

Attitudes De Se and Anaphora: A Presuppositional Account

Hyunoo, Lee * †
Inha University

Hyunoo Lee. 1999. Attitudes De Se and Anaphora: A Presuppositional Account. *Language and Information* 3.2, 51–66. This paper argues that the de se reading is a composite of the corresponding de re one plus the presupposition that there exists a strong epistemic relation between the object of attitude, linguistically realized as an anaphor, and the attitude-bearer, the referent of the anaphor's antecedent. Close examination of the relation between the attitudes de se and anaphora reveals that the notion de se, as currently understood, should be generalized to explain the choice of anaphors. A taxonomy of anaphors in terms of a general notion like presupposition provides further empirical support for the approach taken here. (Inha University)

1. Introduction

The purpose of this paper is twofold. The first goal is to seek a natural, restrictive grammatical model for the semantics of attitudes towards oneself, called attitudes de se by Lewis (1979) or essential indexicals by Perry (1979). Examination of current approaches to de se attitudes shows that no matter how philosophically plausible they are, they fail to receive linguistic supports. This leads me to propose a presuppositional account of the phenomenon in question where the de se reading is treated as a composite of the corresponding de re one plus the presupposition that there exists a strong epistemic relation between the object of attitude, linguistically realized as an anaphor, and the attitude-bearer, the referent of the anaphor's antecedent. The second goal is to investigate the relation between the notion de se and anaphora. Examination of this relation reveals that the notion de se, as currently understood, is not general enough to explain the choice of anaphors. The fact that it is accounted for naturally by a general notion like presupposition hence provides further empirical support for the approach to de se attitudes proposed here.

This paper falls into three major sections followed by some concluding remarks. In section 2, I first discuss what the attitudes towards oneself are and how they are special. This preliminary discussion is followed by a review of Chierchia's (1989) theory of de se which, drawing on Lewis (1979), treats verbs like *believe* as denoting self-ascriptive relations of agents to properties. I argue against this approach from the point of view of compositionality. Building on the ideas advanced in Perry (1979) and Böer and Lycan (1985), in section 3, I propose an alternative account on which the de se reading of an ut-

* For helpful comments and discussions on the subject addressed here, I would like to thank Chungmin Lee, Hee-Don Ahn, and Ed Keenan. I would also like to express my sincere thanks to two anonymous reviewers for their suggestions and comments. The author wishes to acknowledge the financial support of the Korea Research Foundation made in the program year of 1998.

† 402-751 인천 남구 용현동 253번지 인하대학교 영어교육과, E-mail: hylee@dragon.inha.ac.kr

terance is viewed as a composition of the truth-conditional content of the corresponding *de re* reading and the strong epistemic relation of an agent to an object of proposition presupposed by that utterance. To implement this idea formally, I propose a two-tiered semantic model which takes an utterance with the *de se* reading to expressively presuppose the relation mentioned above. It follows from the model proposed here that the *de se* reading asymmetrically entails the *de re* one. In section 4 I devote myself to a discussion of the choice of anaphors or pronouns in relation with the notion *de se*. Drawing on data from Korean, I argue that the short-distance anaphor *susulo* 'self' is an item the use of which presupposes that the referent of its antecedent have access to himself. It will be shown that the existence of such an item provides strong evidence in favor of the approach taken here.

2. De Se and Linguistic Theory

Since Castañeda (1966) brought the special nature of *de se* attitudes to light, it has raised many difficult philosophical problems for the traditional Fregean concept of proposition, especially in the context of belief and other attitudes. Recently, such philosophers as Lewis (1979), Perry (1979), Stalnaker (1981), and Böer and Lycan (1980) and Böer and Lycan (1985) tackle these problems and put forth ways of handling them. As some of these researchers have pointed out, to incorporate their insights into a model of grammar is far from trivial. Fully aware of this, some linguists like Cresswell (1985) and Chierchia (1989) have attempted at a semantic theory which requires a more articulated, rich view of propositions or a grammatical model in which *de se* readings are associated with logical forms that allow certain clauses to be translated into (extensional) properties. Drawing on Chierchia's approach to *de se*, Reeves (1996) proposes a theory which constrains the distribution and interpretation of long-distance anaphora in terms of syntactic devices for individuating utterances, beliefs and other intensional states of linguistic agents.

For the purposes of this paper, I would first like to explicate the privileged nature of *de se* attitudes. Reviewing two competing philosophical approaches to these self-regarding attitudes, I argue against the line of research paradigm advocated in Lewis (1979), Chierchia (1989), and Reeves (1996).

2.1 Attitudes De Se

For concreteness, let us first consider the following examples:

- (1) a. John believes that the president of the United States is in danger.
- b. Bill Clinton is the president of the United States.
- c. John believes that Bill Clinton is in danger.

It has been noted that given (1b), which is true, (1a) and (1c) are not logically equivalent. This is because verbs of propositional attitudes like *believe* create opaque contexts where substituting an expression for another with the same reference leads to change in the truth-condition of the sentence involved. Within the mainstream semantic theory that follows Frege's distinction between reference and sense, (1a) is taken to be ambiguous. On one reading, the sentence reports John's belief that whoever is president of the United States is in danger. What is believed is not simply the proposition of (1b), but in terms of possible world semantics, the set of possible worlds in which the president of the United States is in danger. John may believe that the president of the United States is in danger without believing that Bill Clinton is so. The reading in question is said to ascribe a

de dicto belief. The second reading says that (1a) and (1c) have the same truth value. (1a) is said to report a de re belief that John has about a certain individual in the world where he is. Unlike the de dicto belief, which may be purely conceptual in the sense that he is not directly acquainted with the president of the United States in question, some kind of link directly relates Bill Clinton to the individual who is in fact president of the United States in the de re belief. That is, (1a) says that John believes of the president of the United States, who is Bill Clinton, that he is in danger.

The characteristic feature of ascriptions of attitudes de re is that they induce a strong epistemic relation between belief-bearers and the objects of their beliefs. Philosophers rightly point out that there is another kind of ascription which shares this property with the ascription of de re attitudes but is nevertheless different. Consider the sentences in (2).

- (2) a. John believes that he is in danger.
 b. John believes that he himself is in danger.
 c. John believes, of John, that he is in danger.

According to philosophers such as Castañeda (1966), Lewis (1979), and Perry (1979), (2a) is ambiguous between (2b) and (2c): the latter reports a de re attitude and the former describes a first person, self-regarding, or de se attitude. The de se attitude in (2b) resists assimilation to the de re one. It seems that (2b) entails (2c), but not the converse. Suppose that John sees a man in dangerous situation through a mirror and comes immediately to believe of that man that he is in danger; but unbeknownst to John, it is in fact John himself that he is viewing in the mirror. Then (2c) is true since that man is John; his belief is causally determined by the man he is seeing and his dispositions would be systematically affected if the man's properties were different from what they are. On the other hand, it seems that (2b) might well be false since the mental state of John in (2b) would be the one of someone who says to himself, I am in danger, a circumstance in which his dispositions are quite different from those expected from (2c).

Note also that de se attitudes cannot be reduced to de dicto ones either. Because of the effect of Principle C of the binding theory, it is difficult to construct examples of the form in (2) which exemplify the ascription of de dicto attitudes. For purposes of the present paper, assume that some contextual factors may allow the definite description, the president of the United States, to be coreferential with the proper name, John, in a sentence like (3).

- (3) [John]_i believes that [the president of the United States]_i is in danger.

Then (2b) neither entails nor is entailed by (3). Suppose John suffers from terrible amnesia. John may believe that he himself is in danger but fails to believe that the president of the United States is in danger simply because he does not recognize that he is president of the United States himself. Or, as before, he may believe de dicto that the president of the United States is in danger without believing he himself is in danger.

2.2 Philosophical Underpinnings: Lewis 1979 and Boer and Lycan 1980

The foregoing discussion reveals two things. First, de se attitudes do not fit in de dicto or de re ones. Second, the ascription of de se attitudes has a common property with the ascription of de re attitudes in that both involve a strong epistemic relation of attitude-bearers to the objects of their attitudes. The issue that arises naturally is, then, the question of why the de se attitudes seem to be special and require more than the de re

attitudes. Opinions vary as to how to explain this complexity which the attitudes towards oneself give rise to.

On an approach defended by Castañeda (1966), Lewis (1979), and Perry (1979), attitudes de re and attitudes de se are different but are not entirely independent of each other. That is, sentences like (2a) are ambiguous between the de re and de se readings, as described in the above. The latter reading entails but is not entailed by the former. This is compatible with the fact that the epistemic access we have to ourselves automatically carries attitudes towards a certain res (namely, the self). Thus, if (2a) is true on the de se reading, it is also true on the de re reading. But the truth of (2a) on the de re reading, which might describe the circumstance implied by the above mirror scenario, cannot guarantee that the sentence is also true on the de se reading.

On another approach advocated by Böer and Lycan (1980), the problem is addressed from a very different angle. They claim that de se attitudes are only a slightly special case of de re attitudes. It is possible to have access to a certain entity from different perspectives which are equally direct or vivid and thus are equally fitting to allow beliefs about specific res. But different ways in which an object is accessed can result in one's attribution of contradictory properties to it without thereby necessarily being criticized for having contradictory beliefs.¹ In sentence (2a), John has access to himself in two guises: as the person he would refer to as the first person pronoun *I* and as the person he views in the mirror. Both ways of having access to himself may probably make his attitudes towards himself de re. Under one guise, he believes of John that he is in danger, and under the other he doesn't. Therefore, the problem of (2a) can be reduced to a more general problem which de re attitudes give rise to, and any solution to this more general problem can be applied to the problem in question. The following gives a summary of Böer and Lycan (1980):²

- (4) a. The object of a self-regarding attitude, viz., the proposition believed, as identified by a set of possible worlds, is precisely the same as that of the corresponding de re belief.
- b. Ascriptions of attitudes de se and the corresponding de re ascriptions have the same truth-conditions; any implication apparently carried by de se ascriptions but not by de re ascriptions is pragmatic at best, not semantic.

Thesis (4a) is not likely to be controversial since John's belief about himself and someone else's corresponding de re belief about him have the same truth-condition—they are true under precisely the same conditions and exactly in the same possible world. Thesis (4b), however, is controversial and raises a head-on objection to the position taken by the Castañeda approach. Given these two competing views, a more empirically reliable test is called upon to decide whether sentences such as (2a) are semantically ambiguous or de se ascriptions just bring forth some sort of pragmatic implications. This

1. Consider the example from Quine (1956) in (i).

- (i) a. Ralph believes that the man in the brown hat is a spy.
b. Ralph believes that the man seen at the beach is not a spy.

Suppose Ralph has seen a certain man, called Orcutt, under two types of circumstances, first wearing a brown hat and arousing Ralph's suspicion that he is a spy, second more innocently sitting on the beach; Ralph is unaware that the man seen on these occasions are one and the same. On the de re reading, both sentences in (i) are true, since the two descriptions, the man in a brown hat and the man at the beach, are coreferential for the speaker although Ralph would use them as NPs with different reference.

2. For a more detailed discussion, see Böer and Lycan (1985).

is the issue that I will take up in section 3, where I argue that although de re and de se ascriptions have the same truth-conditions, they carry different non-truth-conditional meanings. Before doing so, I first give a brief review of the linguistic implementation of the Castañedan approach in Chierchia (1989).

2.3 Linguistic Implementation: Chierchia 1989

Building on Lewis's (1979) idea, Chierchia proposes that a linguistic theory of de se take account of the points in (5).

- (5) a. De se readings are associated with specific linguistic structures.
 b. The compositional semantics of such structures crucially resorts to properties rather than propositions.

In order to motivate (5), Chierchia discusses the validity of arguments of the form in (6) and (7) which involves infinitives and gerunds.

- (6) a. Pavarotti { tried/practiced/began } everything that Domingo { tried/practiced/began }.
 b. Domingo tried/practiced/began singing Rigoletto.
 c. Pavarotti tried/practiced/began singing Rigoletto.
- (7) a. The cat wanted to eat the cheese.
 b. The mouse got what the cat wanted.
 c. The mouse got to eat the cheese.

He claims that the validity of argument in (6) is ensured if verbs like *try*, *practice*, and *begin* are assumed to denote relations between individuals and properties which are indirectly attributed to those individuals. He continues to argue that the validity of argument in (7), which was originally discussed by Fodor (1975), also points to the same conclusion. If the verb *want* relates an individual to a proposition in the manner shown in (8), it would incorrectly follow from (9), the translation of the second premise (7b), that what the mouse got is that the cat eat the cheese.

- (8) the cat($\lambda x[x \text{ wants that } x \text{ eats the cheese}]$)
 (9) $\lambda x[\text{the cat wanted } x \wedge \forall y[\text{the cat wanted } y \rightarrow x = y] \wedge \text{the mouse got } x]$

The logic of argument that Chierchia followed is this. The validity of inferences in (7) and (9) supports that verbs like the ones discussed above take property-denoting infinitives or gerunds. The relations denoted by these verbs which yield asymmetric propositional entailment unambiguously characterize attitudes de se. Given that (10a) asymmetrically entails (10b), it is concluded that only the de se reading is available in (10a).

- (10) a. Pavarotti wants very much to get help.
 b. Pavarotti wants very much for Pavarotti to get help.

Chierchia claims that anaphora with proposition-taking verbs exemplified in (11) and (12) provide further evidence in support of (5).

- (11) a. Domingo believes that he is a genius.
 b. Pavarotti believes it too
 c. Pavarotti believes the thing that Domingo believes
 d. Pavarotti believes what Domingo believes
- (12) a. Pavarotti crede che i propri pantaloni siano in fiamme.
 Pavarotti believes that selfs pants are on fire
 ‘Pavarotti believes that his pants are on fire.’
 b. Pavarotti crede che i suoi pantaloni siano in fiamme.
 Pavarotti believes that his pants are on fire
 ‘Pavarotti believes that his pants are on fire.’

He notes that the sentences in (11)b–d are ambiguous between the strict reading and the sloppy one. For example, (11b) means that Pavarotti believes that Domingo is a genius or he believes that Pavarotti himself is a genius. On his account, verbs like believe take propositions or self-ascriptive properties, and thus a strict reading of the object NPs in (11b–d) obtains in case they are anaphorically linked to a proposition and a sloppy one, in case that they are anaphorically linked to a property. In a word, the ambiguity between the strict and sloppy readings depends on whether the pronoun in (11a) is interpreted *de re* or *de se*. Unlike pronouns like the English *he* in (11a) and the Italian *suoi* in (12b), which are interpreted *de re* or *de se*, a long-distance anaphor like *propri* gives rise to only the *de se* reading of the sentence, which Chierchia thought gives empirical support for (5).

Limitations of capacity and space prevent me from discussing details of the formal mechanisms of *de se* which Chierchia introduced, but I only briefly mention the way how his grammar of *de se* works. Consider (13).

- (13) a. Mary thinks that she is in danger.
 b. $Mary_i$ thinks that [she_i is in danger]
 c. $Mary_i$ thinks [O_i that [she_i is in danger]]
 d. λx [thinks(x , x is in danger)](m)
 e. thinks(m , λx [x is in danger])

Sentence (13a) is associated with the two LFs given in (13b) and (13c). In (13b) the pronoun is directly bound by its antecedent, which yields the translation in (13d), hence the *de re* reading of the sentence. In (13c) the pronoun is indirectly bound by its antecedent via the operator O_i adjoined to S and coindexed with that antecedent. This operator is understood as a λ -abstractor, which means that (13c) is translated into (13e), the *de se* reading of (13a).

In the remainder of this section I would like to argue against Chierchia’s approach to *de se* attitudes. For one thing, the fact that pronouns in the opaque context are interpreted *de re* or *de se* in languages like English and Italian does not motivate what his account relies on, viz., (5). Nor does the fact that long-distance anaphors necessarily involve *de se* attitudes in a language like Italian. From a typological point of view, long-distance anaphors in the appropriate context do not necessarily induce *de se* attitudes, and even short-distance anaphors do exhibit something similar to the *de se*

phenomenon—both of which I will show in section 4. Whether pronouns or anaphors are interpreted de se optionally or obligatorily is a purely lexical matter and has nothing to do with the interpretive procedure that requires the *that*-clauses containing them to denote properties which will be self-ascribed to attitude-agents.

Secondly, the fact that anteceded by a singular term, a pronoun in the belief-sentence gives rise to the strict and sloppy readings, as in (11a) does not support his view. Instead, it raises a problem for it. Recall that on his account, the representation of de re attitudes is responsible for the strict reading, and the representation of de se attitudes, for the sloppy one. Consider now (14) (= his (10)).

- (14) a. Everyone in that room thinks that he is Hume.
 b. [[everyone in that room_i][t_i thinks that [O_i [he_i is Hume]]]]
 c. $\forall x_i$ [person in that room(x_i) \rightarrow thinks(x_i , λx [x is Hume])]

As Chierchia himself points out, (14a) is ambiguous between the de se reading given in (14b-c) and the de re reading given in (15).

- (15) a. [[everyone in that room_i][t_i thinks that [he_i is Hume]]]
 b. $\forall x_i$ [person in that room(x_i) \rightarrow thinks(x_i , x_i is Hume)]

Hence, the de se/de re distinction cannot be reduced to the standard kind of bound anaphora. Let us now consider (16).

- (16) a. Everyone in that room thinks that he is genius.
 b. No one in this room thinks so.

In (16b) the pronominal *so* is anaphorically related with the *that*-clause of (16a), but it receives only the sloppy reading. Chierchia's account is problematic since the presence of the de re representation in (15) incorrectly predicts that (16b) should have the strict reading which it lacks. It seems to me that the ambiguity between the strict/sloppy readings of a pronoun in a sentence like (11a) has to do with the options in which it directly picks up the reference of its referential antecedent or it is bound by a λ -abstractor.

Although it is not easy to give an alternative account of the validity of inferences in (6) and (7), it is not necessarily the case that it should count as evidence for the property-denoting status of infinitives and gerunds. The principle of compositionality does not force it to do so, contra Chierchia's argument. And a moment's thought suggests that the opposite is the case. Consider (17).

- (17) a. John believes that he is a genius.
 b. John believes that he is a genius and Mary is a fool.

Like (17a), (17b) is used to report not only attitudes de re but attitudes de se. That is, John believes that he is a genius and Mary is a fool, and is aware that he himself is the object of his belief. On Chierchia's account, the *that*-clause must denote a property to be self-ascribed to John. But this is impossible since the conjunction *and* is interpreted as a boolean meet \wedge and since compositionality disallows the concatenation of a property-denoting expression with a proposition-denoting expression, as indicated by * in (18).

- (18) believes(j , *(λx [x is a genius] \wedge m is a fool))

One might argue that using one of the definitions of self-ascriptivity proposed in Chierchia 1989 is a way out of the problem. Consider the definition in (19).

- (19) x stands in the belief relation with property Q (i.e., x self-ascribes Q , in Lewis's terms) iff x believes (de re) that x has Q and furthermore $K(x, x)$, where K is the cognitive access that we have to ourselves.

Although this definition captures the fact that de se readings (asymmetrically) entail the corresponding de re readings, it cannot give any solution to the problem intrinsic to (18). By virtue of compositionality and the boolean meet \wedge , the self-ascriptive property in (18) should be first converted to a proposition or the converse, but (19) makes reference to a propositional attitude verb and therefore applies to the whole proposition not just part of it.

3. Towards a Presuppositional View of De Se

In the foregoing section I discussed the two competing philosophical views on attitudes de se and argued that the grammar of de se proposed in Chierchia (1989) fails to receive empirical support. In this section I would like to develop a theory of de se which is linguistically motivated. Before pursuing this goal, I first consider whether and how attitudes de se and de re are distinct phenomena from the viewpoint of the truth-conditional semantics in section 3.1. Given that both attitudes give rise to two different readings, I propose a two-tiered semantic theory of de se which follows a line of thought advocated in Kaplan (1979) and Perry (1979), in section 3.2.

3.1 Ambiguous or Not?

As Zwicky and Sadock (1975) demonstrate, it is extremely difficult to test whether an expression is ambiguous when one of the readings which the expression is supposed to have entails the other. In this connection it is natural and meaningful to ask what implications ambiguity tests used by linguists tell concerning the de re/de se distinction.

A type of practice favored by philosophers as well as linguists to argue for ambiguities is a difference in understanding alone. As we have seen, the understanding of a de se attitude is clearly different from that of the corresponding de re attitude. As many researchers have pointed out, a mere difference in the understanding of an utterance is not sufficient to convince ourselves of the ambiguity of the utterance. Keeping this in mind, let us consider the sentences in (19).

- (20) a. That's a dog, but it isn't a dog.
 b. *That's a lion, but it isn't a lion.
 c. That dog isn't a dog; it's a bitch.
 d. *That lion isn't a lion; it's a lioness.

The noun *dog* is used as the technical term dog male canine and the ordinary language term canine, but the noun *lion* does not have such a distinction. Thus, (20a) is not a contradiction, but (20b) is, as the additional information in (20c) and (20d) bears out. The contradiction test above shows that if an expression has two readings one of which entails the other, to assert the latter reading and supply material that implies the denial of the first one does not yield a contradiction. Let us now consider (21).

- (21) a. John believes that he is in danger, but he does not believe that he is in danger.
- b. John believes that he is in danger, but he does not believe that he himself is in danger.
- c. John believes that he is in danger, but he does not believe that he is in danger just because he has been a patient of terrible amnesia.

It seems that people judge (21a) contradictory at first sight. But if they are given some explanation about John's mental state, for example, the one described in (21c), they are willing to accept the sentence. Moreover, sentence (21b) in which the presence of the emphatic reflexive pronoun *himself* forces the sentence to report a de se attitude is more easily accepted than (21a). The contradiction test thus supports the view that the de re/de se distinction is semantic.

Another type of test for ambiguity used in the literature is identity tests like (22).

- (22) a. John and Martha left, and so did Dick and Pat.
- b. John and Martha are married, and so are Dick and Pat.

(22a) allows the crossed understandings—the reading of John left together with Martha and the one it entails—but (22b) does not allow the crossed understandings—the reading of John is married to Martha and the one it entails.³ This amounts to saying that (22a) is vague, but (22b) is ambiguous. Now consider (23).

- (23) a. John believes that he is in danger, and so does Mary.
- b. John believes de se that he is in danger, and Mary believes de se that she is in danger.
- c. #John believes de se that he is in danger, and Mary believes de re that she is in danger.
- d. ?John believes de re that he is in danger, and Mary believes de re that she is in danger.
- e. #John believes de re that he is in danger, and Mary believes de se that she is in danger.

As is well-known from Reinhart (1983), (23a) is ambiguous between the strict reading and the sloppy one. The strict reading is irrelevant to the de re/de se issue, so let us confine ourselves to the sloppy reading. Given this, we may ask whether (23a) yields the crossed understandings. The above paradigm shows that if the first conjunct clause of (23a) is understood to report a de se attitude, the second conjunct must be so. If the first conjunct clause is understood to report a de re attitude, the second conjunct cannot

3. Zwicky and Sadock (1975) argue that the crossed understandings are available for (22b) since anyone who is married is married with someone. I think there is something unpleasant with their reasoning. It is true that anyone who is married is married with someone but it is not necessarily the case that that someone should be one of the referents of the conjunct NPs. If (22b) allows the crossed understandings, the sentence may describe the situation in which John and Martha are a couple and Dick and Pat are married but not a couple. The fact that the sentence is not used to describe such a situation suggests that no crossed understandings are available for (22b).

be understood to report a *de se* attitude. Edward L. Keenan (personal communication) points out to me that if the first conjunct clause is taken to report a *de re* attitude, he has only the strict reading of the sentence. Note also that if the first conjunct is understood to report a *de se* attitude, the sentence is taken to have only the sloppy reading. What is of great interest to the purposes of the present context is that the crossed understandings of *de re/de se* attitudes are not available for (23a), which in turn suggests that the sentence is not vague but ambiguous.

I have so far argued that the two ambiguity tests support the idea that the *de re/de se* distinction involves a difference in meaning. However, I have said nothing about which aspect of meaning this difference belongs to. No clear case is available which enables us to settle the issue of whether it is a truth-conditional or non-truth-conditional one. Nonetheless, I would like to regard it as a non-truth-conditional one. I take the fact that whether a sentence like (21a) sounds contradictory depends on the context of utterance to mean that ascriptions of *de re* attitudes and their corresponding *de se* attitudes have the same truth-conditions, but differ in non-truth-conditional aspects.

3.2 Two-Tiered Semantics of De Se

Given the non-truth-conditional difference between both attitudes, a question that arises naturally is how to formally account for it. Before presenting the formal mechanism for deriving it, it is worthy of note to discuss a line of thought on which my own account is relied.

Perry (1979) proposes to split our pretheoretical notion of belief “content” into two: a proposition believed and a belief state. Consider (2a), repeated as (24).

(24) John believes that he is in danger.

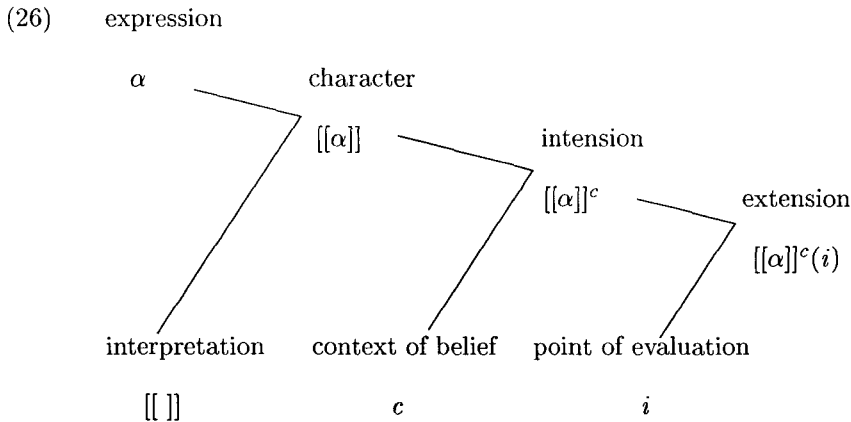
According to him, a proposition believed determines the truth value of a belief ascription like (24) but a belief state is what is going on in a belief-bearer’s head and causally shapes his behavior. Whether (24) is interpreted to ascribe a *de re* or *de se* attitude, his belief object remains the same, but John’s mental state changes. We could say that John believes the proposition $\lambda x[x \text{ is in danger}](j)$ in either ascription, but he believes it in the first person way in the *de se* case and in the third person way in the *de re* case.

An example in point is discussed in Böer and Lycan (1985). Consider (25).

(25) Jones_i believes that he_i (himself) is underpaid, Smith_j believes that he_j (himself) is underpaid, and Brown agrees with Smith that Smith is underpaid.

In one sense Jones and Smith have the same belief—“I am underpaid”—but in another sense Smith and Brown do. Conceptually, functionally, or computationally, the beliefs of Jones and Smith are the same, but truth-conditionally the beliefs of Smith and Brown are the same. Which is the correct sense of same belief depends on whether one is more interested in causes and the explanation of behavior or in semantics, truth, and reliability. For more discussions of the truth-conditional individuation of belief sentences and the computational one, see Böer and Lycan (1985).

Following a suggestion made implicitly in Perry (1979), the *de se* interpretation of a sentence like (24) is obtained in the manner shown by the two-tiered semantic model of interpretation in (26).



The interpretation function $[[\]]$ takes an expression α , and gives back character $[[\alpha]]$. Character then maps contexts of belief, c , which consists of attitude-bearers' mental states, to intensions, $[[\alpha]]^c$, which are 1 (True) or 2 (False) at each points of evaluation i . What is of great importance is the second step from characters to intensions. Like Kaplan's (1979) approach to demonstratives, if no appropriate contexts of belief are provided, I claim, intensions are not arrived at. This means that to be interpreted de se, an utterance of (24) expressively presupposes certain context of beliefs in which John's mental states allow him to be aware that he himself is the object of his belief. Only under such contexts of belief (24) can receive the de se reading. By contrast, the de re reading of (24) does not carry such presuppositions as the corresponding reading requires.

4. De Se and Anaphora

4.1 De Se and Long Distance Anaphora

The presuppositional account of de se ascriptions receives support when we consider the relationship between these phenomena and anaphora. More specifically, it gives a natural explanation of the existence of a strong epistemic relation between the short-distance anaphor *susulo* and its antecedent, which parallels the relation which long-distance anaphors in the propositional attitude context bear with their antecedents.

Before presenting my own explanation of the phenomenon in question, I would first like to show how anaphors and pronouns differ in the belief context with respect to the ascription of de se attitudes. Consider the sentences in (27)-(29).

- (27) a. John-un ku-ka wihem-ey chehay issta-ko mitko issta.
 John-top he-nom danger-to located was-that believes
 'John believes that he is in danger.'
- b. John-un caki-ka wihem-ey chehay issta-ko mitko issta.
 John-top self-nom danger-to located was-that believes
 'John believes that he is in danger.'
- c. John-un wihem-ey chehay issta-ko mitko issta.
 John-top ZerP danger-to located was-that believes
 'John believes that he is in danger.'

- (28) John-un ku-ka wihem-ey chehay issta-ko mitko iss-ciman,
 John-top he-nom danger-to located was-that believes
 wihem-ey cheha-n salam-i cakicasin-ilanun kes-ul moluko issta.
 danger-to located-rel person-nom self-is-that-acc not know
 ‘John believes that he is in danger, but does not know that that man is
 him himself.’
- (29) † John-un caki-ka/Φ wihem-ey chehay issta-ko mitko iss-ciman,
 John-top self-nom/ZerP danger-to located was-that believes
 wihem-ey cheha-n salam-i cakicasin-ilanun kes-ul moluko issta.
 danger-to located-rel person-nom self-is-that-acc not know
 ‘John believes that self/Φ is in danger, but does not know that that man is
 him himself.’

(27a) contains the overt pronoun *ku*, while (27b) and (27c) contain the long-distance anaphor, *caki* and a zero pronoun, written as Φ, respectively. (28) does not sound contradictory, but the sentences in (29) are simply contradictory statements. This fact suggests that unlike overt pronouns, long-distance anaphors and zero pronouns necessarily induce the ascriptions of *de se* attitudes.

4.2 De Se and Short Distance Anaphora

We are now in a position to show that our presuppositional analysis gives a natural explanation of the peculiar aspects of the short-distance anaphor, *susulo*. Before doing so, let me give a brief description of the general properties of this item.

The item *susulo* has two uses. It is used either as an adverbial or as a reflexive pronoun. When it is used as an adverbial, no case-marker is attached to it, as in (30).

- (30) a. John-i susulo ttenassta.
 John-nom for/by oneself left
 ‘John left voluntarily.’
- b. John-i ku il-ul susulo kumantwessta.
 John-nom that job-acc for/by oneself gave up
 ‘John gave up that job voluntarily.’

As with the English emphatic reflexive pronoun, the adverbial use of *susulo* roughly means ‘of one’s own free will’, ‘voluntarily’, or ‘for/by oneself’.

The reflexive use of *susulo* is illustrated by (31).

- (31) John-i susulo-lul thashayssta.
 John-nom self-acc accused
 ‘John volunteered to accuse himself.’

To be used as a reflexive pronoun, it must be marked by the accusative marker *-ul/lul* or the nominative-marker *-ka/i*. As the translation of (31) shows, this reflexive use of *susulo* is correlated with the adverbial use in that both implies the meaning ‘being voluntary’.⁴

As expected from the meaning of the word, the distribution of *susulo* is subject to some restrictions, which is borne out by (32)-(34).

4. An anonymous reviewer correctly points out that an example like (i) is a counterexample to the claim made here.

- (32) Kim_i-un [CP Jim_j-i susulo_{*i/j}-lul thashayssta-ko] malhayssta.
 Kim-top Jim-nom self-acc accused-that said
 ‘Kim said that Jim accused himself.’
 *‘Kim said that Jim accused her (= Kim).’
- (33) Kim-un [CP susulo-ka ku il-ul hayssta-ko] malhayssta.
 Kim-top self-nom that job-acc did-that said
 ‘Kim said that she (= Kim) did that job.’
- (34) Kim_i-un [CP Jim_j-i [CP[NP susulo_{*i/j}-uy calmos-i] khuta-ko]
 Kim-top Jim-nom self-gen fault-nom was big-that
 malhayssta-ko] sayngkakhanta.
 said-that thinks
 ‘Kim thinks that Jim said that his (= Jim’s) fault was big.’
 *‘Kim thinks that Jim said that her (= Kim’s) fault was big.’

In (32) the reflexive pronoun *susulo* is bound by the embedded subject, but it cannot be bound by the matrix subject, conforming to the Specified Subject Condition. (33) shows that *susulo* can itself be used as an embedded subject, violating the Nominative Island Condition. Embedded in a subordinate subject, *susulo* must be bound in the next higher clause, as in (34). What I have so far shown is that *susulo* must be bound in the domain where it is bindable, i.e., in the least domain that contains a referential NP c-commanding it.⁵

Let us now consider (35) and (36).

-
- (i) Chelswu-uy hayngtong-un susulo-lul phamyelsikhinun kes-iessta.
 Chelswu-gen behavior-top self-acc destroy thing-was
 ‘Chelswu’s behavior was the one which destroyed him.’

It is hard to say that the speaker of (i) reports that Chelswu knew that it was himself that he destroyed.

Examples like (i) are problematic in another respect. Consider (ii).

- (i) Chelswu_i-uy hyeng_j-un susulo_{*i/j}-lul pipphanhayessta.
 Chelswu-gen brother-top self-acc criticized
 ‘Chelswui’s brother_j destroyed himself_{*i/j}.’

(ii) shows that *susulo* can be anteceded by the subject NP, but not by its subconstituent. This means that it must be c-commanded. Given the c-command requirement exemplified above, sentences like (i) are somewhat exceptional. It could be reasonably argued that the unusual syntactic behavior of *susulo* in (i) is responsible for the lack of the ‘voluntarily’ meaning of (i).

5. The fact that *susulo* is locally bound might be explained in terms of its intrinsic meaning. Consider (i).

- (i) a. John voluntarily fired Mary.
 b. *John’s scandal voluntarily destroyed Mary.
 c. Mary said that John voluntarily fired Mary.

The contrast between (ia) and (ib) indicates that the adverb *voluntarily* is an agent-oriented one. It requires a co-argument with the Agent role in the same sentence. Note that the meaning of *voluntarily* is related only with the meaning of the embedded verb phrase. There is no possibility whatever that it could be related with the meaning of the matrix verb phrase. This is the reason that *susulo* is locally bound.

- (35) Kim_i-un susulo-uy calmos-i khuta-ko mitko issta.
Kim-top self-gen fault-nom is big-that believes
'Kim believes that his fault is big.'
- (36) † Kim_i-un susulo-uy calmos-i khuta-ko mitko iss-ciman,
Kim-top self-gen fault-nom is big-that believes-but
calmos-i khu-n salam-i cakicasin-ilanun kes-ul moluko issta.
fault-nom was big-rel person-nom was self-that-acc not know
'Kim believes that his fault is big, but he does not know that the person
whose fault is big is him himself.'

The fact that sentence (36) is a contradiction just shows that (35) is understood to report the ascription of a de se attitude.

In this connection, the item *susulo* is contrasted with *(caki)casin*. Consider (37) and (38).

- (37) a. Kim-un (caki)casin-ul piphanhayssta.
Kim-top (self)self-acc criticized
'Kim criticized himself.'
- b. Kim-un Φ caki-in cwu-l moluko (caki)casin-ul piphanhayssta.
Kim-top ZerP self-was that-acc not knowing (self)self-acc criticized
'Kim criticized himself without knowing that he himself was the one criticized
in question.'
- (38) a. Kim-un susulo-lul piphanhayssta.
Kim-top self-acc criticized
'Kim volunteered to criticize himself.'
- b. † Kim-un Φ caki-in cwu-l moluko susulo-lul piphanhayssta.
Kim-top ZerP self-was that-acc not knowing self-acc criticized
'Kim criticized himself without knowing that he himself was the one criticized
in question.'

(37b) perfectly makes sense and thus indicates that the truth-condition of (37a) is simply whether or not John was a self-critic. Whether John knew that it was John himself that he criticized neither affects the truth-conditions of (37a) nor place any felicity condition on it. Interestingly, (38b) seems to be a mere contradiction.⁶ This suggests that in order

6. One reviewer casts a doubt on the grammaticality judgment of (38b), claiming that the sentence sounds acceptable. The judgment is very subtle, and it varies among the native speakers of Korean. Another reviewer reports that the sentence is not natural, but a sentence like (i) sounds quite natural to him.

(i) ?Kim-un Φ caki-in cwu-l moluko susulo-lul piphanha-n kkol-i toyessta.
Kim-top ZerP self-was that-acc not knowing self-acc criticized-rel appearance-nom became
'It happened that Kim criticized himself without knowing that he himself was the one criticized
in question.'

It seems that the presence of the expression, \sim *kkol-i toyta* 'it happens that \sim ' gives rise to the contrast between (i) and (38b). This expression has the potential of reducing the strength of the epistemic relation between linguistic agents to themselves. Thus, the epistemic relation that Kim has with himself by criticizing himself, supposed to be reported by the utterance of (i), is so substantially reduced that the sentence is acceptable.

to know the truth-conditions of (38a), we need to know not only whether John was a self-critic but also whether he knew that it was himself that he criticized. In other words, an utterance of a sentence like (38a) presupposes that the referent of the antecedent of *susulo* has access to himself in terms of the relation expressed by the sentence.

I take the above fact to strongly support the presuppositional account of de se attitudes in that it is naturally extended to explain the phenomenon which might be viewed as another linguistic instance of the key concept of the ascription of de se attitudes, namely, the presence of a strong epistemic relation of linguistic agents to themselves. On the approach taken by Lewis (1979) and Chierchia (1989), such epistemic relations are assumed to exist only in reference to verbs of propositional attitudes. On their approach, there is no way to address the epistemic relation that an utterance of a simple sentence is supposed to exhibit. Given the approach taken here, the choice among anaphors and (zero) pronouns regarding attitude ascriptions is summarized in (39).

- (39) a. *ku*: ascriptions of de re attitudes only
- b. (*caki*)*casin*: ascriptions of de se attitudes when bound in the context of proposition attitude verbs; and ascriptions of de re attitudes when bound in the simple sentence
- c. zero pronoun: ascriptions of de se attitudes only
- d. *susulo*: ascriptions of de se attitudes only

The taxonomy of anaphors and pronouns regarding the type of attitudes thus determines which type of anaphor or pronoun is possible in the context of utterance.

5. Conclusions

As we have seen in the previous sections, the nature of attitudes de se has posed many difficult problems for analytic philosophy as well as for formal semantics. The proposal made in the present study is not intended to solve all these problems, but puts forth some constructive suggestions on the direction of research that practitioners in both fields must take into considerations.

From the philosophical point of view, Lewis (1979) argued the following two theses:

- (40) a. When propositional objects will do, property objects also will do.
- b. Sometimes property objects will do and propositional objects won't.

The asymmetry between propositions and properties in (40) follows from the fact that although there is a one-one mapping from propositions to properties, properties that are supposed to characteristically define attitudes de se cannot be mapped to propositions. Given (40), Lewis proposes that the objects of attitudes be properties. This radical approach, however, contradicts our pretheoretical semantic notion of clauses. Inasmuch as our pragmatic account on which clauses denote propositions proves valid, the theory of attitudes de se can be free from the long-standing philosophical debate on what objects of attitudes are, without necessarily giving up our pretheoretical understanding of clauses and their meanings. Furthermore, the theory can be general enough to dispense with the special requirement, advocated in Chierchia (1989), that clauses of propositional attitude verbs denote properties only when they involve self-ascriptions.

As we have noticed before, there is no compelling reasons to treat clauses involving self-ascriptions specially, and compositionality, in fact, suggests that it is NOT the case.

In addition to compositionality, the linguistic mechanism for expressing self-ascription strongly suggests that its application is not restricted to the contexts of propositional attitude verbs. The choice of anaphors in Korean about self-ascriptions, especially, the short-distance anaphor *susulo*, demands that this relation is generalized to include any anaphor-antecedent relations on which the referent of an antecedent has access to itself. The existence of so-called de se anaphors like *susulo*, which I hope I will be able to convince the reader of, can be accounted for only within the general theory which takes an ascription of de se attitudes to carry an expressive presupposition in the manner specified in the paper.

References

- Böer, Steven E. and William G. Lycan. 1980. Who, me? *The Philosophical Review*, 89:427–466.
- Böer, Steven E. and William G. Lycan. 1985. *Knowing who*. MIT Press, Cambridge, Massachusetts.
- Castañeda, Hector-Neri. 1966. He: a study in the logic self-consciousness. *Ratio*, 8:130–157.
- Chierchia, Gennaro. 1989. Anaphora and attitudes de se. In Renate Bartsch, Johan van Benthem, and Peter van Emde Boas, editors, *Semantics and contextual expression: Groningen-Amsterdam studies in semantics*, volume 11. Foris.
- Cresswell, M. J. 1985. *Structured meaning*. MIT Press, Cambridge, Massachusetts.
- Fodor, Jerry. 1975. *The language of thought*. Harvard University Press, Cambridge, Massachusetts.
- Kaplan, David. 1979. On the logic of demonstratives. In Jr. Peter A. French, Theodore E. Uehling and Howard K. Wettstein, editors, *Contemporary perspectives in the philosophy of language*, pages 401–412, Minneapolis, Minnesota. University of Minnesota Press.
- Lewis, David K. 1979. Attitudes de dicto and de se. *The Philosophical Review*, 88:513–543.
- Perry, John. 1979. The problem of essential indexicals. *Noûs*, 13:3–21.
- Quine, W. V. O. 1956. Quantifiers and propositional attitudes. *Journal of Philosophy*, 53:177–187.
- Reeves, Ruth. 1996. De se and non de se representation. Read at Linguist On-Line Conference: Geometric and Thematic Structure in Binding, October 14–November 4, 1996.
- Reinhart, Tanya. 1983. Coreference and bound anaphora. *Linguistics and Philosophy*, 6:47–88.
- Stalnaker, R. 1981. Indexical belief. *Synthese*, 49:129–151.
- Zwicky, Arnold M. and Jerrold M. Sadock. 1975. Ambiguity tests and how to fail them. In John Kimball, editor, *Syntax and semantics*, volume 4, pages 1–36, New York. Academic Press.

Submitted on: December 6, 1999

Accepted on: January 6, 2000