex was saved>.
active ≠ passive

b. * Alex was saved, but we don't know who <t_{who} saved Alex>.
passive ≠ active

In the broadest terms, the structural view of ellipsis expects syntactic identity to be enforced between antecedent and ellipsis.[2] For the anaphora view, the ungrammaticality of voice mismatch in sluicing presents more of a challenge. Insofar as active and

[1] This is the issue of 'recoverability' as intended by Fiengo and Lasnik (1972).

[2] The grammaticality of voice mismatch in verb phrase ellipsis (i, Hardt 1993: 37) therefore requires further argument, per Merchant (2013) on the position of the Voice head relative to the ellipsis site:

(i) This information could have been released by Gorbachev, but he chose not to <release it>.

passive are truth-conditionally equivalent, the antecedent provides a suitable meaning for *pro* to draw on in (8) just as much as in (7).

Promisingly for the anaphora view, therefore, and deleteriously for the structural view, Nakamura (2013) observes the examples of sluicing in (9):[3]

(9) a. Not so much whether to teach the Bible in public schools, but how? And by whom?

b. GE Capital and Xerox in Stamford responded to inquiries about their use of extended-stay hotels by saying that they use them from time to time, but they were not sure how much or by whom.

The naturally occurring sluices in (9) look to be counterexamples to the requirement for structural identity in voice. In particular, sluicing was said to be ungrammatical in (8a) for attempting to switch from an active antecedent to a passive ellipsis. Yet as laid out in (10), the final sluices in (9) are passive despite the preceding material being active:[4]

(10) a. Active: . . . teach the Bible in public schools . . .
Passive: And by whom <the Bible should be taught>?

b. Active: . . . they use them from time to time . . .
Passive: . . . or by whom <they are used>.

Faced with (9), Nakamura (2013) argues that structural identity conditions on sluicing should be abandoned. On the contrary, I will defend the claim that voice matching is always respected in sluicing by arguing that structural identity holds in (9) – in particular, between the two ellipsis sites.[5]

Notice first that both examples in (9) involve two instances of ellipsis, as set out in (11):[6]

(11) a. . . . , but how? And by whom?

b. . . . how much or by whom.

Such 'double ellipsis' is in fact crucial to (9). Without the intermediate *how (much)* sluices, the single active-passive mismatches laid out in (10) are ungrammatical in (12):

[3] Nakamura's (2013) sources for (9): (a) *Corpus of Contemporary American English*; (b) *The New York Times*, Aug 9, 1998.

[4] See Anand et al (2021: sect. 5.2) on the appearance of the modal in the ellipsis site in (10a).

[5] In lieu of structural identity, Nakamura (2013) adopts Kertz's (2013) information structure account of mismatch effects in verb phrase ellipsis and suggests extending it to sluicing (cf. the equivocal discussion of the prospects for this in Kertz 2013: 422f.). On this view, the crucial factor in an ellipsis voice alternation is not the syntactic mismatch but a clash in topic structure. For instance in (8a), the change from active to passive switches the topic from the 'saver' to the 'saved' in subject position. Nakamura (2013: 650) compares this with (9), where "the *wh*-phrases themselves are contrasted," which he claims neutralises the issue with the clashing topics. In the sense that two *wh*-phrases must be present to contrast with one another, Nakamura (2013) also alights on the importance of double ellipsis to (9). The difference comes with respect to identity, which is preserved as a condition on ellipsis on the present proposal.

[6] See Citko and Gračanin-Yuksek (2020) for robust argumentation that coordinated and disjoined sluices involve two separate instances of clausal ellipsis.

Their paper, like this one, is limited to English. This provides one reason why double ellipsis might not circumvent mismatches in other languages: in a language where coordinated and disjoined sluices do not involve two separate instances of clausal ellipsis, no amelioration is expected.

Another reason why the amelioratory capacity of double ellipsis may vary cross-linguistically could have to do with morphology: in languages with richer morphology – e.g. for case, gender, etc. – it might be harder to conjure the syntax of an ellipsis site absent an overt, structurally identical antecedent.

(12) a. * Not so much whether to teach the Bible in public schools, but by whom?
b. * They use them from time to time, but they were not sure by whom.

To account for (9), I propose that identity is established between the two ellipsis sites, as provided by double ellipsis. As laid out for (9a) in (13), the two ellipsis sites can be analysed as identical (=) with one another in having passive structure. With single ellipsis, this violates identity relative to the overt antecedent material – viz. the ungrammaticality of (8a). But with double ellipsis, the ellipsis sites can relate to one another for the purposes of identity. All the while, choosing passive structure to represent the intermediate ellipsis site preserves truth-conditional equivalence:

(13)
$$\begin{array}{c} \ldots \text{to teach the Bible in public schools} \\ \neq \\ \text{how } \langle \text{the Bible should be taught} \rangle \\ = \\ \text{by whom } \langle \text{the Bible should be taught} \rangle \end{array}$$

The intermediate ellipsis thus provides a slot that can be filled with structure identical to the second ellipsis, resolving the problem of non-identity with the overt material. The two ellipsis sites are then mutually licensing with respect to identity. The same is laid out more compactly for (9b) in (14):

(14) how much <they are used> = by whom <they are used>

In the face of apparent argument structure mismatches like (9), therefore, recognising that the two ellipsis sites can be mutually licensing for identity allows structural identity conditions to be maintained, contra Nakamura (2013).[7]

The empirical point does not depend on any peculiarities of the naturally occurring examples in (9) – viz. intermediate sluices with *how*, and *PRO* (a) and bound *they* (b) subjects. The constructed examples in (15)-(16) avoid these features but pattern the same way.[8] In all cases, the preceding active overt material mismatches in voice with the final passive ellipsis. Without the parenthesised intermediate ellipsis, single ellipsis is bad due to voice mismatch. With double ellipsis, however, acceptability improves, since the two ellipses can match in voice to satisfy identity (=). While these constructed examples have a slightly reduced '?' level of acceptability, the contrast with '*' single ellipsis is appreciable:

(15) ? The university appoints vice chancellors, but the regulations don't say *(when, or) by whom.
when <VCs are appointed> = by whom <VCs are appointed>

[7] 'Short-form' copular clauses have been proposed as solutions to a variety of issues in sluicing (Barros et al 2014; Palaz et al 2025). They are not applicable here, however, to the extent that the candidate copular clause must be grammatical of itself (cf. Abels 2018 on left branch extraction). Copular clauses would not make viable contents for the ellipsis sites in (13) or (14) given their oddity throughout (9), per (i):

(i) a. * Not so much whether to teach the Bible in public schools, but how it is *t*? And by whom it is *t*?
b. * . . . by saying that they use them from time to time, but they were not sure how much they are *t* or by whom they are *t*.

[8] Uncited data were constructed by the author, a native English speaker. Each data point has been checked with at least six other native English speakers and presented to several audiences.

(16) ? Somebody hacked our computer network, but we've no idea *(why, or) by whom.
why <our network was hacked> = by whom <our network was hacked>

Nor is it necessary for the intermediate sluice to be an adjunct, as so far with *how*, *how much*, *when* and *why*. In the constructed (17), the intermediate sluice is an argument:

(17) ? The news was reporting that the exam board had inadvertently revealed the A-level questions to a certain school ahead of time, but they couldn't say *(to which school, or) by whom.
to which school <the A-level questions had been revealed>
= by whom <the A-level questions had been revealed>

It is also possible to construct examples with voice mismatch the other way round. Where in (9) and (15)-(17) mismatch proceeds from active overt material to a passive second sluice, in (18) voice mismatch is mediated in the opposite direction from passive to active:

(18) ? Vice chancellors are appointed, but the regulations don't say *(when, or) which committee.
when <someone appoints VCs> = which committee <t appoints VCs>

The effect of double ellipsis in circumventing mismatches generalises to other argument structure alternations. Further to voice, sluicing disallows ditransitive diathesis (Merchant 2013), as without the parenthesised material in (19). The mismatch is much improved, however, when bridged by an intermediate sluice:[9]

(19) ? They served someone milk, but I don't know *(why, or) to whom.
why <they served milk to someone> = to whom <they served milk t>

As with voice, therefore, analysing double ellipsis as mutually licensing for identity allows structural identity conditions on ellipsis to be maintained in the face of an apparent ditransitivity mismatch like (19).

The circumventory capacity of double ellipsis also generalises beyond sluicing to other kinds of clausal ellipsis. In (20) without the parenthesised material, the single instance of 'stripping' with voice mismatch is bad; whereas with the intermediate ellipsis, the resulting example of 'double stripping' is much improved:

[9] Alternations between null arguments and PPs (i) (Merchant 2013) seem worse (a), especially with adjectives (b):

(i) a. ?(?) John was arguing, but I can't reveal *(when, or) who.
when <John was arguing with someone> = who <John was arguing with t>
b. (?)* John was afraid, but I don't know (why, or) what.
why <John was afraid of something> = what <John was afraid of t>

There may be difficulty in imputing the preposition to the ellipsis site without an overt clue that this is required – cf. pronunciation of *to* in (19). Cf. also the second point in note 6 and Chung's (2006) 'No new words' generalisation.

(20) ? A university committee can appoint visiting lecturers –
*(visiting professors too,) just not by the host department.
(That has to be referred up to the Dean.)
visiting professors too <t can be appointed>
= just not by the host department <visiting professors can be appointed>

Here the interpretation provides further evidence that the second ellipsis is establishing identity with the first and not with the preceding overt material: it's the visiting professors of the intermediate ellipsis, not the lecturers of the overt material, that cannot be appointed by the host department in the final ellipsis.

In sum, this section has mounted a defence of structural identity conditions on clausal ellipsis. Apparent counterexamples in fact exhibit two instances of ellipsis. Double ellipsis mediates mismatches that are impossible with single ellipsis because the two ellipses can satisfy identity for one another. Subsequent case studies will apply this perspective to some puzzles in verb phrase ellipsis. But first, the next section shows that double ellipsis has its limits.

3 Limits on double ellipsis

The previous section showed how double ellipsis can circumvent apparent mismatches with preceding overt material by working as a pair, with each ellipsis satisfying identity for the other. This begs the question of how far removed the ellipses can be from the overt material. The case study in argument structure featured differences along the dimensions of voice and ditransitivity. Does this mean that just any structure can be conjured in an ellipsis site, so long as truth-conditional equivalence is preserved? This section answers in the negative by considering the limits of double ellipsis. Double ellipsis is not an all-powerful amelioratory device; instead, there remain important roles for structural identity, lexical availability and parsability.

Beginning with structural identity, consider (21). Manipulating the *spray∼load* alternation places conflicting requirements on the elided structure. Including the prepositions in the sluicing remnants forces the first ellipsis to have the *onto* structure and the second to have the *with* structure. The ensuing non-identity results in ungrammaticality:

(21) * Mary loaded some stuff onto some vehicle, but I don't know
(onto which vehicle, or) with what stuff.
onto which vehicle <she loaded some stuff t>
*≠ or with what stuff <she loaded the vehicle t>

The sensitivity of double ellipsis to the structure of the two ellipsis sites strongly favours the structural view of ellipsis over the anaphora view. The data in the previous section were analysed in structural terms, shoring up the structural view against apparent counterexamples of argument structure mismatch. Those data could, however, be made to submit to an anaphora account. While such accounts usually draw their support from purported examples of grammatical mismatch, it might still be possible to enforce structural identity on the anaphora view by placing limits on where elided meanings can be sourced from; a *pro*-form embedded in passive structure, for

example, might be required to source its meaning from a passive antecedent. Such an analysis would still not capture the positive effects of double ellipsis on the simplest assumption that each *pro* relates back independently to the overt antecedent material; the sourcing constraint would be broken by a second ellipsis just as much as a lone one. But at a further stretch, the *pro* of the second ellipsis could draw on that of the first ellipsis for its meaning, in which case the sourcing constraint might be voided; that is, relating one *pro* to another might give a free pass for identity. In that case, however, all instances of double ellipsis should be able to evade identity requirements equally. That would include (21), where a second *pro* could pick up relevant meaning about 'loading' from a first *pro* – contrary to fact, since (21) is ungrammatical. In sum, only the structural view can handle the fact that structural identity continues to matter to double ellipsis.

Similar applies to lexical identity, which should not matter on the anaphora view. Yet a special case regarding auxiliaries shows that double ellipsis can be limited by morphology. According to Warner's (1985) Auxiliary Verb Generalization, in verb phrase ellipsis headed by an auxiliary verb, the auxiliary must have the exact same morphological form as its antecedent. Warner's Generalization is violated in (22). Disregarding the parenthesised material first, single ellipsis is bad due to the form of the passive auxiliary mismatching between *was* in the antecedent versus *be* in the ellipsis. Adding in the parenthesised intermediate ellipsis is of no help. The mismatch persists with double ellipsis, since the intermediate sluice would also contain *was*, continuing to mismatch with the finally elided *be*:

(22) * She was appointed to the board, even though it was unclear (for how long, or) whether she was eligible to.
for how long <she **was** appointed to the board>
*≠ whether she was eligible to <**be** appointed to the board>

Again, only the structural view can handle the fact that lexical identity in terms of auxiliary morphology matters to double ellipsis.

Turning to lexical availability, it is not possible for entailments from the overt material to support alternative lexicalisations of the ellipsis sites. Take a pair of relational opposites like *beat* and *lose to*.[10] In (23), someone beating Andy at tennis entails that Andy lost to someone at tennis. With double ellipsis, this presents an opportunity to reconcile the preposition *to* in the final sluice remnant by imputing *lose to* into both ellipsis sites. Yet the example remains unacceptable, apparently because *lose to* cannot be conjured from its relational opposite *beat* (*∼):

(23) * Someone beat Andy at tennis, but I don't know (when, or) to whom.
when <he lost to someone at tennis> = to whom <he lost *t* at tennis>
be beaten *∼ *lose to*

[10] Cf. Hartman (2009), who applies relational opposites to argue against licensing verb phrase ellipsis solely semantically by mutual entailment (Merchant 2001) with examples like (i):

(i) * John will beat someone at chess, and then Mary will <lose to someone at chess>.

Similar applies to (24). Were *borrow* made available on the basis of *lend*, identity would be achievable while reconciling *from* in the final sluice remnant. Instead, (24) is bad because relational opposites are not made lexically available by one another:

(24) * Someone lent John £100, but he won't tell me (on what terms, or) from whom.
on what terms <he borrowed £100 from someone>
= from whom <he borrowed £100 *t*> *lend* *∼ *borrow*

The lexical unavailability of relational opposites highlights a directing role for the preceding overt material. The lexical items it makes available cannot be substituted from the ellipsis sites in pursuit of identity. This is so even in (23) and (24) where identity is within reach of an entailed relational opposite, as suggested by the preposition in the ellipsis remnant. In sum, while in the previous section double ellipsis could manipulate argument structure to achieve identity, it cannot switch out the lexical material provided overtly, even in highly favourable circumstances.

A directing role for the overt material is also at play in a final limit on double ellipsis concerning the parsability of their relative order. Where (25) repeats the crucial parts of Nakamura's (2013) original examples from (9), (26) shows that reversing the order to place the passive sluice first is unacceptable:

(25) a. Not so much whether to teach the Bible in public schools, but how? And by whom?

b. . . . they use them from time to time, but they were not sure how much or by whom.

(26) a. * Not so much whether to teach the Bible in public schools, but by whom? And how?

b. * . . . they use them from time to time, but they were not sure by whom or how much.

This unacceptability can be attributed to local, intermediate ungrammaticality with respect to the preceding overt material. In (25), the combination of active overt material and the first sluice is grammatical. Indeed, on first encounter the first sluice will be parsed as active, matching the overt antecedent. As things turn out, the passive second sluice forces reanalysis of the first to be passive, severing the antecedence relationship with the overt material and instead establishing identity between the two ellipses. In (26), on the other hand, the combination of active overt material and passive first sluice is ungrammatical. Plausibly this ungrammatical mismatch causes the parser to give up at **by whom*. Thus whereas (25) can be parsed as grammatical throughout, local yet fatal ungrammaticality arises at the first sluice in (26).

In support of this explanation, the acceptability of the 'passive sluice first' order in (26) improves when local ungrammaticality is not given the chance to arise. In (27), ellipsis is cataphoric, with the active overt material following both ellipses. With two ellipsis sites in hand on encountering the overt material, it is possible to arrive directly at the grammatical parse without an abortive local mismatch. Identity will be satisfiable between the two ellipses, regardless of any structural mismatch with the subsequent overt material:

(27) ? While they weren't exactly sure by whom < > or how often < >,
the company admitted to using extended stay hotels from time to time.

'Backwards' double ellipsis in (27) is accordingly far more acceptable than (26). Similarly, *either* presages the coming of a second clause which might satisfy identity, somewhat improving acceptability in (28):[11]

(28) ? They use them from time to time, but they were not sure
either by whom < > or how much < >.

Overall, double ellipsis is not an all-powerful amelioratory device. While double ellipsis mediates argument structure mismatches that are impossible with single ellipsis, it does so within the bounds of structural identity, lexical availability and parsability. Preceding overt material determines the available lexical material and causes the parser to abort when it mismatches ungrammatically with the first ellipsis. And favouring the structural view over the anaphora view, a single structure must be available for the ellipsis sites that is compatible with surrounding structural constraints and that respects Warner's morphological identity requirements on auxiliaries. The rest of this paper studies three cases where mutual licensing for identity resolves issues in double verb phrase ellipsis.

4 Dahl's many clauses puzzle

Mutual licensing for identity in double ellipsis can explain away Dahl's many clauses puzzle. To set up the puzzle, consider (29), fixing *he* to mean *John*. This meaning can arise in two ways, with the pronoun either referential or bound. This pronominal ambiguity becomes observable under ellipsis, so that even after fixing *he* to mean *John* the single instance of verb phrase ellipsis has two readings. On the so-called 'strict' reading (a), the pronoun is referential, and so continues to point to John in the ellipsis site. On the 'sloppy' reading (b), the pronoun is bound, covarying with the matrix subject. Thus while meaning John in the first clause, *he* switches to mean Sam in the elliptical clause:

[11] On the point of parsability, it could turn out that double ellipsis is a kind of grammatical illusion. Reviewers point to comparatives like (i) and agreement attraction cases like (ii) as examples that are formally ungrammatical yet relatively, or at least initially, acceptable:

(i) * More people have been to Russia than I have.

(ii) * They key to the cabinets were missing.

Comparatives like (i) appear syntactically faultless while turning out on reflection to be semantically meaningless – *more people* calls for a comparison of two sets of individuals, not just *I*. This is a different empirical signature from double ellipsis, where the meaning is perfectly clear.

That might make (ii) with its coherent meaning a more promising analogue. A key feature of agreement attraction is that many speakers accept sentences like (ii) at first glance but later revise their judgement, as found across speeded vs. untimed acceptability judgement experiments (e.g. Phillips et al 2011; Lewis and Phillips 2015; Parker 2019). If double ellipsis is a grammatical illusion along these lines, then it should exhibit a similar 'two-step' acceptability profile, with initial acceptance later retracted. Future research could pursue this prediction, though double ellipsis is stably acceptable in my judgement, an intuition shared with other native English speakers I have informally consulted.

(29) John realises that he$_{(\text{John})}$ is a fool, though Sam doesn't < >.

a. Strict reading, referential (→) pronoun:
John realises that he$_{\rightarrow \text{John}}$ is a fool, though Sam doesn't
<realise that John is a fool>.

b. Sloppy reading, bound (x) pronoun:
John$_x$ realises that he$_x$ is a fool, though Sam$_x$ doesn't
<realise that x=Sam is a fool>.

c. * Disjoint reference:
* John realises that he$_{(\text{John})}$ is a fool, though Sam doesn't
<realise that Bill is a fool>.

Notice that there is no third reading (c) of (29) that takes the pronoun to mean someone else, e.g. Bill, in the elliptical clause after meaning John in the first clause.

Yet double ellipsis supports just such a third reading (Schiebe 1973, via Dahl 1973). Consider (30), which adds an intermediate elliptical clause to (29). Consistently strict (a) and sloppy (b) readings remain available, as before. But now a third reading is also available (c). On this third 'mixed' reading, the pronoun is apparently read sloppily for the first ellipsis, then strictly for the second. The pronoun first covaries with the subject of the intermediate clause before remaining strict in reference to it:

(30) John realises that he$_{(\text{John})}$ is a fool, and Bill does too < >,
though Sam doesn't < >.

a. Both strict: John realises that he$_{\rightarrow \text{John}}$ is a fool,
and Bill does too <realise that John is a fool>,
though Sam doesn't <realise that John is a fool>.

b. Both sloppy: John$_x$ realises that he$_x$ is a fool,
and Bill$_x$ does too <realise that x=Bill is a fool>,
though Sam$_x$ doesn't <realise that x=Sam is a fool>.

c. Mixed reading: John$_x$ realises that he$_x$ is a fool,
sloppy ↗ and Bill$_x$ does too <realise that x=Bill is a fool>,
strict ↗ though Sam doesn't <realise that Bill is a fool>.

This mixed reading presents a problem on the view that each ellipsis needs to independently establish identity with the overt antecedent material. The structure of that overt material must be fixed as either strict (a) or sloppy (b); it cannot oscillate between its ambiguities, as would apparently be needed for (c) (Hardt 2021: 6).

The mixed reading is not a problem, however, on the view here that ellipsis sites can be mutually licensing for identity. As laid out in (31), the 'Bill meaning' of the first ellipsis can be syntactically represented with a referential pronoun rather than binding, since *[λx. x thinks x is a fool](b) = b thinks b is a fool* (cf. Dalrymple et al 1991: 424f.) This representation with a referential pronoun allows for identity with a strict second ellipsis, thereby yielding the otherwise puzzling mixed reading:

(31) and Bill does too <realise that Bill is a fool>
=
though Sam doesn't <realise that Bill is a fool>

Thus the mixed reading follows naturally from the present proposal that identity in double ellipsis can be established between the two ellipses, without formal reference to the preceding overt material. Furthermore, a 'reverse mixed' reading is correctly predicted to be unavailable. Where the mixed reading (c) of (30) appears to proceed from sloppy to strict, it is not possible to mix readings the other way round from strict to sloppy, as in (32). This follows straightforwardly as a failure of identity, since *John* ≠ *Sam*:

(32) * Reverse mixed: John realises that he$_{\rightarrow \text{John}}$ is a fool,
strict ↗ and Bill does too <realise that John is a fool>,
sloppy ↗ though Sam$_x$ doesn't <realise that x=Sam is a fool>.

In sum, the mixed reading of Dahl's many clauses puzzle ceases to be problematic given the consequences of double ellipsis for identity. Just as double ellipsis mediated otherwise impossible argument structure mismatches in sluicing, here it supports referential possibilities that are not available with single ellipsis. The next section applies this perspective on double ellipsis to make some progress on understanding another puzzling case of verb phrase ellipsis in responses to questions.

5 Elliptical answers

This section shows that double ellipsis can resolve issues with certain elliptical answers that are apparently unable to establish identity with their corresponding question. Consider first the polar and subject question-answer pairs in (33)-(34). Verb phrase ellipsis (a) is perfectly possible in the answers compared to fully pronounced controls (b) (SMALL CAPS = focus):

(33) Did John go shopping?
a. He DID <go shopping>.
b. He DID go shopping.

(34) Who went shopping?
a. SAM did <go shopping>.
b. SAM went shopping.

But in answer to an adjunct question (35), verb phrase ellipsis is bad (a); this despite the fully pronounced (b) proving congruence and the intended meaning of the ellipsis being eminently construable from the question:

(35) Where did John go shopping?
a. * He did <go shopping> in PARIS.
b. He went shopping in PARIS.

It is not altogether clear what the problem is with verb phrase ellipsis in answer to adjunct questions. Certainly, (35) poses an acute problem for any attempt to reduce ellipsis licensing entirely to question-answer congruence. At the same time, it is not obvious why structural identity would go awry given the antecedent material available

in the question. See Kuno (1975), Levin (1979) and Stockwell (2020: sect. 5.7) for further discussion.

Most pertinently here, however, ellipsis becomes good in answer to the same question when accompanied by an additional, contrasting elliptical clause in (36):

(36) Where did John go shopping?
He DIDN'T <go shopping> in PARIS; but he DID <go shopping> in LONDON.

The amelioration from (35) to (36) follows from our present perspective on the role of double ellipsis in identity. Now identity can be established between the two ellipses (37), resolving whatever problem obtained with single ellipsis:[12]

(37) he DIDN'T <go shopping> in PARIS = he DID <go shopping> in LONDON

The same holds of alternative questions (38). Whatever the precise problem with verb phrase ellipsis (a) in an otherwise congruent answer (b) (cf. again Kuno 1975, Levin 1979, Stockwell 2020: sect. 5.7), double ellipsis allows for identity to be established between the two ellipses (c):

(38) Did John recommend Mary with a phone call or with a letter?
a. * He did <recommend her> with a LETTER.
b. He recommended her with a LETTER.
c. He DIDN'T <recommend her> with a PHONE CALL;
he DID <recommend her> with a LETTER.

Thus elliptical answers to adjunct and alternative questions present another case of double ellipsis being good where single ellipsis was bad, with the explanation being that the two ellipses license one another for identity. The final section argues that this finding can shed light on some cases of verb phrase ellipsis that apparently lack antecedents altogether.

6 No (overt) linguistic antecedent

Across the preceding case studies in double ellipsis, we have seen that elided material need not establish identity with *overt* linguistic material. Instead, identity can be established between two ellipsis sites, which license one another for identity. That is, each ellipsis site satisfies identity with respect to other, covert elided material.

Still, in the foregoing there has always been an overt antecedent offering a strong clue as to the meaning of the ellipsis sites. Indeed, section 3 identified limiting roles for the overt material in terms of lexical availability and parsing. But when it comes to evaluating identity, the ellipsis sites ultimately do not relate to the overt material, being free to mismatch with it since they establish identity with each other. Given this lack of a formal role for the preceding overt material, we predict that an overt antecedent should not be necessary for double ellipsis. This section delivers on this prediction by examining cases of double ellipsis without an overt linguistic

[12] The elliptical constituents are identical up to focus: DIDN'T vs. DID, PARIS vs. LONDON. In the terminology of Stockwell (2020, 2022), building on Rooth (1992a,b), the elliptical constituents are 'proper alternatives' to each other.

antecedent. With only a non-linguistic scene to make the intended meaning clear, structural identity can continue to be satisfied between the two ellipsis sites.

For the most part, ellipsis is not possible in the absence of preceding overt material. Even with the supportive context of (39), the attempt at ellipsis in (a) fails (Hankamer and Sag 1976: 392, ex. 6):

(39) (Context: Sag produces a cleaver and prepares to hack off his left hand.)

a. Hankamer: # Don't be alarmed . . . he never actually does < >.
'surface' ellipsis

b. Hankamer: Don't be alarmed . . . he never actually does it.
'deep' pro-form

The context in (39) makes abundantly clear that the intended meaning of the ellipsis is *hack off his left hand*. And indeed, with the addition of *it* to yield a *pro*-form in (b), anaphoric resolution to the non-linguistic scene is successful. This is Hankamer and Sag's (1976) distinction between 'deep' and 'surface' anaphora: in contrast to *pro*-forms like *do it*, ellipsis generally requires a linguistic antecedent.

Yet ellipsis is sometimes possible without a linguistic antecedent. This poses a strong challenge to the structural view of ellipsis; if ellipsis needs to check identity with an antecedent, then there had better be an antecedent to begin with. One such circumstance, however, is double ellipsis.[13] Compared to the single ellipsis in (39), antecedent-less double ellipsis succeeds in (40). Against the same context, double ellipsis is good in (a), in minimal contrast with the attempts at single ellipsis in (b):

(40) (Context: Sag produces a cleaver and prepares to hack off his left hand.)

a. He wouldn't, would he? (cf. Jacobson 2022: ex. 21)

b. # He wouldn't. / # Would he?

c. he wouldn't <cut his hand off> = would he <cut his hand off>

The contrast is accounted for on the view here that ellipsis sites can be mutually licensing for identity (c). The meaning of the ellipses can be found in the non-linguistic scene, as for *do it* in (39). There ellipsis failed identity in the absence of matching linguistic material. With double ellipsis, however, identity can be established between

[13] Double ellipsis is far from the only circumstance where verb phrase ellipsis appears to lack a linguistic antecedent. A full defence of the claim that ellipsis requires a linguistic antecedent would need to address all of them, but is beyond the scope of this paper.

For example, in 'split antecedent' cases like (i) (Webber 1978), there is no single antecedent for the apparently elided verb phrase 'sail around the world or climb Kilimanjaro':

(i) Wendy is eager to sail around the world and Bruce is eager to climb Kilimanjaro, but neither of them can < > because money is too tight.

There are also lexicalised idiomatic cases of verb phrase ellipsis that can be uttered in the absence of a linguistic antecedent. The list in (ii) is compiled from Schachter (1977), Hankamer and Sag (1976: 409f., fn. 19), Hankamer (1978: 69) and Pullum (2000):

(ii) Don't! You didn't! You mustn't! I really shouldn't. Oh, you shouldn't have! Shall we? May I? Please do. How could you? Oh no you don't! You wouldn't! Must you? Should I? Not in my X, you don't!

Miller and Pullum (2013) argue that antecedent-less ellipsis is not limited to fixed idioms. They emphasise the role of p versus $\neg p$ alternatives, whether explicitly stated or raised to salience by contexts of permission or direction. For further discussion, see Poppels (2022: sect. 3.2.1).

the two ellipses. Each ellipsis provides non-overt linguistic material to match the other for identity.

The examples in (41) and (42) work similarly.[14] The context again makes clear the intended meaning. Double ellipsis (a) succeeds where a single ellipsis (b) fails, since only double ellipsis provides linguistic material for each ellipsis site to satisfy identity based on the other:

(41) (Tagline of a Clariol hair dye advert.) (Schachter 1977)

a. Does she or doesn't she?

b. # Does she? / # Doesn't she?

c. does she <colour her hair> = doesn't she <colour her hair>

(42) (Context: I see two people clearly thinking about whether to jump into a very cold pool of water at the bottom of a rock formation while hiking. I turn to you and say:) (Jacobson 2022: ex. 19)

a. You know what? I kind of think that he will if she does.

b. # You know what? I kind of think that he will.

c. he will <jump> = she does <jump>

On this view, such ellipses are not formally 'exophoric' (Miller and Pullum 2013) – that is, pointing outside the linguistic discourse to the world. While the clue as to the meaning comes from a non-linguistic scene, identity is resolved internally to the sentence between the two ellipsis sites. Double ellipsis provides matching, if non-overt, linguistic material to establish identity between the ellipses.

This section started from the fact ellipsis is usually bad without an antecedent. Even though a contextual scenario might make clear the intended meaning, a lone

[14] Other examples of apparently missing antecedents offered by Jacobson also involve double ellipsis; e.g. (i) (cf. Jacobson 2009: 86, ex. 4b) and (ii) (cf. Jacobson 2003: 79, ex. 32; 2008: 58, ex. 41a; 2022: ex. 24):

(i) (Context: I see my friend Chris, about to ski down Inferno on cross-country skis, and say:)
He's not really going to <ski down there>, is he <going to ski down there>?

(ii) (I point to one batch of cookies and say:) These, you may <eat *t*>.
(Pointing to a second batch I say:) Those, you can't <eat *t*> – at least not until they cool down.

The second ellipsis in (ii) might not be necessary, per (iii) (Jacobson 2003: 79, ex. 31):

(iii) (Context: I see you about to grab some cookies:) Not those, you don't < >.

Though again, a second ellipsis might still be present by strong implication; compare tailing off at the semicolon in (36). Alternatively, (iii) could be added to the list of lexicalised idioms in note 13; cf. especially *Not in my X, you don't!*
Others of Jacobson's examples lie further beyond reach. Double ellipsis would have to work across speakers to capture (iv) (Jacobson 2022: ex. 20):

(iv) (Scenario: I see my friend Chris at the top of a double diamond ski slope, and I know he is only a beginner skier. You are concerned and say:) Do you think he really might < >?
(I turn to you and say:) No don't worry. I don't think he really will < >.

The modals in (iv) could be crucial, as also in (v) (cf. Jacobson 2022: ex. 25):

(v) (Scenario: Dad is with two of his kids, Keela and Zack. Keela (the older) has been trying to tell Dad for quite some time that she is very independent and doesn't need help tying her shoes. But Dad is a creature of habit, so he reaches down to help Keela – who says:)
Keela: Dad. Please! I DON'T WANT you to < >!!!

ellipsis fails for lack of a linguistic antecedent with which to establish identity. With double ellipsis, however, there is no formal role for the overt linguistic antecedent, and accordingly no need for one. Instead, the two ellipses can be mutually licensing, each providing linguistic material for the other to satisfy identity.

7 Conclusion

This paper has shored up the view that ellipsis sites contain full syntactic structure. Structure can go unpronounced in the presence of an identical antecedent. Without stipulating against it, that antecedent material should itself be able to be unpronounced. I have argued that this situation obtains in 'double ellipsis', where two ellipsis sites can be mutually licensing in satisfying identity for one another. Per the schematic repeated from (5) and (6), a lone ellipsis that fails identity can become possible when bridged by an intermediate ellipsis with which it establishes identity:

(5)
Overt material
$*\neq$
bad <ellipsis>

(6)
Overt material
$\neq$
intermediate <ellipsis>
$=$
previously bad <ellipsis> becomes good

This perspective on double ellipsis accounts for a range of cases that would otherwise stand as counterexamples to the structural view. For clausal ellipsis, it allowed structural identity conditions to be maintained in the face of apparent argument structure mismatches, within the limits of structural identity, lexical availability, and parsability. And for verb phrase ellipsis, it resolved the puzzle of Dahl's mixed reading and explained why double – if not single – ellipsis is possible in response to adjunct and alternative questions. Most extremely, with no formal role for overt antecedent material, structural identity can hold in double ellipsis even without an overt linguistic antecedent.

Acknowledgements. Many thanks to the editor and anonymous reviewers for this journal. My thanks also to Klaus Abels, Gary Thoms, and Danfeng Wu, and to audiences at the University of Geneva, the Linguistics Association of Great Britain 2022, the Oxford Syntax-Semantics workshop, 'You're on Mute!', Tsuda University, and the 42nd West Coast Conference on Formal Linguistics. This paper supersedes Stockwell (2025).

References

Abels K (2018) Movement and islands. In: The Oxford Handbook of Ellipsis. Oxford University Press

Anand P, Hardt D, McCloskey J (2021) The Santa Cruz sluicing data set. Language 97(1):e68–e88

Barros M, Elliot P, Thoms G (2014) There is no island repair, ms., Rutgers, UCL, University of Edinburgh

Chung S (2006) Sluicing and the lexicon: The point of no return. In: Corver R, Kim Y (eds) Proceedings of the Berkeley Linguistics Society 31, pp 73–91

Chung S, Ladusaw W, McCloskey J (1995) Sluicing and logical form. Natural Language Semantics 3(3):239–282

Citko B, Gračanin-Yuksek M (2020) Conjunction saves multiple sluicing: How *(and) why? Glossa 5:92

Culicover P, Jackendoff R (2005) Simpler Syntax. Oxford University Press, Oxford

Dahl O (1973) On so-called 'sloppy identity'. Synthese 26:81–112

Dalrymple M, Shieber S, Pereira F (1991) Ellipsis and higher-order unification. Linguistics and Philosophy 14(4):399–452

Fiengo R, Lasnik H (1972) On nonrecoverable deletion in syntax. Linguistic Inquiry 3:528

Fiengo R, May R (1994) Indices and Identity. MIT Press, Cambridge, MA

Ginzburg J, Sag I (2000) Interrogative investigations. CSLI, Stanford, CA

Hankamer J (1971) Constraints on deletion in syntax. PhD thesis, Yale University, New Haven, CT

Hankamer J (1978) On the nontransformational derivation of some null VP anaphors. Linguistic Inquiry 9:66–74

Hankamer J, Sag IA (1976) Deep and surface anaphora. Linguistic Inquiry 7:391–426

Hardt D (1993) Verb phrase ellipsis: Form, meaning, and processing. PhD thesis, University of Pennsylvania, Philadelphia, PA

Hardt D (2021) Ellipsis and identity. In: Gutzmann D, Matthewson L, Meier C, et al (eds) The Wiley Blackwell Companion to Semantics, First Edition. John Wiley & Sons, Hoboken, NJ, p 1–15

Hartman J (2009) When eGIVENness overpredicts identity, paper presented at Brussels Conference in Generative Linguistics 4, Brussels.

Jacobson P (2003) Binding without pronouns (and pronouns without binding). In: Kruijff GJM, Oehrle RT (eds) Resource–Sensitivity, Binding and Anaphora. Kluwer, Dordrecht, p 57–96

Jacobson P (2008) Direct compositionality and variable–free semantics: The case of antecedent contained deletion. In: Johnson K (ed) Topics in Ellipsis. Cambridge University Press, Cambridge, p 30–68

Jacobson P (2009) Do representations matter or do meanings matter: The case of antecedent containment. In: Hinrichs E, Nerbonne J (eds) Theory and Evidence in Semantics. CSLI Publications, Stanford, CA, p 81–107

Jacobson P (2019) Ellipsis in categorial grammar. In: Oxford Handbook of Ellipsis. Oxford University Press, p 122–141

Jacobson P (2022) No I'm not on mute: I actually didn't say anything, you're on Mute Workshop, 6 May

Jarvis R (2022) Clausal exceptives and exceptive sluices. In: Özge Bakay, Pratley B, Neu E, et al (eds) NELS 52: Proceedings of the Fifty-Second Annual Meeting of the North East Linguistic Society, Volume 2. GLSA, pp 87–98

Kertz L (2013) Verb phrase ellipsis: The view from information structure. Language 89(3):390–428

Kuno S (1975) Conditions for verb phrase deletion. Foundations of Language 13:161–175

Levin NS (1979) Main-verb ellipsis in spoken English. PhD thesis, The Ohio State University, Columbus, OH

Lewis S, Phillips C (2015) Aligning grammatical theories and language processing models. Journal of Psycholinguistic Research 44:27–46. https://doi.org/http://doi.org/10.1007/s10936-014-9329-z

Lobeck A (1995) Ellipsis: Functional heads, licensing and identification. Oxford University Press, Oxford

Merchant J (2001) The syntax of silence: Sluicing, islands, and the theory of ellipsis. Oxford University Press, Oxford

Merchant J (2013) Voice and ellipsis. Linguistic Inquiry 44(1):77–108

Miller P, Pullum GK (2013) Exophoric VP ellipsis. In: Hofmeister P, Norcliffe E (eds) The Core and the Periphery: Data-driven Perspectives on Syntax Inspired by Ivan

A. Sag. CSLI

Nakamura T (2013) Semantic identity and deletion. English Linguistics 30(2):643–658

Palaz B, Bruening B, Tollan R (2025) Sluicing and subject islands: An experimental approach. Glossa 10:1–30

Parker D (2019) Two minds are not always better than one: Modeling evidence for a single sentence analyzer. Glossa: a journal of general linguistics 4(1):64. https://doi.org/https://doi.org/10.5334/gjgl.766

Phillips C, Wagers MW, Lau E (2011) Grammatical illusions and selective fallibility in real-time language comprehension. In: Runner J (ed) Experiments at the interfaces (Vol. 37). Emerald Publications, Bingley, p 147–180, https://doi.org/http://doi.org/10.1108/S0092-4563(2011)0000037009

Poppels T (2022) Explaining ellipsis without identity. The Linguistic Review 39(3):341–400

Pullum G (2000) Hankamer was! In: Chung S, McCloskey J, Sanders N (eds) Jorge Hankamer WebFest, p https://babel.ucsc.edu/jorgewebfest/pullum.html

Rooth M (1992a) A theory of focus interpretation. Natural Language Semantics 1:75–116

Rooth M (1992b) Ellipsis redundancy and reduction redundancy. In: Berman, Hestvik (eds) The Stuttgart Ellipsis Workshop, p SFB 340

Sag I (1976) Deletion and logical form. PhD thesis, MIT, Cambridge, MA

Schachter P (1977) Does she or doesn't she? Linguistic Inquiry 8:763–767

Schiebe T (1973) Zum problem der grammatisch relevanten identität. In: Kiefer F, Ruwet N (eds) Generative Grammar in Europe. D. Reidel Publishing Company, Dordrecht, p 482–527

Stockwell R (2020) Contrast and verb phrase ellipsis: Triviality, symmetry, and competition. PhD thesis, University of California, Los Angeles, CA

Stockwell R (2022) Contrast and verb phrase ellipsis: the case of tautologous conditionals. Natural Language Semantics 30:77–100

Stockwell R (2025) Recoverability and identity are dissociable in double ellipsis. In: Akolkar S, Galvano A, Ismael A, et al (eds) Proceedings of the 42nd West Coast Conference on Formal Linguistics. Cascadilla Proceedings Project, Somerville, MA, pp 364–373

Warner AR (1985) The Structure of English Auxiliaries: A Phrase Structure Grammar. Indiana University Linguistics Club, Bloomington, Ind

Wasow T (1972) Anaphoric relations in English. PhD thesis, MIT, Cambridge, MA

Webber BL (1978) A formal approach to discourse anaphora. PhD thesis, Harvard University, Cambridge, MA, published 1979 by Garland Publishing, New York

Williams E (1977) Discourse and logical form. Linguistic Inquiry 8:101–139