

# English Transparent Free Relatives: Interactions between the Lexicon and Constructions\*

January 4, 2012

## Abstract

English employs two different types of free relatives: standard and transparent. These two types, different with respect to the syntactic and semantic headedness of the clause, have intriguing similarities and differences. In particular, transparent free relatives (TFRs) are peculiar in the sense that the predicative expression within the relative clause functions as the head of the clause with respect to syntactic as well as semantic properties. In this paper, we provide a non-movement analysis that places an emphasis on the interactions between the lexicon and constructional constraints.

Keywords: transparent free relative, standard free relative, headedness, transparent effect, construction-based

## 1 Introduction

English has two main types of relative clauses, dependent and free relative clauses, as exemplified in the following:

- (1) a. The books [that you read] belong to me.
- b. [What(ever) you read] belongs to me.

In (1a), the relative clause is ‘dependent’ on the referent of the relative pronoun *that* and its antecedent *the books*. Meanwhile, the relative clause in (1b) has no antecedent it modifies and is thus ‘free’ from its antecedent in a sense. What makes this type of ‘free’ relative clause more intriguing is that there is a different type of

---

\*I thank three anonymous reviewers of this journal for comments and suggestions. My thanks also go to Jungsoo Kim, Kyeongmin Kim, Okgi Kim, Hyunwoo Lee, Seulkee Park, and Eun-Jung Yoo for helpful comments. All errors remain mine. This work was supported by the Kyung Hee Best Researcher Program in the year of 2011.

free relative clauses as seen from the following corpus data we extracted from the COCA:<sup>1</sup>

- (2) a. Breast cancer is [what scientists call **multifactorial**].  
b. [What appear to be **trousers**] are/\*is really leggings that only come halfway up each thigh.

Examples like (1b) and those like (2) both involve a *wh*-element in the constituent initial position and have a missing element somewhere in the rest of the constituent. The bracket phrases in both neither have a referent referred to by *what* nor include an antecedent linking the bracket clause. However, there are some important differences between these two types of free relatives. One main difference between examples like (1b) and those like (2) is that the ‘felt’ nucleus of the clause is *what(ever)* in (1b) but *multifactorial* and *trousers* in (2). That is, the syntactic and semantic ‘nucleus’ of the bracketed expression in examples like (1) is the *wh*-phrase, whereas the nucleus in (2) is the expression *multifactorial* and *trousers* located within the relative clause. This can be evidenced from the fact that the NP *trousers* determines the number value of the main verb. In this respect, the free relative clause is ‘transparent’ with respect to the plurality of *trousers*. Following Wilder (1999), we also call examples like (1b) ‘standard’ free relative (SFR, henceforth) and those like (2) ‘transparent’ free relative (TFR, henceforth).

There are many issues related to both of these constructions. In this paper, we review some main grammatical properties of the so called transparent free relative clauses, while comparing them with standard free relative clauses. After discussing three main previous approaches to TFRs in brief, we then provide a construction-based analysis that places strong emphasis on the interactions between the lexicon and constructional properties.

## 2 Transparent Effects and Headedness

Grammatical properties of English free relatives have been discussed in much literature including Kajita (1977), Wilder (1999), Grosu (2003), and Van Riemsdijk (2006) among others. In this section, we will review some main properties that any analysis needs to account for.

**Parenthetical Modifier:** The first difference between SFRs and TFRs we can observe is the parenthetical function of the clause. Consider the following TFR examples we extracted from the COCA:

- (3) a. Do you know (what they call) a musician without a girlfriend?

---

<sup>1</sup>The corpus COCA (Corpus of Contemporary American English) contains 425 million word corpus of American English, dated between 1990 and 2011. Compared to the BNC (British National Corpus) and ANC (American National Corpus), this corpus is large, balanced, up-to-date, and freely-available online.

- b. This project requires (what may appear to be) a lot of extra work.
- c. Obama vowed to end (what he called) the cozy relationship between regulators and the oil industry.

As indicated by the parentheses, the TFR minus the predicate behaves like an optional modifier to the nucleus predicate (Kajita 1977, Wilder 1999, Van Riemsdijk 2006, among others). This parenthetical modifier plays a hedging function, lessening the impact of the utterance in question. This hedging function can be noticed from the possibility of replacing the relative clause minus the nucleus by lexical intensional modifiers like *apparent*, *alleged(ly)*, and *presumed* (Grosu 2003):

- (4) a. [**What appears to be** a pale blue painting] turns into something else entirely.
- b. [An **apparent** pale blue painting] turns into something else entirely.
- (5) a. The most important thing, both in parenting as well as in child care, is a warm and [**what we call responsive**] relationship.  
(COCA:1997:SPOK NBC.Today)
- b. The most important thing, both in parenting as well as in child care, is a warm and [**allegedly responsive**] relationship.

In terms of meaning, we can note that the predication expressed by the TFRs is not necessarily true. For instance, in (4a) the NP to which *what appears to be a pale blue painting* refers may not be *a pale blue painting*, but can turn out to be *a pale green painting*. In a similar fashion, in (5a) the *relationship* that the expression *what we call responsive* modifies may not taken to be a *responsive* one by others. This paraphrase thus supports the TFRs, here lessening the assertion power expressed by the clauses.

Note that SFRs do not have this kind of parenthetical or hedging function:

- (6) a. Kim ate [what Lee offered to her].
- b. We roamed in the streets and ate [whatever possible].

In (6a), no constituents in the bracketed relative clause are parenthetical. Even in non-specific relative clauses like (6b), the *wh*-element is obligatory and the proposition expressed by the clause is meant to be true. This in turn means that the TFR minus its nucleus can be said to have a transparent effect.

**Distributional Properties:** With respect to distributional possibilities, SFRs externally act like nominal clauses. They appear in the contexts where otherwise only NPs can occur (Jacobson 1995, Bresnan and Grimshaw 1978, Quirk et al. 1985):

- (7) a. [Whoever did that] should admit it frankly.  
 b. I took [what they offered me].  
 c. Macy's is [where I buy my clothes].

SFRs in (7) function as the subject, object, or predicative complement. They can serve even as prepositional complements:

- (8) a. I will move to [wherever you want to live].  
 b. John got the tapes from [wherever he keeps his books].

Subject-auxiliary inversion (SAI) also tells us that SFRs are nominal phrases rather than clauses (cf. Huddleston and Pullum 2002). As noted in (9) and (10), neither an interrogative clause nor a CP clause can undergo SAI:

- (9) a. [What books he has written] is certain.  
 b. \*Is [what books he has written] certain?
- (10) a. [That he has written a book] is certain.  
 b. \*Is [that he has written a book] certain?

However, SFRs can freely occur in interrogative and other constructions with SAI, as observed in (11):

- (11) a. [What she told me] is unreasonable.  
 b. Is [what she told me] unreasonable?

All these observations support the fact that SFRs are syntactically NPs in general.

Now consider the distributional properties of TFRs. TFRs can also appear in various syntactic positions including subject, object, and predicative:

- (12) a. [What seems to be a tourist] is lying on the lawn.  
 b. John seduced [what appears to be an underage girl].  
 c. John is [what you might call a fool].

However, unlike SFRs, TFRs can also occur in a modifier position:

- (13) a. The important thing is a [[warm] and [what we call responsive]] relationship.  
 b. This is a [what you might call tricky] example.

The TFRs here are in the prenominal position modifying the noun head.

Note that TFRs have a much wider distributional possibilities than SFRs. The canonical syntactic category of the nucleus is an NP, but other categories like AP, AdvP, PP, and nonfinite VP can appear too, as attested from the following from the COCA:

- (14) a. You're definitely not [what anyone would describe as **ecstatic**].  
(COCA:2007:MAG TownCountry)
- b. In that process I begin to work [what I would call **creatively**].  
(COCA:1992:MAG PsychToday)
- c. She definitely wasn't [what she'd call **in love with Sam Butler**].  
(COCA:1990:FIC Bk:Dazzle)
- d. We started [what we call **picking corn**].  
(COCA:2007:SPOK NPR\_TellMore)
- e. They never become fixed - or [what we call **entrained**] - in a regular cycle. (COCA:1993:SPOK CNN\_News)

In all these examples, the syntactic position of the TFR is identical with the syntactic category of the boldfaced nucleus. For example, in (14b), the TFR occurs in the adverbial position. This is possible since the predicative nucleus is an AdvP *creatively*. We once again observe the transparent effect of TFRs.<sup>2</sup>

**Number Agreement:** The number value of SFRs is singular or determined by the *wh*-phrase while that of TFRs is dependent upon the nucleus. Consider the agreement factors in an interrogative, relative, and SFR clauses, respectively:

- (15) a. **What books he has written** isn't/\*aren't certain.  
b. **The books** Kim has written \*hasn't/haven't been published.  
c. **What(ever) books** he has written \*hasn't/haven't been sold well.

In interrogative constructions (15a), it is the whole interrogative subject NP that determines the grammatical number of the main verb. In the canonical relative clause (15b), however, it is the head NP *the books* that induces number agreement with the main verb. In the SFR (15c), it is not the whole subject phrase but just the *wh*-expression, *what(ever) books*, that determines agreement with the main verb. This implies that the head of the SFR is the nominal phrase with the *wh*-free-relative word as its specifier (Bresnan and Grimshaw 1978, McCawley 1998, Quirk et al. 1985).<sup>3</sup>

<sup>2</sup>Not all syntactic categories can be used as the nucleus. For example, a finite VP cannot be used as a nucleus in the TFR. See section 3.2.

<sup>3</sup>In addition to the differences in the set of available *wh*-words between interrogative and free relative constructions, they differ in that the occurrence of indirect questions depends on the types of predicates, whereas that of free relatives does not.

- (i) a. \*Kim ate which dish Lee served to her.  
b. Kim knew which dish Lee served to her.
- (ii) a. \*Which dish Lee served to her went into the trash.  
b. Which dish Lee served to her was unclear.

As observed, the predicates, *know* and *unclear* allow indirect questions as complements, but *ate* and *went* do not.

Now consider TFR examples from the COCA:

- (16) a. [What appears to be **a pale blue painting**] turns/\*turn into something else entirely. (COCA:1994:NEWS CSMonitor)
- b. [What we call **coincidences**] are/\*is limited to the ones we happen to notice. (COCA:2008:NEWS Chicago)

The number value of the main verb in both examples varies. This variation cannot be accounted for if we identically take *what* as the head of the clause. It is rather the predicate nucleus in the clause that determines the number value of the whole phrase. This observation once again shows us that unlike SFRs, TFRs display transparent effects.

**Preposition restriction:** A transparent effect of the TFRs can also be observed in the selection of a preposition. Consider the following:

- (17) a. He speaks in/\*at/\*on **a Northern dialect**.
- b. He speaks in/\*at/\*on [what linguists call **a Northern dialect**]. (COCA:2001:ACAD AmerScholar)
- (18) a. We live in/\*at/\*on **an age of communication**.
- b. We live in/\*at/\*on [what we call **an age of communication**]. (COCA:1995:ACAD ArtsEduc)
- (19) a. The photos are odd-looking, printed on/\*at/\*in **some sort of cloth**.
- b. The photos are odd-looking, printed on/\*at/\*in [what appears to be **some sort of cloth**]. (COCA:1994:FIC Mov:StarTrek08)

As illustrated here, the bold-faced NP in each of the examples requires a specified preposition in the given context. For example, in (a) examples, we observe that the verb *speaks* requires *in* with the following NP head *dialect* while *live* asks for *in* with the following NP *an age of communication*. These requirements are still effective with the presence of the TFRs in (b) examples. Indicating that the nucleus of the TFRs determines the preposition, this once again displays a transparent effect of the TFRs, supporting the idea that the XP is the head of the entire TFR construction (Kajita 1977, McCawley 1998).

**Coordination Conjunction:** Coordination data also show us the headedness of the nucleus in the TFRs and a transparent effect (Kajita 1977, Van Riemsdijk 1998, 2001, 2006). Observe the following corpus data:

- (20) a. Connelly is pretty excited about [[his life] and [what he considers **the perfect job**]]. (COCA:2002:NEWS Denver)
- b. Omer Stewart counted a number of reasons for [[prehistoric] and [what he described as **primitive**] uses of fire]. (COCA:1999:ACAD GeographRev)

Given the assumption that in canonical coordination structures, like categories are conjoined, the examples in (20) once again show us the transparent effect. In particular, note that the well-formedness of (20b) can be best explained by the assumption that the AP *prehistoric* is conjoined with the AP *primitive* which is the nucleus of the TFR. However, note what happens if the same strings appear in different contexts:

- (21) a. Connelly is pretty excited about [[his life] and [what he considered perfect]]. (SFR reading at best)
- b. Stewart learned [what he described as primitive uses of fire]. (SFR reading only)

In these examples, the free relative clause cannot be a TFR since the NP *his life* in (21a) cannot be coordinated with the AP *perfect*. The only possible way to interpret this is that the free relative is a SFR coordinating with the NP *his life*.

**Extraction Possibility:** Extraction out of the XP in TFRs is possible, whereas it is impossible in SFRs. As noted in Wilder (1999), we can observe that just like complex NPs, SFRs form strong islands (data from Wilder 1999):

- (22) a. \*the student that Mary invited [who likes \_ ]
- b. \*something that Mary invited [whoever is angry about \_ ]

However, TFRs behave differently from SFRs. Consider the following:

- (23) a. Who did she buy a nice portrait of \_ ?
- b. Who did she buy [what seems to be a nice portrait of \_ ?]

As the complement of *portrait* can be *wh*-questioned in the simple case in (23a), we observe the same phenomenon with the TFR as shown in (23b). This again implies the parenthetical function, that is, transparent effect of the clause.

However, a caveat is that extraction out of the TFR is not always possible, as pointed out by Grosu (2003):

- (24) a. \*Who did she draw what no normal person would describe as [a successful caricature of \_ ]?

- b. \*Who did you describe this picture as [a successful caricature of \_\_ ]?

We cannot do justice to the reason of the difference in extraction here since the possibility of extraction is, as is well-known, affected by a variety of factors such as processing or *D(discourse)*-linking property of the NP involved (cf. Grosu 2003).<sup>4</sup> However, we can once again notice the parenthetical behavior of TFRs: the clause minus the nucleus behaves like an optional modifier and its presence does not affect syntactic constraints.

**Definiteness:** Definiteness also shows us a transparency effect of TFRs. The existential ‘*there*’-construction places an indefiniteness constraint on the postcopula subject:

- (25) a. There was [a/\*the long pause] in the incoming fire.  
 b. \*There was [what Jay had cooked] on the table.

The SFR cannot appear in the postcopula position of the *there* construction due to the fact that SFRs canonically evoke a definite interpretation:

- (26) a. \*There was what Jon ate yesterday.  
 b. \*There was what John painted blue on your desk.

However, TFRs can appear in the existential construction when the nucleus is an indefinite NP.

- (27) a. There was [what seemed like **a long pause**] in the incoming fire.  
 (COCA:2007:MAG WashMonth)  
 b. \*There was [what seemed like **the long pause**] in the incoming fire.

This does not of course mean that TFRs cannot be used as definite NPs. As expected from their transparent effects, as long as they are allowed in the predicate position, the definite NP can appear in TFRs:

- (28) a. Here I’m pursuing [what they call **the American dream**].  
 (COCA:2005:NEWS Denver)  
 b. We have achieved [what we consider to be **success**].  
 (COCA:1993:NEWS Atlanta)

---

<sup>4</sup>D-linked wh-phrases do not obey certain syntactic locality constraints. For example, English questions with D-linked wh-phrases can exceptionally violate Superiority:

- (i) a. \*What did you persuade who to read \_\_ ?  
 b. Which book did you persuade which person to read \_\_ ?



- c. Members of his team are now excavating [what he calls **South Street**].  
(COCA:2001:MAG NatGeog)
- d. He looked past faces looking at him or [what they took to be **him**].  
(COCA:1993:FIC Bk:BeamMeUp)

As shown in (27) and (28), TFRs can occur in any type of NP positions, including the indefinite-only position as far as the NP in the TFR satisfies the distributional requirements in context.

**Binding:** The transparency effects can further be supported by binding phenomena (Van Riemsdijk 2000, 2001). As is well known, some expressions must be properly bound by another expression. Consider the following data:

(29) Bound Variable:

- a. Every American president has had [what can be considered **his Monica Lewinsky**].
- b. Every politician will be held accountable for [what we can describe as **his political past**].

(30) Reciprocal:

- a. \*We could sense [whatever feeling you could only describe as **each other's soul**].
- b. \*We could sense [the feeling that you could only describe as **each other's soul**].
- c. At this level of closeness we could sense [what could only be described as **each other's soul**].

(31) Reflexive:

- a. \*He was painting [whoever I took to be **himself**] in the white house.
- b. \*He was painting [the person that I took to be **himself**] in the white house.
- c. He was painting [what I took to be **himself**] in the White House.

In (29), *his* is bound by *every American president* and *every politician*, respectively. The facts that there is no appropriate antecedent to bind *his* within the TFR and that the *every*-NP in the matrix clause binds it can lead us to conclude that the XP is the head of the whole TFR clause. This conclusion can be justified by the bound anaphor facts shown in (29). In a similar manner, the reciprocal pronoun *each other* and the reflexive pronoun *himself* (31a) and (31b) cannot be said to be bound by any other element within the TFR; rather, they are bound by the subject of the matrix clause.

### 3 Previous Analyses

#### 3.1 Parenthetical Placement with Backward Deletion

Capturing the parenthetical function of the TFR clause minus the nucleus, Wilder (1999) initially proposes two steps in generating TFR clauses:

- (32) a. Step 1: Insert a complete relative clause into a matrix clause as a parenthetical expression
- b. Step 2: Delete the predicate in the relative clause using the rule ‘Backward Deletion’

The following illustrates how these two steps are applied:

- (33) I am pursuing [<sub>REL-CL</sub> what they call ~~the American dream~~] the American dream.

As given here, the relative clause *what they call the American dream* is first inserted in the matrix clause, left adjacent to the head NP *the American dream* that the clause modifies. The next step renders the Backward Deletion Rule to delete the predicate *the American dream* in the relative clause which is morpho-phonologically identical with the head NP in the matrix clause.

The gist of this ‘parenthetical’ analysis is to place the predicate phrase in the matrix clause from the beginning, explaining why the predicate phrase determines the number or the definiteness of the clause. Capturing most of the transparent effects of TFRs, this analysis has many advantages, but suffers from one serious problem for examples like the following:

- (34) a. Many churches experienced [what was described as a “**tsunami of giving**” by wealthy dotcom members]. (COCA:2001:MAG Christ-Century)
- b. [What were called **close readings** by star professors] were often bravura performances. (COCA:2005:ACAD AmerScholar)

As also pointed out by Van Riemsdijk (2001), such examples include the TFR in the medial position. This means we cannot apply the Backward Deletion since the TFR clause is not located in the right-end of the clause.

#### 3.2 Shared Structures

As a way of solving problems arising from the TFR examples in the medial position, Van Riemsdijk (1998, 2001) proposes that the TFR-construction is generated by a process in which the structure of a TFR is ‘grafted’ onto another tree structure, that of the matrix clause. The basic idea of this analysis is that a certain constituent can be shared by two different clauses or at two different positions in the given syntactic position. A simple illustration of this analysis can be given in (35):

- (35)  $\left. \begin{array}{l} \text{I am pursuing} \\ \text{What they call} \end{array} \right\} \text{ the American dream.}$

The nucleus *the American dream* here is shared by the matrix as well as the relative clause. By assuming that the same element is present in both clauses, the analysis can thus explain the transparent effects of TFRs with respect to number, definiteness, and so forth.

Attractive though this one seems to be, it also suffers from problems. In particular, it requires additional constraints to block certain phrases from being shared. For example, phrases like finite VPs or auxiliary-headed VPs cannot serve as the nucleus of the TFR:

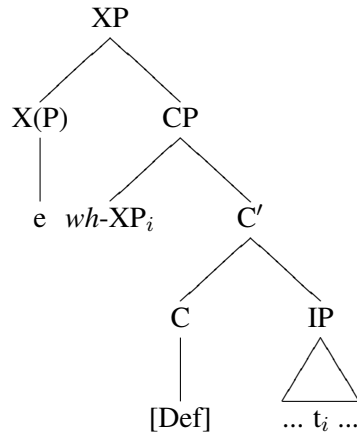
- (36) a. \*John [[what we call] **studies hard**].  
 b. \*John [[what they call] **must be a good student**].  
 c. \*John must [[what they call] **be a good student**].

Note that the shared analysis would generate such structures with the assumption that the bold-faced expression is shared by the subject *John* and *what X call*. However, as shown here, these grammaticality means that TFRs do not allow finite or auxiliary-headed VPs to occur in the XP position though diverse types of major syntactic categories can appear in that position. Issues thus remain of how to block the analysis from generating such cases.

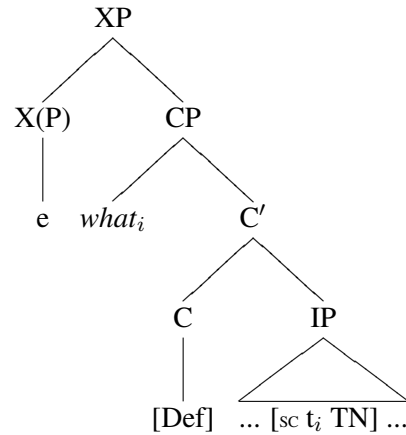
### 3.3 Common Configuration for SFRs and TFRs

Rather than treating SFRs and TFRs in a different way, Grosu (2003) argues that basically TFRs should be analyzed as the same as SFRs, and that they both have the identical configurational structures as given in the following:

(37) a. SFR Structure:



b. TFR Structure:



As represented here, both SFR and TFR structures have identical configurational structures: the *wh* phrase binds a trace in the clause, and the feature [Def] on C introduces the uniqueness effect in semantics. In this analysis, the main difference between SFRs and TFRs comes from the status of *what*: in TFRs *what* is underspecified for syntactic categories as well as for semantics, and that unlike SFRs, it serves as the specifier (subject) of the SC in TFRs.

This indirect analysis can, of course, have merits in providing a uniform analysis for SFRs and TFRs. By assigning no syntactic category to *what*, it provides a way of explaining why the predicate phrase determines the syntactic category of the whole phrase. The analysis captures most of the transparent effects we find in TFRs, but some drawbacks still remain. We thus see why only TFRs can include a small clause. The biggest issue concerns the differences between SFRs and TFRs. In particular, the analysis needs to show why only TFRs have transparent effects while SFRs do not. In addition, this analysis captures transparent effects by means of a highly-underspecified *wh*-chain, but issues remain of why TFRs occur only with raising predicates. The analysis thus requires to explain why the dependency relation in the *wh*-chain is not really long-distance unlike in canonical cases.

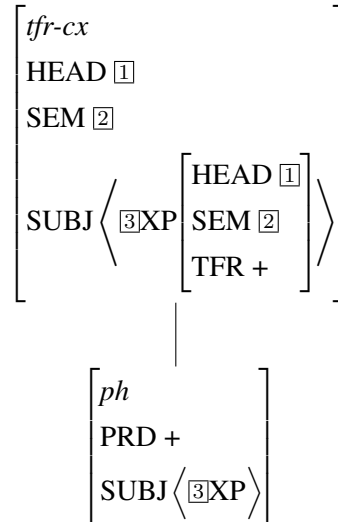
## 4 A Constructional Perspective

One of the main generalizations on TFRs that we have made so far is that the predicative element in the TFR clause functions as the nucleus (syntactic and semantic head) of the TFR clause when its subject is *what*.

Our analysis starts with the assumption that, following Grosu (2003), this *what* is also unspecified for its syntactic and semantic value. As a way of reflecting this in the present context, we assume that there is a construction whose head is a phrasal-level unary construction. This construction functions as a predicative

expression (carrying the feature [PRD +]) and requires a subject with the feature [TFR +] that can be originated only from *what*:

(38) Transparent Free Relative Clause Construction:



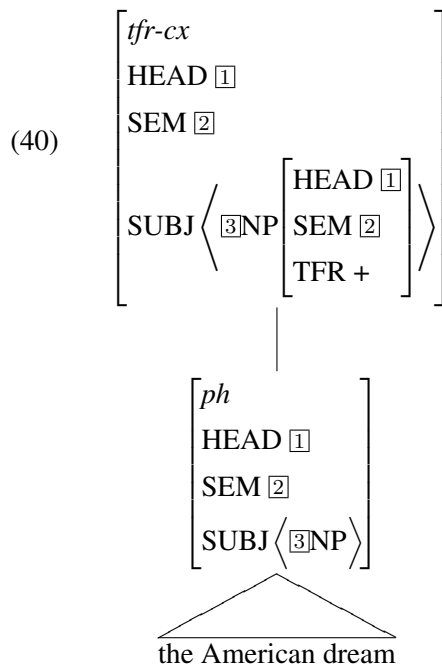
The construction means that a predicative expression selecting a TFR subject can be projected into a TFR construction whose HEAD and SEM values are identical with the selected subject's HEAD and SEM values, respectively.<sup>5</sup> Note that this does not mean that the main properties of the construction can be originated from any predicative NP. Only when a predicative expression selects the TFR expression *what* as its subject, the predicative phrase can be pumped up or projected as a transparent free relative (*tfr-cx*) clause.

For example, the NP *the American dream* can be either realized as a simple NP construction in (39a) or as a TFR construction *tfr-cx* in (39b):

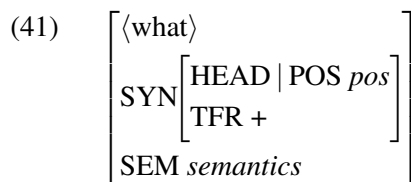
- (39) a. I am pursuing the American dream.  
 b. I am pursuing what they call the American dream.

In (39b), the NP *the American dream* is realized as a predicative expression selecting a TFR subject. This will then allow us to have the following TFR construction:

<sup>5</sup>See Yoo (2008) for a different analysis within a constraint-based perspective. In addition, see Kim (2008), Lee (2009), and Kim and Sells (2011) for the advantages of construction-based perspectives.



Note that the predicative NP *the American dream* selects a [TFR +] subject whose HEAD value and SEM value are identical with its own HEAD and SEM values. This is possible because the phrase is realized as a *tfr-cx* construction (or computationally pumped up to this construction). The structure sharing (or identity) of these two values is made possible since only *what* is lexically specified to have the TFR value and underspecified with the HEAD and SEM values:<sup>6</sup>



As represented here, the expression *what* has no POS and SEM values specified. These two values are determined when this appears as the subject of a predicate. This lexical specification can immediately explain two important constraints in the TFR: *what* being the only possible element in the TFR and its possibility of referring to a human being. The range of possible *wh*-words in SFRs is restricted, but the one in TFRs is much more restricted in that only *what* is licensed. *Wh*-words like *what*, *where*, *when*, but not *who*, *which*, *how* and *why*, can introduce an SFR (Baker 1989, Borsley 1992, Quirk et al. 1985, among others):

(42) a. He got [what/whatever he wanted].

<sup>6</sup>The type values *pos* and *semantics* are topmost values for the feature POS and SEM.

- b. He put the money [where/wherever Lee told him to put it].
  - c. The concert started [when/whenever the bell rang].
- (43) a. \*Lee wants to meet [who Kim hired].
- b. \*Lee bought [which car Kim wanted to sell].

However, in TFRs, only *what* is allowed, regardless of the nature of the nucleus:

- (44) a. A young man is seducing [what/\*who/\*whatever appears to be an underage girl].
- b. The politician persuaded [what/\*who/\*whatever some call his ‘surge’ voters to return to the polls.

In both cases, the nucleus is a human being, but only *what* is licensed here.

Related to the property just observed above, TFRs allow the bare *what* to refer to humans as seen from the following COCA examples:

- (45) a. There had been [what appears to be a murderer on the 5th floor].
- b. Christine witnessed the Hutu killing spree of 1994. She is what Rwandans call a survivor.

On the other hand, the bare *what* cannot be used to refer to humans in SFRs.

- (46) a. \*He seduced what you invited to the party.
- b. \*She officially invited what I persuaded to give a talk at the party.

As is clear from the contrast in (45) and (46), the bare *what* in TFRs functions differently from that in SFRs with regard to the [+human] reference. In the present analysis, this is also expected since the semantic value of *what* is underspecified and eventually identified with the predicate nucleus.

At this point, we need to note that the main verbs in the TFR are all raising verbs including such as those like *appear, seem, to, be, consider, describe, assume, characterize, regard*, and the like. One main property of these raising verbs is that its subject is not determined by itself (not theta-marked by itself) but identified with that of its nonfinite complement (cf. Sag et al. 2003, Kim and Sells 2008):

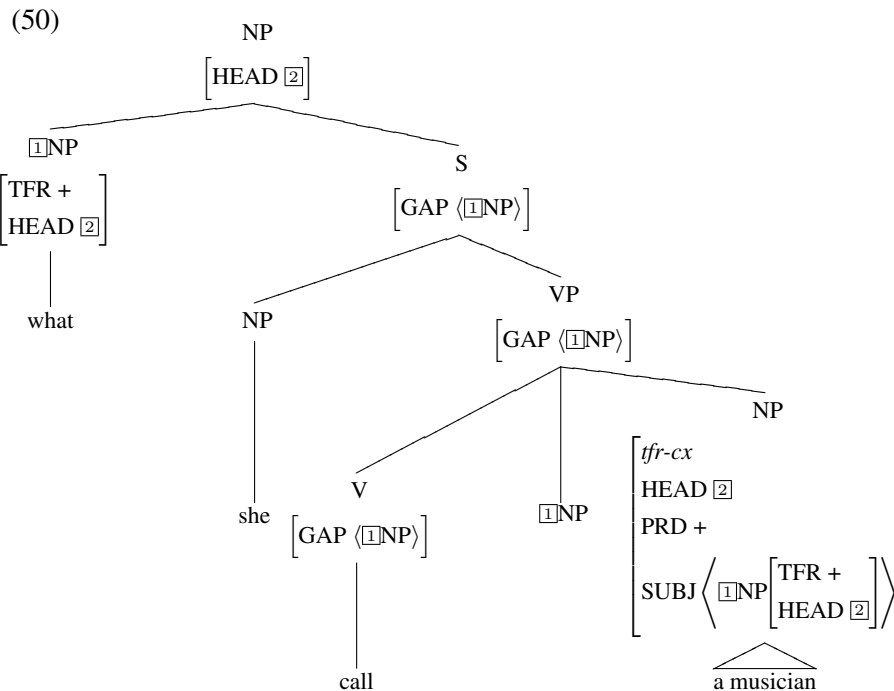
- (47) a. \*John appears to rain.
- b. It appears to rain.
- (48) a. \*John believed Mary to rain.
- b. John believed it to rain.

As represented in Sag et al. (2003) and Kim and Sells (2008), the simplest way to capture this property is to structure-share the subject of a raising predicate with that of its nonfinite VP complement as exemplified by the following:

- (49)
- |    |   |    |  |
|----|---|----|--|
| a. | $\left[ \begin{array}{l} \langle \text{appear} \rangle \\ \text{SYN} \mid \text{HEAD} \mid \text{POS verb} \\ \text{SUBJ} \langle \text{[3]} \rangle \\ \text{COMPS} \langle \text{VP}[\text{SUBJ} \langle \text{[3]} \rangle] \rangle \end{array} \right]$ | b. | $\left[ \begin{array}{l} \langle \text{call} \rangle \\ \text{SYN} \mid \text{HEAD} \mid \text{POS verb} \\ \text{SUBJ} \langle \text{NP} \rangle \\ \text{COMPS} \langle \text{[3]NP}, \text{XP}[\text{SUBJ} \langle \text{[3]} \rangle] \rangle \end{array} \right]$ |
|----|---|----|--|

These lexical entries ensure that verbs like *appear* selects a VP whose subject value is identical with its own value. This system then rules out examples like (47a).

Given the constructional constraints of the construction *tfr-cx* and the lexical entries for *what* and raising predicates, we now can see how the grammar generates a TFR example. Consider the object gapped TFR example first:<sup>7</sup>



The matrix verb in the relative clause *call* selects two NP complements, but its object NP is realized as a gap value ( $\text{[1]}$ ) which is eventually linked to *what*. The predicative complement *a musician* selects a subject which is identified with the object of the raising verb *call*. This gap in turn structure-shares with the [TFR +] *what*. Note that the HEAD value of the predicate NP *a musician* is identified with this *what* in accordance with the constructional constraint of *tfr-cx*, eventually

<sup>7</sup>We represent the gapped element in the tree for ease of exposition. See Kim and Sells (2008) with no terminal for a gapped element.



assigning an NP value to *what* and an indefinite semantic value too.<sup>8</sup> Since the HEAD value includes the syntactic AGR feature and the SEM value includes the indefinite property, the whole phrase is singular and indefinite. This explains the transparent effects of the TFR with respect to number and indefiniteness at least.<sup>9</sup>

There are other welcoming consequences of the present analysis. For example, since the predicative expression eventually determines the syntactic category of the whole clause, we can explain why the distributional possibilities of TFRs are determined by the predicative expression. It further accounts for the preposition restriction as well as coordination facts we have seen earlier. This is possible since the property of the whole TFR in question is determined by the nucleus expression whose syntactic and semantic features are identified with the expression *what* functioning as the head of the clause.

Binding facts also follow since the index value (part of the semantics value) of the whole clause is determined by the predicative expression. Extraction facts also can be expected since the present analysis induces the effect of treating the TRF as a parenthetical expression.

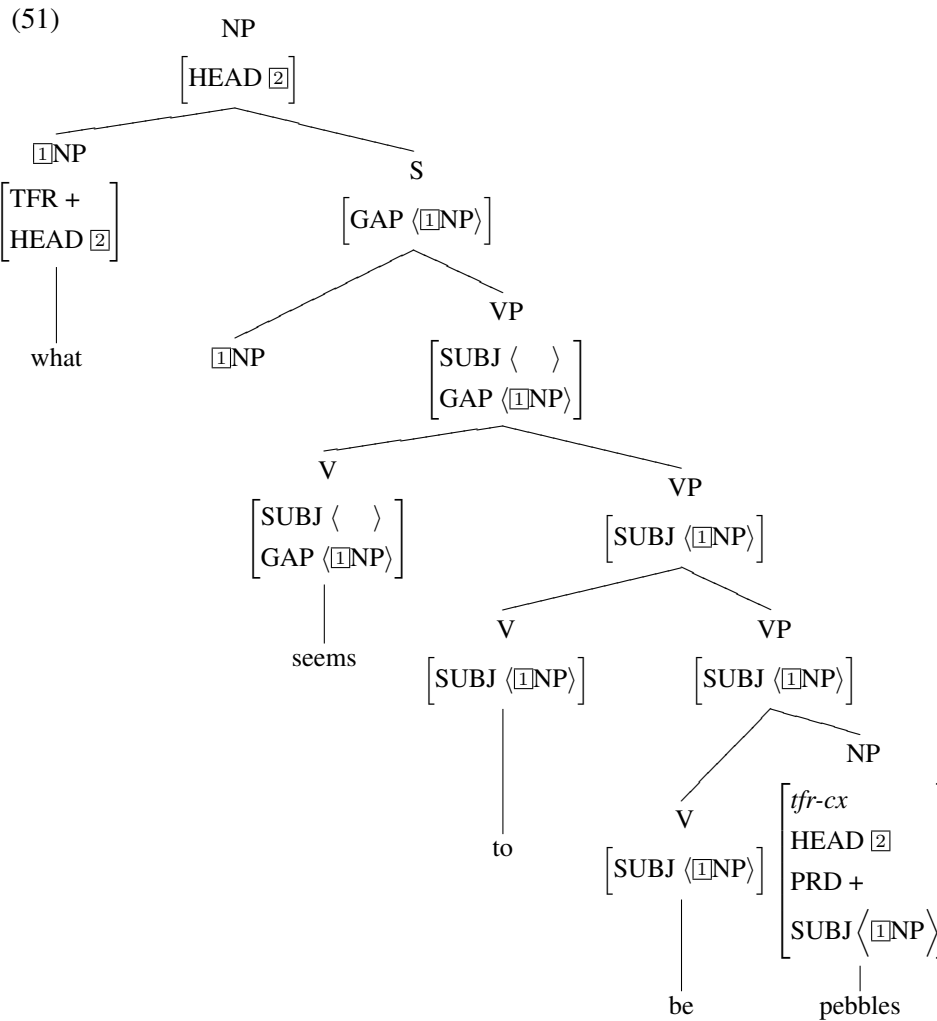
The present analysis can also provide a uniform analysis for the subject gapped TFR example. Consider part of the structure for a sentence like *What seem to be pebbles are strewn across the lawn*:<sup>10</sup>

---

<sup>8</sup>The indefiniteness is originated from the article *a*, and inherited to *what*.

<sup>9</sup>As for the syntactic head of the TFR, following Kim's (2001) analysis where in SFRs the *wh*-expression functions as the syntactic head as an independent constructional constraint, we assume that *what* is the syntactic head. See Kim (2001) and Yoo (2008).

<sup>10</sup>See Kim and Sells (2008) for a detailed analysis of treating the subject-gapped long distance phenomena in English.



The first thing to note here is that *seem*, *to* and *be* are all raising predicates, implying that their subjects are identical with the subject of their complement. The NP *pebbles* functions as a predicate selecting *what* as its subject, indicating the subject of the raising verb *be* is identical with this subject. Since *seem* and *to* are also raising predicates, the subject of *seem*, which is realized as the gap, is in turn identical with that of *pebbles*. In accordance with the *tfr-cx* construction and the lexical entry of *what*, the whole clause in effect inherits its number and definiteness value from the *pebbles*, reflecting the transparent effects of the TFRs.

## 5 Conclusion

English transparent free relatives (TFRs) display many different properties from seemingly similar standard free relatives (SFRs). In particular, they are intriguing with respect to their transparent effects. We have shown that it is not the *wh*-relative

pronoun or the relative clause that determines the clause's number, definiteness, and reference. It is the predicative nucleus that determines such syntactic as well as semantic values of the clause.

We have sketched a construction-based perspective for the analysis of intriguing English transparent relative clauses. Our analysis, adopting Grosu's assumption that *what* in TFRs is underspecified with its syntactic and semantic value, introduces the construction *transparent-relative-clause* with its own constructional constraints. This constructional constraint, interacting with independent properties of raising verbs, can offer us a streamlined analysis for the transparent effects of the construction without resorting to movement operations.

## References

- Baker, C.L. 1989. *English Syntax*. Cambridge, MA: The MIT Press
- Borsley, Robert. 1992. More on the difference between English restrictive and non-restrictive relative clause. *Journal of Linguistics* 28, 139-148.
- Bresnan, J. and J. Grimshaw. 1978. The Syntax of Free Relatives in English. *Linguistic Inquiry* 9. 331-391.
- Bresnan, Joan and Jane Grimshaw. 1978. The Syntax of Free Relatives in English. *Linguistic Inquiry* 9: 331-391.
- Grosu, Alexander. 2003. A Unified Theory of 'Standard' and 'Transparent' Free Relatives. *Natural Language and Linguistic Theory* 21: 247-331.
- Huddleston Rodney, and Geoffrey Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge University Press
- Jacobson, Pauline. 1995. On the Quantificational Force of English Free Relatives. In E. Bach, E. Jelinek, A. Kratzer, and B. Partee (eds.), *Quantification inn Natural Languages* Vol 2: 451-486. Kluwer, Dordrecht.
- Kajita, Masaru. 1977. Towards a Dynamic Model of Syntax. *Studies in English Linguistics* 5: 44-66.
- Kim, Jong-Bok. 2001. Constructional Constraints in English Free Relative Constructions. *Language and Information*, 5.1: 35-53.
- Kim, Jong-Bok. 2008. Grammatical Interfaces in English Object Extraposition. *Linguistic Research* 25.3: 117-131.
- Kim, Jong-Bok and Peter Sells. 2008. *English Syntax: An Introduction*. Stanford: CSLI Publications.
- Kim, Jong-Bok and Peter Sells. 2011. The Big Mess Construction: interactions between the lexicon and constructions. *English Language and Linguistics* 15, 335-362.
- Lee, Nam-Geun. 2009. On the Grammatical Interfaces in English Pseudo-Passive Constructions (In Korean). *Language Research*, 26.3: 21-42.
- McCawley, James. 1998. *The Syntactic Phenomena of English* (2nd Edition).

- Chicago and London: The University of Chicago Press.
- Qiurk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.
- Sag, Ivan A., Tom Wasow, and Emily Bender. 2003. *Syntactic Theory: A Formal Approach*. Stanford: CSLI Publications.
- Van Riemsdijk, Henk. 1998. Trees and Scions – Science and Trees. *Fest-Web-Page for Noam Chomsky*. <http://mitals.mit.edu/chomskydisc/riemsdyk.html>.
- Van Riemsdijk, Henk. 2000. Free relatives Inside Out: Transparent Free Relatives as Grafts. In B. Rozwadowska (ed.), *PASE Papers in Language Studies – Proceedings of the 8th Annual Conference of the Polish Association for the Study of English*, pp. 223–233. University of Wrocław, Wrocław.
- Van Riemsdijk, Henk. 2001. A Far From Simple Matter: Syntactic Reflexes of Syntax-Pragmatics Misalignments. In I. Kenesei and R. Harnish (eds.), *Perspectives on Semantics, Pragmatics and Discourse. A Festschrift for Ferenc Kiefer*, 21–41. Amsterdam: John Benjamins.
- Van Riemsdijk, Henk. 2006. Free Relatives. In M. Everaert and H. Riemsdijk (Eds.), *The Blackwell Companion to Syntax* Vol. II: 338–382. Blackwell.
- Wilder, Chris. 1999. Transparent Free Relatives. In K. Shahin, S. Blake and E-S Kim (eds.), *Proceedings of WCCFL 17*: 685–699. Cambridge: Cambridge University Press.
- Yoo, Eun-Jung. 2008. English Transparent Free Relatives: A Constraint-based Analysis. *Korean Journal of Linguistics* 33: 723–744.