Cleaving the interactions between sluicing and preposition stranding

Cilene Rodrigues  Andrew Nevins  Luis Vicente
Unicamp  Harvard  UC Santa Cruz

Abstract
Merchant (2001) proposes that preposition stranding under sluicing is allowed only in those languages that also allow P-stranding in regular wh-questions. Spanish and Brazilian Portuguese (BP) seem to falsify this generalization, as both are non-P-stranding languages that allow P-stranding under sluicing. Our claim is that, despite initial appearances, Spanish and BP do not constitute counterexamples to Merchant’s generalization. We propose that there are two sources of sluicing in Romance: wh-movement plus IP-deletion (à la Merchant 2001), and clefting plus IP deletion (à la Merchant 1998), the latter being the underlying source for P-stranding sluicing. The apparent P-stranding effect follows from the fact that, as opposed to regular interrogatives, clefts in BP and Spanish do not involve P-stranding at all. We reinforce this conclusion by showing that, in those cases where a cleft base is independently banned, P-stranding under sluicing becomes impossible too.

1 Introduction

Merchant (2001) argues that sluicing involves wh-movement plus IP deletion:

(1) John met someone, but I don’t know \[CP \text{who} [IP \text{John met}]]\]

Considering how this analysis of sluicing interacts with preposition stranding, Merchant (2001, 92) posits the following cross-linguistic generalization:

(2) Form-identity generalization II: P-stranding
A language \(L\) will allow preposition stranding under sluicing iff \(L\) allows preposition stranding under wh-movement.

Although Merchant provides data from many languages in support of (2), recent research has uncovered cases of non-P-stranding languages that, nonetheless, do allow P-stranding under sluicing.\(^1\) In this article, we focus on two

languages that display this tension, namely, Spanish and Brazilian Portuguese (BP). As far as we know, this is the first time the interaction of sluicing and P-stranding is analyzed in Spanish. As for BP, this interaction was noticed first by Almeida and Yoshida (2007), who take it as evidence against the generalization in (2). Here, we re-examine Almeida and Yoshida’s findings, through a comparison of BP and Spanish, and advance an analysis contrary to their claim. We suggest that BP and Spanish have two independent sources of IP deletion: sluicing and pseudosluicing/clefts. Our conclusion is that deletion that involves P-stranding is in fact deletion of a clefted IP whose pivot has wh-moved. Indeed, as Merchant (2001, 101, fn11) remarks, in “languages without overt morphological cases, we may be dealing with a truncation of something like ‘. . . who it is’”. In addition to a wh-pivot and copular verb, we suggest that the cleft also may contain a postcopular relative clause in which the preposition is pied-piped by a relative pronoun.

2 Sluicing with P-stranding in Spanish and BP

Merchant (2001, 99) provides the paradigm in (3) for Spanish, with judgments as indicated (*: ungrammatical, ??: marginal/questionable status).

(3) a. *¿Quién ha hablado Juan con?
who has talked Juan with
b. ??Juan ha hablado con alguien, pero no sé quién
Juan has talked with someone but not know who

Merchant’s claim is that the ungrammaticality of (3) supports the generalization in (2), but the issue is not so clear. For one, there is a difference in acceptability between (3a) and (3b): while Merchant notates (3a) as unacceptaible, (3b) is indicated as only marginal. If P-stranding violations were really unaffected by sluicing, then one would expect (3b) to be as bad as (3a). In addition, the judgement reported in (3b) does not hold among all Spanish speakers. Several speakers of Peninsular Spanish consider (3b) to be only slightly marginal, and some of them report total acceptability. Consider, furthermore, examples such as (4b), where the wh-expression is the D(iscourse)-linked phrase cuál ‘which’. Such examples are acceptable for everybody we have consulted so far, even though their non-elliptical counterparts (4a) are totally ungrammatical.

(4) a. *¿Qué chica ha hablado Juan con?
which girl has talked Juan with
b. Juan ha hablado con una chica, pero no sé cuál
Juan has talked with a girl but not know which

Even though Merchant does not include BP in his corpus, the same contrast can be found in this language: while P-stranding is not allowed under regular wh- movement (5a), it is grammatical when the IP is elided (5b).
Thus, at first sight, one may conclude that the Spanish and BP falsify Merchant’s P-stranding generalization in (2). However, we will demonstrate in what follows that this conclusion is incorrect, and that Merchant’s generalization does in fact hold for the languages under discussion. We will show that apparent counterexamples, such as the ones presented above, do not stem from sluicing, qua a construction involving wh-movement plus IP deletion as in (1). Rather, they are derived from an alternative source: pseudo-sluicing, which consists of a cleft structure whose IP is deleted (see Erteschik-Shir 1977, Merchant 1998 and van Craenenbroeck 2004). Our claim is that BP and Spanish have two sources of IP deletion: sluicing and pseudo-sluicing. However, only pseudo-sluicing results in P-stranding effects. The reason is that, as we will see below, pivots of clefts in Spanish and BP need not be overtly headed by a preposition.

3 The proposal

As mentioned above, we propose that grammatical cases of P-stranding under sluicing stem from an alternative source, one which is different from a regular interrogative. We propose a cleft based on a specificalional copular sentence, as illustrated in (6a). The elided part of the structure is an IP, composed of a copula followed by a DP that is modified by a relative clause. Thus, since the subject of the copula (i.e., the pivot of the cleft) is not introduced by a preposition, ellipsis of the verb and the predicate yields the illusion of P-stranding. Crucially, though, this derivation does not involve P-stranding (6b). As shown below, the proposition appears accompanying the head noun of the relative clause, since deletion targets everything inside the IP domain, the preposition is deleted.

(6) Juan ha hablado con una chica…
Juan has talked with a girl

a. pero no sé cuál es la chica con la que ha hablado Juan
but not know which the girl with the that has talked Juan

As in Spanish, we propose that the input to P-stranding sluicing in Standard BP is also a cleft containing a specificational copular sentence, in which the copular verb is followed by a DP containing a restrictive relative clause. Consider as an example, (7a), which is the Standard BP counterpart of (6a):

(7) O João falou com uma menina...

the João talked with a girl

a. mas eu não sei qual é a menina com a qual João falou

but I not know which is the girl with the which João talked

b. mas eu não sei [IP [

but I not know who is the girl with the

which João falou]]]

At this point it is important to observe that in both BP (8a) and Spanish (8b), the copular verb within the elided cleft may escape deletion, when marked with a certain level of stress. We take the admissibility of an overt copula as initial evidence for a pseudosluicing analysis. 3

(8) a. O João está saindo com alguém, mas eu não sei quem é

the João is going.out with someone but I not know who is

b. Juan está saliendo con una chica, pero no sé qué chica é

Juan is going.out with a girl, but not know what girl is

In this way, we can derive apparent P-stranding effects while preserving the generalization in (2). However, the licensing conditions on this ellipsis will have to be semantic, rather than syntactic, a conclusion that independently supports approaches like Merchant (2001) and especially Potsdam (2007), who argues that sluicing as accomplished by wh- movement from a non-isomorphic underlying structure (namely a pseudocleft, in Malagasy) supports a semantic identity condition for sluicing. Let us now turn to independent evidence favoring the proposed analysis.

3 Almeida and Yoshida (2007) argue against a clefting analysis on the grounds that the copula cannot be retained in sluicing, but subsequently note that the acceptability improves when the copula is stressed relative to the wh- word (cf. their (14e) vs. (14f)), even when P-stranding is not at issue (their (15b) vs. (15c)). These observations follow from the fact that BP (like other Romance languages observing Cinque’s 1993 Nuclear Stress Rule) requires nuclear stress on the most deeply embedded overt constituent, which is the copula, not the wh- that precedes it.
4 Support for the analysis

4.1 Multiple sluicing and P-stranding

Both Spanish (9a) and BP (9b) allow for multiple sluicing, a phenomenon whose relevance for underlying cleft analyses is discussed in van Craenenbroeck (2004, 23):

(9) a. Un estudiante ha leído un libro de Chomsky, pero no sé qué estudiante qué libro
    a student has read a book by Chomsky but not know what student what book

b. O João deu algo para alguém, mas eu não sei o que João gave something for someone but I not know the what para quem
   for who

Interestingly, however, preposition deletion is not possible under multiple sluicing, as shown in (10) and (11). Notice that it does not matter whether it is only the first preposition that is omitted, only the second one, or both: the result is always unacceptable. This is clearly unexpected if the ban on P-stranding is a PF constraint that is avoided only under sluicing by eliding the locus of the violation (as proposed by Almeida and Yoshida 2007). The examples in (10) exemplify cases in which only one of the wh- phrases is a PP, while in (11) we show cases in which both wh- phrases are PPs. In multiple sluicing, the preposition(s) are obligatory.

(10) a. Ella comprou algo para alguém, mas eu não sei o que
    *para quem
    she bought something for someone but I not know the what for who

b. Ela comprou alguma coisa para alguém, mas eu não sei o que
    *para quem
    she bought some thing for someone but I not know the what for who

(11) a. Ella habló con alguien sobre algo, pero no sé *con quien
    *sobre qué
    she talked with someone about something but not know with who about what

b. Ela falou sobre alguma coisa para alguém, mas eu não sei
    *sobre quem *para quem
    sobre the what to who

Given our proposal that P-stranding sluicings stem from underlying clefts, it is tempting to try and attribute the ungrammaticality of these examples to the
fact that pivots of clefts cannot accommodate more than one constituent (12).

(12) a. *No recuerdo [què chica] [què restaurante] es la chica con 
    not remember what girl what restaurant is the girl with 
    la que ha cenado Juan 
    the that has dined Juan 
    b. *Eu não me lembro o que quem é que é a coisa e a 
    I not CL remember the what whom is that is the thing and the 
    pessoa que a Jú falou 
    person that the Jú talked

Nonetheless, this analysis is unlikely to provide the full answer. As pointed out 
to us by Jeremy Hartman (personal communication), even in English (where P- 
stranding is independently allowed), omission of the preposition of the second 
wh-phrase is impossible (13). In other words, under multiple sluicing, the ban 
on deleting prepositions applies even in real cases of sluicing.

(13) Peter talked about something to somebody, but I can’t remember (about) 
what *(to) whom.

It is quite telling that it is only the second wh- phrase in (13) that disallows 
P-stranding. This fact suggests that multiple sluicing is not an exceptional case 
of multiple wh-fronting (cf. Merchant 2001; Richards 2001 for suggestions to 
this effect). Rather, following Lasnik (2006), we propose to analyze (13) as regular 
single wh-fronting plus rightward extraposition of the second wh- phrase, 
followed by IP deletion:

(14) but I can’t remember [CP what [IP Peter talked [about [t] [to whom]]]]

Lasnik supports this analysis by showing that the second wh- phrase in cases 
of multiple sluicing respects the usual constraints on rightward extraposition 
(cf. Ross 1967). To begin with, the ungrammaticality of (13) can be directly 
explained as a consequence of the ban on P-stranding in cases of rightward 
movement.

(15) *Peter talked [PP about t] yesterday [a paper on sluicing].

Second, rightward extraposition obeys the Right Roof Constraint, which bans 
rightward movement from crossing a finite clause boundary. Consider (16), 
where the second wh-phrase (which belongs to the embedded clause) cannot 
extrapose all the way to the matrix domain.4

4Note that, if both wh-phrases are generated in the lower clause, multiple sluicing is reason-
ably acceptable (cf. (i)). Lasnik argues this is because the elided structure is monoclausal: the 
upper clause is “accommodated” outside syntax, and the Right-Roof constraint is respected. We 
avoid this complication by considering only monoclausal examples or examples where each wh-
is generated in a different clause.
Some students said that Mary will speak to some professors, but I can’t remember which students \([_{LP} \text{said} \ [_{CP} \text{that Mary will speak}]]\) to which professors.

Spanish and Brazilian Portuguese show the same behavior as English. First, rightward extraposition is freely available:

(17) a. Juan leyó un libro de Borges ayer
    Juan read a book by Borges yesterday

(18) a. O João leu um livro do Machado ontem
    the João read a book by Machado yesterday

In the same way as in English, rightward extraposition in Spanish and BP respects the ban on P-stranding and the Right Roof Constraint.

(19) a. *María habló \([_{PP} \text{sobre \_ \_ ayer} \ [\text{un libro de Borges}]]\).
    Maria talked about yesterday a book by Borges

(20) a. *O João falou \([_{PP} \text{sobre \_ \_ ontem} \ [\text{um livro do Machado}]]\).
    the João talked about yesterday a book by Machado

We have seen in (10) and (11) above that both Spanish and BP behave in the same way as English (13) in not allowing P-stranding with the second wh-phrase, which suggests that Lasnik’s extraposition analysis can also be extended to these two languages. This conclusion is reinforced by the data below, which are analogous to (16) and show that the Right Roof Constraint prevents the second wh-phrase from originating in an embedded clause. Note, furthermore, that there is no P-stranding in (21), hence its ungrammaticality can only be attributed to a Right Roof Constraint violation.

(i) Peter said that some students will talk to some professors, but I can’t remember which students to which professors.
Thus, we conclude that multiple sluicing involves right extrapolposition of the second wh-phrase in English as well as in Spanish and Brazilian Portuguese. However, multiple sluicing in English differs from its counterpart in Spanish and Brazilian Portuguese with respect to the first wh-phrase: while English forbids P-stranding only in the second wh-phrase (13), Spanish and Brazilian Portuguese forbid it on both wh-phrases. This was exemplified above in (11); another example is provided below.

The pseudo-sluicing analysis we are proposing explains the ungrammaticality of (11) and (22). P-stranding on the second wh- is banned simply because, as in English, rightward extrapolposition is incompatible with P-stranding. What about P-stranding in the first wh-phrase? Suppose that we created a cleft to license P-stranding in the first wh-. Given that clefts are bi-clausal structures, the second wh- could only survive ellipsis if it moved out of the embedded relative clause. This, however, would violate the Right Roof Constraint. We submit, therefore, that the ungrammaticality of such examples is not due to P-stranding constraints, but rather to an illicit movement of the second wh-phrase.  

Note that, on top of the Right Roof Constraint, the second wh- phrase also incurs a potential island violation (extraction out of a relative clause). However, it is unlikely that this is a problem, given the discussion of island repair in Merchant (2001). What is crucial for us, though, is that Right Roof Constraint violations cannot be repaired by ellipsis. This is evidenced by the ungrammaticality of (16) and (21), where there are no island or P-stranding factors that could contribute to their unacceptability.
In sum, the lack of P-stranding in multiple sluicing in Spanish and Brazilian Portuguese follows from the biclausal nature of the cleft structures that form the source of these examples. In order to license multiple sluicing at all, wh- movement of the 1st wh- plus rightward extraposition of the 2nd wh- are needed, neither of which license P-stranding.

4.2 Else modification

Merchant (2001, 122) uses else modification (24a) to argue against a clefting analysis of sluicing in English. He points out that else modification is not possible with clefts (24b), but it is possible with regular interrogatives (24c). Hence, he concludes that the elided substructure in (24a) is not a cleft, but a regular interrogative.6

(24) a. Harry was there, but I don’t know who else.
   b. …but I don’t know who *(else) it was that was there.
   c. …but I don’t know who else was there.

Therefore, given our analysis, we would expect else modification to be impossible in Spanish and BP sluices displaying P-stranding effects. Somewhat unexpectedly, though, this prediction is not correct for BP.

(25) O João saiu com a Maria mas eu não sei quem mais.
the João went out with the Maria but I not know who else

This, however, is not a counterexample to our analysis. As shown in (26), BP allows else modification in clefts. Therefore, the grammaticality of (25) is compatible with a cleft analysis of P-stranding sluicing.7

---

6 However, it might be important to observe that examples like (i) are available on the internet, which suggests that some speakers of English do allow clefts with else modification.

(i) I’d like to know what else it was that Clyde Tombaugh was looking for when he found it.
(http://www.bautforum.com/archive/index.php/t-1762.html)

7 Notice that Almeida and Yoshida (2007) presents the sentence in (i) as ungrammatical, arguing that BP does allow else modification in clefts.

(i) O Pedro estava aqui, mas eu não sei quem mais *(é).
the Pedro was here but I not know who else is

However, our example (26) shows that else modification in clefts is possible in certain contexts in BP. As an anonymous reviewer points out, the acceptability of else modification may be mediated by the presence of a postcopular constituent: notice that, in (26), there is a postcopular constituent, in the same way as in the English example in this footnote.
Me fala quem mais é que você quer convir para sua festa.

In Spanish, on the other hand, P-stranding sluices are incompatible with else modification (27a). Similarly, else modification is impossible in clefts (27b), but allowed in regular interrogatives (27c). This shows that P-stranding sluices in this language stem from an underlying cleft.

a. Juan ha hablado con una chica rubia, pero no sé *(con)
   Juan has talked with a girl blonde but not know with qué chica más.
   what girl else
b. *No sé qué chica más es la chica que ha hablado
   not know what girl else is the girl with that has talked
   Juan.
   Juan
c. ¿Con qué chica más ha hablado Juan?
   with what girl else has talked
   Juan

Given the parallelisms above, it is quite reasonable to assume that P-stranding sluices stem not from regular interrogatives, but from clefts. Spanish and BP independently have different restrictions on else-modification in clefts, that pattern along with their corresponding P-stranding sluices. This difference is arguably due to the fact that clefts do not impose an exhaustivity restriction in BP, as negative quantifiers can serve as their pivots, in contrast to Spanish and English:

a. Não foi ninguém que bateu na porta
   Not was nobody that knocked on the door
b. *No fue nadie que golpeó en la puerta
   Not was nobody that knocked on the door

c. *It was nobody that knocked on the door

In sum, the (un)availability of else modification in Spanish and BP sluicings pattern exactly with clefts, and this difference is arguably related to the bleached nature of clefts in the latter.

4.3 Aggressively non-D-linked wh-phrases

Another diagnostic that Merchant (2001, 123) uses to tease sluicing and pseudosluicing apart is the licensing of aggressively non-D-linked (the hell) wh-phrases such as what the hell, who the f**k, when the devil, que diabos, que ostia, quién cojones, que porra. As the distribution and dependency of these phrases has been linked to polarity items (den Dikken and Giannakidou, 2002), and as their expressive content involves either religion or reproduction (see Pinker (2007) for discussion of why these topics might accompany expressive language), we henceforth refer to such phrases as RPIs. Merchant offers the paradigm in (29)
to illustrate that while RPIs cannot occur in sluicing, they are perfect as pivots of clefts, being thus allowed in pseudosluicing:

(29) Someone dented my car last night.
   a. *I wish I knew who the hell!
   b. I wish I knew who the hell it was!

Almeida and Yoshida (2007) apply this diagnosis to BP in order to argue against a pseudosluicing analysis of IP deletion with P-stranding effects. As they observe, BP is similar to English in that RPIs (such as que porra in BP and quién cojones in Spanish) are perfect cleft pivots (30), but are not allowed in syntactic contexts involving IP deletion with P-stranding (31).8

(30) a. Que porra é que a Maria viu?
    what f**k is that the Maria saw
   b. ¿Quién cojones es la persona a la que ha visto María?
    who testicles is the person to the that has seen Maria

(31) a. *A Maria tá gostando de alguma coisa naquela casa, mas eu
    the Maria is liking of some thing in.that house but I
    ainda não sei que porra!
    yet not know what f**k
   b. *María ha salido con alguien, pero no sé quién cojones.
    Maria has gone.out with someone, but not know who testicles

We would like to point out, however, that the unacceptability of (31) might not constitute strong evidence against a pseudosluicing analysis of IP deletion with P-stranding effects in BP and Spanish. To begin with, note that these examples are ungrammatical even if the preposition is not omitted.9

(32) a. *A Maria tá gostando de alguma coisa naquela casa, mas eu
    the Maria is liking of some thing in.that house but I
    ainda não sei de que porra!
    yet not know of what f**k
   b. *María ha salido con alguien, pero no sé con quién
    Maria has gone.out with someone, but not know with who
testicles

Moreover, as shown in (33), the hell wh-phrases are compatible with pseudosluicing as long as the copular verb is not deleted. Notice that, as mentioned in section 3, the copular verb receives stress.

---

8 Note that a in (30b) is not a preposition, but a differential case marker (cf. Torrego 1998; Cuervo 2003).
9 Sáez (2006) is the first to note this for Spanish.
(33)  a. Maria tá gostando de alguma coisa naquela casa, mas eu
       ainda não sei que porra é!
       Maria is liking of some thing in that house but I
       yet not know what f*ck is
       b. Maria ha salido con alguien, pero no sé quién cojones
       Maria has gone out with someone, but not know who
       is
testicles

These facts are somewhat similar to the observation in Merchant (2002) that
English disallows the hell wh-phrases in regular cases of sluicing (29a)/(34a)
while allowing it when swiping occurs (34b). Merchant also notes that in swip-
ing the preposition carries some level of stress.

(34)  a. *They were arguing, but I don’t know about what the hell.
       b. They were arguing, but I don’t know what the hell ABOUT.

Following the theory of Gussenhoven (1984) for the placement of sentence ac-
cents in English, Sprouse (2006) argues that the contrast between regular sluic-
ing (29a)/(34a) and swiping (34b) with respect to RPI-licensing results from a
combination of following phonological facts:

(35)  a. Each focus domain must receive its own sentence accent (in ac-
cordance with Gussenhoven’s Accent Projection Principle).
       b. Displaced wh-words form their own focus domain.
       c. The non-focused material that follows the wh-word is involved in
       the nearest focus domain.
       d. RPI expressions like the hell are not allowed to have an accent, as
       only an argument, a predicate or a preposition can bear accent
       (in accordance with Gussenhoven’s Sentence Accent Assignment
       Rule).

Thus, applying (35), IP deletion gives us the following possible accented focus
domains.10

(36)  a. …[but I don’t know][what]
       b. *…[but I don’t know][what][the hell]
       c. …[but I don’t know][what][about]
       d. …[but I don’t know][what][the hell about]

As required by (35a) and (35b), the wh-word in (36a) forms a focus domain,
thus receiving an accent. (36c) and (36d) are cases of swiping, in which the
wh-word receives its accent and material that follows it forms another focus
domain receiving its own accent. In (36d), the the hell phrase is not accented
because its syntactic status does not satisfy the Sentence Accent Rule (35d).

10 We concentrate on the wh-word and the constituent that follows it, putting aside the accents
of other constituents.
The ungrammaticality of (36b) follows from a combination of the constraints in (35).

Turning now to BP and Spanish, let us assume that sentence accents in these languages are also subject to the constraints presented in (35). This allows us to derive the fact that RPIs are not allowed in cases of pseudo-sluicing unless they are followed by some material, such as the copula.\(^\text{11}\) Thus, (37) and (38), which are cases of pseudo-sluicing in BP and Spanish, are parallel to (36). In the same way as (36b), (37b) and (38b) are blocked by a combination of all the constraints in (35).

(37)  
  a. [mas eu não sei] [o que]
      but I not know the what
  b. *[mas eu não sei] [que] [porra]
      but I not know what f**k
  c. [mas eu não sei] [que] [porra é]
      but I not know what f**k is

(38)  
  a. [pero no sé] [qué]
      but not know what
  b. *[pero no sé] [qué] [cojones]
      but not know what testicles
  c. [pero no sé] [qué] [cojones es]
      but not know what testicles is

In sum, the non-licensing of RPIs in P-stranding sluices in BP and Spanish should not be taken as an argument against a pseudo-sluicing analysis. As we have shown, this type of wh-phrase is disallowed only when everything but wh-the-hell phrase survives deletion. We have argued that his may follow from the fact that, due to restrictions on sentential accents, an RPI (such as the hell in English, que porra in BP, or qué cojones in Spanish) cannot be the last overt element prior to an ellipsis site.

4.4 Complementizers under sluicing

In BP, the complementizer que ‘that’ can appear in structures in which a wh-phrase occupies the spec of CP, as shown in (39):

(39)  
    Quem que você viu?
    who that you saw

Our next argument for a clefting source of P-stranding under sluicing in BP relates to the presence of this complementizer in sluiced structures.\(^\text{12}\) As shown in (40), in some dialects of this language, que can appear in a P-stranding

---

\(^{11}\)Unfortunately, neither BP nor Spanish allow swiping, thus we cannot test if a following preposition is also able to license the-hell wh- phrases.

\(^{12}\)Complementizer retention of this sort is routinely used in Southern-Central dialects of Brazilian Portuguese. We thank Paulo Medeiros for assistance in these judgements.
sluiced clause (see Merchant 2001, 74-82 for discussion of complementizer retention in ellipsis as related to prosodic cliticization).

(40)  a. O João falou com alguém.
     the João talked with someone
b. Será (com) quem que?
     will.be with who that

The retained complementizer, unlike the wh-word for ‘what’, undergoes obligatory vowel reduction in these contexts, and induces nasal place assimilation, both of which are evidence that it cliticizes and forms a prosodic word with the element to its left, yielding [kẽiŋki] for quem que. This encliticization of a complementizer under ellipsis is not unique to wh-clefts, and occurs under finite forms of the verb ter to have, wherein tem que surfaces as [tẽĩki]:

(41) O João não quer fazer isso, mas ele tem que [fazer isso]
     the João not wants to.do this, but he has that to.do this
     ‘John doesn’t want to do this, but he has to’

We propose that examples like (39) are clefts, which entails that (40b) is a case of pseudo-sluicing. (The preposition com is optional in (41) because PPs can be the pivots of clefts.) As we argue throughout this paper, there are two sources of sluicing in Spanish and BP: one is regular wh-movement plus IP deletion (that is, cases of “proper” sluicing), while the other is pseudosluicing, with deletion of a cleft consisting of a copula and predicate. We argue that que in (40b) cannot be related to an underlying regular interrogative, as it allows P-stranding. Rather, it must be part of an underlying cleft. This is therefore an instance of retention of the complementizer in sluicing. Under this analysis, (40b) receives the following parse. 13

(42) Será quem que [IP é a pessoa com a qual o João falou]
     will.be who that is the person with the who the João spoke

Since the complementizer can be retained under pseudo-sluicing, an RPI may now appear, given that it has prosodic support to its right:

(43) O João comeu alguma coisa estragada, mas eu não sei que
     the João ate some thing rotten, but I not know what
     f**k that

The hypothesis that complementizer retention in sluicing diagnoses an underlying cleft is supported by the fact that que cannot follow both wh-words in cases of multiple sluicing:

13 van Craenenbroeck (2004) contains an extensive analysis of a similar construction in Dutch, in which da ‘that’ survives sluicing along with the wh-phrase. In the same way as we do, he takes this fact as evidence in favor of an underlying cleft. Although the conclusion is the same, it should be noted that the Dutch data are somewhat different from BP, in that da is not a complementizer, but a demonstrative pronoun; see van Craenenbroeck (2004, 14-16) for justification.
Our analysis correctly predicts that complementizer retention should be possible within multiple sluicing only when the second wh-phrase follows the complementizer. This is because, as discussed in 4.1, the second wh-phrase is extraposed to the right, hence it must necessarily follow que.14

(45) a. O Paulo deu algo para alguém, mas eu não sei o que que para quem.
   the Paulo gave something for someone but I not know the what for who
   que para quem
b. *O Paulo deu algo para alguém, mas eu não sei o que que para quem.
   the Paulo gave something for someone but I not know the what for who
   que para quem

c. mas eu não sei o que que [que o Paulo deu [para quem]

In summary, the fact that complementizer-retaining sluicing allows apparent P-stranding is explained by an underlying cleft analysis, which in turn correctly predicts the distribution of multiple sluicing in these configurations.

4.5 Split questions

Arregi (2007) examines the phenomenon of split questions in Spanish, in which the answer to the question is apparently added to the question itself as a tag.

(46) ¿Qué libro ha leído Juan, Guerra y Paz?
   what book has read Juan War and Peace

Arregi shows that examples like (46) are actually composed of two syntactically independent sentences. The first one is a regular wh- question, whereas the second one is a yes-no question that undergoes clausal ellipsis (stripping). Arregi’s analysis is schematized in (47), reflecting the assumption that stripping also requires movement to the left periphery prior to ellipsis (Merchant 2004). We will assume that this analysis can also be extended to BP.

(47) ¿Qué libro ha leído Juan? ¿[Guerra y Paz], [Juan War and Peace has read J

When the tag of a split question is a PP, we find the following pattern: if the first sentence is a regular wh- question, then the preposition in the tag cannot be dropped.

---

14 It should be recalled in (44) that que ‘what does not undergo vowel reduction, surfacing as [ke], while que the complementizer must undergo vowel reduction and surface as [ki].
(48)  a. ¿Con qué chica ha salido Juan, *(con) Elena?
with what girl has gone out Juan with Elena
b. Com qual menina ele saiu, *(com) a Elena?
with what girl he went out with the Elena

However, if the first sentence is a cleft-based question, then the preposition in the tag must be omitted.

(49)  a. ¿Cuál es la chica con la que ha salido Juan, *(con) Elena?
which is the girl with the that has gone out Juan, with Elena
b. Qual é a menina com quem ele saiu, *(com) a Elena?
which is the girl with who he went out with the Elena

This paradigm follows from the assumption that ellipsis in the tag is licensed under parallelism with the antecedent. Thus, if the antecedent is a regular (non-cleft) wh-question, then the tag must also have a non-cleft structure. This blocks P-stranding. On the other hand, if the antecedent is a cleft, then the tag must also be a cleft, licensing P-stranding.

5 Refining aspects of the analysis: Empty-Headed and Free Relatives

Thus far, we have seen that there are good reasons to treat IP deletion plus P-stranding effects as being derived from an underlying cleft. However, Spanish and BP show subtle differences that suggest they differ in the way in which this cleft is syntactically composed.

5.1 Types of wh-pivots in Spanish

We have seen that P-stranding sluices in Spanish are derived from a copular sentences. However, this hypothesis needs some refining, as not every copular sentence can result in a well-formed sluice. Consider the following contrast. In (50a), we see a well-formed copular sentence, which nonetheless becomes ungrammatical if IP is elided (50a). This example contrasts with (6), repeated here as (51).

(50) Juan ha hablado con una chica...
    Juan has talked with a girl
    a. pero no sé quién es la chica con la que Juan ha hablado
    but not know who is the girl with the that Juan has talked
    b. *pero no sé quién.
    but not know who.

(51) Juan ha hablado con una chica...
    Juan has talked with a girl
a. pero no sé cuál es la chica con la que ha hablado Juan
   but not know which is the girl with the that has talked Juan
b. pero no sé cuál
   but not know which

Spanish P-stranding sluicing is preferred with D-linked wh-phrases (the same is noted for Polish by Szczegielniak 2006). This may be related to the fact that Spanish clefts impose a stronger exhaustivity requirement than BP (cf. section 4.2). In addition, we would like to raise the question of what the relative clause within the cleft is headed by. In Brazilian Portuguese, there need not be an overt head for the unsluiced cleft:

(52) O João dançou com alguém, mas eu não sei quem é com a
    the John danced with someone, but I not know who is with the
    qual o João dançou
    that the John danced

While in Spanish, an empty headed relative along with the wh-word quién is marginal:

(53) a. *Juan ha hablado con una chica, pero no sé quién es con
    Juan has talked with a girl, but I not know who is
    la que ha hablado Juan
    with the that the John danced
b. ?Juan ha hablado con una chica, pero no sé cuál es con
    Juan has talked with a girl, but I not know who is
    la que ha hablado Juan
    with the that the John danced

Thus, the unavailability or dispreference for sluicing with quién may be related to its inability to license an empty headed-relative following the copula.

5.2 Sluicing within free relatives

In colloquial BP, relative clauses optionally allow dropping of the preposition, as also noted by Almeida and Yoshida (2007, 359-360):

(54) A menina (com) quem o João dançou na festa estava bêbada
     the girl with who the John danced at the party was drunk

Therefore, in colloquial BP, deletion of an IP might actually be erasing a relative clause that does not have a preposition to begin with, cf. also (55).

(55) O João dançou com alguém …
    the John danced with someone
    a. mas eu não sei quem é que o João dançou
       but I not know who is that the John danced
b. mas eu não sei \([CP \text{ quem } [IP \circ [RC \text{ quem } \text{ Joâo dançou}]])\]
but I not know who is that the Joâo danced

In addition, colloquial BP allows P-drop in free relatives as well as in headed relative clauses.

(56) Eu vi quem você dançou na festa
I saw who you danced at the party

With these facts in mind, let us turn now to a case study in Brazilian Portuguese, lending additional support to the analysis proposed in this paper. In Colloquial Brazilian Portuguese, the verb conhecer, which translates into English as *to know/to be acquainted with*, does not subcategorize for clausal complements (57a), but for DPs (57b).

(57) a. O Joâo não conhece quem saiu com a Maria
the Joâo not is.acquainted who went.out with the Maria
b. O Joâo conhece o noivo da Maria
the Joâo is.acquainted the boyfriend of the Maria

Interestingly, conhecer may be used to embed a sluiced constituent:

(58) O Joâo beijou alguém, mas eu não conheço quem
the Joâo kissed somebody, but I not am.acquainted who

However, in this case, preposition stranding is obligatory:

(59) Eu deveria falar com alguém lá na administração, mas eu não
I should talk with someone there at the administration but I not
conheço (*com) quem
am.acquainted with who

English allows the ‘acquainted with’ reading of *know* when an overt free relative is used:

(60) I’m supposed to talk with someone in the administration, but I don’t
know whoever it is

(59) poses two questions about the nature of sluicing. First, according to the literature, sluicing targets IPs. Thus, if conhecer selects DPs, what category has been elided in (58) and (59)? Second, what is forcing preposition stranding in (59)? In answering these questions, we propose that the underlying verbal complements in (58) and (59) are free relatives. Following Donati (2006), we assume that free relatives are structures in which a wh-head moves to the CP domain and reprojects. In this way, it forms a clause headed by a DP. This analysis follows the general conclusion that free relatives involve a CP-internal wh-phrase but behave categorically as a DP discussed in Grosu (2002).

To illustrate this, consider the free relative in (61a) and its derivation sketched in (61b)/(61c).
a. I ate [what you ate]
b. \([CP_{TP} you [VP ate what]]\)
c. \([DP_{TP} what [TP you [VP ate]]]\)

We suggest extending this analysis to (58) and (59), which we take to be cases of sluicing within free relatives. Doing so, (58), repeated here as (62), results from derivation in (63), with ellipsis of IP within the embedded DP. Elision, which is an optional process, is not obligatory in these cases either.

\[(62) \text{O} \text{ João beijou alguém, mas eu não conheço quem the J kissed somebody, but I not am.acquainted who} \]

\[(63) \text{a. } [TP_{TP} eu não conheço [DP_{TP} o João beijou quem]]]) \]
\[\text{b. } [TP_{TP} eu não conheço [DP quem [TP_o João beijou J]]] \]
\[\text{c. } [TP_{TP} eu não conheço [DP quem [TP o João beijou J]]] \]

This analysis allows us to preserve the idea that sluicing targets IP. Given that \(conhecer\) selects only for DP, in (59) the embedded domain cannot be headed by a prepositional phrase. Since Brazilian Portuguese allows preposition dropping within relative clauses, in (59) the wh-phrase is allowed to move alone, without pied-piping the preposition:

\[(64) \text{Eu deveria falar com alguém lá na administração, mas eu não I should talk with someone there at.the administration but I not know who I should talk there at.the administration} \]

Donati observes that in free relatives, the wh-item can reproject because it moves as a head. Thus, the examples in (65) are ungrammatical because they involve phrase-movement.

\[(65) \text{a. } *I ate what fruit you ate.} \]
\[\text{b. } *Eu comi qual qual fruta você comeu I ate which fruit you ate} \]

As our analysis predicts in cases of sluiced free relatives embedded under \(conhecer\), the wh-item cannot be a phrase either (66). This also explains why the preposition must be dropped, as shown in (59).

\[(66) *Eu deveria falar com alguém lá no prédio da I should talk with someone there at.the building of.the administration, mas eu não conheço com qual pessoa know who I should talk there at.the administration but I not am.acquainted with which person} \]

In contrast to Brazilian Portuguese, Spanish cannot resort to a free relative to feed sluicing with \(conocer\), regardless of whether the preposition is present or not:
As discussed before, in Spanish, sluicing with preposition stranding occurs only when the moved wh-item is a D-linked element, having thus a phrasal status. Given that wh-items must be bare heads in free relatives, Spanish does license sluicing within a free relative and, as a consequence, any attempt to embed a sluiced clause under *conocer* will be ungrammatical.

We note that the occurrence of sluicing underneath a verb that does not allow a clausal complement in BP is not unique to *conhecer*. The verb *arrumar*, meaning ‘to fix’, does not take CP complements, but allows a sluiced wh-phrase as its complement:

(68) Eu tenho que dançar com alguém na festa, e eu já arrumei quem fixed who

'I have to dance with someone at the party, and I’ve already fixed who it will be’

In sum, the occurrence of sluicing within free relatives shows that sluicing may apply to structures other than those with regular wh-movement. This type of sluicing is licensed in Brazilian Portuguese arguably because preposition deletion occurs within relative clauses. This strengthens our claim that Brazilian Portuguese p-stranding under sluicing is to be related to the availability of preposition-dropping.

6 Conclusion

BP and Spanish are non-P-stranding languages that display P-stranding under IP-deletion. Hence, these two grammars might be taken as prima facie evidence against the P-stranding generalization in (2) posited by Merchant (2001) – as done, for instance, by Almeida and Yoshida (2007). However, as we have shown above, in analyzing the robustness of (2) against BP and Spanish, one must consider that these grammars have two sources of IP deletion: sluicing, which conforms to Merchant’s analysis in (1), and pseudosluicing, which involves deletion of a clefted IP. Crucially, only pseudosluicing presents P-stranding effects. Therefore, contrary to Almeida and Yoshida’s claim, neither BP nor Spanish can be taken as evidence against the generalization (2).

The strongest implication of this analysis is that all languages that appear to violate this generalization (see footnote 1) should be reducible to a pseudosluicing analysis. English may lack the pseudosluicing derivations for one of three reasons: (i) pseudosluicing derivations are only available as a ‘last resort’ (ii) pseudosluicing derivations are available in languages that have *pro*
subjects for clefts (iii) pseudosluicing derivations are available in languages
with bleached clefts that do not impose exhaustivity.\footnote{Note, however, that in the course of our discussion we have shown that at least two of Merchant’s original arguments, namely impossibility of else modification and impossibility of the hell wh-phrases, may not go through entirely for English. In addition, as noted by Fortin (2007, 215ff), English disallows P-stranding in questions with the idiom against X’s wishes, but allows P-drop in sluicing with it:}

Whichever of these turns out to be right for English, with respect to Romance, the conclusion is that (2) should be reformulated as (69), with the availability of P-stranding under sluicing relativized not to languages, but to individual syntactic configurations.

\begin{equation}
\text{(69) Form-identity generalization II: P-stranding (revised)}
\end{equation}

For any syntactic configuration $C$, if P-stranding is banned in $C$ in non-elliptical environments, it will also be banned in $C$ under sluicing.

The revision in (69) constitutes a confirmation of the underlying intuition in Merchant (2001) – namely, that the repair effect of sluicing is selective: while strong islands virtually disappear under ellipsis, P-stranding violations remain unaffected. As a consequence, (69) lends support to theories of locality that treat strong islands and P-stranding violations as independent phenomena.\footnote{In Fox and Lasnik (2003) and Lasnik and Park (2003), island amelioration under sluicing is attributed to one-fell-swoop movement across intermediate islands and deletion of otherwise unlinearizable chain copies, respectively. Neither of these strategies would be applicable to ameliorate P-stranding, a welcome conclusion. Merchant (2001) does not offer an account for why Preposition stranding is illicit in certain languages, but his claim throughout the book is that movement violations alleviated by sluicing are due to PF violations that are redeemed by deletion, which presumably means that P-stranding is not a PF constraint. One possibility for why P-stranding is disallowed in BP and Spanish, following Bayer (1996), could be that it is actually due to an LF-movement violation: notably, focus sensitive particles such as só only’, which arguably require LF movement, cannot appear between Preposition and DP (in contrast to English):}

\begin{enumerate}
\item \begin{enumerate}
\item *Whose wishes did he get married against?
\item John got married against someone’s wishes, but I don’t know whose.
\end{enumerate}
\end{enumerate}

Note that else modification is possible with (ib), both in sluices and in clefts:

\begin{enumerate}
\item \begin{enumerate}
\item John got married against his parents wishes, and God knows who the hell else’s.
\item It wasn’t your fault? Then who the hell else’s was it?
\end{enumerate}
\end{enumerate}

Thus, even English, under certain circumstances, may employ a cleft to circumvent P-stranding violations (see van Craenenbroeck 2007 for an elaboration of this idea). However, our focus in this paper is limited to Romance.

\begin{enumerate}
\item Eu falei só com a Maria
I spoke only with the Maria
\item *Eu falei com só a Maria
I spoke with only the Maria
\end{enumerate}

If só and its complement need to move at LF, they will leave a trace next to the Preposition, and the ill-formedness of (ii) can be stated as a condition against stranded prepositions at LF, possibly supporting a broader conclusion that P-stranding is due to constraints on movement and not a PF filter.
References


Szczegielniak, Adam. 2006. All sluiced up, but no alleviation in sight. Ms., Boston College.

Acknowledgements

Many thanks to Karlos Arregi, Jeroen van Craenenbroeck, Jeremy Hartman, Mary Kato, Hironobu Kasai, Marcus Lunguinho, Jason Merchant, David Pesetsky, Maria Polinsky, Eric Potsdam, Johan Rooryck, Benjamin Spector, Adam Szczegelniak, Guillaume Thomas, and three anonymous Going Romance reviewers.

Appendix: French Data

French disallows P-stranding in questions but allows apparent P-stranding in sluicing (cf. also Merchant (2001, 98, fn7)). Like Spanish, French prefers D-linked wh-phrases for sluicing:

(70) *Qui tu as dansé avec?
    Who you have danced with

(71) *Laquelle tu as dansé avec?
    Which you have danced with

(72) Jean a dansé avec quelqu'un, mais je ne sais pas qui
    Jean has danced with someone, but I know who

(73) Jean a dansé avec une des filles, mais je ne sais pas laquelle
    Jean has danced with one of the girls, but I know which

French also does not allow P-stranding in cases of multiple sluicing:

(74) Jean a mangé avec une des filles dans un des restaurants mais
    Jean has eaten with one of the girls at one of the restaurants but
    je ne sais pas *(avec) laquelle *(dans) lequel
    I know who *(with) which *(at) which

French also allows an overt copula. In fact, it offers overt evidence of a cleft possibility, given use of c'était 'it was'.

(75) Jean a dansé avec une des filles, mais je ne sais pas laquelle
    Jean has danced with one of the girls, but I know which
    c'était it was

(76) Jean a dansé avec une des filles, mais je ne sais pas laquelle
    Jean has danced with one of the girls, but I know which
    c'était la fille avec qui il a dansé
    it was the girl with who he has danced

French allows else-modification in the P-stranded sluice, and also has a weaker exhaustivity requirement in clefts:
(77) Jean a dansé avec Marie, mais je ne sais pas (avec) qui d’autre
Jean has danced with Marie, but I don’t know who else

(78) C’est personne qui a frappé la porte
It is nobody that has knocked the door