

Case-Mismatching in Urdu Sluicing*

SANA KIDWAI

UNIVERSITY OF CAMBRIDGE

ABSTRACT In this paper, I present novel data from two types of exceptional case-mismatching in Urdu sluicing. In the first, structurally licensed case alternations allow grammatical case-mismatching between the remnant (overt wh-word) and the correlate (phrase in the antecedent which is targeted by the wh-word) but only as far as the meaning remains the same. In the second, using cleft sources as an island evasion strategy results in nominative case on the remnant, leading to mismatching between non-nominative correlates and nominative remnants. The data clearly indicates that syntactic structure must be present at the e(llipsis)-site (the missing material) and that this structure is under semantic rather than syntactic identity with the antecedent. The exact nature of the semantic identity condition remains inconclusive.

1 INTRODUCTION

Sluicing is a type of clausal ellipsis in which only the wh-phrase of a question is pronounced overtly (Ross 1969). Consider (1, 2):

- (1) I'm planning something but I can't tell you what --
- correlate
remnant e-site
- antecedent
sluice

- (2) A: I met someone.
B: Who?

Let us first establish the terminology that will be used in this paper. *Ellipsis site* or *e-site* refers to the missing structure. The *remnant* is the wh-phrase that is pronounced overtly. Together, the remnant and e-site form the *sluice*. The *antecedent* is the clause from which the meaning of the missing structure is derived. Within the antecedent, the *correlate* is the phrase to which the remnant refers.

This paper tackles three interlinked research questions. The primary research question which forms the basis of most sluicing literature is as follows: *is there syntactic structure in the e-site?* In other words, is ellipsis a PF-phenomenon in

* This work is largely a re-write of work supervised by Klaus Abels during my undergraduate degree at University College London. Thanks also to the audience at the COPiL Poster Event 2020 for their feedback.

which syntactic structure is present and simply deleted at PF, or is ellipsis an LF-phenomenon in which there is no syntactic structure to begin with and meaning is recovered through LF-copying? In case of the former, we must then ask: *is this syntactic structure under syntactic or semantic identity with the antecedent?* In other words, is the deleted structure syntactically identical to the antecedent or is it only semantically, but not necessarily syntactically, identical to the antecedent? And finally, *how can we formulate the identity condition?*

It has long been established that sluicing shows strong case-matching effects between the correlate and the remnant (Ross 1969). Various exceptions to this generalisation have been brought forward by the recent literature. In this paper, I present novel data from Urdu sluicing which exhibits two types of exceptional case-mismatching.¹ I show that the data can only be accounted for by a hybrid approach with syntactic structure in the e-site governed by a semantic identity condition. However, although both types of case-mismatching can be neatly explained by a hybrid approach, slightly different assumptions must be made for each. This makes for an interesting discussion when both must be accounted for within the same language.

The structure of the paper is as follows. In section 2, I outline relevant properties of sluicing and the three main approaches that have been taken in the literature. In section 3, I examine case-mismatching due to case alternations, followed by discussion of case-mismatching due to island evasion in section 4. In both types of mismatching, we see that syntactic structure with semantic identity is necessary to explain the full set of facts. Section 5 discusses the implications of this data in constraining and formalising the semantic identity condition which remains an open issue for further research on sluicing.

2 BACKGROUND

In this section, I give some background on sluicing, specifically laying out the two key properties that form the basis of this paper. I then provide an overview of the three main theoretical approaches towards sluicing.

Ross (1969) noted four key properties of sluicing (3). It is the latter two that are of interest in this paper.

- (3) Key properties of sluicing:
- (i) The sluice behaves like a CP.
 - (ii) Remnant size follows from the constraints on regular wh-movement in a language.
 - (iii) There is robust case-matching between the correlate and remnant.

¹ The findings and discussion in this paper may also apply to Hindi. Indeed, the two are often referred to as ‘Hindi-Urdu’ in the literature. I refrain from doing so as the data is very nuanced and likely to be easily affected by dialectal differences. That being said, much of the literature cited is based on Hindi or Hindi-Urdu. Here, I rely on my own native speaker intuition to identify what is applicable to the dialect of Urdu reported in this paper.

- (iv) Sluicing is island-insensitive.

Firstly, in languages with morphological case like Urdu, strong case-matching is seen between the correlate and the remnant. For example, in (4), the remnant must have instrumental case to match the correlate and is ungrammatical with any other case, such as nominative. This is despite the fact that the verb *pata* (know) assigns nominative case to its object (5). This shows that the case on the remnant is not assigned by the verb that embeds the sluice.²

- (4) *Wo kisi=se baath kar- raha he lekin mujhe*
 He.NOM someone.OBL=INS talk do-PROG.M.SG be.PRES.3.SG but I.OBL.DAT
*pata nahi kis=se / *kaun.*
 know.PFV.M.SG not who.OBL=INS / *who.NOM

‘He’s talking to someone but I don’t know who.’

- (5) *Mujhe jawaab / *jawaab=se nahi pata.*
 I.OBL.DAT answer.NOM / *answer=INS not know.PFV.M.SG

‘I don’t know the answer.’

Secondly, sluicing often appears to be island-insensitive. Sentences like (6a) are grammatical, but their non-elided counterparts (6b) are not. This is unsurprising as the non-elided version requires *wh*-movement out of an island, if the sluiced structure is identical to the antecedent.

- (6) a. Alex ate the instant noodles that someone gave me but I don’t remember who.
 b. *Alex ate the instant noodles that someone gave me but I don’t remember who₁ Alex ate the instant noodles that t₁ gave me.

There are three main approaches to sluicing as summarised in Table 1 and discussed below.

	Syntactic structure	Syntactic identity
Purely syntactic approaches	✓	✓
Purely semantic approaches	×	×
Hybrid approaches	✓	×

Table 1 Theories of sluicing.

In his seminal work on sluicing, Ross (1969) argued in favour of a ‘move-and-delete’ approach according to which there is syntactic structure in the *e*-site and it is

² All Urdu examples in this paper are my own and have been corroborated with other native speakers.

syntactically identical to the antecedent. Wh-movement in the sluice results in the wh-word being moved to CP. The TP is then deleted, leaving behind the wh-word as a remnant. The deleted structure is referred to as the *pre-sluice*. Case-matching occurs because the same verb is found in both the antecedent and sluice prior to deletion, and so the same case is assigned to both the correlate and remnant. The real challenge for purely syntactic approaches lies in the apparent island-insensitivity of sluicing. Ross proposed that islands may be a PF-phenomenon, and thus, ‘repair’ is possible through deletion. However, there are attested examples of island-sensitivity in sluicing, specifically in contrast sluicing (7), which indicate that deletion cannot be responsible for island-insensitivity.

- (7) *Alex ate the instant noodles that Li gave me but I don’t remember who else.
Intended: Alex ate the instant noodles that Li gave me but I don’t remember who else gave me instant noodles that Alex ate.

On the other end of the spectrum are purely semantic approaches (Chung, Ladusaw & McCloskey 1995, Culicover & Jackendoff 2005) which posit that there is no syntactic structure in the e-site, i.e. the sluice consists simply of a wh-word and an empty TP, and that meaning is recovered under semantic identity with the antecedent. Case-matching is the result of a case-copying mechanism which copies the case of the correlate onto the remnant. Island-insensitivity in sluicing is unsurprising as there is no structure in the e-site and therefore no island violation to begin with. Once again, the island-sensitivity of some types of sluicing remains unexplained.

Occupying the middle ground are hybrid approaches (Merchant 2001) which argue for syntactic structure in the e-site but under semantic rather than syntactic identity with the antecedent. This allows for paraphrased sources in the sluice (8, c.f. 6b).

- (8) Alex ate the instant noodles that someone gave me but I don’t remember who₁
(t₁ gave me the instant noodles.)

Island-insensitivity is seen due to the availability of paraphrases in the e-site which do not incur island violations as in (8, c.f. 6b). Island-sensitivity is seen when there is no grammatical paraphrase available (see section 4).

Although the hybrid approach has lately received considerable empirical support from cross-linguistic data (Abels 2015, Barros, Elliott & Thoms 2014, Rodrigues, Nevins & Vicente 2009, Rudin 2019), it remains unclear exactly how the semantic identity condition should be formalised. While paraphrases are wanted and even needed in some instances, their availability must be constrained to prevent widespread case-mismatching. I return to this issue in more detail in section 5.

3 CASE-MISMATCHING WITH CASE ALTERNATIONS

Urdu has a highly productive case system (Butt & King 2004, Mohanan 1994). There are seven cases, generally realised as post-nominal clitics. These are summarised in Table 2.

Case	Clitic
Nominative	-∅
Ergative	-ne
Dative	-ko
Accusative	-ko
Instrumental	-se
Genitive	-k-
Locative	-mein/par/tak/∅

Table 2 Urdu case system.

Crucial to this paper is the fact that case alternations are possible in certain environments, i.e. in a given environment, both case A and B are grammatical on the nominal in question. There are four such pairs of alternating cases:

- (9) Case alternation pairs in Urdu:
- (i) Ergative-Nominative
 - (ii) Ergative-Dative
 - (iii) Nominative-Accusative
 - (iv) Accusative-Instrumental

I test for case-mismatching in sluicing by alternating between each pair on the correlate and remnant. In other words, in an environment where both case A and B are grammatical, is it possible to have case A on the correlate and case B on the remnant and vice versa? Purely syntactic approaches predict grammatical mismatching between all four pairs as the same verb is present in both the antecedent and sluice, and the alternation is structurally licensed. Purely semantic approaches do not predict any grammatical mismatching at all since feature-copying from the correlate to the remnant results in strict case-matching only. Finally, hybrid accounts predict structurally licensed case-mismatching to be grammatical but only as far as the meaning of the antecedent and sluice is the same.

3.1 Ergative-Nominative

Urdu is generally classified as a split-ergative language. In transitive clauses, ergative case (-ne) is found on subjects of perfective present and past tense clauses (10a), while nominative case (-∅) is found on subjects elsewhere (10b, 10c), barring dative subjects.

- (10) a. *Sana=ne chawal khaiy.*
 Sana=ERG rice.NOM eat.PFV.M.PL
 ‘Sana ate (the) rice.’
- b. *Sana chawal khaiy-gi.*
 Sana.NOM rice.NOM eat.PFV.3.SG-FUT.F.SG
 ‘Sana will eat (the) rice.’
- c. *Sana chawal khathi he.*
 Sana.NOM rice.NOM eat.IPFV.F.SG be.PRES.3.SG
 ‘Sana eats rice.’

There is also a subset of unergative verbs that allow ergative subjects under the same tense-aspect conditions as transitive verbs. There is some speaker variation regarding the acceptability of ergative case in intransitive clauses, but for those speakers who accept it, it is in this environment that ergative case alternates with nominative case (11).

- (11) *Omar / Omar=ne chillaya.*
 Omar.NOM / Omar=ERG yell.PFV.M.SG
 ‘Omar yelled.’

Ergative case is generally associated with volitionality or agency. Butt & King (2004) cite work by Bashir (1999) on use of ergative case in Urdu TV drama. They establish that ergative case has no semantic contribution when it is obligatory, as in transitive clauses, but is associated with volitionality when it is optional, as in unergative clauses. The semantic contribution of ergative case becomes more apparent with an appropriate modifier. For example, using ergative case with *ghalti=se* (‘by mistake’) is significantly worse than using nominative case (12). This supports the idea that ergative subjects are interpreted as having control over the event.

- (12) *Omar / ?Omar=ne ghalthi=se chillaya.*
 Omar.NOM / ?Omar=ERG mistake=INS yell.PFV.M.SG
 ‘Omar yelled by mistake.’

We now move onto the sluicing data. The ergative-nominative alternation does not allow grammatical case-mismatching in either direction.

- (13) a. *Koi khaansa tha lekin mujhe*
 Someone.NOM cough.PFV.M.SG be.PST.M.SG but I.OBL.DAT
*pata nahi kaun / *kis=ne.*
 know.PFV.M.SG not who.NOM / *who.OBL=ERG
 ‘Someone coughed but I don’t know who.’

- b. *Kisi=ne khaansa tha lekin mujhe*
 Someone.OBL=ERG cough.PFV.M.SG be.PST.M.SG but I.OBL.DAT
*pata nahi *kaun / kis=ne.*
 know.PFV.M.SG not *who.NOM / who.OBL=ERG

‘Someone coughed but I don’t know who.’

This is expected under the hybrid approach because the semantic identity condition is not satisfied in sluices with ergative-nominative mismatch due to the different semantic contributions of the two cases.

3.2 Ergative-Dative

Ergative (-ne) and dative (-ko) case alternate on subjects of infinitival *be* clauses (14).

- (14) *Sana=ne / Sana=ko paRhai karni he.*
 Sana=ERG / Sana=DAT study.NOM do.INF.F.SG be.PRES.3.SG

‘Sana has to study.’

Based on the discussion in the previous section, ergative case should have a semantic contribution when used in this environment as it is optional. Butt & King (2004) identify the ergative-dative alternation as ‘wants-to vs must-do.’ This can be illustrated through use of a modifier such as *zabardasti* (‘forcefully/unwillingly’) which clashes with the semantic interpretation of ergative case.

- (15) *Kisi=ko / ?kisi=ne zabardasti school*
 Someone.OBL=DAT / ?someone.OBL=ERG unwillingly school.OBL.LOC
jaana he.
 go.INF be.PRES.3.SG

‘Someone has to go to school unwillingly.’

Butt and King go on to comment that dative case is unmarked in this environment so while the ergative subject always has control over the action, the dative subject may or may not.³ So while there is a distinction in the meaning of ergative and dative case, there is an overlap for at least some speakers. The distinction may also be more or less rigid for different speakers.

The hybrid account then predicts that ergative-dative case-mismatching in sluicing should be acceptable for speakers such as myself who do not associate mutually exclusive semantics with this case pair. As shown below (16), mismatching is acceptable: the remnant can have either ergative or dative case regardless of which case is found on the correlate.

³ Davison (2015) claims the opposite: dative subjects are consistent with non-volitional readings whereas ergative subjects are interpretable either way. Regardless of the correct semantic contributions of each of the cases, it is clear that there is an overlap in meaning between ergative and dative case.

- (16) *Kisi=ne* / *kisi=ko* *school* *jaana he*
 Someone.OBL=ERG / someone.OBL=DAT school.OBL.LOC go.INF be.PRES.3.SG
lekin mujhe pata nahi kis=ne / kis=ko /
 but I.OBL.DAT know.PFV.M.SG not who.OBL=ERG / who.OBL=DAT /
kissey.
 who.OBL.DAT

‘Someone has to go to school but I don’t know who.’

Firstly, it should be noted that despite the grammaticality of case-mismatching, case-matching is still the preferred option (see [section 5](#)). Secondly, there are some factors which can improve the acceptability of mismatching although it is unclear why. For example, using the fused dative pronoun (*kissey*)⁴ rather than the overt dative case marker (*-ko*) improves the mismatch significantly. In addition, the nature of the predicate embedding the sluice also seems to affect acceptability, although I have yet to find a predicate which disallows the mismatch entirely. Mismatching with simplex verbs (e.g. *pata* ‘know’) is straightforward (16). Mismatching with complex N-V (e.g. *faraq-paRhna* ‘effect-fall’) or V-V (e.g. *bathaya-gaya* ‘told-went’) predicates is borderline acceptable but improves significantly by using the fused dative pronoun or by including focus markers in the antecedent (17).

- (17) a. *Kisi=ne* / *kisi=ko* (to) *school* *jaana*
 Someone.OBL=ERG / someone.OBL=DAT (FOC) school.OBL.LOC go.INF
 (hi) *he* *lekin mujhe faraq nahi paRh-raha*
 (FOC) be.PRES.3.SG but I.OBL.DAT effect not fall-PROG.M.SG
kis=ne / kis=ko / kissey.
 someone.OBL=ERG / someone.OBL=DAT / someone.OBL.DAT

‘Someone has to go to school (for sure) but I don’t care who.’

- b. *Kisi=ne* / *kisi=ko* (to) *school* *jaana*
 Someone.OBL=ERG / someone.OBL=DAT (FOC) school.OBL.LOC go.INF
 (hi) *he* *lekin abhi tak bathaya nahi gaya*
 (FOC) be.PRES.3.SG but now till tell.PFV.M.SG not go.PST.M.SG
kis=ne / kis=ko / kissey.
 who.OBL=ERG / who.OBL=DAT / who.OBL.DAT

‘Someone has to go to school for sure but we haven’t yet been told who.’

On the other hand, mismatching can be made worse by making the semantic information of the case markers more salient. In (18), the modifier *zabardasti* (‘forcefully/unwillingly’) is used in the antecedent with a dative correlate, highlighting its ‘must-do’ interpretation. Mismatching with an ergative remnant is now infelicitous as there is no longer an overlap between the semantics of the correlate and remnant.

⁴ Pronouns marked with accusative or dative case can be realised one of two ways: pronoun + *-ko* (e.g. *mujh=ko* ‘1.SG=ACC/DAT’) or a fused form (e.g. *mujhe* ‘1.SG.ACC/DAT’) (Butt & King 2004).

- (18) *Kisi=ko zabardasti school jaana he lekin*
 Someone.OBL=DAT unwillingly school.OBL.LOC go.INF be.PRES.3.SG but
mujhe pata nahi #kis=ne / kis=ko /
 I.OBL.DAT know.PFV.M.SG not #who.OBL=ERG / who.OBL=DAT /
kissey.
 who.OBL.DAT

‘Someone has to go to school unwillingly but I don’t know who.’

Therefore, case-mismatching in sluicing is grammatical with this case alternation pair, as long as the semantic information of the two cases can overlap in the given sentence.

3.3 Nominative-Accusative

The nominative-accusative alternation (\emptyset vs *-ko*) in Urdu is the result of differential-object-marking (DOM) on direct objects.

- (19) *Sana kelay / kelon=ko khaey-gi.*
 Sana.NOM bananas.NOM / bananas.OBL=ACC eat.PFV.3.SG-FUT.F.SG

‘Sana will eat (the) bananas.’

DOM in Urdu is conditioned by definiteness and specificity. Butt (1993) shows that marked nominals receive only a specific interpretation. She gives the example below to illustrate this. (20a) provides a context which is compatible only with a non-specific interpretation. Using accusative case (20c) is infelicitous showing that it is incompatible with a non-specific interpretation.

- (20) a. *Adnan aaj raat=ke saalan=ke =liye*
 Adnan.NOM today night.OBL=GEN.OBL curry.OBL=GEN.OBL =LOC
murghi chahtha tha.
 chicken.NOM want.IPFV.M.SG be.PST.M.SG
 ‘Adnan wanted chicken for tonight’s curry.’
- b. *Us=ke khaansame=ne bazaar=se murghi*
 He.OBL=GEN.OBL cook.OBL=ERG market.OBL=INS chicken.NOM
kharidi.
 buy.PFV.F.SG
 ‘His cook bought chicken from the market.’
- c. *#Us=ke khaansame=ne bazaar=se murghi=ko*
 He.OBL=GEN.OBL cook.OBL=ERG market.OBL=INS chicken.OBL=ACC
kharida.
 buy.PFV.M.SG
 ‘His cook bought a particular chicken from the market.’

On the other hand, unmarked nominals are compatible with both specific and non-specific interpretations. The non-specific interpretation has already been shown in (20b). Dayal (2003) argues that bare nominals in Urdu can give rise to specific interpretations when they refer to a contextually available antecedent. This is illustrated in (21). The context supports a specific interpretation of *guRiya* (doll), i.e. the doll under discussion. Although there is a preference to use accusative case (21a), nominative case is also acceptable (21b).

(21) Maariyah is telling her friends about her new doll which she then had to give away. Her friends are discussing whether or not they saw the doll before she gave it away. One of her friends says:

- a. *Mein=ne guRiya=ko dekha tha.*
 I.OBL=ERG doll.OBL=ACC see.PFV.M.SG be.PST.M.SG

‘I saw the doll.’

- b. *Mein=ne guRiya dekhi thi.*
 I.OBL=ERG doll.NOM see.PFV.F.SG be.PST.F.SG

‘I saw the doll.’

We can now identify an overlap in the semantics of nominative and accusative case in this context, similar to the overlap between ergative and dative case discussed in the previous section. The contextually salient antecedent in sluicing should give rise to specific interpretations of nominative remnants, overlapping with the specific interpretation of accusative case. Once again, grammatical case-mismatching is acceptable. ‘Which-NP’ type sluices are used as they directly inquire about specific objects, further prompting a specific reading of the bare nominal in question.

- (22) a. *Us=ne kisi khelonay=ko toRa tha*
 He/she.OBL=ERG some.OBL toy.OBL=ACC break.PFV.M.SG be.PST.M.SG
lekin mein=ne dekha nahi kaunse khelonay=ko /
 but I.OBL=ERG see.PFV.M.SG not which.OBL.M.SG toy.OBL=ACC /
kaunsa khelona.
 which.M.SG toy.NOM

‘He/she broke some toy but I didn’t see which toy.’

- b. *Us=ne koi khelona toRa tha lekin*
 He/she.OBL=ERG some toy.NOM break.PFV.M.SG be.PST.M.SG but
mein=ne dekha nahi ?kaunse khelonay=ko /
 I.OBL=ERG see.PFV.M.SG not which.OBL.M.SG toy.OBL=ACC /
kaunsa khelona.
 which.M.SG toy.NOM

‘He/she broke some toy but I didn’t see which toy.’

Note that mismatching from a nominative correlate to an accusative remnant (22b) is worse than the opposite (22a). One possible explanation is that the specific reading of the bare antecedent is not salient enough. This seems plausible as mismatching is improved by using the more specific *koi eik NP* (some one NP / a certain NP) frame for the correlate (23).

- (23) *Us=ne koi eik khelona toRa tha lekin*
 He/she.OBL=ERG some one toy.NOM break.PFV.M.SG be.PST.M.SG but
mein=ne dekha nahi kaunse khelonay=ko / kaunsa
 I.OBL=ERG see.PFV.M.SG not which.OBL.M.SG toy.OBL=ACC / which.M.SG
khelona.
 toy.NOM

‘He/she broke a certain toy but I didn’t see which toy.’

The acceptability of this type of case-mismatching in Urdu sluicing has also been noted by Bagasur (2014). They also mention that mismatching in the accusative-nominative direction is better than the nominative-accusative direction. They explain this by claiming that mismatching from all non-nominative correlates to nominative remnants is possible but not vice versa. I do not agree with their judgments; for example, mismatching from ergative correlates to nominative remnants is unacceptable for me (section 3.1). However, I return to Bagasur’s idea in section 5.

Finally, using focus markers once again improves mismatching, including mismatching from a nominative correlate to an accusative remnant.

- (24) *Us=ne koi khelona / kisi khelonay=ko to*
 He/she.OBL=ERG some toy.NOM / some.OBL toy.OBL=ACC FOC
toRa hi tha lekin mein=ne dekha nahi
 break.PFV.M.SG FOC be.PST.M.SG but I.OBL=ERG see.PFV.M.SG not
kaunsa khelona / kaunse khelonay=ko.
 which.M.SG toy.NOM / which.OBL.M.SG toy.OBL=ACC

‘He/she broke some toy for sure but I didn’t see which toy.’

To summarise, case-mismatching between nominative and accusative case is grammatical because both allow specific interpretations. Mismatching from accusative case to nominative case is better than the opposite although this can be improved by making the specific reading of nominative case more prominent or by using focus markers. This asymmetry is discussed further in section 5.

3.4 Accusative-Instrumental

Accusative (-*ko*) and instrumental case (-*se*) alternate on causees of some causative verbs.

- (25) *Sana=ne Omar=ko / Omar=se khaana chakh-vaya.*
 Sana=ERG Omar=ACC / Omar=INS food.NOM taste.PFV-CAUS.M.SG
 ‘Sana made Omar taste the food.’

Saksena (1980, 1982) provides a detailed discussion of this alternation and shows that it signals a contrast between target and non-target semantics. Most causative verbs select either an accusative or an instrumental causee in which case the two are in complementary distribution. Accusative case is consistently found on *affected agents*, i.e. agents which are affected by the activity; for example, the act of drinking affects the agent (26a). Instrumental case, on the other hand, is found on agents that are not affected by the event they participate in; for example, the act of cutting does not affect the agent but rather the patient (26b).

- (26) a. *Sana=ne Omar=ko / *Omar=se paani pil-aya.*
 Sana=ERG Omar=ACC / *Omar=INS water.NOM drink.PFV-CAUS.M.SG
 ‘Sana made Omar drink water.’
 b. *Sana=ne Omar=se / *Omar=ko peR kat-vaya.*
 Sana=ERG Omar=INS / *Omar=ACC tree.NOM cut.PFV-CAUS.M.SG
 ‘Sana made Omar cut the tree.’

Saksena argues that accusative causees are semantically complex: they have both an agent and patient interpretation. The latter allows them to be interpreted as targets or affected agents. This is reflected in the negation facts. In (27), the second clause is the non-causative counterpart of the first and is negated. In (27a), the first clause contains an accusative causee and negation of the second clause is felicitous. Saksena explains this by claiming that the first clause refers to the causee’s patient interpretation while the second negates its agent interpretation. Therefore, there is no contradiction. In contrast, instrumental causees do not allow negation of the non-causative clause as they only have an agentive interpretation and so negation is contradictory (27b).

- (27) a. *Sana=ne Omar=ko paani pil-aya lekin*
 Sana=ERG Omar=ACC water.NOM drink.PFV-CAUS.M.SG but
us=ne nahi piya.
 he.OBL=ERG not drink.PFV.M.SG
 ‘Sana made Omar drink water but he didn’t drink it.’
 b. #*Sana=ne Omar=se peR kat-vaya lekin us=ne*
 Sana=ERG Omar=INS tree.NOM cut.PFV-CAUS.M.SG but he.OBL=ERG
nahi kaata.
 not cut.PFV.M.SG
 ‘Sana made Omar cut the tree but he didn’t cut it.’

Accusative and instrumental case alternate on causees of verbs which are compatible with both target and non-target semantics. The alternation makes use of the contrastive semantics of the two cases. When the agent carries accusative case, the objective of the action is interpreted as completion of the activity by the agent. This is expected as the agent is the target of the action. Conversely, when the agent carries instrumental case, the objective is only completion of the activity with the agent merely being an instrument in achieving this.

Considering the clearly distinct semantic contributions of the two cases in this environment, it is predicted that this alternation will not induce case-mismatching in any type of sluicing. This is indeed the case (28).

- (28) a. *Sana=ne kisi=ko khaana chakh-vaya lekin*
 Sana=ERG someone.OBL=ACC food.NOM taste.PFV-CAUS.M.SG but
*mujhe pata nahi kis=ko / *kis=se.*
 I.OBL.DAT know.PFV.M.SG not who.OBL=ACC / *who.OBL=INS
 ‘Sana made someone taste the food but I don’t know who.’
- b. *Sana=ne kisi=se khaana chakh-vaya lekin*
 Sana=ERG someone.OBL=INS food.NOM taste.PFV-CAUS.M.SG but
*mujhe pata nahi *kis=ko / kis=se.*
 I.OBL.DAT know.PFV.M.SG not *who.OBL=ACC / who.OBL=INS
 ‘Sana made someone taste the food but I don’t know who.’

Once again, a difference in semantic information prevents case-mismatching in sluicing.

3.5 Interim Summary

Based on the data and discussion above, we see that case-mismatching is grammatical in Urdu sluicing and that it is highly constrained. Out of the four case alternation pairs, only two allow mismatching under sluicing (Table 3).

Case Alternation Pair	Case-Mismatching
Ergative-Nominative	×
Ergative-Dative	✓
Nominative-Accusative	✓
Accusative-Instrumental	×

Table 3 Case-mismatching in Urdu sluicing.

The pattern is very clear. Case-mismatching is only allowed where the semantic contribution of the case on the remnant is interpreted as being the same as that of the case on the correlate. Ergative-nominative and accusative-instrumental alternations

have distinctly different semantic content and disallow mismatching. The distinction between the contribution of the ergative-dative pair is not as rigid and we find grammatical mismatching. Similarly, nominative-accusative cases have an overlap in their semantic information and also allow grammatical case-mismatching.

We can identify the following two constraints on grammatical case-mismatching in Urdu sluicing:

- (29) Constraints on grammatical case-mismatching in Urdu sluicing:
- a. The verb in the antecedent (and the sluice) licenses both the cases found on the correlate and the remnant.
 - b. The case-marking on the correlate and remnant has the same overall meaning.

Purely syntactic accounts are easily able to capture the first constraint as there is structure in the e-site and it is identical to the antecedent under syntactic identity. It follows naturally from this that the case of the remnant must at least be licensed by the verb in the antecedent for mismatching to be possible as the same verb is found in the sluice. The problem here is that such approaches do not cover the constraint on the semantic contribution of the two cases, consequently incorrectly predicting that all four of the case alternation pairs should give rise to grammatical case-mismatching in sluicing. However, as discussed, the mismatching is extremely sensitive to the meaning of the case markers. Thus, syntactic accounts predict more case-mismatching than is found.

On the other hand, purely semantic accounts of sluicing predict less case-mismatching than is found. While these approaches can account for the second constraint (29b) through the semantic identity condition, they cannot account for the first constraint (29a) as they cannot refer to syntactic structure. More importantly, as mentioned earlier, these accounts resort to a case-copying mechanism to derive case connectivity; thus, such approaches do not predict case-mismatching at all and would have to make arbitrary stipulations to capture the data.

Conversely, hybrid approaches are able to account for both constraints. The first constraint follows from the presence of syntactic structure in the e-site, while the second follows from the semantic identity condition. This predicts exactly what we find in Urdu sluicing: case-mismatching is allowed between pairs licensed by the same verb as long as the semantic content of the antecedent and the sluice is identical.

Finally, it should be noted that a significant amount of speaker variation is expected with these judgements. The semantic contribution of the case markers can at times be quite subtle and difficult to pinpoint. It may also be more or less salient for individual speakers. A systematic survey of the semantic contribution of the Urdu case markers is needed. In general, the prediction is that speakers who have overlapping semantics for structurally licensed case pairs will allow grammatical case-mismatching in sluicing; speakers who do not have overlapping semantics will not.

4 CASE-MISMATCHING IN ISLAND REPAIR

In this section, I discuss the second type of case-mismatching in Urdu sluicing. I first summarise the island evasion strategies proposed by the hybrid approach (section 4.1), and then present supporting evidence from Urdu sluicing (section 4.2). As we shall see, grammatical case-mismatching between a non-nominative correlate and a nominative remnant is a clear diagnostic for the presences of a copular verb in the e-site. In fact, case-matching is categorically ungrammatical in these sentences, reinforcing the need for cleft sources. We also see that case-mismatching is not allowed when an isomorphic or short source is available. This suggests that there may be some sort of hierarchical availability of paraphrases which I return to in section 5.

4.1 Island Evasion Strategies

Under the hybrid approach, island-insensitivity in sluicing is due to the availability of alternative paraphrased sources. Barros et al. (2014) identify three types of paraphrases of which I discuss two in this paper: short and truncated cleft sources.

- (30) Alex ate the instant noodles₁ that someone gave me t₁ but I don't remember who₂ ...
- a. *〈 Alex ate the instant noodles₁ that t₂ gave me t₁. 〉 - isomorphic source
 - b. 〈 t₂ gave me the instant noodles. 〉 - short source
 - c. 〈 it was t₂. 〉 - truncated cleft source

The isomorphic source is syntactically identical to the antecedent and, in these instances, contains an island violation. The fully pronounced version of the sentence with the isomorphic source is ungrammatical. Unlike isomorphic sources, paraphrased sources do not incur island violations. It is, therefore, more appropriate to speak of *island evasion* rather than island repair, as there are no islands to repair if paraphrased sources are used.

Short sources were first proposed by Merchant (2001) as evasion strategies for islands which form propositional domains. Short sources select this smaller domain instead of the entire matrix clause as their antecedent; for example, in (30), the antecedent of the short source is *someone gave me t₁* rather than *Alex ate the instant noodles that someone gave me t₁*. Independent evidence for the availability of short sources comes from cases such as (31) where only the short source is compatible with the correct reading (Barros et al. 2014, Merchant 2001).

- (31) We need to know what he is doing, and why ...
- a. 〈 he is doing it. 〉
 - b. #〈 we need to know what is doing. 〉 (Merchant 2001: 201-207)

Truncated cleft sources consist of an expletive pronoun, *it*, followed by the copula, *be*, and the XP pivot, *t*. The relative clause that modifies the pivot in a full cleft is

absent in truncated clefts.⁵ The term *cleft sources* will be used to refer to truncated cleft sources in the rest of this paper. Evidence for cleft sources comes from p-or-q sluices in which the antecedent is a disjunction of propositions (32) (Barros et al. 2014). The only plausible source for such sluices is a cleft source, which is further supported by the fact that such sluices are only found in languages with cleft constructions.

- (32) Either something's on fire or Sally's baking a cake, but I don't know which₁
 ⟨ it is t₁. ⟩

Neither short nor cleft sources contain islands. Using either of these paraphrases in the sluice results in an apparent island-insensitivity when in fact there is no island violation in the first place. Conversely, when no paraphrase is available, we predict that sluicing should be ungrammatical if paraphrases are indeed responsible for the island-insensitivity of sluicing. This is correct as shown in the following section.

4.2 Evidence for Island Evasion from Urdu

The literature on Hindi-Urdu sluicing gives somewhat mixed reports for island violations in sluicing. Bhattacharya & Simpson (2012) obtained judgements from native speaker linguists for sluicing out of complex NPs, adjuncts, wh-islands and coordinated-DP structures. The results of this investigation were inconclusive. Gribanova & Manetta (2016) conducted the same survey with 10 non-linguist native speakers. They made slight changes to Bhattacharya and Simpson's data, using more colloquial vocabulary and providing contexts for each test sentence. They also included sentences with island violations but no sluicing and sentences with sluicing but no island violations to establish a baseline. They reported that sluicing out of complex NPs and coordinated-DPs was deemed acceptable by at least 8 out of 10 speakers. Bagasur (2014) also reported grammatical sluicing out of these islands as well as adjunct islands based on their own native speaker intuitions. My own judgement is also that sluicing out of islands is acceptable in Urdu.

In addition, previous accounts of Urdu sluicing (e.g. Bhattacharya & Simpson 2012, Gribanova & Manetta 2016, Manetta 2013) have proposed a move-and-delete analysis along the lines of Ross (1969), and have noted that Urdu sluices generally pattern with wh-questions as opposed to cleft sources, indicating that clefts are not the default pre-sluice in Urdu. Therefore, any account of Urdu sluicing which posits syntactic structure in the e-site faces the island repair problem.

In this section, I present new evidence for island evasion from Urdu sluicing. Using case to probe the nature of the e-site, I show that the island-insensitivity of sluicing can be directly attributed to the availability of a grammatical paraphrase. The predictions are as follows. Purely syntactic approaches claim that islands are a PF-phenomenon and therefore can be repaired through deletion. Consequently, all sluicing should be island-insensitive. Moreover, there should be strict case-matching

⁵ See Barros et al. (2014) for discussion on full cleft vs truncated cleft sources.

across the board as the same verb is found in the antecedent and sluice and so the correlate and remnant should always receive the same case. Purely semantic approaches make the same predictions but for different reasons. All sluicing should be island-insensitive as there is no structure in the e-site and therefore no island violations. There should be strict case-matching as case is copied from the correlate to the remnant. Finally, hybrid approaches predict that sluicing should be island-insensitive when there is a source available that does not incur island violations; when there is no such source available, sluicing should be island-sensitive. Case-mismatching should be grammatical where the verb in the paraphrase assigns a different case from the verb in the antecedent. This is particularly relevant in identifying cleft sources. The copular verb in Urdu assigns nominative case. Since Urdu generally exhibits strong case-matching effects in sluicing (e.g. [Bagasur 2014](#), [Bhattacharya & Simpson 2012](#)), grammatical case-mismatching between a non-nominative correlate and nominative remnant can be directly attributed to the presence of the copular verb in the pre-sluice.⁶

4.2.1 Short sources (relative clause, CP-complement, coordinated-CPs)

Sluicing out of relative clauses (33a), CP-complements (33b) and coordinated-CPs (33c) is grammatical in Urdu.

- (33) a. *Sana kisi aisay bandey=ko nokri*
 Sana.NOM some.OBL like.this.OBL person.OBL=ACC job.NOM
dena chahthi he jo kisi
 give.INF.M.SG want.IPFV.F.SG be.PRES.3.SG REL.NOM some.OBL
Pakistani zabaan=se waqif ho lekin mujhe
 Pakistani language.OBL=INS familiar be.FUT.3.SG but I.OBL.DAT
pata nahi kaunsi Pakistani zabaan=se /
 know.PFV.M.SG not which.OBL.F.SG Pakistani language.OBL=INS /
zabaan.
 language.NOM

‘Sana wants to give a job to someone who is familiar with a Pakistani language but I don’t know which Pakistani language.’

- b. *Sana=ne mujhe bathaya tha keh Hira*
 Sana=ERG I.OBL.DAT told.PFV.M.SG be.PST.M.SG that Hira.NOM
kisi=se naraz he lekin mujhe yaad nahi
 someone.OBL=INS angry be.PRES.3.SG but I.OBL.DAT memory not
*kis=se / *kaun.*
 who.OBL=INS / *who.NOM

‘Sana told me that Hira is angry at someone but I don’t remember who.’

⁶ Verbs that license multiple cases - and therefore may allow case-mismatching for other reasons (see [section 3](#)) - have been omitted from the data in this section to ensure that any case-mismatching seen is only due to the presence of the copular verb.

- c. *Sana=ne ghar saaf kya aur kisi=se baath*
 Sana=ERG house.NOM clean do.PFV.M.SG and someone.OBL=INS talk
ki lekin mujhe pata nahi kis=se /
 doPFV.F.SG but I.OBL.DAT know.PFV.M.SG not who.OBL=INS /
 **kaun.*
 *who.NOM

‘Sana cleaned the house and spoke to someone but I don’t know who.’

In each of these sentences, the isomorphic source contains an island violation but there is a smaller propositional domain that can be used as an antecedent by a short source. Case-mismatching is ungrammatical in (33b, 33c) indicating that a cleft source is not used. (Grammatical case-mismatching in (33a) is discussed in section 4.2.2.)

Moreover, contrast sluicing out of these islands is also grammatical (34), contra the general island-sensitivity of contrast sluicing.

- (34) a. *Sana kisi aisay bandey=ko nokri*
 Sana.NOM some.OBL like.this.OBL person.OBL=ACC job.NOM
dena chahthi he jo Urdu=se
 give.INF.M.SG want.IPFV.F.SG be.PRES.3.SG REL.NOM Urdu=INS
waqif ho lekin mujhe pata nahi aur
 familiar be.FUT.3.SG but I.OBL.DAT know.PFV.M.SG not and
kaunsi Pakistani zabaan=se < wo waqif
 which.OBL.F.SG Pakistani language.OBL=INS < he/she.NOM familiar
ho. >
 be.FUT.3.SG >

‘Sana wants to give a job to someone who is familiar with Urdu but I don’t know which other Pakistani language < he/she should be familiar with. >’

- b. *Sana=ne mujhe bathaya tha keh Hira*
 Sana=ERG I.OBL.DAT told.PFV.M.SG be.PST.M.SG that Hira.NOM
Nabeel=se naraz he lekin mujhe yaad nahi aur
 Nabeel=INS angry be.PRES.3.SG but I.OBL.DAT memory not and
kis=se < wo naraz he. >
 who.OBL=INS < Hira.nom angry be.PRES.3.SG >

‘Sana told me that Hira is angry at Nabeel but I don’t remember who else < Hira is angry at. >’

- c. *Sana=ne ghar saaf kya aur Omar=se baath*
 Sana=ERG house.NOM clean do.PFV.M.SG and Omar=INS talk
ki lekin mujhe pata nahi aur kis=se
 do.PFV.F.SG but I.OBL.DAT know.PFV.M.SG not and who.OBL=INS
 ⟨ *Sana=ne baath ki.* ⟩
 ⟨ Sana=ERG talk do.PFV.F.SG ⟩
- ‘Sana cleaned the house and spoke to Omar but I don’t know who else
 ⟨ Sana spoke to. ⟩’

The grammaticality of contrast sluicing can be attributed to the grammaticality of the short source in these cases. This supports the idea that the island-(in)sensitivity of various types of sluicing can be explained by the availability of grammatical paraphrases rather than by stipulating that merger sluicing is island-insensitive while contrast sluicing is island-sensitive.

4.2.2 Cleft sources (relative clause, coordinated-DPs, adjunct island, wh-island)

Sluicing out of coordinated-DPs (35a), adjuncts (35b) and wh-islands (35c) is also grammatical in Urdu.

- (35) a. [*Omar aur kisi laRke*] =*ki kal laRai*
 [Omar and some.OBL boy.OBL] =GEN.F.SG yesterday fight
hui thi lekin mujhe pata nahi
 happen.PST.F.SG be.PST.F.SG but I.OBL.DAT know.M.SG not
*kaunsa laRka / *kis laRke=ki.*
 which.M.SG boy.NOM / *which.OBL boy.OBL=GEN.F.SG
- ‘Omar and some boy had a fight yesterday but I don’t know which boy.’
- b. *Omar naraz ho-ga agar Sana kisi*
 Omar.NOM angry be.FUT.3.SG-FUT.M.SG if Sana.NOM some.OBL
laRke=se baath karay-gi lekin humein
 laRke.OBL=INS talk do.FUT.3.SG-FUT.F.SG but we.OBL.DAT
*yaad nahi kaunsa larka / *kis laRke=se.*
 remember not which.M.SG boy.NOM / *which.OBL boy.OBL=INS
- ‘Omar will be angry if Sana talks to some boy but we don’t remember which boy.’

- c. *Mein dekh-rahi thi keh kaunsa laRka*
 I.NOM watch-PROG.F.SG be.PST.F.SG that which.M.SG boy.NOM
kisi laRki=se baath karay-ga lekin
 which.OBL girl.OBL=INS talk do.FUT.3.SG-FUT.M.SG but
*mein=ne bathaya nahi kaunsi laRki / *kis*
 I.OBL=ERG tell.PFV.M.SG not which.F.SG girl.NOM / *which.OBL
laRki=se.
 girl.OBL=INS

‘I was watching which boy would talk to some girl but I didn’t say which girl.’

Once again, the isomorphic source is ungrammatical in these cases. Short sources are unavailable as these islands are not propositional. We also see that case-matching on the remnant is ungrammatical, giving further proof that neither an isomorphic nor short source is used here. Instead, the remnant must have nominative case, indicating that a cleft source is used instead. In addition to having nominative remnants, these sluices also pattern with clefts in other ways. For example, contrast sluicing out of these islands is ungrammatical (36) because *else*-modification is ungrammatical with clefts in Urdu.

- (36) a. * [*Omar aur Nabeel*] =ki kal laRai hui
 [Omar and Nabeel] =GEN.F.SG yesterday fight happen.PST.F.SG
thi lekin mujhe pata nahi aur kaunsa
 be.PST.F.SG but I.OBL.DAT know.PFV.M.SG not and which.M.SG
laRka < *wo tha.* >
 boy.NOM < it be.PST.M.SG >

≠ ‘Omar and Nabeel had a fight yesterday but I don’t know which other boy < it was. >’

- b. * *Omar naraz ho-ga agar Sana Nabeel=se*
 Omar.NOM angry be.FUT.3.SG-FUT.M.SG if Sana.NOM Nabeel=INS
baath karay-gi lekin humein yaad nahi aur
 talk do.FUT.3.SG-FUT.F.SG but we.OBL.DAT remember not and
kaunsa larka < *wo he.* >
 which.M.SG boy.NOM < it be.PRES.3.SG >

≠ ‘Omar will be angry if Sana talks to Nabeel but we don’t remember which other boy < it is. >’

- c. **Mein dekh-rahi thi keh kaunsa laRka*
 I.NOM watch-PROG.F.SG be.PST.F.SG that which.M.SG boy.NOM
Sana=se baath karay-ga lekin mein=ne bathaya
 Sana=INS talk do.FUT.3.SG-FUT.M.SG but I.OBL=ERG tell.PFV.M.SG
nahi aur kaunsi laRki < *wo he.* >
 not and which.F.SG girl.NOM < it be.PRES.3.SG >
- ≠ ‘I was watching which boy would talk to Sana but I didn’t say which other girl < it was. >’

Recall that both case-matching and mismatching between the correlate and remnant are grammatical in sluicing out of relative clauses (33a). This can be explained by the availability of both short and cleft sources where the former leads to case-matching and the latter to case-mismatching. Only case-matching is possible in contrast sluicing (34a, c.f. 37) because although the short source is grammatical with *else*-modification, the cleft source is not. As a result, case-mismatching is ungrammatical in contrast sluicing out of relative clauses.

- (37) **Sana kisi aisay bandey=ko nokri dena*
 Sana.NOM some.OBL like.this.OBL person.OBL=ACC job.NOM give.INF.M.SG
chahthi he jo Urdu=se waqif ho
 want.IPFV.F.SG be.PRES.3.SG REL.NOM Urdu=INS familiar be.FUT.3.SG
lekin mujhe pata nahi aur kaunsi Pakistani zabaan
 but I.OBL.DAT know.M.SG not and which.F.SG Pakistani language.NOM
 < *wo ho.* >
 < *it be.FUT.3.SG* >
- ≠ ‘Sana wants to give a job to someone who is familiar with Urdu but I don’t know which other Pakistani language < it is. >’

Sluices with relative clauses, coordinated-DPs, adjuncts and *wh*-questions also pattern with clefts in regards to mention-*some* modification. Urdu clefts do not allow mention-*some* modification and neither do sluices with these islands (38 - 41).

- (38) A: *Sana kisi aisay bandey=ko nokri*
 Sana.NOM some.OBL like.this.OBL person.OBL=ACC job.NOM
dena chahthi he jo kisi
 give.INF.M.SG want.IPFV.F.SG be.PRES.3.SG REL.NOM some.OBL
Pakistani zabaan=se waqif ho.
 Pakistani language.OBL=INS familiar be.FUT.3.SG
- ‘Sana wants to give a job to someone who is familiar with a Pakistani language.’

Case-Mismatching in Urdu Sluicing

B: *Maslan, kaunsi zabaan=se / *kaunsi*
 Example, which.OBL.F.SG language.OBL=INS / *which.F.SG
zabaan?
 language.NOM

‘For example, which language?’

(39) A: [*Omar aur kisi laRke*] =*ki kal laRai*
 [Omar and some.OBL boy.OBL] =GEN.F.SG yesterday fight
hui thi.
 happen.PST.F.SG be.PST.F.SG

‘Omar and some boy had a fight yesterday.’

B: **Maslan, kaunsa laRka?*
 Example, which.M.SG boy.NOM

‘For example, which boy?’

(40) A: *Omar naraz ho-ga agar Sana kisi*
 Omar.NOM angry be.FUT.3.SG-FUT.M.SG if Sana.NOM some.OBL
laRke=se baath karay-gi.
 laRke.OBL=INS talk do.FUT.3.SG-FUT.F.SG

‘Omar will be angry if Sana talks to some boy.’

B: **Maslan, kaunsa laRka?*
 Example, which.M.SG boy.NOM

‘For example, which boy?’

(41) A: *Mein dekh-rahi thi keh kaunsa laRka*
 I.NOM watch-PROG.F.SG be.PST.F.SG that which.M.SG boy.NOM
kisi laRki=se baath karay-ga.
 which.OBL girl.OBL=INS talk do.FUT.3.SG-FUT.M.SG

‘I was watching which boy would talk to some girl’

B: **Maslan, kaunsi laRki?*
 Example, which.F.SG girl.NOM

‘For example, which girl?’

Finally, Urdu clefts can only take one argument, therefore, sluices with cleft paraphrases are predicted to disallow multiple sluicing. This is shown with sluicing out of coordinated-DPs in (42).⁷

⁷ Multiple sluicing out of relative clauses, adjuncts and wh-islands is independently predicted to be ungrammatical due to violations of the clause-mate condition (Abels & Dayal 2017).

- (42) **Har laRki aur Ahmed restaurant jaathe*
 Every girl.NOM and Ahmed.NOM restaurant.OBL.LOC go.IPFV.M.PL
hein lekin mujhe pata nahi kaunsi laRki
 be.PRES.3.PL but I.OBL.DAT know.PFV.M.SG not which.F.SG girl.NOM
kaunse restaurant.
 which.OBL.M.SG restaurant.OBL.LOC
- ≠ ‘Every girl and Ahmed go to restaurants but I don’t know which girl which restaurant.’

To summarise, sluicing out of relative clauses, coordinated-DPs, adjuncts and wh-questions seems to be island-insensitive in Urdu due to the availability of a cleft source which allows for island evasion. This is supported by the fact that the sluice patterns with clefts rather than wh-questions in these cases: the remnant must always be nominative, and contrast sluicing or *else*-modification, mention-*some* modification and multiple sluicing are ungrammatical.

4.2.3 Isomorphic sources (left branch constraint, subject island)

Sluicing out of a left branch (43a) and subject (43b) is also grammatical in Urdu.

- (43) a. *Us=ne kisi=ki kitaab=ka page*
 He.OBL=ERG someone.OBL=GEN.F.SG book.OBL=GEN.M.SG page.NOM
phaaRa tha lekin mein=ne dekha nahi
 tear.PFV.M.SG be.PST.M.SG but I.OBL=ERG saw.PFV.M.SG not
*kis=ki / *kaun.*
 who.OBL=GEN.F.SG / *who.NOM
- ‘He tore someone’s book’s page but I didn’t see whose.’
- b. *Kisi actor=ke dost=ne eik restaurant*
 Some.OBL actor.OBL=GEN.OBL.M.SG friend.OBL=ERG one restaurant.NOM
khola he lekin mujhe pata nahi
 open.PFV.M.SG be.PRES.3.SG but I.OBL.DAT know.PFV.M.SG not
*kis actor=ke / *kaunsa actor.*
 which.OBL actor.OBL=GEN.OBL.M.SG / *which.M.SG actor.NOM
- ‘Some actor’s friend has opened a restaurant but I don’t know which actor’s.’

Interestingly, neither a short source nor a cleft source can be the pre-sluice here. There is no propositional domain within the first clause for the short source to use as an antecedent. A cleft source would result in case-mismatching, which is ungrammatical in these sentences. Moreover, contrast sluicing is grammatical in these cases (44) which we have seen is not possible with cleft sources.

- (44) a. *Us=ne Omar=ki kitaab=ka page*
 He.OBL=ERG Omar=GEN.F.SG book.OBL=GEN.M.SG page.NOM
phaaRa tha lekin mein=ne dekha nahi
 tear.PFV.M.SG be.PST.M.SG but I.OBL=ERG saw.PFV.M.SG not
aur kis=ki.
 and who.OBL=GEN.F.SG
 ‘He tore Omar’s book’s page but I didn’t see who else’s.’
- b. *Mahirah Khan=ke dost=ne eik restaurant*
 Mahirah Khan=GEN.OBL.M.SG friend.OBL=ERG one restaurant.NOM
khola he lekin mujhe pata nahi
 open.PFV.M.SG be.PRES.3.SG but I.OBL.DAT know.PFV.M.SG not
aur kis actor=ke.
 and which.OBL actor.OBL=GEN.OBL.M.SG
 ‘Mahirah Khan’s friend has opened a restaurant but I don’t know which other actor’s.’

The grammaticality of these sentences is not at all puzzling when we take into account the simple fact that the left branch and subject are not islands in Urdu (45).

- (45) a. *Kis=ki tum=ne kitaab=ka page phaaRa?*
 Who.OBL=GEN you.OBL=ERG book=GEN.M.SG page.NOM tear.PFV.M.SG
 ‘Whose book’s page did you tear?’
- b. *Kis actor=ke tumhara khyaal he*
 Who.OBL actor=GEN.OBL you.OBL.GEN thought.NOM be.PRES.3.SG
dost=ne restaurant khola?
 friend.OBL=ERG restaurant.NOM open.PFV.M.SG
 ‘Which actor’s friend do you think opened a restaurant?’

Thus, the grammaticality of sluicing out of a left branch or subject is due to the availability of isomorphic sources which in these cases do not incur any island violations.

4.2.4 Interim summary

Returning to the predictions made by each of the approaches, we can see that only the hybrid approach is able to account for the data. All sluicing is not island-insensitive; deletion cannot be responsible for repairing islands.⁸ Lack of structure and therefore lack of islands also cannot be at the root of the island-insensitivity of some sluicing. Furthermore, there is not always strict case-matching between the

⁸ There is also other evidence to suggest that sluicing is subject to locality conditions (Abels & Dayal 2017). It would not make sense for sluicing to be subject to some locality constraints but not others, i.e. islands.

correlate and the remnant. This indicates that the same verb is not always present in the antecedent and e-site, nor is case copied from the correlate to the remnant. Instead, what we see is that island-sensitivity of sluicing follows directly from the availability of grammatical sources. The island evasion strategies used in Urdu sluicing are summarised in Table 4.

	Island	Pre-sluice	Case-mismatching	Contrast sluicing
Left branch	×	Isomorphic	×	✓
Subject	×	Isomorphic	×	✓
CP-complement	✓	Short	×	✓
Coordinated-CPs	✓	Short	×	✓
Relative clause	✓	Short	×	✓
		Cleft	✓	×
Coordinated-DPs	✓	Cleft	✓	×
Adjunct	✓	Cleft	✓	×
Wh-question	✓	Cleft	✓	×

Table 4 Island evasion strategies in Urdu sluicing.

The left branch and subject are not islands in Urdu and so isomorphic sources can be used in sluicing. CP-complements and coordinated-CPs are predicational islands and allow use of a short source. Case-matching is seen because the same verb is found in the antecedent and sluice. The short source is compatible with *else*-modification, and so contrast sluicing is grammatical with these islands. On the other hand, cleft sources are used with coordinated-DPs, adjuncts and wh-questions. The remnant must always be nominative as the copular verb assigns nominative case in Urdu. This leads to case-mismatching with non-nominative correlates. These sluices also pattern with clefts in other ways: contrast and multiple sluicing are ungrammatical, as is mention-*some* modification. Relative clauses are an interesting case for which both short and cleft sources are available. Sluices with case-matching and mismatching show the same behaviour as other sluices with short and cleft sources respectively.

5 DISCUSSION

In the preceding sections, we have seen two types of case-mismatching in Urdu sluicing: case-mismatching due to structurally licensed case alternations where the cases in question have overlapping semantics, and case-mismatching due to using cleft sources in the pre-sluice. We have already established that purely syntactic and semantic approaches are unable to account for the data, and that a hybrid approach with syntactic structure and semantic identity is necessary. A major outstanding issue for hybrid approaches is capturing the semantic identity condition in precise

and formal terms. In this section, I discuss the implications of the data in light of this issue.

A basic problem for hybrid approaches is accounting for robust case-matching. There is nothing to prevent case-mismatching across the board through widespread use of paraphrased sources. Under the hybrid approach as it stands, the following two types of sentences should be grammatical. In the first instance, we have two synonymous verbs, *pata* and *jaantha* (know); the former assigns dative case to its subject (46a), while the latter assigns nominative case (46b). The sluiced version with *pata* in the antecedent and *jaantha* in the e-site is ungrammatical (47b).

- (46) a. *Sana=ko pata he keh Omar aa-raha*
 Sana=DAT know.PFV.M.SG be.PRES.3.SG that Omar.NOM come-PROG.M.SG
he.
 be.PRES.3.SG

‘Sana knows that Omar is coming.’

- b. *Sana jaanthi he keh Omar aa-raha*
 Sana.NOM know.IPFV.F.SG be.PRES.3.SG that Omar.NOM come-PROG.M.SG
he.
 be.PRES.3.SG

‘Sana knows that Omar is coming.’

- (47) a. *Kisi=ko pata he keh Omar*
 Someone.OBL=DAT know.PFV.M.SG be.PRES.3.SG that Omar.NOM
aa-raha he lekin mujhe yaad nahi
 come-PROG.M.SG be.PRES.3.SG but I.OBL.DAT remember not
kis=ko < pata he ... >
 who.OBL=DAT < know.PFV.M.SG be.PRES.3.SG ... >

‘Someone knows that Omar is coming but I don’t remember who.’

- b. **Kisi=ko pata he keh Omar*
 Someone.OBL=DAT know.PFV.M.SG be.PRES.3.SG that Omar.NOM
aa-raha he lekin mujhe yaad nahi
 come-PROG.M.SG be.PRES.3.SG but I.OBL.DAT remember not
kaun < jaantha he ... >
 who.NOM < know.IPFV.M.SG be.PRES.3.SG ... >

‘Someone knows that Omar is coming but I don’t remember who.’

Second, we have a sentence which does not require sluicing out of an island and therefore has a grammatical isomorphic source (48a). In (48b), a cleft source has been used resulting in case-mismatching between the instrumental correlate and nominative remnant. The elided version is ungrammatical.

- (48) a. *Sana=ne kisi=se sabzi katvai lekin*
 Sana=ERG someone.OBL=INS vegetables.NOM cut.CAUS.F.SG but
mujhe pata nahi kis=se < *Sana=ne*
 I.OBL.DAT know.PFV.M.SG not who.OBL=INS < Sana=ERG
sabzi katvai. >
 vegetables.NOM cut.CAUS.F.SG >

‘Sana made someone cut vegetables but I don’t remember who.’

- b. **Sana=ne kisi=se sabzi katvai lekin*
 Sana=ERG someone.OBL=INS vegetables.NOM cut.CAUS.F.SG but
mujhe pata nahi kaun < *vo tha.* >
 I.OBL.DAT know.PFV.M.SG not who.NOM < it be.PST.M.SG >

‘Sana made someone cut vegetables but I don’t remember who.’

The fact that case-mismatching is an exception and not the norm indicates that while paraphrased sources are available, there is a strong preference for isomorphic sources. It has been suggested in the literature that there may be a hierarchical availability of sources for sluicing. Van Craenenbroeck (2010) examines the diagnostics Merchant (2001) uses to argue that pre-sluices are always wh-questions and shows instead that the diagnostics are also compatible with wh-questions being used most of the time and cleft sources being used as a last resort when no other source is available. Similarly, Barros et al. (2014) propose that pre-sluices are available in a preferential order based on economy principles. Isomorphic sources require least effort, followed by short sources and finally cleft sources.

This sort of hierarchical availability is seen most clearly in island evasion. As we saw in section 4.2, case-mismatching is ungrammatical where a short source (33b, section 4.2.1) or an isomorphic source (43b, section 4.2.3) can be used. This indicates that cleft sources are unavailable here.⁹

- (33b) *Sana=ne mujhe bathaya tha keh Hira*
 Sana=ERG I.OBL.DAT told.PFV.M.SG be.PST.M.SG that Hira.NOM
kisi=se naraz he lekin mujhe yaad nahi
 someone.OBL=INS angry be.PRES.3.SG but I.OBL.DAT memory not
*kis=se / *kaun.*
 who.OBL=INS / *who.NOM

‘Sana told me that Hira is angry at someone but I don’t remember who.’

⁹ We also need diagnostics to differentiate isomorphic and short sources which may then shed light on the availability of short sources when they are in competition with isomorphic sources.

- (43b) *Kisi actor=ke dost=ne eik restaurant*
 Some.OBL actor.OBL=GEN.OBL.M.SG friend.OBL=ERG one restaurant.NOM
khola he lekin mujhe pata nahi
 open.PFV.M.SG be.PRES.3.SG but I.OBL.DAT know.PFV.M.SG not
*kis actor=ke / *kaunsa actor.*
 which.OBL actor.OBL=GEN.OBL.M.SG / *which.M.SG actor.NOM

‘Some actor’s friend has opened a restaurant but I don’t know which actor’s.’

An interesting challenge posed by the Urdu data to this otherwise simple idea is the grammaticality of both case-matching and mismatching in relative clause islands (33a, section 4.2.1). The version with case-matching is better than the version with case-mismatching, indicating that use of the short source is preferred over use of the cleft source. However, if pre-slucices do indeed become available in the specified order and cleft sources are unavailable if short sources are, then case-mismatching should not be possible here at all.

- (33a) *Sana kisi aisay bandey=ko nokri dena*
 Sana.NOM some.OBL like.this.OBL person.OBL=ACC job.NOM give.INF.M.SG
chahthi he jo kisi Pakistani
 want.IPFV.F.SG be.PRES.3.SG REL.NOM some.OBL Pakistani
zabaan=se waqif ho lekin mujhe pata
 language.OBL=INS familiar be.FUT.3.SG but I.OBL.DAT know.PFV.M.SG
nahi kaunsi Pakistani zabaan=se / zabaan.
 not which.OBL.F.SG Pakistani language.OBL=INS / language.NOM

‘Sana wants to give a job to someone who is familiar with a Pakistani language but I don’t know which Pakistani language.’

Similarly, we see that case-matching is preferred in the data with case alternations, even where case-mismatching is grammatical. Again, we can think of this in terms of a preference for isomorphic sources. As we saw in section 3, some verbs are able to license more than one case in specific environments. For a verb to check a certain case feature, it must have the relevant features. Therefore, there must be two forms of a verb that is able to license two cases: a form which has the relevant features to check case A, and a form which has the relevant features to check case B. Since there is a preference to use isomorphic sources, there is naturally a preference to have the same form of the verb in the antecedent and e-site. However, we can ask the same question as with sluicing out of relative clauses: why should it be possible to use a different form of the verb at all when the same form can be used, i.e. when an isomorphic source can be used? I have no answer to this question at present.

We also see that while using what we might call a paraphrase with a different form of the verb is grammatical, using a cleft source is not in these environments, as indicated by ungrammatical case-mismatching (16, section 3.2).

- (16) *Kisi=ne* / *kisi=ko* *school* *jaana he*
 Someone.OBL=ERG / someone.OBL=DAT school.OBL.LOC go.INF be.PRES.3.SG
lekin mujhe pata nahi kis=ne / kis=ko /
 but I.OBL.DAT know.PFV.M.SG not who.OBL=ERG / who.OBL=DAT /
*kissey / *kaun.*
 who.OBL.DAT / *who.NOM

‘Someone has to go to school but I don’t know who.’

It seems then that we want the pre-slucose to be as structurally similar to the antecedent as possible. Capturing this intuition, as well as the hierarchical preference of different paraphrases, in formal terms is the real challenge. There have been several attempts at formulating the semantic identity condition (Abels 2015, Barros et al. 2014, Merchant 2001, Van Craenenbroeck 2010) but it seems almost that we need a combination of syntactic and semantic identity. Syntactic identity will ensure that structural similarity is prioritised; where syntactic identity is violated, we can resort to a semantic identity condition.

A classic situation in which we would expect the above is sluicing out of islands. The isomorphic source is ungrammatical due to an island violation, and so syntactic identity cannot be maintained. Use of a short or cleft source is therefore possible. Where the latter is used, the remnant must always be nominative leading to case-mismatching with non-nominative correlates (section 4.2.2). However, as we have seen, it is not enough for the general meaning of the pre-slucose and antecedent to be the same, as the case markers can also carry subtle but important semantic information. It is unclear how the semantic contribution of the nominative remnant and non-nominative correlate overlap and therefore how the semantic identity condition is met. It has been proposed by Bagasur (2014) that mismatches from non-nominative correlates to nominative remnants may be because nominative case can be conceptualised as a subset of the other cases. We can assume something along the lines of a nanosyntax case tree (Figure 1) where the cases are built on top of one another, with nominative case being the smallest.

We can also assume that such a representation means that nominative has a subset of features of the other cases, and is thus, able to satisfy the semantic identity condition. However, this raises two immediate issues. First, is having a subset of features enough to satisfy the semantic identity condition or is a full match necessary? The former seems counter-intuitive. One possibility we can consider is that the semantic identity condition weakens as we go down the hierarchy of paraphrases. In fact, this may be necessary independent of the case-mismatching facts. While using cleft sources results in the same overall meaning, one does wonder how semantically identical these are to the antecedent. Furthermore, although the short source and isomorphic source are identical to their respective antecedents, there is still one difference: wh-movement and the presence of the trace. Again, although the overall meaning is identical, the exact semantics of the pre-slucose with a trace and of the antecedent without one cannot be exactly identical. This issue

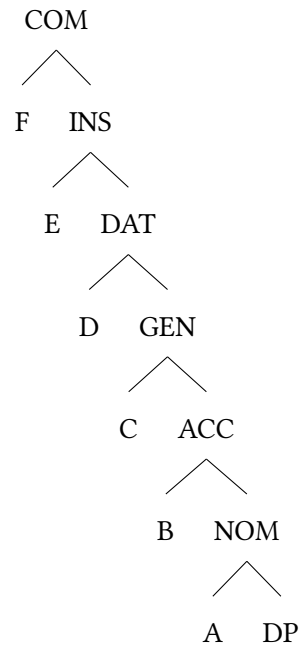


Figure 1 Nanosyntax case tree (Caha 2009).

remains inconclusive until we have a more precise formulation of the semantic identity condition.

A more concrete problem is the fact that this conceptualisation makes incorrect predictions. First, as mentioned previously, mismatching from non-nominative correlates to nominative remnants is not grammatical across the board according to the judgements presented in this paper. In fact, it is only grammatical with the nominative-accusative alternation or when a cleft source is used to avoid island violations in sluicing. Second, if we allow nominative case to meet the semantic identity condition by virtue of having a subset of features of the other cases, then we also predict that mismatching from other “smaller” cases to “bigger” cases should be possible. For example, we would expect mismatching from instrumental correlates to dative remnants or mismatching from genitive correlates to accusative remnants, and so on. This is incorrect.

The discussion in this section has shown us two things. Firstly, there is a hierarchical order in which pre-sluice structures become available. Structures with greater syntactic similarity are preferred over structures with less syntactic similarity. It may be possible to capture this by including both a syntactic *and* a semantic identity condition. Secondly, formulating these conditions, especially the semantic identity condition, is a big challenge. There are many contradictory intuitions: for example, nominative remnants should only be able to satisfy the semantic identity condition when due to cleft sources but not otherwise. Thus, it remains to be seen how the full set of data can be captured.

6 CONCLUSION

In this paper, I have presented novel data from Urdu sluicing which illustrates two types of exceptional case-mismatching. In the first, structurally licensed case alternations with overlapping semantics allow case-mismatching between the correlate and remnant. In the second, use of a cleft source for island evasion results in nominative case on the remnant, leading to mismatching between a non-nominative correlate and a nominative remnant.

On the basis of this data, we can answer our first two research questions. *Is there structure in the e-site?* Yes, without syntactic structure we cannot explain why structurally licensed case alternations should make grammatical case-mismatching possible or why some types of sluicing are island-sensitive. *Is this structure under syntactic or semantic identity with the antecedent?* Semantic identity is more plausible. Under syntactic identity we cannot explain why all case alternation pairs do not lead to grammatical case-mismatching or why all types of sluicing are not island-insensitive.

Our final research question remains open: *how can we formulate the identity condition?* This is a challenging issue as several almost contradictory facts need to be accounted for. While we want paraphrases to be available, we also do not want them to be freely available. A hierarchical availability of paraphrases with a combination of syntactic and semantic identity seems like the right direction at this point.

REFERENCES

- Abels, K. 2015. On the interaction of P-stranding and Sluicing in Bulgarian. In O. Mueller-Reichau & M. Guhl (eds.), *Aspects of Slavic Linguistics: Formal Grammar, Lexicon and Communication*, chap. 1, 1–22. Berlin: De Gruyter.
- Abels, K. & V. Dayal. 2017. On the Syntax of Multiple Sluicing. In A. Lamont & K. Tetzloff (eds.), *Proceedings of NELS47*, vol. 1, Page numbers unknown. Amherst, MA: GLSA publications.
- Bagasur, MJ. 2014. Sluicing in Hindi-Urdu. Handout of workshop presentation, LISSIM 8.
- Barros, M., P. Elliott & G. Thoms. 2014. There is no island repair. Unpublished ms. Rutgers University, University College London, University of Edinburgh.
- Bhattacharya, T. & A. Simpson. 2012. Sluicing in Indo-Aryan: An investigation of Bangla and Hindi. In J. Merchant & A. Simpson (eds.), *Sluicing: Cross-linguistic perspectives*, vol. 38, chap. 9, 183–218. New York: Oxford University Press. doi:10.1093/acprof.
- Butt, M. 1993. Object Specificity and Agreement in Hindi/Urdu. In K. Beals, G. Cooke, D. Kathman, S. Kita, K. E. McCullough & D. Testen (eds.), *Papers from the 29th Regional Meeting of the Chicago Linguistic Society*, 89–103. Chicago: Chicago Linguistic Society.
- Butt, M. & T. H. King. 2004. The Status of Case. In Veneeta Dayal & Anoop Mahajan (eds.), *Clause Structure in South Asian Languages. Studies*

- in *Natural Language and Linguistic Theory*, 153–198. Dordrecht: Springer.
doi:https://doi.org/10.1007/978-1-4020-2719-2_6. https://link.springer.com/chapter/10.1007/978-1-4020-2719-2_{_}6{#}citeas.
- Caha, P. 2009. *The nanosyntax of case*: University of Tromsø dissertation.
- Chung, S., W. Ladusaw & J. McCloskey. 1995. Sluicing and Logical Form. *Natural Language Semantics* 3(3). 239–282.
- Culicover, P. & R.; Jackendoff. 2005. *Simpler Syntax*. New York: Oxford University Press.
- Davison, A. 2015. Hindi-Urdu : Central Issues in Syntax. In Artemis Alexiadou & Tibor Kiss (eds.), *Syntax - Theory and Analysis: An International Handbook*, chap. 42, 1478–1518. Berlin: De Gruyter Mouton. <https://linguistics.uiowa.edu/sites/clas.uiowa.edu.linguistics/files/HBCS2013.pdf>.
- Dayal, V. 2003. A Semantics for Pseudo Incorporation. Unpublished ms. Rutgers University.
- Gribanova, V. & E. Manetta. 2016. Ellipsis in Wh-in-Situ Languages: Deriving Apparent Sluicing in Hindi-Urdu and Uzbek. *Linguistic Inquiry* 46(4). 631–668.
doi:[10.1162/ling_a.00225](https://doi.org/10.1162/ling_a.00225).
- Manetta, E. 2013. Copy theory in wh-in-situ languages: Sluicing in Hindi-Urdu. *Journal of South Asian Languages* 6. 3–24.
- Merchant, J. 2001. *The syntax of silence: Sluicing, islands, and the theory of ellipsis*. New York: Oxford University Press.
- Mohanan, T. 1994. *Argument Structure in Hindi*. Stanford: CSLI Publications.
- Rodrigues, C., A. Nevins & L. Vicente. 2009. Cleaving the interactions between sluicing and P-stranding. *Romance languages and linguistic theory* 2. 1–25.
- Ross, H. 1969. Guess who. In *5th Annual Meeting of the Chicago Linguistic Society (CLS 5)*, 252–286. Chicago.
- Rudin, D. 2019. Head-Based Syntactic Identity in Sluicing. *Linguistic Inquiry* 50(2). 253–283. doi:https://doi.org/10.1162/ling_a.00308.
- Saksena, A. 1980. The Affected Agent. *Language* 56(4). 812–826.
- Saksena, A. 1982. Case Marking Semantics. *Lingua* 56. 335–343.
- Van Craenenbroeck, J. 2010. Invisible Last Resort: A note on clefts as the underlying source for sluicing. *Lingua* 120(7). 1714–1726. doi:[10.1016/j.lingua.2010.01.002](https://doi.org/10.1016/j.lingua.2010.01.002).
<http://dx.doi.org/10.1016/j.lingua.2010.01.002>.

Sana Kidwai
University of Cambridge
sak83@cam.ac.uk