On Perception Verb Complements

Tsutomu Matsunami

1. Introduction

In this article I will discuss the problems with regard to the analyses of perception verb complement constructions in English. Many proposals have been made in the past on the various problems of perception verb complements.[1] None of these proposals seems to have provided a fully satiafactory solution to the problems. In section 2 I will discuss some of the problems in the recent analyses. Section 3 will outline the analysis of the relevant constructions in the theory of Marantz (1984). His theory provides a clue to a possible solution to the problems. In section 4 I will propose a plausible solution to the problems in the framework of Marantz. I assume the readers' familiarity with the theory of Marantz.

2. Problems

2.1. Perception verbs occur with *that*−clause complement :

(1) a . I saw that John stole the car.

b . I heard that they came up the stairs.

They also occur with *to*−infinitival complement and bare VP complement. But they show the following contrasts in grammaticality :

(2) a .*I saw John to steal the car.

b . I saw John steal the car.

(3) a .*I heard them to come up the stairs.

b . I heard them come up the stairs.

(4) a . We saw John to be an obnoxious person.

b .*We saw John be an obnoxious person.

(5) a .*We've never seen John to be so happy before.

b . We've never seen John be so happy before.

As sentences (2a, b) and (3a, b) show, the tenseless form of verbs cannot occur in *to* -infinitival complement of perception verbs, but must occur in bare VP complement.

Sentences (4a, b) and (5a, b) show that in some cases, the VP complement must be *to* -infinitive, as in (4a), and in other cases, it must be bare infinitive, as in (5b).

In the previous analyses perception verbs with *to*-infinitival complement have been treated as one of "raising to object" verbs. The complement of raising to object verbs are assigned the same categorial status as that of the *that*-clause complement ; i. e., these verbs subcategorize for a $\bar{S}$ complement (See Akmajian (1977), Williams (1980) and Stowell (1981)) :

(6) a . I saw $_{\bar{S}}$ [that John stole the car]

b . I saw $_{\bar{S}}$ [John to be an obnoxious person]

An advantage of this analysis is that it simplifies the subcategorization of raising to object verbs-these verbs always take a sentential complement.

This analysis is supposed to reflect semantic differences between the perception verbs with *that*-clause or *to*-infinitival complement and those with bare VP complement. Akmajian (1977) and Kirsner and Thompson (1976), for example, claim that perception verbs with sentential complement express "indirect perception"; i. e., in sentence like (4a), we can assert the truth of the complement proposition with indirect evidence, without actually seeing *John*. Perception verbs with bare VP complement, on the other hand, express "direct perception".

This direct/indirect distinction, however, is too simple. It cannot account for the grammaticality difference between sentences (1a, b) and sentences (2a) and (3a). Since all of these sentences have a sentential complement, they should be able to express an indirect perception. However, only those sentences with a *that*-clause complement are grammatical. Perception verbs with *to*-infinitival complement do express indirect perception, as in (4a). Typical raising to object constructions have the same restrictions :

(7) a . I consider that John stole the car.

b .*I consider John to stole the car.

The *to*-infinitival complement of raising to object verbs is generally restricted to a state predicate or a perfective predicate (See Palmer (1974) and Borkin (1984)). The question here is why the *to*－infinitival complement of raising to object verbs is restricted to have only such a class of predicates. The clausal analyses do not seem to give a satisfactory answer.

2.2. Since perception verbs with bare VP complement express direct perception, the complement must be a directly perceivable event or temporal state. However, this direct perceivability does not seem to account for the differences of sentences (4a, b) and (5a, b)

in grammaticality. To judge whether a person is happy or not, we usually have to see him or his behavior. Similarly we need to observe a person or his behavior to judge whether he is an obnoxious person or not. So both cases are a report of direct perception. It might be argued that the *be* in the complement of (5b) ascribes "intentionality" to *John*, which makes the complement a perceivable temporal state. On the other hand, the complement in (4a) expresses a generic state, which makes sentence (4b) ungrammatical. See Williams (1984) for intentionality of *be* verb. Still the ungrammaticality of (5a) is not accounted for.

Furthermore, perception verbs take AP and PP complements :

(8) a . I've never seen John so sick.

b . I saw John inside the box.

These sentences are grammatical without the intentional *be*. The complements of these sentences express a temporal state. That is, (8a), for instance, means that John was sick when I saw him, not at other time. Now consider the differences of the following sentences in grammaticality :

(9) a . I saw John dead.

b . *I saw John be dead.

(10) a . *I've never seen John intelligent.

b . I've never seen John be intelligent.

Sentence (9b) is ungrammatical because *dead* is a durative state not under the control of *John*. (10a) is ungrammatical because *intelligent* is not a directly "seeable" state and we need to observe *John*'s intentional behavior to directly perceive his intelligence. So *be* is required to express his temporal state. The problem here is the grammaticalness of sentences like (8a, b) and (9a). *Dead*, for example, is a durative state predicate, not a temporal state predicate. The AP and PP complements in these constructions, however, say something about the temporal state of the NP object at the time of perception ; i. e., sentence (9a), for instance, means that I saw John when he was dead. The analysis of direct perceivability has nothing to say about why the AP and PP complements always express the temporal state of the object NP.

2.3. What is the categorial status of the complements of sentences like (2b), (5b), and (8a, b) ? The Small Clause analysis in the framework of Government and Binding theory (Chomsky (1981) and Stowell (1981)) would assign the following small clause categories to the complements (See the arguments of Safir (1983) for the Small Clause theory and the arguments of Williams (1983) against it) :

(11) a. I saw $_{V}$ [John steal the car]

b. I've never seen $_{A}$ [John so sick]

c. I saw $_{P}$ [John inside the box]

This clausal structure is implied by the Projection Principle of GB theory. The small clause is assigned the same θ-role with its *to*-infinitival counterpart. The main difference between *to*-infinitival complement and the small clause is that only *to*-infinitival complement has "tense feature". Thus in GB theory, the verbs of sentences (1)−(10) all have a clausal complement. This makes it difficult to explain the various problems we have seen in the previous sections.

As is well known, raising to object constructions permit a quantifier to have narrow scope with respect to the raising verb as well as wide scope :

(12) I consider someone to be intelligent.

=I consider [∃x [x is intelligent]]

=∃x [I consider x is intelligent]

Williams (1984) points out that perception verbs with a bare infinitival complement can have only wide scope :

(13) John saw someone leave.

=∃x [John saw x leave]

≠John saw [∃x [x leave]]

Under the Small Clause theory there is no reason for a narrow scope reading to be missing.

3. The Theory of Marantz

Here I will outline Marantz's analysis of English sentences relevant to the present discussion. He proposes three levels of sentence representations : the logico-semantic (l-s) structure, the syntactic (s) structure, and the surface structure. The l-s structure provides a representation of the semantic relations among the constituents of a sentence relevant to syntax. The s structure of a sentence displays the grammatical relations among constituents, such as the relation between a verb phrase and its subject or the relation between a verb and its object. The surface structure represents the surface phrase structure constituents of the language. The relations in the surface structure include the structural relations of constituent structure government and string adjacency and the morphology-connected relations of case marking and agreement.

At the l-s structure the argument-taker-argument relation and semantic-role-assigner -semantic-role-assignee relation are distinguished. The argument-taker-argument relations

are the relation between the function and its arguments. The role-assigner-assignee relations hold between a constituent bearing a semantic role and the item that determines which semantic role the constituent will bear.

3.1. First, let us see the analysis of the canonical predication constructions in the framework of Marantz :

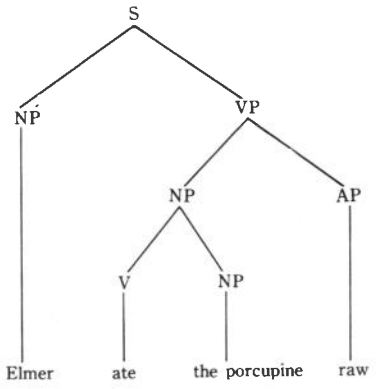
(14) a. Elmer ate the porcupine raw.

b. The porcupine arrived at the party terribly drunk.

In these sentences *the porcupine* bears two semantic roles at l-s structure. In sentence (14a), for example, it bears the eaten role, which is assigned by the verb *eat*, and the raw role, which is assigned by the predicate *raw*. Marantz says that *raw* in (14a) says something about the temporal location of eating of the porcupine event ; the porcupine was eaten at the time when it was raw.

Thus the NP object receives its semantic roles from the main verb and the predicate. It is the argument of the main verb, but not the argument of the predicate *raw*. The predicate *raw* is analyzed as a predicate modifier, taking the verb phrase *ate the porcupine* as the argument[2]. The l-s structure of (14a) is represented as in (15).

(15)



3.2. In a canonical raising sentence, a constituent bears a grammatical relation outside the clause in which it bears its semantic relations. The theory of Marantz prohibits analyzing sentences like (16) as containing some sort of propositional argument to the raising verb as the Small Clause theory does. Nor can the phrase *fond of porcupines* be analyzed as a predicate as the Predication Theory of Williams (1980, 1983) does.

(16) Hortense considers Elmer fond of porcupines.

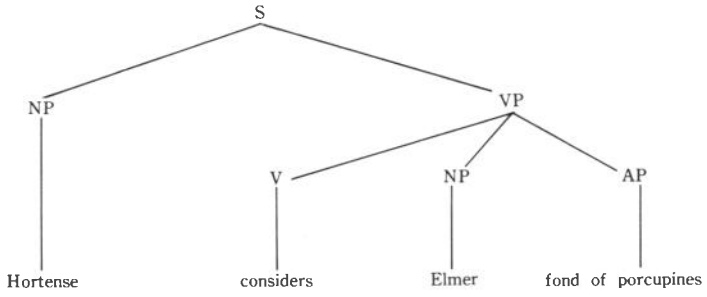
Marantz claims that sentences (16) and (17) must have different l-s structures, for the verb

consider have different meanings in (16) and (17).

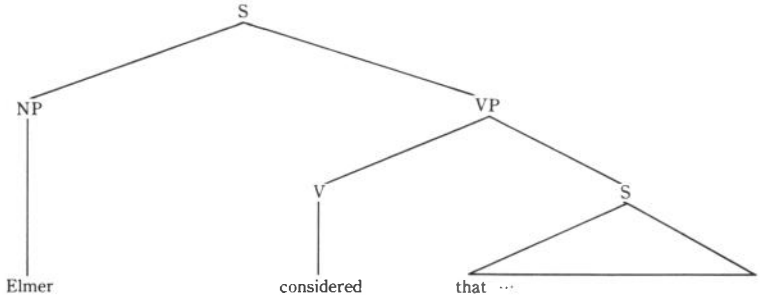
(17) Elmer considered that Hortense had been with his firm for five days.

When *consider* occurs with *that*-clause, it has meaning like "think over" ; when it appears in a construction like (16), its meaning is close to that of *believe*. Sentences (16) and (17) are assigned the l−s representations as in (18) and (19), respectively.

(18)



(19)



In (19), *that*-clause is a propositional argument of the matrix verb. In (18), *Elmer* bears no semantic relations with respect to *consider*. This NP appears in the VP of the sentence as the semantic role assignee, that is, *fond of porcupines* assigns *Elmer* a semantic role. However, the AP *fond of porcupies* does not take *Elmer* as an argument. This AP functions as a modifier of the NP *Elmer*. Unlike a regular modifier, which takes an argument as a modifiee but does not assign a semantic role to the modifiee, this modifier does not take the modifiee as its argument but only assigns a semantic role to the modifiee. This sort of modifier is defined as "predication" modifier. Thus at 1-s level the NP *Elmer* bears no semantic relations with respect to the main verb. At s level it is assigned a grammatical relation with respect to the main verb.

Marantz claims that raising to object sentences lik (20) share the semantics of raising consrtructions with a predication modifier like (16), not the semantics of raising constructions with *that*-clause[3].

(20) Hortense believes Elmer to have sold his last porcupines.

Consider the following sentences :

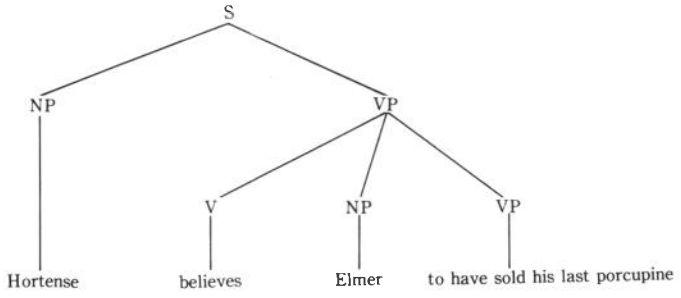
(21) a . Hortense considered Elmer to have sold his last porcupine.

b .*Hortense considered Elmer to sell his last porcupine (today).

c . Hortense considered that Elmer sells his last porcupine today.

In (20a), *Hortense* is said to believe a predication, that is, the property of having sold his last porcupine applied to *Elmer*. In (20c), the considerer is thinking over some statement. (20b) shows that the VP complement to *consider* must contain some tense/aspect (usually perfect, progressive, or habitual) that allows it to take on a property reading. Thus Marantz proposes that English allows some verbs to take predicates as if they were modifiers and turn them into predications. This is the marked property of raising verbs, for verbs usually do not take nonmaximal projections of VP predicates as arguments. Raising verbs with modifiers are unmarked cases, since modifiers are maximal projections. Thus the *to*-infinitival complement of (20) is treated as a predication modifier, having the same 1-s structure with sentences like (16) :

(22)



4. Solutions

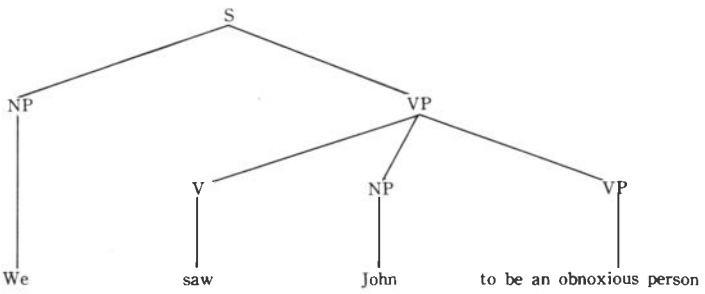
In this section I will try to solve the problems discussed in section 2 in the framework of Marantz we have seen in section 3.

4.1. I assume that perception verbs with *to*-infinitival complement have the same 1-s structure with raising to object verbs.

(23) We saw John to be an obnoxious person.

In sentence (23) *John* bears no semantic relations with respect to the verb *saw* and its semantic role is provided by the infinitival complement *to be an obnoxious person*. The infinitival complement is an argument of the verb and functions as a predication modifier of *John*, expressing a property of the object NP. The 1-s structure of (23) is :

(24)



We have seen that the direct/indirect perception distinction cannot explain the ungrammaticality of the following sentences :

(25) a . *I saw John to steal the car.

b . *I heard them to come up the stairs.

As we have seen, the infinitival complement of raising to object verbs, being a predication modifier of the object NP, must express a property of the NP. This explains the ungrammaticalty of sentences (25a,b), since the complements express activities.

4.2. How can the perception verb constructions with bare predicate complement be represented at 1-s level? We have seen that they have different meanings from those constructions with *to*-infinitival complement. This implies that a bare predicate perception construction has different 1-s structure from its *to*-infinitival counterpart.

Remember that the predicate modifier in sentences like (26) specifies the temporal location of the event described by the main verb :

(26) Elmer ate the porcupine raw.

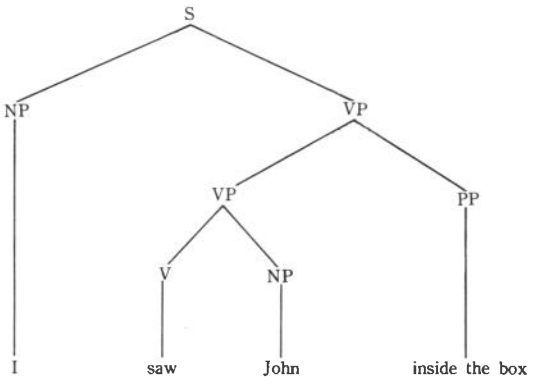
Now consider the perception verbs with AP or PP complements :

(27) a . I saw John dead.

b . I saw John inside the box.

As we have seen, these AP and PP complements also specifies the temporal location of perceiving event. So I conclude that the verbs in sentences like (27a, b) also take a predicate modifier ; i. e., the complement predicate serves as the predicate modifier of the main predicate *saw John*. The object NP bears two semantic roles : one from the main verb and the other from the complement predicate. It is the argument of the main verb but not of the complement predicate. The 1-s structure of (27b), for example, will be represented as in (28).

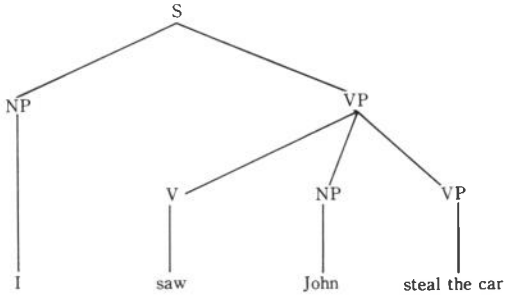
(28)



4.3. Perception verb constructions with bare VP complement do not seem to have the same l-s stucture with those of sentences like (27a, b), since the bare VP complement does not serve as predicate modifier, specifying the temporal location of the main verb event. Rather, in these constructions the direct perception of the event or state described in the complement is reported. They also do not seem to have the same l-s structure with raising to object constructions, since these complements do not express a property of the object NP.

Thus I am led to conclude that in sentences like (2b) and (3b), the VP complement acts as argument of the main verb and receives a semantic role from the main verb. The VP complement, however, does not serve as the modifier of the object NP as in the case of raising to object constructions. In bare VP complement constructions the direct perception of the event or temporal state is reported. I think that the object NP receives a semantic role from the main verb and acts as an argument of the main verb. It also receives a semantic role from the complement predicate, but does not act as the argument of the complement predicate[4]. Although the l-s structure of perception verb constructions with bare VP complement is represented like that of raising to object constructions, the semantic and the grammatical relations of the constituents are different. The l-s structure of (2b) is ;

(29)



4.4. As we have seen, perception verb constructions with bare VP, AP or PP complement allow the object quantifier to have only wide scope with respect to the main verb. This is probably bacasue the object NP does not act as the argument of the complement predicate but acts only as the argument of the main verb. On the other hand, the object quantifier of the raising to object verbs can have both narrow and wide scope because it acts as the argument of the complement predicate and also obtains a grammatical relation with the main verb at s structure level.

I have concluded that at l-s level the complement of perception verbs have four different functions. The *that*-clause complement is the propositional argument of the vetb. The *to* -infinitival complement functions as the predication modifier of the object NP. The bare VP complement is an argument of the main verb. The AP and PP complements function as the predicate modifier of the main predicate. This classification of perception verb complements accounts for the differences of sentences (2)-(5) and (8)-(10) in grammaticality. Since the *to* -infinitival complement is a predication modifier, it has to express a property of the object NP. This explains the ungramaticality of sentences (2a) and (3a). Sentence (5a) is ungrammatical because the complement expresses a temporal state of the object NP, not its property. Since the bare VP complement is a direct argument of the main verb, I assume that the complement must express a directly perceivable event or a directly perceivable temporal state of the object NP. This explains the ungrammaticality of sentences (4b) and (9b). The complement of (4b) expresses a generic property of the object NP. The complement of (9b) expresses a durative state. The bare VP complement cannot be a predicate modifier, either. For the AP and PP complements as predicate modifer specify the temporal location of the main predicate event. This also explains the ungrammaticality of sentence (10a) because the complement *intelligent* does not specify the temporal location of the main predicate event.

Footnote

1. The discussion here is limited to sentences like (1)−(5) and (8)−(10). Although I will not treat perception verbs with verb-*ing* complement, a similar observation could be made about them. See Gee (1975, 1977) for various syntactic phenomina of perception verb complements.
2. Modifier is defined as a function from predicates of type X to constituents of type X.
3. Williams (1980) claims that sentences like (i) and sentences like (ii) have different structures because the interpretion of them is different :
   (i) John considers Bill silly.
   (ii) John considers Bill to be silly.
   I think that the difference could be attributed to the presence and absence of *to*.

4. This analysis of perception verb construction with bare VP complement is another marked property of English in the framework of Marantz, since it violates the Nonmaximal Projection Principle.

References

Akmajian, A. (1977), "The complement structue of perception verbs in an autonomous syntax framework". in P. W. Culicover, T. Wasow and A. Akmajian eds, *Formal Syntax*, Accademic Press, New York, 427-460.

Borkin, A. (1984), *Problems in Form and Function*, Ablex Publishing Corporation, New Jersey.

Chomsky, N. (1981), *Lectures on Government and Binding*, Foris Publications.

Gee, J. (1975), *Perception, Intentionality and Naked Infinitives : A Study in Linguisitcs and Philosophy*, Doctoral dissertation, Stanford University.

Gee, J. (1977), "Comments on Akmajian", in P. W. Culicover, T. Wasow and A. Akmajian eds, *Formal Syntax*, Academic Press, New York, 461-481.

Kirsner, R. S. and Thompson, S. A. (1976), "The role of pragmatic inference in semantics : a study of sensory verb complements in English", *Glossa*, 10, 200-240.

Marantz, A. P. (1984), *On the Nature of Grammatical Relations*, The MIT Press.

Palmer, F. R. (1974), *The English Verb*, Longman.

Safir, K. (1983), "On Small Clause as Constituents", *LI*, 14 (4), 730-735.

Stowell, T. (1981), *Origins of Phrase Structure*, Doctoral Dissertation, MIT, Cambridge, Massachusetts.

Williams, E. (1980), "Predication", *LI*, 11 (1), 203-238.

Williams, E. (1983), "Against Small Clause", *LI*, 14 (2), 237-308.

Williams, E. (1984), "There-Insertion", *LI*, 15 (1), 131-153.