

Why-stripping in English: A corpus-based perspective*

Jong-Bok Kim · Anne Abeillé
(Kyung Hee University · Université de Paris)

Kim, Jong-Bok and Anne Abeillé. 2019. Why-stripping in English: A corpus-based perspective. *Linguistics Research* 36(3), 365-387. Why-stripping in English, a type of elliptical constructions, consists of the expression *why* and a focus bearing remnant (e.g., *Why me?*). The construction shares certain properties with related constructions such as Sluicing and Stripping but at the same time has its own independent peculiarities. This paper first reviews some key properties of the construction and investigates its uses with attested corpus data. In accounting for how unexpressed expressions in such an elliptical construction can be resolved, there have been two competing trends, movement-and-deletion and direct interpretation (DI) approaches. The former postulates unpronounced syntactic representations at the ellipsis site while the latter consults the semantic/discourse information present in the antecedent. A variety of the corpus data we have identified as Why-stripping challenge the movement-and-deletion approach that heavily depends upon the syntactic information in the antecedent in order to resolve the elided parts. The paper briefly sketches how the attested data can be accounted for within a DI approach. (Kyung Hee University · Université de Paris)

Keywords Why-stripping, corpus, connectivity, deletion, direct interpretation, construction-based

1. Introduction

Stripping or bare argument ellipsis is an elliptical construction in which only a single constituent remains and everything else in a clause is taken to be deleted or unexpressed (Hankamer and Sag 1976):

* An earlier version of this paper was presented at the two workshops, Corpus-based and Experimental Approaches to Ellipsis held at at University of Kentucky during 2017 LSA Institute in July 29, 2017, and Approaches to Fragments and Ellipsis in Spoken and Written English held at Universidade de Vigo in September 28-30, 2017. We thank audiences of these two workshops as well as reviewers of this journal for questions and helpful comments. Our thanks also go to Jungsoo Kim, Jonathan Ginzburg, Philip Miller, Joanna Nykiel, and Jeffrey Runner for constructive feedback and suggestions. This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2017S1A2A2041092).

- (1) a. Kim left Seoul and Lee ~~left Seoul~~ as well.
 b. I met Kim yesterday, but I ~~did not meet~~ Bill.

Similar to this, English allows the so-called Why-stripping construction where the adverbial *why* occurs with a focal-stressed non-*wh* remnant, as illustrated by the following attested examples:

- (2) a. I'm sure a big chunk of it was cash that he kept. Why cash?
 b. The round began with a simple question: Why golf?

The remnant (*cash* in (2a)) can have an overt correlate in the preceding sentence. But the remnant (*golf* in (2b)) can also have a covert correlate provided by the context.

One main question for the account of such an elliptical construction is how we obtain the semantically propositional meaning from what appears to be syntactically less than sentential structures. That is, the non-sentential Stripping constructions in (2) would induce propositional meanings like the following, constructed from the linguistic or contextual environments:

- (3) a. Why did he keep cash?
 b. Why do we play golf?

In answering the question of how to resolve a propositional meaning for non-sentential utterances, there have been two different trends for the analysis of Stripping examples: movement-and-deletion and direct interpretation (DI) approaches. The movement-and-deletion approach assumes that the strips are canonical utterances of the type S (see, among others, Merchant 2003; Weir 2014; Yoshida et al. 2015). This syntax-based view suggests that examples like (2) include an ellipsis site that has an internally structured material through the derivation, and PF deletion renders some of it unpronounced. The meaning composition in this view is then dependent upon the derivational source. Meanwhile, the DI approach assumes that there is no ellipsis involved in such Stripping examples (see, among others, Ginzburg and Sag 2000; Culicover and Jackendoff 2005; Jacobson 2016; Kim 2017). That is, there is no syntactic structure

at the ellipsis site other than the base-generated remnant with *why*. Within this view, the interpretation depends on the surrounding discourse.

This paper reviews some of the key properties of the Why-stripping construction that the previous literature has discussed. The main goal of this paper is to examine these properties with attested examples and evaluate the arguments (mainly related to connectivity effects) for the movement-and-deletion approach. In doing so, the paper refers to attested data extracted from corpora such as COCA (Corpus of Contemporary American English, Davies 2008). The paper also briefly sketches a DI approach that can avoid the issues that movement-and-deletion approaches encounter.

2. Some basic properties

There are two key constructions related to Why-stripping, Sluicing and Stripping, as exemplified by the following:

(4) a. Sluicing:

John ate something, but I don't know what (~~John ate~~).

b. Stripping:

John ate something, but not an apple ~~John ate~~.

As illustrated here in (4a), Sluicing is a type of ellipsis where everything except the *wh*-expression is taken to be elided from the clause. Meanwhile, Stripping, also known as bare argument ellipsis, leaves one remnant (constituent) behind, as illustrated by the example (4b).

The Why-stripping construction we are dealing with here is the combination of Sluicing and Stripping, in the sense that the construction has a *wh*-expression (*why*) and one focal remnant constituent (Merchant 2012; Ortega-Santos et al. 2014; Yoshida et al. 2015). The construction, however, has its own independent properties, differentiating itself from its family of constructions like Sluicing and Stripping. For instance, unlike Sluicing, Why-stripping allows only *why* as the remaining *wh*-expression:

- (5) Sluicing:
 A: Kim made a secret ingredient.
 B: How/when/where?
- (6) Why-stripping:
 A: Here is the secret ingredient – vinegar.
 B: Why vinegar?
 B': *How/*when/*where vinegar?

Sluicing occurs in both matrix and embedded environments as in (7), but Why-stripping is mainly a matrix phenomenon.¹ Even though Why-stripping is not impossible in embedded environments, it is quite restricted. For instance, the idiomatic expression like *why the long face* cannot occur in the embedded clause (see Yoshida et al. 2015 also):

- (7) a. You always whisper when you talk? Why?
 b. I had a sick feeling in my stomach. I didn't understand why.
- (8) a. I felt a little lost and unsure of myself. "Why the long face?" said Mum.
 b. Here's your beer, buddy. *I don't understand why the long face.

Another intriguing property of Why-stripping is that a variety of categorial expression can function as the focal remnant (Yoshida et al. 2015: 328). Not only

1 Yoshida et al. (2015) suggest that Why-stripping can also occur in embedded environments. However, it seems that Why-stripping is mainly for main-clause environments. The corpus data give us only two main types of remnants in the embedded clause. The first main type of Why-stripping in the embedded clause is *why not* (total 185 tokens), while the second type is *why* and an adverb (39 tokens):

- (i) a. It might be possible for us to use the library facilities. I don't see why not.
 b. ... it doesn't really matter anymore. You'll see why shortly.

However, note that adverbs like *shortly* can be linked not to *why* but to the matrix verb *see*. We have not been able to identify other categorial types of the remnant in the embedded clause. As shown in (8), Why-stripping is much restricted in the embedded clause, whose reason is unknown at this point.

phrasal constituents (as in (9) and (10)) but also lexical expressions (as in (11) and (12)) can serve as the non-*wh* remnant as long as they get a focal stress (Ginzburg and Sag 2000):

(9) A: John danced with Mary.

B: Why [_{PP} (WITH) MARY]?

(10) A: All she wanted to do was sit alone in her close.

B: Why [_{AP} SO GLOOMY]?

(11) A: We come up with four categories. B: Why [_{num} FOUR]?

(12) A: I feel like asking questions.

B: Then, why [_V WAIT]?

As seen in (11) and (12), numerals like *four* or bare verbs like *wait* alone are not targets for syntactic operations, but they can serve as the remnant in Why-stripping.

3. Arguments for movement-and-deletion approaches

As noted earlier, one main trend for the account of Why-stripping is a movement-and-deletion analysis. Consider the following examples from Yoshida et al. (2015):

(13) A: John was eating natto.

B: Why NATTO (and not another food)?

Yoshida et al. (2015) argue that, as illustrated in (14), such a Why-stripping example involves a base generation of *why* in the Spec of CP with movement of the focused remnant to the specifier position of a Focus projection as well as clausal ellipsis of the remaining parts (see Ortega-Santos et al. 2014 for a similar account).

(14) [_{CP1} Why [_{CP2} NATTO [_{TP} he was eating t_{natto}]]]?

The motivation for such a movement-and-deletion analysis comes from connectivity effects (Yoshida et al. 2015: 331-337). First, consider the restriction on the possible preposition in Why-stripping.

(15) Matrix Why-stripping

A: John relies on Mary.

B: Why ON/*OF MARY (but not others)?

(16) Embedded Why-stripping

John relies on Mary, but I don't understand why ON/*OF MARY.

The data here indicate that the only possible preposition of the remnant is the one that is required by the verb occurring in the putative source sentence. The movement-and-deletion approach attributes this connectivity to the postulation of the sentential sources constructed from the antecedent clause:

(17) a. Why does John rely on/*of Mary?

b. ... but I don't understand why John relies on/*of MARY.

The unacceptability of having the preposition *of* in (15) and (16) thus can be attributed to that of the source sentences.

Binding facts also illustrate connectivity effects in Why-stripping. Consider the following examples in the so-called picture-noun constructions:

(18) A: He_i is selling all of these pictures.

B: Why (even) PICTURES OF HIMSELF_i?

B': *Why (even) PICTURES OF him_i?

As illustrated here, there is a grammatical difference in using the anaphor *himself* and the pronoun *him*. This contrast can be easily expected given their full-fledged source sentences:

- (19) a. Why is he_i selling pictures of himself_i?
 b. *Why is he_i selling pictures of him_i?

The example (19a) is legitimate since the anaphor is bound by the subject (Binding Principle A), but (19b) violates the condition that a pronoun must be free in the same clause (Binding Principle B).

In addition to these connectivity effects, facts concerning prepositional stranding and voice matching conditions may also support a movement-and-deletion approach. Note that the PP remnant can optionally omit the preposition in both matrix and embedded Why-stripping:

- (20) A: John was talking to Mary.
 B: Why (to) Mary?

Within the movement-and-deletion analysis, this optionality can be attributed to the optionality of moving the focused phrase, as illustrated in the following:

- (21) a. [Why [to MARY]_i [John was talking ____ i]].
 b. Why [[MARY]_i [John was talking to ____ i]].

As shown here, either the focused PP *to Mary* or the NP only *Mary* (stranding *to* behind) can undergo movement, and the ellipsis of the remaining clause would license the two options as in (20).²

The requirement of voice matching also supports movement-and-deletion approaches. The following examples illustrate that voice matching is required in Sluicing, but not in VP-ellipsis (Merchant 2004):

- (22) a. Someone must remove the trash, and it is apparent *who by/*by who
 [~~the trash must be removed~~]. (Sluicing)
 b. The janitor must remove the trash whenever it is apparent that it
 should be [~~removed~~]. (VP-ellipsis)

2 This optionality is referred to as the P-stranding generalization by Merchant (2004). The generalization has been challenged by non-P-stranding languages like Spanish and Brazilian Portuguese that nonetheless licenses P omission under sluicing. See Almeida and Yoshida (2007).

Similar to Sluicing but different from VP-ellipsis, Why-stripping disallows voice mismatching as shown in the following:

(23) A: Max brought the roses.

B: Why Max?

(24) A: Max brought the roses.

B: *Why by Max?

This effect can also be accounted for if we resort to the putative sources of the Why-stripping. (24b) is derived from the source *Why were the roses brought by Max?*, and this violates the syntactic identity requirement.

In sum, connectivity effects, optionality in the preposition stranding, and voice matching appear to support movement-based approaches. The movement-and-deletion view, at first glance, seems to gain strong support from such phenomena, but as we will discuss in what follows the syntax-based view may cover only parts of the attested data.³

4. Corpus findings

4.1 Corpus used and search methods

To see the authentic uses of Why-stripping, we have investigated corpora including COCA (Corpus of Contemporary American English). COCA is the largest structured corpus of American English that continues to be updated and which is based on a variety of genres (e.g., spoken, fiction, magazines, newspapers, and academic), and it contains 560 million words of text from 1990-2017 (about 90 million from each of the five genres just mentioned). To extract Why-stripping examples from the corpus, the research used simple string searches with some regular expressions, as illustrated by the following:

³ As suggested by Kim (2016), Kim and Nykiel (2019), case mismatches in sluicing and fragment answers in Korean and Polish can also be problematic in movement-deletion approaches.

- (25) a. why * [y*] (where [y*] means quotation marks)
 b. why * * [y*]

When necessary, we have also tried to use more complex searches, as exemplified by the following:

- (26) a. why det|art * [y*]
 b. why [nn*][v*][j*][r*] [y*] (j = adjective, r = adverb)

Using such string searches, we have identified total 1285 tokens for the qualitative study of Why-stripping. One thing to note here is that the investigated tokens exclude one key type of Why-stripping:

- (27) He can't read it to me. Why not?

Why-stripping with the remnant *not* is the most frequent type, but we have removed these tokens from our investigation in order to avoid any biased interpretation from one-dominant type.

4.2 Types of the remnant and relations with correlates

We have noted that there are two main types of Why-stripping with respect to the status of correlate: overt or covert correlate. Consider the following Sluicing examples first (Chung et al. 1995):⁴

- (28) a. They were firing, but at what was unclear.
 b. He served the soup, but I don't know to whom.

In these examples, the remnant of ellipsis (*at what* and *to whom*) has no overt correlate within the antecedent clause but it is linked to *firing* and *served*,

4 There are two different types of sluicing: merger and sprouting. In the merger type, the remnant *wh*-phrase has an overt correlate in the antecedent, while in the sprouting type (e.g., (28)), the remnant *wh*-phrase has no overt correlate within the antecedent. See Chung et al. (1995) and Kim (2015a).

respectively. In a similar manner, in Why-stripping, the remnant can have a covert correlate:

- (29) a. I'm not into all that X-Files bullshit... but it was a psychic connection.
 Why **to me**? (COCA 2002 FIC)
 b. It asks the questions. What has happened? Why has it happened? Why
 to me? (COCA 2010 ACAD)

In each of these examples, there is an implicit correlate in the preceding context. For instance, the remnant *to me* in (29a) is linked to the expression *a psychic connection* while *to me* in (29b) is linked to *happened*.

Different from such examples where the remnant functions as an argument or a modifier to a linguistic antecedent, there are examples where the remnant has no link to a linguistic expression, but is conjectured from the discourse:

- (30) Tonight he wore a new silk tie and a velvet-collared wool overcoat tailored to his four-foot height. She'd never seen him like this... "Yet look who came out in the cold," she said, wiping the snow from her collar. "Why **so nervous**?" (COCA 2012 FIC)

The remnant *so nervous* is describing the man's posture in the context and is not linked to any linguistic expression.

To further understand what kind of correlate is used in Why-stripping, we have classified the remnants by category and identified the types of their correlate. The main type of remnants include indefinite NP, definite NP, bare V, AP, and PP remnants.⁵ For each of these remnant types, we have checked if it is linked to an overt or a covert correlate. Let us consider each type in more detail, beginning with indefinite-NP remnants which can have either overt or covert correlates:

- (31) a. Derrick 'Starfire' Dukes, who'd had a few professional boxing matches, but made his living mostly as a professional wrestler. Why a **wrestler**?

5 We have excluded examples with an Adv remnant as in *They want to make inroads here. Why here?* Similar to *Why not?* such examples are also prevalent with more than 700 tokens.

(COCA 1994 SPOK)

- b. Why did someone want this house to burn? To scare me? To kill me?
 ... There was nothing hit-and-miss about the gunshots that ended Patty
 Kay's life. So why **a fire**? (COCA 1994 FIC)

Examples like (31a) have an overt correlate for its remnant, while those like (31b) has a covert correlate. The remnant *a fire* is semantically linked to the covert correlate *to burn*. Our data show that 10 out of 32 (31.2%) indefinite NP remnants have no overt correlate.

When the remnant is an indefinite NP, it typically has an overt correlate, but when the remnant is a definite-NP, it dominantly has no overt correlate. Total 173 out of 190 tokens (91%) have no overt correlates:

- (32) a. Her pastor was leaving town. And soon. She frowned. Why **the rush**?
 (COCA 2011 FIC)
 b. I was wondering if you might be able to meet us at the ER in about
 ninety minutes or so. Why **the delay**? (COCA 2016 FIC)

As illustrated here, the context provides a link between the remnant and its putative source sentence. The remnant *the rush* is linked to the situation of the pastor's leaving town soon. The remnant *the delay* is associated with meeting in one and half hours later. The major difference in having covert correlates between the indefinite and definite NP remnants seems to be related to their difference in their reference properties. A definite NP refers to an individual whose reference denotes the common knowledge shared between the speaker and the hearer while an indefinite NP has no shared knowledge between the interlocutors (Huddleston and Pullum 2002: 368). Since the definite NP assumes a shared knowledge for its reference, it may allow covert correlates.

Examples with bare V remnants also predominantly have no correlates. Total 57 tokens out of 64 (89%) have no overt linguistic correlates:

- (33) a. Heaven forbid the voters should elect anyone else. They've always
 voted for a Magetry. Why **change**? (COCA 2002 FIC)
 b. I want to turn it around now. Why **wait**? (COCA 2004 SPOK)

Both remnants *change* and *wait* have links to the previous situations: no linguistic expressions are linked to them. The key reason for the high frequency of having covert correlates seems to do with the fact that the situation described by the remnant verb is tightly connected to the situation described in the previous context. This allows having no overt correlates.

As for the adjectival remnant, when it is a single Adj, it typically has a correlate, but with the AP remnant, more than 70% have no linguistic correlate.

- (34) a. In Italy, Lieutenant Robert Martin's hand-me down P-51 Mustang, 'Queen Cole,' got a fresh coat of red paint on the tail section. Why **red**? (COCA 2003 MAG)
- b. The Sterling Tigers begin the season ranked No. 1 despite finishing 12-16 last year. Why **so high**? (COCA 2003 NEWS)

In (34a), the bare Adj remnant *red* has an overt correlate in the preceding sentence, but the AP remnant *so high* is evoked from the situation of its being number one. The degree expression helps the interlocutors locate the covert correlate of the remnant.

The final type we have identified is PP remnants. With the PP remnants, out of a total 52 tokens, 38 (73%) have no correlates.

- (35) a. "Maybe they wish to sponsor your spaceship-" "Why should they do that? Why now? Why **in person**?" (COCA 2000 FIC)
- b. This is her first introduction to the American people, right? Why now? Why **in this manner**? (COCA 2011 SPOK)

Most of the PP remnants, as shown here, function as an implicit (temporal/locational) or unexpressed adjunct argument in the previous situation.

The following summarizes the findings with respect to the frequency of covert and overt correlates depending on the type of remnants as shown in Figure 1.

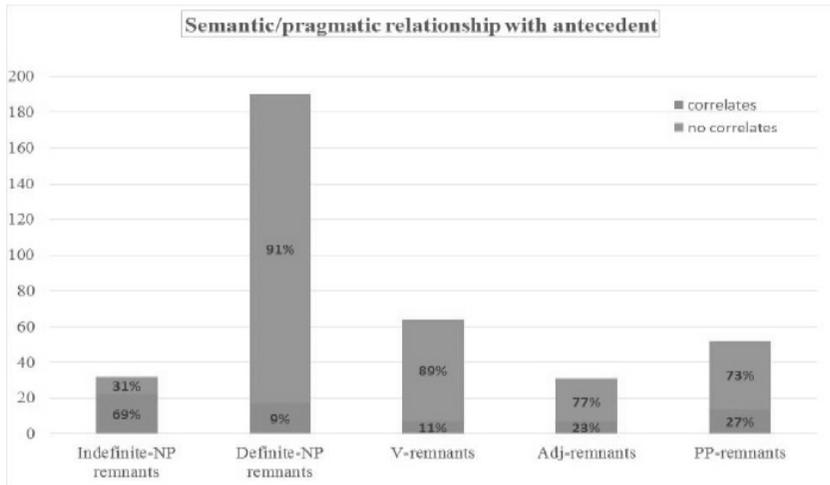


Figure 1: Frequency of the overt/covert correlate with phrasal remnants

As the figure shows, the NP remnant type has the highest frequency. The key observation we can make here is that each type has a high degree of covert correlate. In particular, the definite NP remnants have the highest percentage of covert correlate.

4.3 On connectivity effects

As we have discussed, connectivity effects in Why-stripping seem to support the existence of full-fledged source sentences in the construction. However, the corpus data we have identified indicate that connectivity effects can often be overridden.

The movement-and-deletion account predicts the preposition identity between the remnant and its correlate (Yoshida et al. 2015), but the corpus yields examples violating this morphosyntactic identity constraint:

- (36) a. John relies on Mary, but I don't understand why/how come **on**/*of **Mary**.
 b. Robin's body had been found at the golf course just a half mile from her home under some leaves. Why **on the golf course**? (COCA 2005 SPOK)

Such examples indicate that there is no need to have complete syntactic identity in constructing the putative source sentence. What matters is thus not syntactic but semantic or pragmatic information.

Within the movement-and-deletion approach, case matching is expected between the correlate and the remnant (see Yoshida et al. 2015). However, note that attested Why-swiping examples often have no case matching between the two:

(37) Of course this is a war against Islam. They say that Saddam is a dictator, but why **him**? (COCA 2003 SPOK)

(38) A: You, the boy, Carrier, and two carpenters. Go in, negotiate a purchase. Tell them we've no interest in them beyond buying what we need.
B: But why **us**? (COCA 2004 FIC)

These examples reflect the general constraint that when a pronoun alone functions as a fragment answer, it is required to be accusative (Ginzburg and Sag 2000). This implies that that we cannot enforce syntactic identity in building putative sentential sources for the remnant in Why-stripping.

If Why-stripping involves movement plus deletion, then we expect it to be subject to constraints on movement. Instead, island insensitivity is an often observed in the construction, as seen from the following corpus examples (see Merchant 2003 for sluicing and Yoshida et al. 2015 for a similar point too):

(39) a. Well, I think what's happening is both risky and immoral. Why **immoral**? (COCA 2009 SPOK)
b. A pack of lean dogs, all different colors, loped across the street far ahead of her. Why **dogs**? (COCA 1992 FIC)

Examples like (39a) would violate the strong island constraint, the Coordinate Structure Constraint, while those like (39b) would violate the Complex NP Constraint. This implies that syntax-based views alone are not satisfactory.

4.4 Grammatical and discourse relations between the remnant and its correlate

As we have seen, there is no need to have category identity between the remnant and its putative correlate. The corpus data offer various non-identity types of the linkage between the remnant and its correlate.

First, consider the following example:

- (40) a. What are these white men so angry about? ... Why **the anger**? (COCA 2010 SPOK)
 b. But you changed the policy, saying that a gun used once in a crime should be destroyed, not recycled. Why **the change**? (COCA 2000 NEWS)

The attested data often allow such a categorical mismatch between the remnant and the putative correlate. Once again, we observe that the relationship between the remnant and the putative correlate is not syntactic identity but semantic associations between the two.

Another often found type is a synonymous relation. Observe the following:

- (41) a. Well, why are tomatoes so complex? ... the tomatoes have all these different things about them. Why **so complicated**? (COCA 2012 SPOK)
 b. We've got to get over to your apartment fast. Why **the hurry**? (COCA 2008 FIC)

In these examples, the remnant and the correlate are synonymous. The adjective *complicated* is semantically similar to the putative correlate *complex* while *hurry* is synonymous to *fast*. The data once again suggest that the syntactic identity between the remnants and the correlates is not a strong requirement in Why-stripping.

The remnant can have a semantic/pragmatic relationship with the antecedent. The remnant *at that age* in (42) is anaphorically refer to the expression *9 years old* in the preceding sentence.

- (42) And as I understand it, Nicole, that's what drew you into the movement when you were 9 years old. Why **at that age**? (COCA 1999 SPOK)

A similar situation can be found from examples like the following:

- (43) a. In fact, they reviewed and approved our press release. So we're a bit scratching our heads to figure out why **this**. (COCA 2004 FIC)
b. But to actually go overseas and, you know, it's an investment of time. There is safety issues as well. Why **that**? (COCA 2004 SPOK)

In these examples, the remnants *that* and *this* refer to the situations described by the previous sentence.

- (44) a. ... figure out why this happened.
b. ... Why is that the case?

The putative source sentence thus has no syntactic relation with the previous linguistic environment.

Another type of semantic/pragmatic relation can be found from the following examples:

- (45) a. She raised her fist and punched the air. "Why **the anger**?" (COCA 2004 FIC)
b. ... So I let him babble on about strange happenings and dramatic rescues. Why **argue**? (COCA 2008 FIC)

In both examples, the remnants *the anger* and *argue* describe the possible results coming out from the preceding situations. For instance, raising one's fist and punching the air is a typical act of anger.

As illustrated by such examples, in many Why-stripping, the key for the resolution of propositional meaning depends on contextual clues, rather than syntactic identity. Consider more examples where the discourse plays a key role in figuring out the meaning of the remnant:

- (46) Robert Shapiro sent me a copy of this book when it was in Galley. I got a chance to read it. I really loved it. For a first effort, I thought it was extraordinary. Why **fiction**? (COCA 2001 SPOK)

In this example, the remnant *fiction* is cued by the expression *this book* and the context, whose genre is conjectured from the context. A similar situation can be observed from examples with bare verb fragments:

- (47) a. "You worked there, didn't you?" 'You know the answer, so **why ask**?'
(COCA 2009 FIC)
b. We still love this place. So **why go**? (COCA 2001 MAG)

In sum, authentic uses of the Why-stripping construction tell us that there need not be a morphosyntactically identical correlate for the remnant. In many attested examples we have seen so far, there is no syntactically identical correlate. It is the semantic or surrounding contexts that provide clues for the meaning resolution. The movement-and-deletion approaches, requiring syntactic relations with the preceding context, are clearly challenged by such semantically or pragmatically controlled Why-stripping.

5. A direction for a direct interpretation approach

We have seen that attested corpus data challenge any movement-and-deletion approaches. The data further indicate that the complete syntax of the fragmental remnant in Why-stripping is just the categorial phrase projection of the fragment itself. Following the DI (direct interpretation approach) adopted by Ginzburg and Sag (2000), Culicover and Jackendoff (2005), Kim (2017), and others, we also allow any maximal projection to function as a fragment answer, as illustrated by the following example:

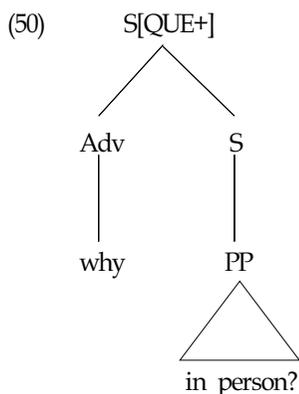
- (48) A: How did Paul tell you that?
B: In person.

Within the DI approach, there is no syntactic structure at the ellipsis site and fragments are the sole daughter of an S-node, directly generated from the following construction:

(49) Head-Fragment Construction:

Any category can be projected into a NSU (non-sentential utterance)
when it functions as a focus establishing constituent (FEC).

All the fragments in Why-stripping belong to this Head-Fragment construction, as seen from the following structure.



The *wh*-expression *why* combines with a sentence expression projected from the remnant *in person*. The interpretation of this Why-stripping depends on the context.

The Head-Fragment Construction allows any focal expression (FEC) to be projected to a sentential expression. This FEC is associated with the *wh*-expression *why* which assigns a focus to the remnant (Ginzburg 2012). Note the role of DGB (dialogue-game-board) here where the contextual parameters are ‘anchored’ and where there is a record of who said what to whom, and what/who they were referring to (see Ginzburg 2012). Uttering a question or an utterance can introduce the information about QUD (Question-Under-Discussion) as well as FEC.

For an illustration, consider a typical example with an overt matching

correlate:

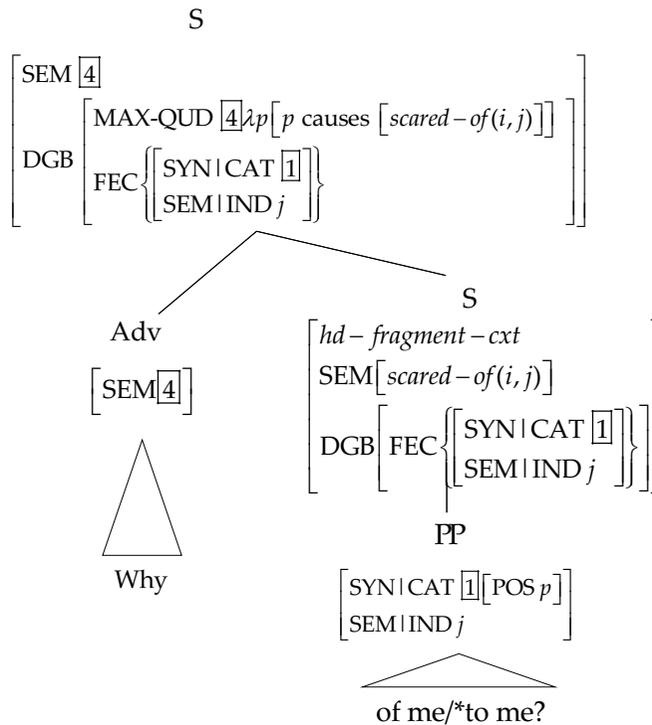
(51) I finally said, “You. I’m scared of you.” Why of me? (COCA 2011 FIC)

When uttering the declarative sentence *I’m scared of you*, the DGB, where discourse is structured around QUD, can introduce the following:

(52) QUD: $\lambda p[p \text{ causes } [\textit{scared-of}(i,j)]]$

In this QUD, triggered by the assertion (assertions are always addressing some QUD), what is questioned is the reason why the speaker is scared of the hearer. The Why-stripping sentence *Why of me?* is where the PP of *me* functions as a FEC, licensing the following structure:

(53)



As represented here, uttering the Why-stripping construction in the dialogue introduces the information about the FEC. This information is linked to the QUD introduced from the declarative sentence. The QUD evoked is that the speaker is scared of the hearer, and the hearer is asking the reason for this, in particular, with the focus being placed on *of me*. Note that, through the index value (j), the evoked QUD information is also linked to the argument structure of the verb *scare* which requires a theme subject and an experiencer complement with the preposition marking value *of*. This in turn means that the fragment *to me* would result in the mismatch in the syntactic information. However, having the NP *me* would not violate this matching requirement.

The present system assumes that any remnant can be projected into an S as long as a proper context is given. This is why not only a phrasal constituent but also a single expression can be the remnant in Why-Stripping, as attested by the corpus examples:

- (54) a. When she's in New York, she enjoys relaxing with us. Why **us**?
(COCA 2008 FIC)
b. It was confusing! Confusing? Why **confusing**? (COCA 2003 SPOK)
c. But they pursued me here. Why **here**? (COCA 2007 FIC)

Now consider examples with no linguistic correlate at all.

- (55) Japan came in first at 61 mbps; the average speed for all 30 nations considered was 9 mbps. *Why the difference?* (COCA 2008 MAG)

The MAX-QUD introduced here is from the context: there is no linguistic correlate here. The QUD could vary but it is linked to the question why the difference:

- (56) a. ... why the difference exists
b. ... why the difference occurs between the two
c. ... why the difference happens
d. ...

As long as context allows, there could be another putative source. Even though we have sketched a DI approach, it could tell us that a DI approach can avoid problems of any analysis requiring syntactic identity between the remnant and the putative correlate. It thus can serve as a feasible alternative to the movement-and-deletion approach.

6. Conclusion

English Why-stripping is a type of elliptical constructions that require a special mapping relation from form to function. In resolving the mapping relation from the non-sentential utterance (fragment) to a propositional meaning, there have been two main trends: movement-and-deletion and direct interpretation (DI) approaches. The former has enjoyed some popularity, mainly supported by phenomena displaying connectivity effects.

In this paper, we first reviewed some key properties of the construction that the literature has pointed out. To check the feasibility of the movement-and-deletion approach, we have investigated authentic uses of the construction with online corpora like COCA. The corpus data challenge many facets of the movement-and-deletion approaches that typically require syntactic identity between the remnant and the putative source sentence. Meanwhile, the DI approach introduces no additional syntax: fragments are mapped into non-sentential utterances and induce sentential interpretations from the enriched discourse. Why-stripping is simply the projection of combining the focus marking *why* with such a fragment. The DI approach is further supported by the robust account of flexible connectivity effects, discourse initial fragments, and island repair in English.

References

- Almeida, Diogo and Masaya Yoshida. 2007. A problem for the preposition stranding generalization. *Linguistic Inquiry* 38: 349-362.
- Chung, Sandra, William A. Ladusaw, and James McCloskey. 1995. Sluicing and logical

- form. *Natural Language Semantics* 3: 239-282.
- Culicover, Peter and Ray Jackendoff. 2005. *Simpler syntax*. Oxford: Oxford University Press.
- Davies, Mark. 2008. The Corpus of Contemporary American English (COCA): 560 million words, 1990-present. Available online at <https://www.english-corpora.org/coca/>.
- Ginzburg, Jonathan. 2012. *The interactive stance: Meaning for conversation*. Oxford: Oxford University Press.
- Ginzburg, Jonathan and Ivan A. Sag. 2000. *Interrogative investigations*. Stanford: CSLI Publications.
- Huddleston, Rodney and Geoffrey K. Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- Jacobson, Pauline. 2016. The short answer: Implications for direct compositionality (and vice versa). *Language* 92(2): 331-375.
- Kim, Jong-Bok. 2015a. Syntactic and semantic identity in Korean sluicing: A direct interpretation approach. *Lingua* 66: 260-293.
- Kim, Jong-Bok. 2015b. Fragments in Korean: A direct interpretation approach. *Studies in Generative Grammar* 25(3): 703-733.
- Kim, Jong-Bok. 2016. *The Syntactic structures of Korean: A construction grammar perspective*. Cambridge: Cambridge University Press.
- Kim, Jong-Bok. 2017. Why-stripping in English and Korean: A direct interpretation approach. *Studies in Generative Grammar* 27(4): 733-751.
- Kim, Jong-Bok and Joanna Nykiel. 2019. Fragments and structural identity on a direct interpretation approach. Ms. Kyung Hee University.
- Merchant, Jason. 2004. Fragments and ellipsis. *Linguistics and Philosophy* 27: 661-738.
- Merchant, Jason. 2012. Ellipsis. In Artemis Alexiadou, Tibor Kiss, and Miriam Butt (eds.), *Handbook of contemporary syntax*. Berlin: Walter de Gruyter.
- Nakao, Chizuru, Masaya Yoshida, and Ivan Ortega-Santos. 2013. On the structure of Japanese 'Why'-stripping. In Nobu Goto, Koichi Otaki, Atsushi Sato, and Kensuke Takita (eds.), *Proceedings of GLOW in Asia IX*. Tsu, Japan: Mie University.
- Ortega-Santos Iván, Massaya Yoshida, and Chizuru Nakao. 2014. On ellipsis structures involving a *wh*-remnant and a non-*wh*-remnant simultaneously. *Lingua* 138: 55-85.
- Phillips, Colin and Dan Parker. 2014. The psycholinguistics of ellipsis. *Lingua* 151: 78-95.
- Ross, John Robert. 1969. Guess who? In Robert I. Binnick, Alice Davison, Georgia M. Green, and Jerry L. Morgan (eds.), *Proceedings of the Fifth Annual Meeting of the Chicago Linguistics Society*, 252-286. Chicago, Illinois: University of Chicago.
- Sag, Ivan A. and Joanna Nykiel. 2011. Remarks on Sluicing. In Stefan Müller (ed.), *Proceedings of the 18th International Conference on Head-Driven Phrase Structure Grammar, University of Washington*, 188-208. Stanford, CA: CSLI Publications.
- Weir, Andrew. 2014. Why-stripping targets Voice Phrase. In Hsin-Lun Huang, Amanda Rysling, and Ethan Poole (eds.), *Proceedings of NELS*, Vol. 43, 235-248. Amherst, Mass:

GLSA.

Yoshida, Masaya, Chinzuru Nakao, and Ivan Ortega-Santos. 2015. The syntax of *Why*-stripping. *Natural Language and Linguistic Theory* 33(1): 323-370.

Jong-Bok Kim

Professor
Dept. of English Linguistics and Literature
Kyung Hee University
26, Kyungheedaero, Dongdaemun-gu,
Seoul, 02447, Rep. of Korea
E-mail: jongbok@khu.ac.kr

Anne Abeillé

Professor
LLF, CNRS
Université de Paris
12-22, Rue Marguerite Duras,
Paris, 75013, France
E-mail: abeille@linguist.univ-paris-diderot.fr

Received: 2019. 11. 03.

Revised: 2019. 11. 30.

Accepted: 2019. 12. 07.