1. Introduction

Since Castaneda (1966) first brought the phenomenon to light, de se interpretation has represented somewhat of a difficulty. Analyses of the phenomenon exist, of course. However, no analysis given to date has been free of serious problems. In this paper I aim to advance the analysis of Tancredi (1997), overcoming one such problem.

The phenomenon of de se interpretation is readily enough observable, as in the following modified example from Castañeda (1966). We are to consider a war hero named Quintus who is suffering from amnesia and has forgotten who he is. As part of his therapy he is given a book to read that happens to be about his own exploits in the war. He is unaware, however, that the book is about him. Through reading, he forms the opinion that the main character is a hero, and based on this opinion we are licensed in uttering (1a).

(1)  a. Quintus believes that he is a hero.
    b. Quintus does not believe that he is a hero.

When asked directly about whether he himself is a hero, however, Quintus will vehemently deny it, based on which denial we are licensed in uttering (1b). While the two sentences in (1) have the appearance of contradicting one another in terms of their overt form, we can easily understand them as being non-contradictory. The standard explanation for this fact is that the pronouns in the two examples are given distinct interpretations – a non-de se interpretation in (1a) and a de se interpretation in (1b). Indeed, examples like this were adduced precisely for the purpose of establishing the reality of such a distinction in interpretation.

That the explanation given for the non-contradictoriness of (1a,b) is on the right track can be seen by comparing (1) with the minimally different sentences in (2).

(2)  a. Quintus believes that Quintus is a hero.
    b. Quintus does not believe that Quintus is a hero.

Unlike with the pair of sentences in (1), the pair in (2) is felt to be a contradiction. Since the only difference between the two sentences lies in the use of a pronoun in (1) vs. a name in (2), this must be the locus of the
distinction. The obvious conclusion to draw from this distinction is that a name, unlike a pronoun, cannot give rise to a *de se* interpretation.

While the conclusion drawn about names above is correct, some care must be taken in distinguishing two notions. The examples above are all concerned with what counts as a possible interpretation of a sentence. However, there is a related question of when these sentences can be appropriately uttered. Let’s distinguish between what can be called a *de se* thought and a non-*de se* thought. Quintus’s thoughts about his being a hero will count as non-*de se* thoughts during his amnesia when they are based solely on the story he has read, though once he is cured of his amnesia we can imagine these thoughts becoming *de se* thoughts, where he identifies himself as the main character of that story. We can then ask which of the sentences in (1) and (2) can be used appropriately when he has *de se* thoughts and which when he has non-*de se* thoughts. We have already examined the latter case and found (1a,b) and (2a) to all be appropriate. However, once Quintus is cured of his amnesia and all his thoughts about his being a hero are, let’s assume, *de se*, it is still appropriate to utter not only (1a) but also (2a). It is entirely unsurprising, of course, that (1a) would be appropriate in these circumstances since (1a) is presumed to encode *de se* interpretation directly and hence exactly matches Quintus’s newfound *de se* thoughts. The more surprising case is (2a), since this sentence was seen not to be able to encode *de se* interpretation directly. We can see what’s happening here, though, once we realize that the *de se* interpretation of (1a) entails the (non-*de se*) interpretation of (2a). This means that whenever the former is true the latter will be true as well. If we are interested in whether the thought in question is *de se* or not, (1a) will obviously be the better sentence to choose after Quintus’s recovery. If we are not interested in that question, though, (2a) can function equally well, and in any case will always be true when (1a) is (under the appropriate interpretation).

The upshot of the above discussion is that a sentence can be compatible with a *de se* thought without having a *de se* interpretation. This is an important conclusion to keep in mind when thinking about *de se* interpretation. The contrast between (1b) which is true of the amnesic Quintus and (2b) which is false of him is what shows that there is a real distinction in the potential for *de se* interpretation of a name and a pronoun. Focusing on the positive examples in (1a) and (2a) alone, however, could misleadingly suggest that the distinction is not one of semantic interpretation but rather simply one of pragmatic preference.
2. **Past Analyses of De Se interpretation**

I turn now to semantics analyses of *de se* interpretation. These largely fall into two categories. On the one hand are analyses based on Kaplan’s (1977/89) notion of a context parameter, including most prominently recent analyses by Schlenker (2003) and Anand (2006). These analyses equate *de se* interpretation with reference via an Agent role. To distinguish *de se* interpretation, which is possible for 1st, 2nd or 3rd person pronouns, from first person reference, which is restricted to 1st person pronouns, an extra index parameter is introduced, following Lewis (1970), in addition to the context parameter from Kaplan. First person reference is then analyzed as reference via the Agent of the Context, while *de se* interpretation is analyzed as reference via the Agent role of an Index. These analyses contrast with those of Tancredi (1997) and Higginbotham (2003) which seek to derive *de se* interpretation by interpreting pronouns as dependent on thematic roles that attribute thoughts to a subject. The basic idea underlying these analyses is that a propositional attitude predicate assigns a thematic role to its subject, and a *de se* pronoun refers to that subject via that role.

As Higginbotham has argued persuasively, the Kaplan-based analyses fail in that they do not capture the essence of *de se* interpretation. As Shoemaker (1968) and later Evans (1982) have shown, one of the hallmarks of *de se* interpretation is that it exhibits the property of immunity to error through misidentification. This property follows directly from the analyses of Tancredi (1997) and Higginbotham (2003), but needs to be stipulated under the analyses of Schlenker (2003) and Anand (2006), as argued by Higginbotham (2003). I take this to be sufficient grounds for rejecting this approach to *de se* interpretation.

Though the analyses of Tancredi (1997) and Higginbotham (2003) fare better with respect to immunity to error through misidentification, they suffer instead from an overgeneration of *de dicto* readings, a problem brought to light by Oshima (2001). The generation of *de se* interpretation requires a *de se* pronoun to be in the scope of the attitude predicate assigning a thematic role to the pronoun’s antecedent. Since *de se* interpretation comes from a complex interpretation of a pronoun, however, this complex interpretation is expected to be available for generating non-*de se, de dicto* interpretations as well, interpretations which manifestly fail to be observed. Though I will make no attempt at saving the Kaplan-based analyses, I will propose an
extension of the analyses in Tancredi (1997) and Higginbotham (2003) that overcomes the *de dicto* problem by refining the analysis of *de dicto* interpretation and of *de se* pronouns.

3. Revising the Analysis

Tancredi (1997) analyzes *de se* pronouns as having an interpretation that is distinct from that of non-*de se* pronouns. In particular, a *de se* pronoun is analyzed as referring to the individual having a token thought \( t \), where formally \( t \) is taken to be a variable that must be bound. The relevant binder for \( t \) is taken to be introduced by a propositional attitude predicate. Applying the analysis to (1a) under the *de se* analysis of the pronoun gives the following:

(3) Quintus believes that he is a hero.

Quintus has a thought \( T \), characterized as a belief, the content of which is that the person having \( T (= t) \) is a hero.

Here the occurrence of the variable \( t \) inside the pronoun *he* is bound by (the existentially quantified thought) \( T \). Since identifying a thought as one’s own is not a process of identification of two things conceptualized as distinct, there is no room for misidentification. And though the referent of *he* does become identified with the referent of *Quintus*, the identification proceeds exclusively through links that are themselves immune to error through misidentification. The pronoun refers to the individual having thought \( T \). There is no identification in this reference, and hence no room for misidentification. \( T \) is identified through variable binding with the belief attributed to Quintus, again not a process that involves a speaker taking two conceptually independent objects to be one and the same and hence not a process open to misidentification. The belief is attributed to Quintus through the thematic role of the verb *believe*. Again, this is simply the operation of non-conscious syntactic mechanisms which involve no conscious process of identification and hence no room for misidentification. This last step, however, has as a result that the reference of the pronoun *he* is Quintus.

The above analysis is essentially duplicated in Higginbotham (2003), though the clothing differs. Higginbotham analyzes the pronoun as referring via an event role rather than explicitly via a thought, though otherwise the analysis is identical to that of Tancredi (1997). The only event roles that
actually give rise to de se interpretation are those that attribute thoughts. This can be seen by considering the contrast of (1a,b) with the following:

(4)  a. John has an immune system that protects him from disease.
    b. John doesn’t have an immune system that protects him from disease.

In (a) (from David Lewis, p.c.), him cannot be given a de se interpretation. This shows up in the fact that (a) and (b) are obligatorily contradictory, a situation that we already saw not to arise in sentences containing a pronoun that can potentially be interpreted de se, such as those in (1). Though this does not refute Higginbotham’s analysis, it does require addition of a restriction on which specific thematic roles do and which do not give rise to de se interpretation, an addition which will essentially make the analysis indistinguishable from the analysis of Tancredi. I will thus restrict myself to using Tancredi’s analysis for illustration below, though the conclusion will be equally applicable to Higginbotham’s analysis as well.

The analysis under consideration essentially treats de se interpretation as a subclass of de dicto interpretation: the attitude holder – in (1) Quintus – is taken to identify the referent of the pronoun he as the individual having the very token thought that he is entertaining. While this works well for generating the de se interpretations that we want, it introduces the possibility of such pronouns being interpreted de dicto with respect to some other individual in the sentence. That this is problematic can be seen by considering the following examples.

(5)  Quintus believes that the doctor thinks he is a hero.

The problem this example poses is that while it does give rise to a potential de se interpretation with respect to Quintus, it fails to give rise to a de dicto interpretation with respect to the doctor. In particular, the pronoun cannot be interpreted as referring de dicto to the person that the doctor takes to have thought t. If it did, the sentence would be false if Quintus doesn’t think the doctor has access to Quintus’s token thoughts, regardless of whether Quintus believes that the doctor thinks that Quintus is a hero. This can be made perspicuous by using the following pseudo-analysis of the sentence:

(6)  Quintus has belief B: [the doctor has thought T: [the person with belief B is a hero]]
By giving the pronoun lowest scope, it should under a scope analysis of attitudes be interpreted \textit{de dicto} with respect to the doctor. Furthermore, under such an interpretation there is nothing formally wrong with the sentence – all variables that need to be bound are bound, and the interpretation that results is perfectly well-formed. The absence of such an interpretation thus points to a serious problem in the analyses that gave rise to it.

I take the problem illustrated above to arise from an insufficient analysis of \textit{de dicto} interpretation and a corresponding shortcoming in the interpretation proposed for a \textit{de se} pronoun. It has long been noted that \textit{de dicto} interpretation of an expression can be with respect to the subject of any attitude predicate that embeds the expression. What is less commonly noted is that \textit{de dicto} interpretation is at least potentially with respect to a unique expression. To see this, consider the following scenario. I the speaker do not believe in witches, and neither does John. John, however, thinks that Mary does. I utter the following:

(7) John thinks Mary believes the wicked witch of Tokyo can fly.

This sentence can well be true. In particular, if according to John Mary has the belief that she might put into words as “The wicked witch of Tokyo can fly,” the sentence is true. Under such an interpretation, \textit{the wicked witch of Tokyo} is interpreted \textit{de dicto} with respect to Mary. Note, however, that by being in the scope of \textit{believes} this definite description also ends up in the scope of \textit{thinks}. If being in the scope of an attitude predicate were sufficient for being interpreted \textit{de dicto} with respect to the subject of that predicate and if multiple \textit{de dicto} interpretation were possible, it would follow that the sentence would give rise to a presupposition violation. This is so since according to John there are no witches, and so the expression \textit{the wicked witch of Tokyo} fails to denote any individual for John. In order for the sentence to be interpreted as true, \textit{the wicked witch of Tokyo} has to be interpreted \textit{de dicto} with respect to Mary \textit{and not with respect to John}. Since we can easily take the sentence to be true in the situation described, it follows that what I will call \textit{targeted de dicto} interpretation is possible – \textit{de dicto} with respect to exactly one individual among a larger number of legitimate potential options. Given that targeted \textit{de dicto} interpretation is possible, the default linguistic assumption should be that that is the only option made available by the grammar.
Though I will adopt the assumption that all *de dicto* interpretation is targeted, I should note that it is extremely difficult to actually establish this assumption with an argument. The best argument I can give is to show that this assumption overcomes the problems facing the analysis of *de se* pronouns in Tancredi (1997) and is hence empirically justified. However, the force of that argument rests on the prior assumption that the approach of Tancredi (1997) is itself correct in broad outline. If a Kaplan-based analysis is ultimately shown to be preferable this argument will be correspondingly diminished. Since my goal in this paper is only to show that it is possible to overcome the challenge posed by the absence of non-*de se* interpretations for *de se* pronouns, I will satisfy myself with this argument, together with its limitations.

So, what is *de dicto* interpretation? Intuitively *de dicto* interpretation is interpretation of a referring expression from the perspective of an individual to whom some attitude is being attributed. To say that the wicked witch of Tokyo is interpreted *de dicto* with respect to Mary is then to say that it is interpreted from her perspective within a sentence that attributes some belief, doubt, knowledge, etc. to Mary. If we are to capture this intuition in our formal analysis, we need some way of introducing perspectives. The easiest way to do this is through the attitude verbs themselves. To say that an individual x has an attitude R toward some content c is to say that c is the content of R for x, i.e. with respect to x’s interpretational apparatus. Building on the core insights of Tancredi (200?), I will identify this apparatus with a set of models, Mx. This implicit reference to the attitude holder’s interpretational apparatus is what makes *de dicto* interpretation possible. *The wicked witch of Tokyo* will not be given a referent by either the speaker’s (i.e. my) or John’s referential apparatus, but it will be given a referent by Mary’s interpretational apparatus, and this is all that is required for the sentence to be interpretable. I formalize the analysis of *believe* in (8) below.

(8) \[
[\text{he}_\text{de se}]^M = \Lambda S. \lambda x. \exists T \, (\text{Belief}(T,x) \land \text{Content}(t,[S]^M_x))
\]

I use the capital lambda (\(\Lambda\)) for Syntactic Function Application, applying to a syntactic structure rather than the interpretation thereof. Thus, *believe* takes a sentence S and an individual x and returns truth iff x has a belief T whose content is given by the interpretation of S with respect to x’s interpretational apparatus.
With the above analysis of belief in place, we can now address the formal analysis of *de dicto* interpretation. The analysis at this point is straightforward: an expression $e$ is *de dicto* with respect to an individual $x$ iff $e$ is interpreted with respect to $\text{M}_x$. This makes it possible now to give a refined analysis of a *de se* interpreted pronoun. As before I take *de se* analysis to involve reference via a token thought $t$ introduced as a free variable in the interpretation. I add here, however, a restriction that ensures that $\text{he}_{de\;se}$ will be interpreted *de dicto* with respect to the individual it refers to.

(9) \[ [\text{he}_{de\;se}]^M = \lambda x: \text{x has t and } \exists S \text{ Content } (t, [S]^M) \]

This interpretation requires that the same interpretational apparatus used to interpret the pronoun also be used to give the content of the thought that secures the reference for the pronoun.

It is straightforward to see how this combination of a *de dicto* analysis for attitude predicates and direct incorporation of a *de dicto* restriction into *de se* pronouns avoids Oshima’s problem. Like all expressions, the pronoun *he* in (9), repeated below, can only be interpreted *de dicto* with respect to one individual.

(10) Quintus believes that the doctor thinks he’s a hero.

Furthermore, given the semantic interpretation of $\text{he}_{de\;se}$ in (9) above, the pronoun has to refer to the individual with respect to whom it is interpreted as *de dicto*. This makes it impossible for the pronoun to refer to Quintus and be interpreted *de dicto* with respect to the doctor, as desired.

The analysis just given works best with a scopal treatment of *de dicto* interpretation, one where the individual with respect to which a given expression is interpreted *de dicto* is determined by the syntactic position of that expression. Getting the relevant *de se* interpretation of *he* referring to Quintus in (10) under these assumptions requires syntactically raising the pronoun out of the complement of *think* and into the complement of *believe*. This requirement follows from the particular technical implementation of *de dicto* interpretation that was adopted. Other options are equally available that would keep all the benefits of the proposed analysis without relying on syntactic movement for restricting *de dicto* interpretation, though I leave the
details of such an analysis to the interested reader.

BIBLIOGRAPHY


