



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Zhou, Lian

Title:
De Se Communication: language, thought and co-aboutness

Date:
2019

Persistent Link:
<https://hdl.handle.net/11343/241574>

Terms and Conditions:
Terms and Conditions: Copyright in works deposited in Minerva Access is retained by the copyright owner. The work may not be altered without permission from the copyright owner. Readers may only download, print and save electronic copies of whole works for their own personal non-commercial use. Any use that exceeds these limits requires permission from the copyright owner. Attribution is essential when quoting or paraphrasing from these works.

De Se Communication:
Language, Thought and Co-aboutness

by

Lian Zhou

A dissertation submitted in total fulfillment of the requirements
of the degree of Doctor of Philosophy

ORCID: 0000-0001-7710-8069

School of Historical and Philosophical Studies

University of Melbourne

July 2020

Abstract

This dissertation is about the co-aboutness problem of de se communication. An essential requirement of successful communication is that participants of communication must talk about the same subject matter. I call this requirement the co-aboutness condition of communication. According to the traditional picture of communication, what is communicated is transmitted between participants of that communication in accordance with the replication transmission model. In that case, the co-aboutness condition is satisfied by same-saying, which is the replication of one other's linguistic expressions. The basic idea behind the same-saying strategy is that we talk about the same subject matter if we use the same words. However, for an idiosyncratic type of communication, which is de se communication, same-saying simply does not work. De se communication is the communication of de se thought. In de se communication, we will definitely talk past each other if we just replicate one another's words. For example, if I say, 'I am happy', and you reply, 'I am happy', we are not talking about the same subject matter because I am talking about my feeling and you are talking about your feeling. So, how can we satisfy the co-aboutness condition in de se communication? This question is the central question of this dissertation. I argue that in de se communication, what is communicated is transmitted in accordance with a new model, the translation transmission model. In order to satisfy the co-aboutness condition of communications which conform to the translation transmission model, a new strategy must be developed. I call the new strategy 'translation'. Participants of a de se communication talk about the same subject matter by means of translating one another's words. The most important part of this translation strategy is the translation of first person terms. The singular first person indexical 'I' is representative of first person terms. So, in this dissertation I will focus on the translation of 'I'. For the purpose of fulfilling the co-aboutness condition in de se communication, which term is a good translation of 'I'? In order to answer this question, I introduce Worsnip's analogy between the intrapersonal incoherence and interpersonal disagreement (Worsnip, 2019, pp. 252-259). Based on

this analogy, I develop a further analogy between the intrapersonal de se inference and the interpersonal de se communication. This analogy gives us an important clue for looking for a good translation of 'I': if 'X' is a good translation of 'I', then the co-referential relation between 'X' and 'I' must closely imitate the co-referential relation between two tokens of 'I'. So, the term whose co-referential relation with 'I' best imitates the co-referential relation between two tokens of 'I' is the translation of 'I' for which I am looking. This term, I argue, is 'you'. The analogy between the 'I-I' same-saying and 'I-You' translation is sustained by three similarities. The first similarity is that the 'I-I' co-reference in de se inference and the 'I-You' co-reference in de se communication are both explained by explanations of the externalist approach, and their externalist explanations are analogous to one another. This similarity is the most important one of the three similarities. Why is there 'I-You' co-reference in de se communication? How does my 'I' co-refer with your 'you'? I argue that the co-reference of 'I-You' translation is established by the match between the mutual recognition relation and the 'I-You' reciprocity in de se communication. The second similarity is that the 'I-I' co-reference is transparent to the producer of de se inference, just as the 'I-You' co-reference is transparent to the participants of de se communication. The third similarity is that both the 'I-I' co-reference and the 'I-You' co-reference are immune to the effect of misrepresentation. I articulate those three similarities. My articulation justifies the legitimacy of the analogy between 'I-I' same-saying and 'I-You' translation. In this articulation, I examine the 'I-You' translation by three criteria for determining the proper connection between translation and co-aboutness. After this examination, I conclude that the 'I-You' translation satisfies all three criteria. At the end of the dissertation, I provide my answer to the central question of this dissertation: when participating in a de se communication, we fulfill the co-aboutness requirement by using 'I-You' translation.

Declaration

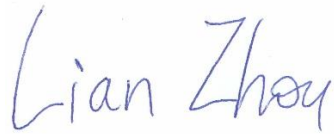
I declare that

(I) the dissertation comprises only my original work towards the degree of Doctor of Philosophy;

(II) due acknowledgement has been made in the text to all other material used
and

(III) the dissertation is fewer than the maximum word limit in length, exclusive of tables, figures, bibliographies and appendices.

Signature:



Name in Full: Lian Zhou

Date: 22/07/2020

Acknowledgments

First of all, I would like to express my most sincere appreciation to my principal supervisor, Prof. Greg Restall. He gave me the opportunity to pursue my PhD degree at University of Melbourne under his direction. I sincerely appreciate him for his continuous support, guidance and encouragement during my entire candidature period. Without his supervision, I would not have been able to write this dissertation. Similarly, I would like to express my deep gratitude to my co-supervisor Dr. Laura Schroeter. She provided much valuable advice for my research, and her profound knowledge of philosophy is always a source of inspiration for me.

Also, I would like to thank Dr. Shawn Standefer and Dr. Holly Lawford-Smith for their generous help and encouragement when I experienced difficult times during my candidature period. I thank my fellow colleagues for all the stimulating discussions and joyful chats we had together.

Last but not least, I would like to thank my parents. Their support is priceless for me.

Table of Contents

Abstract	i
Declaration	iii
Acknowledgments.....	iv
Table of Contents	v
Introduction	1
A. De Se Thought and De Se Communication	3
B. The Problem of Co-aboutness.....	10
C. Content of Each Chapter	15
Chapter I The Peculiarity of De Se Communication	20
1.1 Three Proposals on the Real Problem of De Se Attitude.....	22
1.1.1 Dilip Ninan on the Conflict between Agreement and Explanation	23
1.1.2 Matheus Valente on the Incompatibility between Principles	29
1.1.3 Recanati on De Se Communication	31
1.2 De Se Communication vs Other Types of Communication.....	37
1.2.1 The Idiosyncrasy of De Se Communication	39
1.2.2 Presupposition and Co-aboutness	40
1.2.3 Replication Transmission Model	42
1.2.4 Translation Transmission Model.....	47
1.3 Characterizations of Translation Transmission Model.....	51
1.3.1 Frege on Private Sense	52
1.3.2 Perry on Structured Proposition	56
1.3.3 Recanati on Systematic Transformation.....	61
1.3.4 Preliminary Characterization of Translation Transmission Model.....	67
Chapter II De Se Thought and Identification.....	71
2.1 Essential Indexical	74
2.2 Two Criticisms of Essential Indexical	77
2.2.1 Generic Opacity instead of Essential Indexicality.....	78

2.2.2 De Dicto/De Re and De Se.....	84
2.3 Defenses of the Essential Indexical	96
2.3.1 Matthew Babb on Indexical Intention.....	97
2.3.2 Juliana Lima on Principles of Action Explanation.....	99
2.4 De Se Thought and IEM.....	104
2.4.1 Immunity to Error through Misidentification.....	105
2.4.2 IEM and First Person Thought	108
2.4.3 IEM and De Se Thought.....	116
2.4.4 The Defense of the Definition Account.....	121
2.4.5 A Possible Objection.....	133
Chapter III Interpersonal Co-aboutness in De Se Communication.....	139
3.1 Co-aboutness and ‘I-You’ Translation	142
3.1.1 Criteria of Co-aboutness	143
3.1.2 Analogy between Intrapersonal and Interpersonal	144
3.1.3 ‘I-You’ Translation	148
3.2 ‘I-You’ Translation and ‘I-I’ Same-saying.....	151
3.2.1 Co-referentiality of Tokens of ‘I’	152
3.2.2 Externalist Explanation of ‘I-I’ Co-reference.....	156
3.2.3 The ‘External/Internal’ Distinction.....	164
3.2.4 Internalist Explanation of ‘I-You’ Co-reference.....	166
3.2.5 The Defect of the Internalist Explanation	168
3.2.6 Externalist Explanation of ‘I-You’ Co-reference.....	170
3.2.7 Four Possible Objections.....	179
3.2.8 Transparency and Immunity.....	187
3.3 Summary	194
Chapter IV Conclusion	197
Bibliography	205

Introduction

The relation between language and thought is a philosophically deep question that has been discussed by philosophers for many generations. Usually, philosophers discuss the relation between thought and language in general.¹ but interestingly, there are still some specific types of thoughts which relate to corresponding linguistic expressions which attract philosophers' particular attention. In this dissertation, I focus on one of them.

The topic of my dissertation is the relation between the de se thought (roughly, thought about myself as myself²) and first person linguistic expressions, of which a possible representative is the first person pronoun 'I'. There have already been many studies on both de se thought.³ and the first person pronoun.⁴ but in the current

¹ For a comprehensive review of this issue, please refer to the first chapter of Christian Barth's *Objectivity and the Language-Dependence of Thought* (New York: Routledge, 2011)

² Some philosophers use this definition as an informal definition of de se thought. For example, García-Carpintero says: 'Now, in the 1960s and 1970s Castañeda, Perry and Lewis argued that thoughts about oneself 'as oneself' — de se thoughts — raise special issues.' (García-Carpintero, 2013, p. 191) It should also be noted that here I choose the phrase 'de se thought', rather than the phrase 'first person thought', to designate the official object of my discussion. Those two phrases are interchangeable in a wide range of cases, but not in all cases. This is because there is a subtle, but non-trivial, difference between 'de se thought' and 'first person thought' within a highly special type of communication context. I will take this difference into consideration and specify what it is in Chapter 2. However, in other parts of the dissertation, I take 'de se thought' and 'first person thought' (also 'de se communication' and 'first person communication') to be interchangeable, unless otherwise specified.

³ For example, *Attitudes De Se: Linguistics, Epistemology, Metaphysics*, edited by Neil Feit and Alessandro Capone, CSLI Publications, 2013.

⁴ There are at least two monographs devoted to the studies of semantics of I, see Maximilian de Gaynesford, *I: the meaning of the first person term* (Oxford: Clarendon Press, 2006) and Ingmar Brinck,

literature, the two issues are often discussed separately, without taking the study of one of them to be a default background for the study of the other. This does not imply that the connection between the two issues is external and superficial. On the contrary, the connection is so immediate and intrinsic that sometimes *de se* thought is just called 'I'-thought⁵. Some philosophers even believe that the most accessible way to define *de se* thought is to define it by means of its linguistic expression: for them, a *de se* thought is a thought in which linguistic expression involves the use of a first person pronoun.⁶ (I do not agree with them on this definition of *de se* thought). As this chapter is just an introduction of the dissertation, I do not propose to elaborate here how strong the connection between the *de se* thought and the first person linguistic expression is, because that can be manifested only at the end of the entire dissertation: the significance of the connection will be unfolded gradually through each chapter. However, there are three issues that need to be addressed in the introduction. The first issue is to introduce the model that I adopt to articulate the relation between *de se* thought and first person linguistic expression (and also the reason why I adopt it).

The Indexical 'I': the First Person in Thought and Language (Kluwer Academic Publishers, 1997)

⁵ De Gaynesford complains about philosophers' use of the phrase 'I-thought', because this phrase 'is potentially misleading: it smudges the distinction between the thoughts expressed by *I* and their expression using *I*.' (De Gaynesford, 2006, p. 3) I also think that the 'I-thought' is not an accurate label, because it blurs my distinction between the *de se* thought and non-*de-se* first person thought (please refer to Section 2.4.3 for more information about my distinction), although I think in most cases, 'I-thought' means the same as '*de se* thought'.

⁶ Manuel García-Carpintero claims that

Similarly, let us take *de se* attitudes to be those typically expressed by means of sentences including first-personal pronouns, such as (here paradigm examples should be mentioned, like those in Perry's Lingers', Heimson's, Messy Shopper's, and Lewis's Two Gods thought-experiments), so that philosophers including Perry and Lewis have argued that traditional theories of content could not properly characterize them. (Manuel García-Carpintero, 2017, p. 255)

Please also refer to Dilip Ninan, 2016, p. 88 and Komarine Romdenh-Romluc, 2008, p.146.

The second issue is to point out the angle from which we can properly enter the relevant articulation. The third issue is to clarify the core problem of the dissertation, which serves as a thread combining all the chapters of this dissertation together as a whole. I will address the first and second issues in the first section of the introduction. Following this, I will precede to introduce a clear articulation of the core problem of this dissertation, which is the problem of how to satisfy the co-aboutness condition in the case of de se communication. The satisfaction of the co-aboutness condition in de se communication is puzzling, because the traditional way of satisfying the co-aboutness condition in communication in general (namely same-saying), does not work in de se communication. At the end of this chapter, I will provide an outline of the main argument of each chapter.

A. De Se Thought and De Se Communication

De se thought can be studied from several different points of view. According to philosophers who believe in the *essential indexicality thesis*, for example John Perry⁷, the characteristic feature of de se thought is that it is a challenge to the traditional doctrine of propositional content. The traditional doctrine does not distinguish between content and a thinker's way of accessing that content. But Perry argues that the distinction between the content and one's way of accessing that content is crucial for a correct understanding of the difference between a de se thought and a non-de-se thought that share the same truth condition⁸. However, according to the interpretation of the 'essential indexicality thesis', which is the most popular one currently, the thesis instead posits an essential connection between the de se thought

⁷ See John Perry, 'The Problem of the Essential Indexical', *Noûs* 13, pp. 3-21, 1979.

⁸ For example, Smith's thoughts 'I am happy' and 'Smith is happy' share the same truth condition (both are true if, and only if, Smith is actually happy), but the former is a de se thought, and the latter is a non-de-se thought.

and the motivation of our intention driven behavior. Human beings must formulate their intentions for motivating certain behaviors in the form of de se thought, so that they can successfully initiate the performance of those behaviors⁹. Therefore, from this point of view, to study de se thought, first and foremost, is to study the relation between de se thought and rational agency. There is also another tradition in the study of de se thought, proposed by David Lewis, which focuses on the study of the content of de se belief. Before the publication of David Lewis' 1979 essay¹⁰, philosophers generally shared the consensus that the content of a propositional attitude is a proposition with a definite truth value. Under Lewis' possible world semantic framework, this is equivalent to saying that the content of a propositional attitude is a set of possible worlds. Philosophers at that time did not regard de se belief as an exception to that standard account of propositional content. However, Lewis argues that de se belief is indeed an exception. It is highly implausible, according to him, to assume that the content of de se thought can be something with a definite truth value, since the truth value of a same de se belief varies from one case to another, depending on which individual is believing that belief. Therefore, Lewis suggests that the content of de se belief must be something to which we can ascribe a flexible truth value. The option that he provides is to regard the content of de se belief as a type of property, rather than a proposition. Under his possible world semantic framework, a property is a set of possible objects. On this account, a de se belief can get a truth value only if the content of that belief, i.e., a property, can be ascribed to an individual, which, in the case of de se belief, is always the owner of that belief. This is what distinguishes de se belief from all the other kinds of belief. So, for philosophers who work in the Lewisian tradition, to study de se thought is to study the peculiarity of the content of de se belief, i.e., the content of a de se attitude is a property (or a set of centred worlds,

⁹ See the third chapter 'Indexicality, the De Se, and Agency' of Herman Cappelen and Josh Dever's book *The Inessential Indexical: On the Philosophical Insignificance of Perspective and the First Person* (Oxford University Press, 2013).

¹⁰ See David Lewis, 'Attitudes de dicto and de se', in *Philosophical Review* 88 (4):513-543 (1979).

in Lewisian terms)¹¹.

However, I will not adopt either Perry's or Lewis' approach. I have two reasons for this decision. The first reason is the widely known challenge made by Robert Stalnaker in his 1981 paper 'Indexical Belief'¹². I call his challenge a communication challenge. Stalnaker's main argument is that neither Perry's distinction between belief content and the way of accessing belief content, nor Lewis' centred world theory of belief content can account for the interpersonal information transferred in de se communication. To illustrate Stalnaker's argument, let us imagine a scenario like this:

Suppose I just finished my shopping at a supermarket, and I am waiting in a queue to pay for my purchase. A member of the supermarket staff comes to me and tell me, 'This is not a queue for payment. You picked the wrong queue.' I then realize that I am standing in the wrong queue and need to look for a correct one to finish my purchase.

In this scenario, there is a communication between the supermarket staff and me, which leads me to acquire a new de se thought ('I am in the wrong queue.') The acquisition of this de se thought is the result of my reception of information conveyed by the member of staff's utterance 'You picked the wrong queue'. However, the proposition expressed by that utterance is not a de se proposition. This gives rise to the question of how the reception of a non-de-se proposition can result in an acquisition of a de se thought. Stalnaker insightfully points out that neither Perry's proposal nor Lewis' theory can answer this question properly. Why?

Now let us first examine a possible answer to this question according to Perry's proposal. Perry distinguishes the belief content (which in this scenario is a traditional

¹¹ For a more detailed discussion of Perry's and Lewis's respective proposals, please refer to the section 2.2 of the Chapter 2.

¹² See Robert Stalnaker, *Indexical Belief*, in R. Stalnaker (ed.), *Context and Content*, pp. 130-149, Oxford: Oxford University Press, 1999.

de re proposition with definite truth condition), from the belief state (i.e. the way of accessing belief content). On Perry's account, only the belief content gets transmitted in a communication. The supermarket staff's belief state is a second personal belief state ('you picked the wrong queue. '), and her belief content is a traditional de re proposition ('That guy picked the wrong queue. '), which is equivalent to saying that the information transmitted in this communication, according to Perry's proposal, is the de re proposition. However, based on the transfer of a de re proposition in the communication, I acquire a de se belief ('I am standing in the wrong queue '). A de se belief is not a de re belief. So why does the transfer of a de re content generate a de se way of accessing that de re content, i.e. a de se belief state.? Perry's proposal provides no satisfactory answer to this question.

It seems that David Lewis' theory that the content of de se belief is a property can avoid the question which Perry's proposal has to face, because Lewis' theory does not rely on the distinction between belief state and belief content. But this does not mean that Lewis' theory is not troubled by the communication challenge. To reveal why the theory of content as property fails to explain the information transferred in a de se communication, we just need to compare my belief content and the staff member's belief content, when both are formulated according to Lewis' theory. On the one hand, the staff member's belief in the scenario is 'You picked the wrong queue', and the content of her belief is a property of *talking to a person who picked the wrong queue*. On the other hand, my belief, which I acquire as a result of the communication, would be expressed as 'I am standing in the wrong queue', and the content of my belief is a property of *standing in the wrong queue*. The content of my belief is not the content of the staff member's belief (while in Perry's case, the content of my belief and the content of the staff's belief are identical). Here we encounter a problem which is a bit similar to the problem we encountered in Perry's case: how could it be possible that my recognition of the speaker's content *talking to a person who picked the wrong queue* brings out the production of a different content *standing in the wrong queue* in the belief I form? Lewis' theory does not provide a satisfactory answer to this

question¹³. Now I have tested both Perry's and Lewis' theories of de se attitudes (and contents) against Stalnaker's communication challenge and found that neither of them passes the test. This explains in part why I do not adopt either Perry's or Lewis' approach.

My second reason for rejecting both Perry's and Lewis' approaches is that neither approach adequately takes into account the philosophical significance of first person linguistic expression. In other words, the connection between de se thought and first person linguistic expression is neglected in both approaches. In my opinion, this is a serious defect. This reason can be quite cogent, especially when considered against the background of Stalnaker's communication challenge. People may think that simply entertaining a de se thought in one's mind does not necessarily involve the relevant linguistic expressions that would be used to express that thought. However, the interpersonal (verbal) communication of de se thought put the connection between thought and language under the spotlight because de se thought is communicated by means of first person pronouns¹⁴. If we examine Lewis' centred world semantics from

¹³ The claim that I believe Lewis' account of de se thought fails to provide a satisfactory response to Stalnaker's communication challenge does not mean that I do not find any insight concerning de se thought in Lewis' account. On the contrary, I do find a valuable insight concerning de se thought in Lewis' property account of de se content. I will address this insight in Section 2.2.2.

¹⁴ There is even a stronger claim: the connection between de se thought and first person linguistic expression can only be properly revealed when de se thought and first person language are both embedded into the context of communication. This stronger claim, according to Evans, is supported by both Geach and Strawson. Evans says,

It is worth clearing away, at this early stage, a curious idea, expressed in different ways by Geach and by Strawson, to the effect that the interest of 'I' is exclusively the interest of a communicative device—that is, in effect, that there are no Ideas corresponding to the pronoun and available to be exercised in thinking. According to Geach, Descartes in his solitary meditations had no need of 'I' in such judgements as 'I'm getting into an awful muddle'; he could have judged 'This is really a dreadful muddle!' Similarly, Strawson suggests that it is right to speak of self-ascription, for instance of pain, only because one tells others that one is in pain; otherwise one's judgement can simply be 'There is a pain' Both these philosophers are preoccupied with the fact that there is no need for me to tell myself who it is who is

this angle, an interesting contrast can be discovered: Lewis tries to provide a semantic explanation of the content of de se attitudes as property (which can be modeled as a set of possible objects or a set of centred worlds) but it seems that he has no concern in supplying a semantic explanation for the de se indexical pronoun 'I'. This neglect of first person language and the further neglect of the connection between first person language and de se thought, I believe, are the reasons why both Perry's and Lewis' proposals fail to respond to Stalnaker's challenge.

In order to study the relation between de se thought and first person pronouns, I need a new perspective that highlights the importance of the first person linguistic expression in the study of de se thought, and vice versa. Inspired by Stalnaker, I believe that there is indeed a perspective that can fulfill this requirement. This is the perspective of (verbal) communication. There are various types of communication in our daily life. We communicate with each other by means of speech, gesture, facial expression, or even winking. However, I think there could be no doubt that the most informative and complex communication type is the verbal communication, and also the most important aspect of the verbal communication is that a content, which is encoded in the semantic content (and in some cases, pragmatics,) of relevant language expression, is transferred interpersonally. In this sense, verbal communication intrinsically binds thought and language together, and it provides a good perspective for the study of the relation between language and thought. As my goal in the dissertation is to study a specific connection between the first person linguistic expression and the de se thought, rather than the relation between language and thought in general, I will mostly focus on a special kind of communication, which is the communication of de se thoughts by means of first person linguistic expressions. This kind of communication is called de se communication.

getting into a muddle, or who is in pain. (Evans, 1982, p. 208)

This is a very interesting claim, although in this dissertation, I do not need to accept it. In fact, I do not think it is true.

Since I have decided to study the relation between de se thought and first person language from the angle of de se communication, there are at least two options that are available to me in choosing which side should be the starting point of my study. On the one hand, I could study the relation by first introducing the linguistic instrument that we use in de se communication, and then proceeding to articulate what is communicated in the de se communication, i.e. the de se thought. On the other hand, I could articulate the relation between de se thought and first person pronouns in the reverse order: first introducing what is communicated, and then proceeding to the study of the linguistic instrument of communication. I choose the first option. This choice does not imply that I have a commitment to the priority of language over thought. On the contrary, no argument in this dissertation relies on (or even involves) any assumption about the priority of language over thought or the priority of thought over language.¹⁵ I choose the first option mostly because of a practical concern: there have already been many philosophical works devoted to de se communication,¹⁶ but unfortunately most of them concentrate on what is communicated and only a few of them pay attention to the linguistic instrument by which we conduct de se communication.¹⁷ So according to my understanding of the current literature, the first option is much less popular than the second one. This, however, by no means implies that the approach associated with the first option is less rewarding than the approach

¹⁵ In Chapter 1 and Chapter 3, I discuss the relation between same-saying and same-thinking, and I also discuss the relation between translation of language and translation of thought. But I do not include in my discussions the the priority of language over thought or the priority of thought over language.

¹⁶ See *About Onself: De se thought and communication* (edited by Manuel García-Carpintero and Stephan Torre, Oxford University Press, 2016), Chapter 6 to Chapter 12. Also see Dilip Ninan, 'De se attitudes: Ascription and Communication', in *Philosophy Compass* 5 (7):551-567, 2010.

¹⁷ For some noticeable examples, see Komarine Romdenh-Romluc, 'First-person thought and the use of "I"', in *Synthese* (2008), 163, pp. 145-156, and Stefano Predelli, 'Intentions, indexicals, and communication', *Analysis* (2002), 62, pp. 310-316, and also Maximilian de Gaynesford, *I: the meaning of the first person term* (Oxford: Clarendon Press, 2006), Chapter 11.

associated with the second one. On the contrary, I believe this could imply that there may be some unexplored landscapes on the path from first person language to de se thought, while the path which starts from the de se thought has possibly already been adequately explored.

B. The Problem of Co-aboutness

Now it is clear the de se communication will be the focus of this dissertation. The next step is to articulate the questions that are to be covered. There are several questions that will be addressed in the dissertation. But among all of them, there is a single question that constitutes the central theme of this dissertation and it is the 'glue' of my research in the sense that it 'glues' every chapter of this dissertation together. This question is the problem of co-aboutness¹⁸.

An essential requirement of a successful communication is that participants in the communication must talk about the same subject matter. This is what I call the co-aboutness requirement of communication. In the case of communication of singular thought, for example de se communication, the co-aboutness requirement requires the participants to focus on the same individual. If participants in a communication fail to satisfy the co-aboutness requirement, for example, if they talk past each other, they fail to conduct a successful communication. Philosophers generally have no

¹⁸ There may be a concern about my terminology choice here: the reason why I do not adopt the widely accepted term 'co-reference', but instead insist on using the term 'co-aboutness' which is less often used, is that 'co-reference' is often adopted to indicate the co-referential relation between proper names, indexical terms and kind terms, which are all words, but the scope of my discussion is not limited to words: I include sentences and even inferences, which are groups of sentences, into the scope of my study. Therefore, I choose the term 'co-aboutness' to indicate the situation of 'being about the same subject matter' which takes place at the level of both words and sentences (and even inferences). But for me, the term 'co-reference' can also mean 'being about the same subject matter'. In some sections, I also use the term 'co-reference' for the purpose of improving intelligibility.

disagreement about the connection between co-aboutness and successful communication being essential, though they have different ideas on how to characterize this connection. For example, Gareth Evans argues as follows,

Now, when we discussed referential communication involving the one-off devices, in chapter 9, we noticed quite considerable restrictions on the ways in which hearers must think of referents in order to understand remarks. We traced those restrictions to the requirement that the hearer, if he is to understand, must know of some object that it is the referent. But we observed that this source of restriction would lapse if there were other participants in the communicative transaction, on whom hearers could rely to tell them which individual a speaker is referring to (though the fact that one-off referential communication does not exploit such a mechanism is not an accident). Proper names, however, are precisely not one-off referential devices; and we can think of the person (or persons) who introduced the audience to the name in question, and reinforced his pattern of use of it, as just such an authoritative third party - letting the audience know (something that can be done in countless different ways) which individual a speaker is referring to. (Evans, 1982, pp. 399-400)

In this paragraph, Evans clearly states that satisfaction of the co-aboutness requirement is a necessary condition for successful communication ('the hearer (...) must know of some object that it is the referent'). But it is unclear if Evans believes that the satisfaction of co-aboutness is also a sufficient condition for successful communication. Robert May's standpoint on this issue, on the other hand, is clearer. Let us take a look at his comment on a paragraph he quotes from Frege's paper 'On Sense and Reference':

What Frege is saying here seems little different than what he says in "The Thought." If our gaze is limited to sentences, and hence just to S-equivalent languages, then different speakers may associate different senses with the proper name "Aristotle" (although they in fact speak different languages), and

this may be "tolerated" so long as all of these senses determine the same reference. But in circumstances in which what is paramount is not the exigencies of communication but rather the structure of the language, such as in demonstrative science such as mathematics, we must meet a more stringent criterion of uniqueness of sense for proper names. We cannot, so to speak, play along as if there were just one language; rather we must insure that there is only one language in question, and that it meets the relevant standards of sense invariance. (May, 2006, p. 127)

May's view here is that satisfaction of the co-aboutness requirement is a necessary condition for a successful communication, but not a sufficient one. By fulfilling the co-aboutness condition, participants of communication ensure that the linguistic expressions which they use to communicate share the same Fregean referent. But at the same time, participants might have different Fregean senses associated with the same linguistic expression in everyday contexts. May argues that 'we must meet a more stringent criterion of uniqueness of sense for proper names' when we are trying to be precise. This indicates that according to May, making sure that the same linguistic expression is used by participants to communicate the same Fregean sense is also a requirement for successful communication in the strict sense.

Richard G. Heck, Jr agrees with May on this point. He says,

In any event, what the discussion shows is that utterances of demonstrative expressions can differ, in ways relevant to understanding and communication, even when they do not differ in their referential properties: the successful communication of information from one speaker to another depends not only upon speakers' identifying the right objects as the demonstrata, but also upon their thinking of these objects in the right sorts of ways, although it does not depend upon their thinking of the objects in any particular ways. (Richard G. Heck, Jr., 2002, p. 25)

Heck argues that successful communication not only requires participants to 'identify

the right objects', but also requires them to 'think of those objects in the right sort of ways'. I think that by the phrase 'identifying the right objects', Heck means something similar to my 'satisfaction of the co-aboutness condition'. But by the phrase 'thinking of these objects in the right sorts of ways', I think Heck means we should not only put some restrictions on the Fregean referent, which is the subject matter of the communication, but also put some restrictions on the Fregean sense, which is the 'mode of presentation' of the subject matter.¹⁹ This means that according to Heck, a successful communication not only requires participants in that communication to focus on the same subject matter, but also requires that the subject matter must be presented to those participants in a certain way.

Based on my analysis of May's and Heck's ideas, it seems that once we adopt the Fregean semantic framework of sense and reference, we are inclined to take the satisfaction of the co-aboutness requirement as a necessary condition for communication, but not a sufficient one. But in this dissertation, I am not going to discuss whether the achievement of the co-aboutness requirement is merely a necessary condition for successful communication or whether it is both a necessary and sufficient condition. What I want to emphasize here is that Evans, May and Heck all agree that co-aboutness is an essential requirement for a successful communication.

The next question is, how do we achieve the co-aboutness condition in a communication? A common answer to this question is that, we achieve the co-aboutness condition by using the same linguistic expression. This is the so-called 'same-saying' strategy. The basic idea behind the 'same-saying' strategy is that, in a communication, if we are using the same words, we are talking about the same subject matter. This idea is widely accepted. For example, Recanati says:

As Prosser emphasizes, use of the same word by the interlocutors typically triggers a presupposition of coreference. In a conversation in which different

¹⁹ Frege characterizes 'sense' as the mode of presentation of reference.

tokens of a given word are produced by the speech participants, the tokens are de jure coreferential (unless something defeats the presupposition). It is presupposed that all the users use the word in the same way. (Recanati, 2016, p. 123)

So, in the communication of singular thought, if participants want to fulfil the co-aboutness condition, they can use the same referential expressions to refer to the same individual about whom they wish to talk. As Recanati suggests in this paragraph, 'use of the same word by the interlocutors typically triggers a presupposition of coreference'. But this is still not a comprehensive characterization of the same-saying strategy. We need to find the criteria which can be used to determine whether the same-saying strategy is effective or not.

In Laura Schroeter's writings, I can find at least three criteria that we can use to determine whether the co-aboutness condition is properly fulfilled by means of the same-saying strategy:

A primary semantic function of proper names, common nouns or predicates is to ensure that all parties to a conversation pick out the very same subject matter with these words. Sameness of meaning must somehow guarantee sameness of subject matter – and it must do so in a way that is accessible to ordinary speakers. At the same time, however, we normally intend such words to pick out objectively interesting and important features of the world, not just arbitrary stipulative categories. To vindicate this epistemic ambition, sameness of meaning must allow for speakers' fallibility about the precise nature of the subject matter and about the best ways to identify it. Sameness of meaning must not close off legitimate inquiry or debate into the nature of the subject matter. (Schroeter, 2012, p. 177)

and

On an acceptable account of meaning identity, sameness of meaning must

guarantee sameness of subject matter, while allowing for fallibility about the nature of that subject matter. (Schroeter, 2012, p. 178)

The three criteria that I find in those two paragraphs are:

First, the connection between same-saying and co-aboutness should be tight enough to ensure that participants in a communication refer to the same subject matter when they use the same linguistic expressions.

Second, the connection between same-saying and co-aboutness should also be loose enough to leave open the possibility that participants in the communication can have disagreements over the nature and properties of the subject matter.

The third criterion is that the same-saying 'must do so in a way that is accessible to ordinary speakers', which means for competent participants in communication, the connection between same-saying and co-aboutness should be transparent to them: they know that if they are using the same linguistic expression, they are talking about the same subject matter.

However, when we try to satisfy the co-aboutness requirement for de se communication, we find that the same-saying strategy simply does not work. If someone tells me 'I am lucky today', and I want to express agreement with her, I cannot do this by simply replicating her words. If I reply to her by saying 'Yes, I am lucky', what I am talking about is my luck, not her luck. This is a clear demonstration that in de se communication, the same-saying strategy leads to the failure of fulfilling the co-aboutness condition.

So, how can we satisfy the co-aboutness requirement of a de se communication? This is exactly the central question of my study.

C. Content of Each Chapter

In this section, I summarize the content of each chapter of this dissertation.

The first chapter is an introductory chapter. I begin this chapter by asking what is the alleged 'problem of de se attitude'. I introduce Dilip Ninan's, Matheus Valente's and François Recanati's answers to this question. Dilip Ninan argues that the real problem concerning the de se attitude is that there is a conflict between multiple people's agreement on a certain doxastic attitude and an explanation of their actions based on that attitude, when the attitude in question is a de se attitude. Matheus Valente argues that the real problem associated with the de se attitude takes place when we try to share our de se attitudes with other people. He argues that there are normally three principles to which an attitude and content normally conforms. However, when I try to share a de se attitude with others, Valente argues, I have to reject at least one of those principles. François Recanati argues that the communication of indexical attitudes raises a challenge for the traditional understanding of how a (verbal) communication works, due to what he calls the indexical attitudes' *limited accessibility*. I agree with Recanati on his general strategy for answering this question, but I do not agree with the specific answer that he provides.

For me, one of the real problems of de se attitude is how to satisfy the co-aboutness requirement of a de se communication. For communication that conforms to the replication transmission model, the co-aboutness condition is achieved by utilizing the same-saying strategy. But in my view, de se communication is distinguished from other types of communication because the replication transmission model is incompatible with de se communication. Instead, de se communication necessarily conforms to the translation transmission model. I then introduce Frege's, Perry's and Recanati's characterizations of the translation transmission model. Based on my analysis of those three proposals of the translation transmission model, I propose my own characterization of that model. I argue that in translation transmission model, not every type of translation is evaluated as being equally useful or valid, especially when we use that model to explain de se communication. There is a certain type of

translation which can better secure the achievement of the co-aboutness condition in de se communication, compared with other types of translation. I call this type of translation the 'X translation'. In Chapter 3, I will reveal what the 'X translation' actually is.

The second chapter is not about de se communication *per se*, but it is about what is communicated in de se communication, i.e. de se thought. In order to express my de se thoughts and communicate them to someone else, I must first be able to produce de se thoughts about myself. In Chapter 2, I try to articulate one aspect of the production of de se thought, which concerns whether or not our production of de se thought involves identification (and/or mis-identification). I call this 'the identification problem of de se thought'. In order to clarify whether a process of identification is involved in the production of de se thought, I introduce two theses as the background of my clarification. They are the essential indexical thesis and the immunity to error through misidentification thesis (IEM). A thesis implied by the essential indexical thesis concerning identification may seem to be incompatible with another thesis implied by the IEM thesis concerning identification. This seems to be a very problematic situation. However, I argue that the alleged incompatibility is apparent, and de se thoughts are both essentially indexical and immune to error through misidentification.

After discussing Cappelen & Dever's criticism of the essentiality of indexicality *per se* and Magidor's criticism of the property account of de se content, I argue that the essential indexical thesis is closely related to an identification thesis, which is that in entertaining a de se thought, I do not need to appeal to any representation of myself for the purpose of identifying myself. This thesis reveals a close connection between de se thought and (cognitive) identification. However, this connection may seem to be incompatible with another epistemic property of de se thought, which is the immunity to error through misidentification. According to the Simple Account of IEM, the de se thought is immune to the error of misidentification because entertaining a de se thought does not involve identification. This *identification-free* condition seems to be incompatible with the identification thesis implied by the essential indexical thesis.

However, I argue that this apparent incompatibility can be resolved if we adopt a new account to explain the connection between de se thought and the IEM thesis, which does not contain the *identification-free* condition. This new account is my Definition Account. What does ‘being immune to the error of misidentification’ really mean for de se thought? I argue that the attribution of IEM property to de se thought is a corollary of the definition of de se thought. I call this the Definition Account of the relation between de se thought and the IEM thesis. The basic idea of Definition Account is that de se thought has the property of IEM simply because of its definition. Since there is no ‘identification free’ condition in the Definition Account, there is no incompatibility between the essential indexical thesis and the Definition Account of IEM.

The third chapter is the conclusion part of my dissertation. In this chapter, I present my articulation of the fourth feature of the translation transmission model. I argue there is a specific type of translation, which can guarantee the co-aboutness for de se communication better than other types of translation. This is the ‘X translation’. In this chapter, I will reveal what ‘X translation’ actually is. First, I introduce three normative theses concerning the connection between translation and co-aboutness in communications that conform to the translation transmission model. Second, I introduce Alex Worsnip’s analogy strategy as a background for my subsequent discussions. Worsnip argues that there is an analogy between the interpersonal disagreement and the intrapersonal incoherence. I adopt his approach to draw an analogy between the intrapersonal de se inference and the interpersonal de se communication. Let us call the co-reference of multiple tokens of ‘I’ in a de se inference the ‘I-I’ co-reference. I argue there is a certain type of translation in de se communication, whose co-referentiality is analogous to the ‘I-I’ co-reference. This is the ‘I-You’ translation. ‘I-You’ translation is the ‘X translation’ that I am looking for.

The analogy between ‘I-You’ co-reference and ‘I-I’ co-reference is plausible, because there are some non-trivial similarities between them. The first similarity is that both of them are established without appealing to any specific representation of the

relevant subjects, and therefore are not affected by any misrepresentation of those subjects. The second similarity is that the co-referentiality of 'I-I' same-saying is transparent to the producer of de se inference, just as the co-referentiality of 'I-You' translation is transparent to the participants of de se communication. The third similarity, which is the most important similarity, is that the 'I-I' co-reference and the 'I-You' co-reference are both explained by the externalist explanations. Here I introduce Jenann Ismael's externalist explanation of the co-reference of I-thoughts held by a single person, and based on Ismael's externalist approach, I develop a similar externalist explanation of 'I-You' co-reference. Why does my 'I' co-refer with your 'you' in a de se communication? According to my externalist explanation, the 'I-You' co-reference is explained and secured by factors of the external aspect of communication. Those factors are the mutual recognition between participants of a communication and the 'I-You' reciprocity as a norm which governs the use of 'I' and 'You'. The match between the mutual recognition and the 'I-You' reciprocity not only explains why the co-reference of the 'I-You' translation can be secured without appealing to specific representation of either 'I' or 'You', but also explains why the 'I-You' co-reference is recognizable for participants of de se communication, i.e. it explains the transparency of the co-referentiality of 'I-You' translation. Once the externalist explanation of 'I-You' co-reference is articulated, I can continue to explain why other types of possible translation for de se communication do not surpass the 'I-You' translation. At the end of this chapter, I will examine whether the 'I-You' translation follows the three criteria that I propose. The result of my examination shows that 'I-You' translation does follow those criteria, so I think it is safe to say that 'I-You' is the 'X translation' that I am searching for. Since 'X translation' is a crucial part of the translation transmission model, by clarifying what 'X translation' is, we make substantial progress in articulating the translation transmission model.

Chapter I The Peculiarity of De Se Communication

There is no doubt that philosophers have been enthusiastic about studies on the first person perspective (including the de se attitude and its relevant linguistic expressions) for decades since John Perry and David Lewis published their memorable papers (Perry 1979, Lewis 1979) in the late 1970s. But after years of enquires and debates, philosophers are in a situation which is a little embarrassing: they rarely have any consensus on what the real problem (or real problems) of the de se attitude is supposed to be, even after many explanations and hypotheses being proposed during the past forty years²⁰. For philosophers who want to make their own contributions to the current debates on the study of de se attitude, they must answer a crucial question: what are the real problems (or the real problem, if someone argues there is only one single problem) of the de se attitude? By the phrase ‘real problem’, I mean a problem which is peculiar to the de se attitude and which differentiates the de se attitude from the rest of our mental and linguistic lives.

²⁰ See Dilip Ninan’s complaint about the situation:

But despite widespread agreement that there is a problem of de se attitudes (‘the problem of the essential indexical’ to use Perry’s term), the literature on these topics has been less than completely clear on just what that problem is supposed to be. More specifically, what is unclear is what the distinctive problem of de se attitudes is, what problems such attitudes raise over and above other more familiar problems facing theories of propositional attitudes (e.g. Frege’s Puzzle). (Dilip Ninan, 2016, p. 86)

and also see Matheus Valente’s criticism of a once popular characterization of the essential indexical problem:

In the following sections, we will see that demarcating the “special indexical challenge” is far from an easy task. Indeed, many have failed to see that in order to prove that one such challenge really exists, it is not enough to show that one cannot fully characterize indexical attitudes neither by means of de re nor de dicto propositions, since the same predicament is true of other singular attitudes. (Matheus Valente, 2017, p.2)

In this chapter, I will propose my own characterization of a problem that I believe is peculiar to the de se attitude. It is the problem of de se communication. I will argue that our (verbal) communication of the de se attitudes creates a unique problem for any explanation of communication because of the de se attitude being involved. I will continue to argue that this problem is not found in non-indexical communication and not necessarily found in non-first-person indexical communication, therefore it differentiates the de se communication differentiated from other types of communication. I call this problem 'the necessary application of the translation transmission model', which will be explained in the latter sections of this chapter. However, it should be noted that I do not argue that the real problems of de se attitude are exhausted by my characterization of the de se communication problem. There is more than one real problem associated with the de se attitude, but there is also no way for me to discuss all of them in a single chapter. I will show at the end of this chapter that an advantage of my strategy of concentrating my attention solely on de se communication is that an important question, which is the question of semantic tension, i.e. a tension between the same-saying and the co-aboutness, can be properly revealed by means of the discussion in this chapter.

First, I introduce three recently proposed theories on what is the real problem of the de se attitude. After a careful examination of them, I discover that all three proposals focus on the same aspect of the first person problem, which is the problem of the de se communication, although two of them focus on this problem implicitly, while the third focuses on it explicitly. What is more, there is a single question which is essentially implied in their discussions on the de se communication. This is the question of semantic tension, i.e. a tension between the same-saying and co-aboutness. I then proceed to introduce my own characterization of the problem of de se communication. The crucial difference between my approach and the approaches contained in those three proposals is that I pay attention not only to what is communicated, i.e., the content of communication, but also to the vehicle by means of which we do the communication, i.e., the first person linguistic expression. I

discover that the alleged translation transmission model is necessarily applied to the de se communication, which means I have to translate what I received in a de se communication in order to successfully participate in that communication. I call this the peculiarity (or idiosyncrasy) of de se communication. Since we (verbally) communicate our thoughts by means of language, the translation takes place at the level of thought as well as the level of language. So, there are semantic and cognitive translations. However, my focus in this chapter is on the language. The necessary applicability of translation transmission model becomes a challenge to any explanation of communication because it destroys a fundamental presupposition of communication, which is that all participants of a communication can talk about the same topic using the same linguistic expressions, and, what is more, their talking about the same topic is a default implication of them using the same linguistic expressions. This presupposition of consistency and assurance between the same saying and the co-aboutness is in most cases taken for granted in communications to which the replication transmission model is applied, but the employment of translation transmission model breaks both the consistency and the assurance. A large part of my dissertation can be regarded as my response to the semantic tension challenge.

1.1 Three Proposals on the Real Problem of De Se Attitude

In this section, I put under scrutiny three different proposals that have been recently proposed about the real problem of the de se attitude. The first one is Dilip Ninan's account of a conflict between the principle of agreement and the principle of explanation in the case of the de se communication (Dilip Ninan, 2016). The second one is Matheus Valente's account that my sharing someone else's de se attitude leads to the rejection of certain theses about propositional attitude which are otherwise plausible in the case of non-de se communication (Matheus Valente, 2017). The third one is Recanati's account of the incompatibility between the de se communication and

the presupposition of the naïve conception of communication (Recanati, 2016). I argue that all three proposals share the philosophical insight that turning our attention to de se communication is the right direction for discussing the real problem of the de se attitude. Two of them contain this insight implicitly (in the cases of Ninan and Valente), while the third one exhibits it explicitly (in the case of Recanati). What is more, there is a single question which gets implied in all three proposals. This question is what I call the semantic tension, i.e. a tension between the same-saying and the co-aboutness. I argue that the shared implication of the question of semantic tension is not accidental for those proposals. It is not accidental at all that a serious discussion on de se communication should take the question of semantic tension as one of its top concerns, because, as I will show later in this chapter, the necessity of the semantic tension problem is what makes de se communication unique.

1.1.1 Dilip Ninan on the Conflict between Agreement and Explanation

Dilip Ninan (2016) argues that the real problem concerning the de se attitude is that there is a conflict between multiple people's agreement on a certain doxastic attitude and an explanation of their actions based on that attitude, when the attitude in question is a de se attitude. Ninan argues that in the case of non-de-se doxastic attitude, the principle of agreement and the principle of explanation can be consistent with each other, but the intervention of the de se attitude breaks the harmony. To elaborate his argument, let us first introduce his characterization of the agreement principle:

(4) AGREEMENT

Agreement is a two-place relation between a group of individuals and a content.

A group of individuals stands in the agreement relation to a content p iff all the members of the group believe p .

Let B_{xp} and B_{yq} be token beliefs belonging to x and y respectively. Then x and y agree on something in virtue of x 's having token belief B_{xp} and y 's having token belief B_{yq} iff $p = q$. (Dilip Ninan, 2016, p.100)

It can be clearly seen that according to Ninan's characterization of agreement, Perry and Lewis agree on P if both of them believe that P , and Perry and Lewis both believe the same thing, i.e. that P , if they agree with each other on P . Besides the principle of agreement, Ninan also introduces a principle of explanation. By the term 'explanation', he means a relation between someone's mental attitudes on the one side, and someone's action on the other side. His characterization of explanation is this:

(5) EXPLANATION

Suppose the fact that x performed action α is explained by the fact that x has beliefs B_{xp1}, \dots, B_{xpn} and desires D_{xq1}, \dots, D_{xqk} .

Then, if y has beliefs B_{yp1}, \dots, B_{ypn} and desires D_{yq1}, \dots, D_{yqk} , then, other things being equal, y will also perform α .

(...) I take it that other things are not equal if:

- y is suffering from psychological deficiencies such as weakness of will; or
- y , for whatever reason, is unable to perform action α or
- y is able to perform α but performing α , is not the most natural, normal, or efficient way for y to reach y 's goal.

But, of course, this list is not exhaustive. (Dilip Ninan, 2016, p.102)

According to Ninan's characterization, if Perry and Lewis perform the same action, when other things are equal, their actions can possibly be explained by the same set of beliefs (and/or desires), and if Perry and Lewis share the same set of beliefs (and/or desires), when other things are equal, the set of belief (and/or desires) will lead them to perform the same action. Now we have Ninan's characterizations of agreement and

explanation at hand, and it can be easily discovered from those characterizations that the principle of agreement and the principle of explanation are consistent with each other, and they can be connected. On the one hand, if Perry and Lewis agree with each other on something, they share the same set of beliefs about that thing, and if they share the same set of beliefs, given other things being equal, it can be rationally expected that they perform the same action. On the other hand, if Perry and Lewis perform the same action, when other things are equal, it can be rationally assumed that they share the same set of beliefs (and/or desires) about something, and this in turn means that Perry and Lewis agree with each other on that thing. Therefore, the agreement can be rationally assumed by the explanation, and the explanation can be rationally expected based on the agreement. So far so good.

However, Ninan argues that the consistency between the principle of agreement and the principle of explanation is merely a limited consistency, and it will vanish once the *de se* attitude is taken into consideration. To illustrate this point, let us imagine a scenario like this:

Imagine that you and I are walking in the woods when a bear begins to chase me. Now suppose that you and I agree on all of the relevant facts of the case, and we also agree on how we'd like the situation to turn out. For example, suppose that we agree that I'm being chased by a bear. That is, suppose I believe *de se* that I'm being chased by a bear and that you believe *de te* that I am being chased by a bear, i.e. you have a belief you could express to me by saying, "You are being chased by a bear". (Dilip Ninan, 2016, p. 103)

In this scenario, it can be rationally expected that I will curl up into a ball to avoid being mauled by a bear once I notice that I am the one who is being chased by the bear. However, it can also be rationally expected that you will not curl up into a ball because it is not you who is being chased by any bear in that scenario. According to the principle of explanation, our choosing the different actions implies that we do not share the same set of beliefs. You and I agree that I am being chased by a bear and you

are not being chased by any bear. This belief is shared by both of us. Here the consistency between the principle of explanation and the principle of agreement disappears: in the scenario, when the belief that we agree on is a de se belief like 'I am being chased by a bear' (or at least it is a de se belief for one of us), we cannot rationally expect that both of us are going to perform the same action.

Ninan argues that the inconsistency between agreement and explanation is the real problem of the de se attitude. I find that his discovery is full of philosophical insight, but it seems to me that he does not exhaust the philosophical insight contained in his own discovery. There is at least one more question that we can continue to inquire into after the inconsistency has been realized: where does this inconsistency come from? Since the inconsistency is caused by the application of de se attitude, and the inconsistency is between the principle of agreement and the principle of explanation, the question can be further specified in this way: which one of the two principles, the agreement or the explanation, is affected by the impact of de se attitude to the extent that it fails to be in accordance with the other principle, or perhaps both of them? I believe that the answer to this question is the agreement principle. To begin my argument, let us first take a close look at the explanation principle. Basically, what the explanation principle says is that the same action can be explained by the same set of beliefs (and/or desires), and the same set of beliefs (and/or desire) result in the same action. When de se beliefs (and/or de se desires) are included in the set of beliefs (and/or desires), it does not affect the criteria which we adopt to examine the sameness of actions. However, the inclusion of de se attitudes does have an effect on the way in which we examine the sameness of attitudes. Which standard shall we adopt to decide what can be counted as being the same belief, when the beliefs in question are de se beliefs? This is exactly the question that the agreement principle is supposed to answer. So, the examination of the explanation principle leads us to a further examination of the agreement principle. Let us recall Ninan's formulation of the criterion about believing the same belief in his agreement principle:

Let B_{xp} and B_{yq} be token beliefs belonging to x and y respectively.

Then x and y agree on something in virtue of x 's having token belief B_{xp} and y 's having token belief B_{yq} iff $p = q$. (Dilip Ninan, 2016, p.100)

What does the equation ' $p=q$ ' in Ninan's formulation mean? I think there are at least two options for understanding the meaning of ' $p=q$ '. ' $p=q$ ' can be understood as ' p and q are the same proposition', or it can also be understood as ' p and q are about the same thing'. At first glance, there seems to be no conflict between the two understandings of ' $p=q$ ', because in the case of non-perspective ordinary propositional attitude, if ' $p=q$ ' this means that p and q are the same proposition, then p and q are also about the same thing; if p and q are about the same thing, then p and q can possibly be the same proposition, although not necessarily the same proposition. In other words, in the case of non-indexical communication, the same-saying implies the co-aboutness by default, and the co-aboutness can be compatible with the same-saying. This conclusion can be easily drawn from a classical Fregean framework, according to which two propositions are about the same thing, if they share the same truth condition (let us call this the 'sameness condition'), and two propositions are identical if it is impossible for a qualified epistemic agent to confirm an application of a certain propositional attitude to one proposition but deny an application of the same attitude to the other proposition (let us call this the 'identity condition'). So far so good. But when it comes to the pair of beliefs mentioned in the scenario described above, the relation between the sameness condition and the identity condition do not seem to be as consistent as it was. Let us take a look at the pair of beliefs:

I: 'I am being chased by a bear!' You: 'You are being chased by a bear!'

According to the scenario, the truth conditions of the two beliefs are the same, which means that they are about the same thing. But it seems that the propositional contents of the two beliefs are not identical, because I can confirm my belief that 'I am being chased by a bear' and at the sametime deny my belief that 'You are being chased by a bear'. So, we have no reason to deny that the two beliefs are not identical within the Fregean framework. What is more, it can be very tricky to argue that it is possible for

two attitudes to be identical within the Fregean framework when one of them is a de se attitude. Let us take a look at the two examples below:

1. I: 'I am being chased by a bear!' You: 'Perry is being chased by a bear!'

2. I: 'I am being chased by a bear!' You: 'I am being chased by a bear!'

In the first example, it is easy to claim that those two propositional attitudes are not identical, because I, as a single epistemic agent, can plausibly confirm one of them but deny the other, as long as I do not have enough evidence to believe that I am Perry. In the second example, it is obvious that I must confirm both attitudes at the same time. It is at least equally obvious, if not more obvious, that those two attitudes differ in their truth conditions. This confusion is caused by the Fregean framework itself, since the results we get from applying that framework to de se communication betray our intuition. One possible protest is that what gets involved in the scenario mentioned above is not my belief that 'You are being chased by a bear!', but rather, it is your belief that 'You are being chased by a bear' that gets involved. Therefore, when examining the identity condition of those two beliefs, it should be my belief 'I am being chased by a bear' and your belief 'You are being chased by a bear' that get examined, not my belief 'I am being chased by a bear' or my belief 'You are being chased by a bear'. We are not justified to say that the two beliefs are not identical, if the protest is cogent. However, I argue that this protest is not only an objection to the conclusion that those two beliefs are not identical, but also an objection to the Fregean framework itself. What is requested by the protest is a criterion which can be adopted in judging whether multiple propositional attitudes are identical or not when those attitudes are maintained by different epistemic agents. This request can hardly be fulfilled within the Fregean framework, because in the Fregean framework, we can judge whether two propositional attitudes are identical or not only when those two attitudes are maintained by a single epistemic agent. So, to fulfill the request, a new framework, which allows the examination of identity (or identities) of attitudes maintained by multiple epistemic agents, must be introduced. A typical situation, in which multiple

epistemic agents evaluate the identity or difference between their attitudes, is the situation of communication. The new framework, as I will argue in subsequent parts of this chapter, is related to the communication of the de se attitude.

Now it is explicit why Ninan's conclusion that the real problem of the de se attitude is an inconsistency between the principle of explanation and the principle of agreement does not exhaust the philosophical insight of his own argument. From an examination of the two principles, I find that the agreement principle is the one which is deeply affected by the impact of de se attitude, and the real impact on the agreement principle caused by the de se attitude is that the criterion of two attitudes' being about the same thing and the criterion of two attitudes' being identical are no longer consistent. To put it in other words, it is an inconsistency between the co-aboutness and the same-saying. Therefore, after a careful discussion of Ninan's discovery, I find it basically to be a problem of semantic tension.

1.1.2 Matheus Valente on the Incompatibility between Principles

Matheus Valente (2017) argues that a real problem associated with the de se attitude takes place when we try to share our de se attitudes with other people. He argues that there are three principles to which a (propositional) attitude with its content normally conforms. However, when I try to share a de se attitude with others, Valente argues, I have to reject at least one of those principles. To introduce his argument, let us first take a close look at the three principles. This first principle is the principle of explanation. Valente takes the idea of explanation from Dilip Ninan, and he clearly refers to Ninan when he gives out his formulation of that principle:

The principle (let us call it, following Ninan, 'Explanation') that two agents who are doxastically identical should be disposed to perform the same actions. (Matheus Valente, 2017, p.12)

The second principle is the principle of absoluteness, which claims that if an attitude is true for one agent, it should be true for other agents as well (Matheus Valente, 2017, p.13). The third principle is the principle of shareability, which states that an attitude should be capable of being shared by different agents. For a non-indexical belief, it seems that it can conform to all the three principles easily. But Valente invites us to consider a situation in which we are facing a sharing of a de se belief:

Take some arbitrary agent Amelia who has a belief she expresses by uttering 'I am an aviation pioneer'. What would it take for Berthold, who is distinct from Amelia, to hold the same belief that she expresses by means of that utterance? (Matheus Valente, 2017, p.11)

There are three possible answers, according to Valente, to explain how Berthold can share Amelia's de se belief:

My argument, which I take to be a development of an argument found in Ninan (2016),⁶ is that the three most plausible answers to that question each lead to the rejection of a different but independently plausible thesis about attitudes. Here are the three possible replies:

Option #1: For Berthold to share Amelia's first-person attitude, he would need to form a belief (perhaps in response to Amelia's utterance) which he would express by means of 'You [pointing at Amelia] are an aviation pioneer'.

Option #2: For Berthold to share Amelia's first-person attitude, he would need to have himself a first-person attitude about himself, one which he would express by means of 'I am an aviation pioneer'.

Option #3: It is impossible to share someone else's first-person attitudes (e.g. they are private, or limited accessibility, etc.) (Matheus Valente, 2017, p.11)

If I choose the first option, it turns out I have to reject the principle of explanation, because the way of sharing Amelia's de se belief as shown in the first option leads Berthold to plausibly perform actions which are different from the actions Amelia

performs based on her de se belief. If I choose the second option, it results in my rejection of the principle of absoluteness, because that way of sharing a de se belief as shown in the second option makes it possible that the same belief is true for one agent, but false for another agent. If I adopt the third option, I am obviously in contradiction with the principle of shareability. Therefore, no matter which option I adopt, Valente argues, I have to reject at least one of the three principles summarized above. This is a challenge to the traditional theory of attitude (and its content), and since the challenge is raised because of the impact of de se attitude, Valente argues that this is a real problem of the de se attitude.

One thing that I think is important in Valente's characterization of a problem of the de se attitude is that he clearly refers to the shareability of de se attitude. To share a de se attitude, which is mine, with other people, is to communicate that attitude to them. Although Valente sticks to his choice of terminology and talks only about shareability of de se belief thorough his entire argument, I think that his discussion makes implicit reference to the communication of de se attitudes, and therefore, makes implicit contribution to my discussion of de se communication.

1.1.3 Recanati on De Se Communication

Between all three proposals discussed in section 1.1, Recanati's proposal is the only one that focuses explicitly on the communication aspect of the problem of indexical thought (of which the de se thought is a subset). François Recanati (2016) argues that due to what he calls the limited accessibility, the communication of indexical attitudes raises a challenge for the traditional understanding of how a (verbal) communication works²¹. To discuss this challenge, let us first introduce the traditional understanding

²¹ 'This property of indexical thoughts creates an obvious problem with respect to communication.'
(Recanati, 2016, p. 149)

of communication, which Recanati calls the ‘Naïve conception of communication’:

The Naïve Conception of Communication rests on the idea that communication is the replication of thoughts: the thought the hearer entertains when he understands what the speaker is saying is the very thought which the speaker expressed. (Recanati, 2016, pp. 149-150)

According to this definition of the Native Conception of Communication, it can be analyzed into two tenets:

The Naïve Conception comports two tenets, which Weber calls the ‘mind-to-speech principle’ and the ‘speech-to-mind principle’ (Weber 2013: S208):

(Mind to speech) Speaker’s thought = thought expressed by the utterance

(Speech to mind) Thought expressed by the utterance = hearer’s thought

(Recanati, 2016, p.152)

Similar to the equations I discussed in section 1.1.1, here another pair of equations is presented in the Naïve Conception of Communication. In section 1.1.1, we saw that there is an ambiguity in understanding what an equation between two propositions means. That equation can be understood as indicating an identity relation between two propositions (I call this same-saying), or it can be understood as indicating two propositions’ sharing the same truth condition (I call this co-aboutness). One similar ambiguity takes place here as well: the pair of equations in the Naïve Conception of Communication can be understood as showing an identity relation between the speaker’s thought, the thought expressed by the utterance, and the hearer’s thought, or it can be understood as showing that those three thoughts are employed to keep track of the same thing. This ambiguity is actually noted by Recanati himself. He makes a distinction between the subjective content of a thought and the objective truth conditional content associated with the same thought. According to him:

The vehicles (i.e. thoughts) are supposed to do certain things, which they can do only if the context in which they are deployed satisfies certain conditions.

So, the vehicles carry presuppositional content: their deployment means something, namely, that the context satisfies the conditions, whatever they are, on which their correct functioning relies. That aspect of the overall content of the vehicle is context-independent, in contrast to the truth-conditional content, which constitutively depends upon what I have called the external factor. (Recanati, 2016, pp. 145-146)

Recanati's distinction here is very similar to my distinction: the subjective content of a thought corresponds to what I call the identity of thought, and the truth conditional content of a thought corresponds to what the thought is keeping track of. The distinction's two sides, as Recanati argues, go hand in hand without any problem when we apply the Naïve Conception of Communication to understand a non-indexical communication, which means that the truth conditional content of the speaker's thought, the truth conditional content of the thought expressed by speaker's utterance, and the truth conditional content of the hearer's thought are the same, so are their subjective contents. When the Naïve Conception of Communication is adopted to explain a communication of indexical thoughts, the consistency between the preservation of the subjective content and the preservation of the truth conditional content simply vanishes. Recanati notices that:

This property of indexical thoughts creates an obvious problem with respect to communication. How can we manage to communicate such thoughts to those who are not in the right context? If the hearer entertains a thought with the same subjective content as the speaker's thought, that thought will not retain the truth-conditions the speaker's thought originally had. Were the hearer, upon understanding the speaker's first person utterance, 'I am thirsty', to entertain the subjective content of the speaker's thought, he would thereby ascribe to himself the property of being thirsty (he would think 'I am thirsty'). Communication would fail dramatically, for the speaker meant to communicate that she was thirsty, not that the hearer was. To keep the truth-conditions constant across contexts as we shift from the speaker's to the

hearer's point of view, we have to adjust the internal content of the thought, in the same way as Frege says we do in speech with the indexicals when the context changes. (Recanati, 2016, p. 149)

As Recanati shows in this paragraph, when communicating our indexical thoughts to each other, we may find it difficult to preserve both the subjective content and the truth conditional content at the same time. For example, my belief 'I am being chased by a bear' and your belief 'I am being chased by a bear', according to Recanati, share the same subjective content, but do not share the same truth-conditional content, and my belief 'I am being chased by a bear' and your belief 'You are being chased by a bear' may share the same truth condition, but they differ in subjective content. If I try to communicate one of my indexical beliefs by saying that 'I am being chased by a bear' and you understood it as 'You are being chased by a bear', the truth condition of our beliefs is preserved during the communication, but there is no preservation of subjective content here. If you try to preserve the subject content of our indexical beliefs in that communication, you have to fabricate a belief such as 'I am being chased by a bear' based on your acceptance of my utterance that 'I am being chased by a bear'. However, this is obviously a misunderstanding of both my utterance and my thought.

The fact that we have to choose between the preservation of subjective content of our beliefs and the preservation of truth conditional content of those beliefs is a real challenge for the Naïve Conception of Communication, according to which both contents can be preserved smoothly. This is a challenge which takes place when we try to explain the indexical communication using the Naïve Conception of Communication. Recanati's response to this challenge is that the entire project of Naïve Conception of Communication should be rejected. Instead, he argues that we need a new conception of communication which involves a transformation, not a replication, of what is communicated to explain communications that involve the transmission of indexical attitudes:

The Naïve Conception of Communication rests on the idea that

communication is the replication of thoughts: the thought the hearer entertains when he understands what the speaker is saying is the very thought which the speaker expressed. But in the indexical case, there is no replication, but some kind of systematic transformation. (Recanati, 2016, p149)

I agree with Recanati that a new conception of communication is needed, and I also agree with him that the characteristic of the new conception should be a transformation of what is communicated, instead of a replication. However, there are also three disagreements (which are not trivial) between Recanati's proposal and my proposal. First, in his characterization of the Naïve Conception of Communication, Recanati notices there is a distinction between the subjective content of thought and the truth conditional content of thought. This distinction is closely related to the distinction between same-thinking and co-aboutness that I mentioned in Section D of the Introduction. However, I am not satisfied by Recanati's description of the relation between the subjective content and the truth conditional content. According to Recanati, in the Naïve Conception of Communication, the preservation of subjective content and the preservation of truth conditional content are consistent, which means they take place at the same thing during a communication without bothering each other. However, the preservation of subjective content does not only go hand in hand with the preservation of truth conditional content, but rather, it reveals and ensures the preservation of truth conditional content. To put it in my terminology, in the Naïve Conception of Communication, the same-thinking condition reveals and ensures the co-aboutness condition by default: it can rationally be assumed by default that the all the participants of a communication, which conforms to the Naïve Conception of Communication, are talking about the same thing if they are deploying the same thought. The second disagreement between my proposal and that of Recanati's concerns what a proper understanding of the systematic transformation should emphasize. Both Recanati and I agree that the alleged systematic transformation is the most important part of the new conception of communication. According to Recanati, we should focus on the transformation of the subjective content. I do not take his

approach. I argue that a more fruitful approach is to focus on the transformation of the linguistic expression. This disagreement can be easily discovered from the earlier discussions in this section. Although I am well aware of the close connection between Recanati's 'subjective content/truth conditional content' distinction and my 'same-thinking/co-aboutness' distinction, I never adopt his terminologies, because my 'same-thinking/co-aboutness' distinction is developed as a preparation for my characterization of the cognitive tension, but for me, what is involved in *de se* communication first and foremost is not the cognitive tension, but rather the semantic tension. This disagreement between my proposal and Recanati's can be traced back to a more elementary disagreement on what the vehicle of communication is: Recanati takes thought to be the vehicle of communication²², but I believe that linguistic expressions, which are deployed to express thought, are the real vehicle of (verbal) communication. This point will be further elaborated in subsequent sections of this chapter. The third disagreement between Recanati and myself is about the target of the discussion. Recanati clearly states that his discussion on both the Naïve Conception of Communication and the new conception of communication takes the indexical communication in general as its target.²³ But the target of my research in this dissertation is solely the *de se* communication. Of course, the *de se* communication is a subspecies of the indexical communication, but focusing on the indexical communication in general makes it difficult to reveal the peculiarity which can be found only in *de se* communication. Therefore, in order to focus on *de se* communication, I will distinguish it from other types of indexical communication in the latter of this chapter (see section 1.2.4).

²² 'The vehicles (i.e. thoughts) are supposed to do certain things, which they can do only if the context in which they are deployed satisfies certain conditions.' (Recanati, 2016, pp. 145-146)

²³ 'This property of indexical thoughts creates an obvious problem with respect to communication. How can we manage to communicate such thoughts to those who are not in the right context?' (Recanati, 2016, p. 149)

1.2 De Se Communication vs Other Types of Communication

In the previous section, I discussed three recent proposals on what is the real problem of the de se attitude. Although there are some differences between those proposals, it can be summarized that there are at least three consensuses between them based on my analysis. First, they all focus, implicitly or explicitly, on a special aspect of the de se attitude problem, which is the de se communication. Second, they all agree that the de se communication raises challenges to the traditional theory of the explanation of communication (although their characterizations of the challenge differ in detail) and that those challenges can be raised only by de se communication. Third, it is shown in my discussions that all the challenges described in those proposals are either caused by what I call the semantic tension, which is a tension between same-saying and co-aboutness, or caused by what I call the cognitive tension, which is a tension between same-thinking and co-aboutness. The three consensuses, especially the third one, provide important materials for my next discussions.

In this section, I try to elucidate my own characterization of a special problem in de se communication. For me, de se communication is the communication of de se thoughts and de se attitudes (and their contents) using the first person linguistic expressions. Here the de se communication is an umbrella concept, under which a wide variety of states and contents can be included, for example de se belief communication, de se desire communication, de se imagination communication, and so on. For the sake of focusing on the essential peculiarity of de se communication, I limit the scope of my discussion in this section only to de se belief communication.²⁴.

24. I use the phrase 'de se belief communication' in this chapter only because I, following Perry and Lewis, take belief to be a common object of de se communication. But this by no means implies that the domain of my discussion is limited to belief. I am quite confident that my discussion here can be applied to communication of other types of first person attitude, such as thought or judgement as

For the current purpose, I propose a definition of *de se* belief communication that is relatively narrow: by ‘*de se* belief communication’ I mean the communication of *de se* belief by means of first person sentences. The narrowness of the definition comes from the narrowness of my definition of *de se* belief.²⁵ By ‘*de se* belief’ I mean belief about oneself as oneself, i.e. based on one’s first person perspective. Introducing some examples may help explicate what ‘about oneself as oneself’ means. Let us suppose there is only one person whose name is Peter on the earth and Peter has a belief in his mind that ‘Peter feels good’. In this case Peter has a belief about himself since he is Peter, but his belief is not about himself as himself since there is no evidence showing that Peter believes that the ‘Peter’ involved in this belief is his own name. However, suppose that now Peter has another belief that ‘I feel good’, this belief contrasts with the previous belief sharply as this belief is not only about Peter the person, but also about Peter himself as himself. So, in a nutshell, the characterization ‘about oneself as oneself’ can be regarded as being composed of two parts: the ‘about oneself’ part indicates the object about which the *de se* belief should be, and the ‘as oneself’ part indicates the proper form of *de se* belief, which is the form of first person belief. I hope those examples help to clarify what I refer to as *de se* belief. By ‘first person sentence’, I mean sentences which contain the first person indexical terms like ‘I’, ‘my’, ‘me’ or ‘mine’.

well.

25. A wider conception of *de se* belief can possibly include the communication of ‘you’-belief by means of second person proposition. For instance, in his paper ‘Why My *I* Is Your *You*’, Emar Maier asserts ‘I focus on the speaker’s expression of first and second person *de se* beliefs, and the hearer’s subsequent interpretation of utterances containing *I* and *you*.’ (Emar Maier, 2015, pp.236-237). This wider conception of *de se* communication is relevant to the subsequent discussions in my thesis, as ‘I-You’ translation is an important topic of my argument. But I will not discuss this wider conception in detail here.

1.2.1 The Idiosyncrasy of De Se Communication

The aim of this section is to explain the necessity of idiosyncrasy of de se communication by appealing to the combination of the overall structure of de se communication and some semantic features of the first person indexical. By 'idiosyncrasy of de se communication', I mean the generally necessary applicability of Translation Transmission Model to the de se communication, rather than 'Replication Transmission Model'. In communications of beliefs which do not involve indexical elements, the audience just need to replicate the speaker's belief, and in belief communications involving indexical elements other than the first person indexical, the hearer sometimes needs to translate the speaker's belief in order to accept it, although sometimes she does not have to do so. However, in the case of the de se belief communications, the hearer always has to translate the speaker's belief, otherwise the communication fails. The idiosyncrasy of de se belief communication has evoked numerous discussions on the nature of the content of de se belief and the content of belief in general, and also discussions about the shareability and/or the un-shareability of de se belief.

In the following sections, I aim to clarify the key problem with which I will deal in the rest of this chapter. My strategy for clarification is a three-step strategy. First, I introduce the traditional conception of the general structure of communication. Then I argue that many philosophers currently agree that there is a translation transmission model based on which a de se belief is transmitted from one participant to other participants of a de se communication by means of a certain type of translation. The third step, which is the most important one, is to point out there are at least two questions entangled in the discussion of the translation transmission model and they should be separated from each other carefully. The first of the two questions is how the translation transmission is operated, and the second is why the translation transmission operation is necessary. Let us begin by sketching the outline of the traditional conception of communication.

1.2.2 Presupposition and Co-aboutness

The widely accepted tradition conception of communication, which is also known as the Naive Conception of Communication.²⁶, was proposed by Stalnaker in his 1978 article '*Assertion*':

In "Assertion" Robert Stalnaker provides a model of communication that has been widely accepted within philosophy of language. (...) Stalnaker's account for how the content of assertion affects the context relies heavily on the notion of presupposition. (...) Stalnaker's proposal for how the content of an assertion affects the context is that successful assertion involves adding propositions to the stock of propositions presupposed by the conversational participants. (Torre, 2015, p. 12)

I roughly agree with Stalnaker on his idea that the common presuppositions of the participants of a communication grounds the possibility of that communication being successfully operated, because it is almost self-evident that in order to successfully communicate with each other, participants of a communication must share some consensuses from the very beginning, and what must be included in those consensuses is a common focus, i.e. they must talk about the same thing. Also, I agree with him that the enlargement of the common presuppositions can be taken as a standard to evaluate the success of a relevant communication, which also implies that if the common presuppositions do not enlarge at all, or even decreases, then the relevant communication should be viewed as having failed. However, as to the detailed definition of 'presupposition', I do not entirely agree with Stalnaker, according to whom:

²⁶ Please refer to section 1.1.3.

Roughly speaking, the presuppositions of a speaker are the propositions whose truth he takes for granted as part of the background of the conversation. (Stalnaker, 1999, p. 84)

And

Presuppositions are what is taken by the speaker to be the common ground of the participants in the conversation, which is treated as their common knowledge or mutual knowledge. (Stalnaker, 1999, p.84)

And also

In the simple picture, the common ground is just common or mutual belief, and what a speaker presupposes is what she believes to be common or mutual belief. The common beliefs of the parties to a conversation are the beliefs they share, and that they recognize that they share: a proposition ϕ is common belief of a group of believers if and only if all in the group believe that ϕ , all believe that all believe it, all believe that all believe that all believe it, etc. (Stalnaker, 2002, p.704)

If I take Stalnaker's three definitions of 'presupposition' to be coherent with each other, which I think is a very reasonable move, I am then led to the conclusion that only propositional presuppositions can be commonly presupposed by all the participants of a conversation. I am not satisfied with this conclusion. In Section 3.2, I will explore the possibility that something, which is not propositional, needs to be included in the common ground for the sake of making a communication successful.²⁷ That new element conforms to Stalnaker's second definition, but not his first or third one. For my current purpose, Stalnaker's second definition is more relevant and useful than his first and third definitions. So, I will use the term 'common ground' which appears in Stalnaker's second definition, rather than the term 'common

²⁷ Although that is similar to the common knowledge/mutual knowledge as Stalnaker mentions in the second and third quotations.

presuppositions’, in the rest of this chapter.

In the previous sections, when discussing Ninan’s and Recanati’s proposals on what the real problem of the *de se* attitude is, I mentioned the concept of ‘co-aboutness’. This is a very flexible concept, that can be applied to indicate that two names refer to the same referent, or that two propositions share the same truth condition, or two thoughts keep track of the same thing. In the case of communication, I deploy the concept of ‘co-aboutness’ to capture an essential requirement for a successful communication, which is that participants of a successful communication must talk about the same topic. If participants of a communication do not share a common topic, they either fail to communicate with each other or they simply talk past each other. In the traditional conception of communication, the fulfillment of ‘co-aboutness’ requirement is closely related to the common ground. But how can we make sure that in a certain communication co-aboutness is successfully included in the common ground, or to put it in another way, how can participants of a certain communication be assured that they are not talking past each other, according to the traditional conception of communication? To answer this question, we need to move on to investigate how the content of a communication is delivered between its different participants.

1.2.3 Replication Transmission Model

Following the traditional picture of communication, I agree that successful communication means adding new propositions to the set of propositions presupposed by the conversational participants. It can be easily discerned that in this process the idea of ‘belief transmission’ is involved:

This model of assertion involves the transmission of belief from speaker to hearer....I follow Dirk Kinderman in calling this model the “Simple Transfer

Model” of communication. Following Clas Weber, we can take the Simple Transfer Model of communication to involve commitment to the following two principles:

The Mind-to-Speech Principle: the content of the utterance is identical to the content of a belief the speaker expresses.

The Speech-to-Mind Principle: the content of the utterance is identical to the content of a belief the hearer acquires. (Torre, 2015, pp. 12-13)

Here we can find that the model of the belief transmission equipped in the traditional conception of communication is very simple: the content of the speaker’s belief, the content of the speaker’s utterance and the content of the belief that the hearer acquires from hearing the speaker’s utterance are all identical, which means that there is only a single content which undergoes all those three stages of transmission and remains identical during the whole process of communication. This is why Stephan Torre calls this model the ‘simple transfer model’. Recanati has a more condensed characterization of this model:

The Naïve Conception of Communication rests on the idea that communication is the *replication* of thoughts: the thought the hearer entertains when he understands what the speaker is saying is the very thought which the speaker expressed. (Recanati, 2015, pp. 149-150)

By the phrase ‘Naïve Conception of Communication’ Recanati means the same thing that I mean by ‘traditional picture of communication’. Recanati’s characterization of the model of belief transmission in traditional picture of communication is more advanced than Torre’s characterization because it not only points out the sameness of what is transmitted (i.e. belief) during the process of transmission, but also (more importantly) captures the way in which the transmission is done: according to Recanati’s characterization, in the traditional picture of communication, a belief is transmitted from the speaker to her hearer(s) by means of ‘replication’. Therefore, I

call this model of belief transmission the replication transmission model. The next task I face is to specify what is replicated in the replication transmission model. Intuitively, the answer to this question can be that it is the thought, or the content of the speaker's mental (or propositional) attitude, that is replicated during a communication. However, as I noted at the beginning of this chapter, the focus of my study is the role that the (first person) linguistic expression plays in de se communication. This focus reminds us of an often neglected fact: in (verbal) communications, the primordial vehicle deployed by people is language, not thought. Therefore, what is replicated in the replication transmission model, first and foremost, is the linguistic expression. Since not only the thought, but also the language that is employed to express thought, are replicated, we are now in front of an important question: what is the relation between the replication of language and the replication of thought in a communication? My answer is that for replication transmission model, the replication of language is more important than the replication of thought. In Section B of the Introduction, I first introduced the connection between same-saying and co-aboutness, I quoted an important paragraph from Laura Schroeter, who argues:

A primary semantic function of proper names, common nouns or predicates is to ensure that all parties to a conversation pick out the very same subject matter with these words. Sameness of meaning must somehow guarantee sameness of subject matter – and it must do so in a way that is accessible to ordinary speakers. (Schroeter, 2012, p. 177)

It should be noted that Schroeter here is talking about the 'semantic function' of linguistic expressions, not their 'cognitive function', and she also claims that it is the 'sameness of meaning', not the 'sameness of thought', which is supposed to guarantee the sameness of subject matter. I agree with Schroeter on this point. I believe that replication of linguistic expression implies, and also makes explicit the replication of thought in the traditional picture of communication. For participants of a communication which is compatible with the traditional picture of communication, their thinking about the same subject matter is implied and manifested by their talking

about the same subject matter, and their talking about the same subject matter is guaranteed by their replications of linguistic expressions, i.e. they use the same words/sentences. Since we human beings have no capacity to read other people's minds directly, we cannot confirm that we are thinking about the subject matter directly. We have to adopt an indirect strategy: we confirm that we are thinking about the subject matter by confirming that we are replicating one another's linguistic expressions. I call the replication of language the 'same-saying', and I call the replication of thought the 'same-thinking'. Based on my characterizations of the relation between same-saying and same-thinking, I think it is safe to claim that same-saying plays a more important role than same-thinking with regard to guaranteeing the convergence of subject matter. In this section, I focus on articulating the guarantee relation between same-saying and the fulfillment of the co-aboutness condition, and in subsequent sections, I will articulate the incompatibility between same-saying and co-aboutness in de se communication. However, I will not discuss same-thinking specifically in this chapter. This is because, first, as I have argued, same-thinking is not as important as same-saying with regard to the guarantee of co-aboutness, and second, I believe some parts of my discussions on same-saying can also be applied to same-thinking, therefore same-thinking has been covered by my discussions on same-saying.

So far, we have a sketch of how the traditional picture of communication is constructed. It consists of two important elements: the common ground and the replication transmission model. The former includes the co-aboutness requirement that participants of a communication must share a common subject matter, as they do not want to talk past each other, and the latter explains how the content of a communication is delivered between its different participants. I call the former 'co-aboutness' element and the latter 'same-saying' element. Therefore, the traditional picture of communication can be formulated as:

Traditional Picture of Communication (version 1) = Common ground (co-aboutness)
+ Replication Transmission Model (same-saying).

However, this formulation is not detailed enough, because the relation between the co-aboutness and the same-saying has not been clarified in it. A proper characterization of that relation is illuminated by the paragraph from Laura Schroeter that I quoted above. I take Schroeter's point. What she means by the phrase 'sameness of subject matter' is the same as what I mean by 'co-aboutness'. Although there is a slight difference between what she means by 'sameness of meaning' and what I mean by 'same-saying', this difference does not obstruct my appropriation of her idea. The relation between the two elements of the traditional conception of communication can be characterized as a guarantee relation: Replication Transmission Model guarantees Common Ground. Therefore, we have a revised version of the formulation:

Traditional Picture of Communication (version 2) = Replication Transmission Model (same-saying) → Common Ground (co-aboutness)

in which the '→' stands for a guarantee relation. What is more, based on the second version of the formulation of the traditional picture of communication, I propose a criterion which can be used to examine the validity and effectiveness of the replication transmission model, or to put it in another way, to examine whether participants following the replication transmission model in a communication leads to an enlargement of their common ground or not. I call this criterion the Repetition Principle:

Repetition Principle: a belief is successfully transmitted through the replication transmission model if and only if the hearer(s) can express the same belief as she acquires from the speaker simply by repeating the speaker's utterance.

I think the Repetition Principle is quite plausible because it is in fact a successive multiple application of the replication transmission model. Let us suppose there are several successive communications between multiple participants A, B, C and D. First, A delivers a message to B, and then B delivers the message she gets from A to C, and so on. In this scenario, B (and also C) is both a hearer and a speaker. According to the guarantee relation between same-saying and co-aboutness, when communicating

with C, B can simply repeat A's utterance if B wants to make sure that what C receives from B is the same as what B receives from A. If what C receives is identical with what A expresses, the replication transmission model works smoothly in that communication. This is an instance of what is indicated in the repetition principle. One thing which is worthy emphasizing here is that although in my formulation of Repetition Principle I mention the utterance only, this does not imply that only the utterance is repeated according to Repetition Principle. Just like the same-saying / same-thinking entanglement in the replication transmission model, what is repeated according to the repetition principle is not limited to the linguistic instrument for communication, but also what is communicated, i.e. the thought delivered during a communication. Now, if the replication transmission model is as universally valid, as those proponents of the traditional picture of communication claim it to be, it should be able to be adopted in the explanation of every communication that happens in the world. This, in turn, implies that every communication that happens in the world should be able to pass the examination of the repetition principle. However, that is simply not the case. I find a notable exception.

1.2.4 Translation Transmission Model

Imagine a scenario like this:

One day Peter met John on the street and greeted John, 'Hi, John! It is sunny today!' John replied, 'Yes, it is sunny today'. Then Peter said, 'I feel good!' and John replied, 'I feel good as well!'

In their conversation, John repeats Peter's utterances twice:

The first repetition:	Peter 'It is sunny today.'	John 'It is sunny today.'
The second repetition:	Peter 'I feel good.'	John 'I feel good.'

I argue that the second repetition differs from the first repetition at an important aspect. In the first repetition, the truth conditions of the two sentences are the same. The two sentences are both true if and only if the weather of the area where Peter meets John is sunny or not when they meet. That means Peter and John express the same belief about the same thing, and they manage to do so simply by repeating each other's utterances (in this scenario it was John who did the repetition). However, in the second repetition, the truth condition of Peter's sentence is apparently different from the truth condition of John's sentence. Peter's sentence is true if and only if Peter felt good and John's sentence is true if and only if John felt good. That means when John repeats Peter's utterance, he expresses a different belief about a different object compared with what Peter expresses. In addition, I think it is obvious that if John wanted to express a similar belief about the same object when he hears Peter's utterance 'I feel good', he should say 'you feel good' rather than repeating 'I feel good'. What John did in the first and the second repetitions are operations of the same kind (both repeating), but the consequences of his operations are entirely different: the former smoothly *passes* the examination of the repetition principle, but the latter does not. The collision between the second repetition and the repetition principle reveals that the replication transmission model is not universally valid and effective: there is a certain type of communication which does not conform to the replication transmission model, and therefore cannot be included into the traditional picture of communication. This exceptional type of communication is *de se* communication.

Although *de se* communication does not conform to the replication transmission model, this does not imply that it is also incompatible with the other component of the traditional conception of communication, i.e. the common presupposition as a realization of co-aboutness. On the contrary, the common presupposition still gets involved in the *de se* communication: if Peter tells John that 'I am feeling good' and John replies 'you are feeling good', it is clear that they are talking about the same topic, and they reach a consensus which can be a part of the common presupposition without any doubt for their relevant communications in the future. The real question

unanswered here is how the consensus is reached, or to put it another way, how the co-aboutness, between a *de se* belief (I am feeling good) and a non-*de-se* belief (You are feeling good), is reached in this conversation. In the traditional picture of communication, the co-aboutness is guaranteed by the same-saying, as I discussed in the last section. But in the conversation that we are currently discussing, there is apparently no same-saying at all: Peter said: 'I am feeling good' but John replied: 'You are feeling good'. The co-aboutness lost its guarantee there. To solve this problem, I think, first we need to find out how the *de se* belief is transmitted from the speaker to the hearer and then gets added into the common presupposition, i.e. we need a new transmission model for *de se* belief, which is different from the replication transmission model. So, what is this new transmission model?

When Peter told John that 'I am feeling good' and John replied 'Oh, you are feeling good', not only can it be discovered that a belief was transmitted from Peter to John, but we can also notice that John translated Peter's utterance, rather than repeating it: he substituted his 'You' for Peter's 'I'. The translation of utterance reflects the transformation of belief expressed by the relevant utterance. This is why Recanati asserts that:

But in the indexical case (i.e. indexical communication), there is no replication, but some kind of systematic transformation. The truth-conditions can be preserved across contexts only if the internal content of the thought is modified or adjusted so as to compensate for the shift in point of view from the speaker to the hearer. It follows that the Naive Conception is false. (Recanati, 2016, p. 150)

What Recanati means by the phrases 'systematic transformation' and 'modified or adjusted' is roughly the same as what I mean by 'translation'. In the case of *de se* belief communication, beliefs are transmitted by means of a certain type of translation. Here translation plays the same role, which is the specific way of transmission, as replication plays in cases that can be embedded into the traditional picture of communication. If

we replace the ‘replication’ in the replication transmission model with the ‘translation’, we can get the new belief transmission model that we need, which I call the translation transmission model.

Now it seems that the idiosyncrasy of *de se* belief communication can be characterized as adopting the translation transmission model, rather than the replication transmission model. However, I do not think that this is a comprehensive characterization. Readers may have already found that in the paragraph cited above when Recanati talks about the ‘systematic transformation’, he is talking about it in the context of indexical communications in general. But what I am talking about up till now is a ‘translation’ taken place within the context of *de se* communication. The reason why I limit my discussion of the translation transmission model specifically to *de se* communication is that I believe the modalities of applicability of translation transmission model in *de se* cases and non-first-person indexical cases are different. The translation transmission model necessarily applies to *de se* communication, but not necessarily to indexical communications other than *de se* communication. In the scenario described above, Peter and John have two exchanges of belief, which are both about indexical beliefs. The difference between the two exchanges is that the first one is about a non-*de-se* indexical belief (It is sunny today), while the second is about a *de se* belief (I feel good). The first exchange conforms to the traditional picture of communication, reaching a consensus (co-aboutness) by means of the replication transmission model (same-saying). This is a typical example to show that indexical communication rather than *de se* communication does not necessarily involve the translation transmission model (this point is further elaborated in section 1.3.2). In short, the distinctions I try to draw can be shown in a table form as listed below:

Non-indexical communication	The replication transmission model necessarily works.
Indexical communication other than <i>de se</i>	The replication transmission model

se communication	works on some occasions, and the translation transmission model works on other occasions.
De se communication	The translation transmission model necessarily works.

Therefore, a comprehensive characterization of the idiosyncrasy of de se belief communication must exhibit the modality of applicability of translation transmission model, which is that the translation transmission model necessarily applies to de se belief communication.

1.3 Characterizations of Translation Transmission Model

I have shown in the last section that the distinctive feature of the translation transmission model is that it includes in itself an operation of translation. But this is still a very vague description, especially when compared with the characterization of two principles ('Mind-to-Speech' and 'Speech-to-Mind') in the replication transmission model mentioned in section 1.1.2. In order to get a more detailed characterization of the translation transmission model, I think there are at least three questions that need to be answered. The first question is: what is translated in the translation transmission model? The second question is: where and when does the translation operation take place? The third question is: how the translation is operated? There have already been several answers to those three questions in current literature, but it seems that philosophers have no consensus in their discussions of indexical communication. In this section I discuss three proposals to explain indexical communication. Each proposal contains its own answers to those questions mentioned above. Since one of my major tasks in this chapter is to answer the question

‘what is my characterization of the translation transmission model in *de se* communication?’. This is different from discussions about indexical communication in general. I will show at the end of this section that a scrutiny of the three proposals offers important insight in answering my major question.

1.3.1 Frege on Private Sense

In his famous essay *Thought*, Frege discovers a puzzling problem which bothers him. The problem is about the alleged ‘I-thought’:

Now everyone is presented to himself in a particular and primitive way, in which he is presented to no one else. So, when Dr. Lauben thinks that he has been wounded, he will probably take as a basis this primitive way in which he is presented to himself. And only Dr. Lauben himself can grasp thoughts determined in this way. But now he may want to communicate with others. He cannot communicate a thought which he alone can grasp. Therefore, if he now says ‘I have been wounded’, he must use the ‘I’ in a sense which can be grasped by others, perhaps in the sense of ‘he who is speaking to you at this moment’. By doing this he makes the conditions accompanying his utterance serve towards the expression of a thought. Yet there is a doubt. Is it at all the same thought which first that man expresses and now this one? (Frege, 1997, pp. 333-334)

Perhaps this was the first time in the history of analytic philosophy that the problem of *de se* attitude was raised. Also, it should be noted that from the very beginning of his discussion, Frege embeds the problem of *de se* attitude into a larger context of *de se* communication. Perhaps that is because he had an intuitive insight that only within the context of communication can the difference between *de se* attitude and other types of (indexical or non-indexical) be properly revealed,

although he did not give us a discursive articulation of that insight. As widely acknowledged, communication essentially involves expression, of which the most important part is the linguistic expression. So, the importance of embedding the discussion of *de se* attitude into a context of *de se* communication lies in the fact that once a belief is communicated, rather than simply being kept in the speaker's mind, the belief is immediately connected to its linguistic expression, which allows us to know what the speaker believes: it is possible to grasp the content of a belief by means of grasping the meaning of its linguistic expression. This is the approach Frege takes in the cited paragraph.

It is widely known that according to the Fregean semantic framework, each linguistic expression is associated with a sense by means of which something is connected to the expression as its referent. But Frege admits in the paragraph cited above that it is difficult to apply this semantic framework to the expression of *de se* attitude, because in the case of the first person sentence, which is usually adopted to express *de se* attitude, it seems there are at least two different senses associated with the same first person sentence, each of which is in accordance with a specific situation. When Dr. Lauben keeps the belief 'I have been wounded' in mind, the referent of the term 'I', which is Dr. Lauben himself, only needs to be 'presented to himself in a particular and primitive way', therefore the sense associated with the term 'I' here is a particular, primitive and (most importantly) private presentation of Dr. Lauben. However, this is a sense which only he (Dr. Lauben) can grasp and can be presented to no one else. Therefore, if he wants to convey this belief to somebody else, there must be another different sense associated with the indexical term 'I' that is available to him, and he can use the sense in his expression 'I have been wounded': this has to be a sense which can be grasped by others, i.e. a public and shareable sense. The option Frege provides for the public sense of 'I' in Dr. Lauben's case is 'he who is speaking to you at this moment'.

However, it is important to note that according to Frege, when being kept in the speaker's mind, the first person sentence has no public sense, merely a private one.

A public sense is to be associated with a first person sentence when and only when the person who keeps that belief in mind tries to communicate that belief with other people using linguistic expressions. Given that one linguistic expression can have only one sense at a time according to Frege, it seems that the first person sentence must drop its private sense when being ascribed the public sense. So, an operation of substituting the public sense for the private sense is found here, and I think this substitution operation can be seen as a counterpart of the translation operation in the translation transmission model. Now in the Fregean proposal we find answers for the three questions I raised at the beginning of this section:

Question 1: what is translated in a de se communication?

Fregean Answer: it is the sense of first person sentence that gets translated.

Question 2: where and when does the translation operation take place?

Fregean Answer: the operation takes place in the speaker's mind when she tries to communicate her de se belief with someone else.

Question 3: how is the translation operation done?

Fregean Answer: the operation is done by substituting a public sense of a first person sentence for its private sense.

A criticism that may quickly come to readers' minds is that the translation operation is misplaced in Frege's proposal. In the translation transmission model, the translation operation takes place in the hearer's mind: it is those hearers' responsibility to translate what they get from the speaker's utterance in order to successfully conduct a de se communication. However, according to Frege's proposal, the translation operation is done by the speaker. This does not conform to our intuition on this issue. What makes the situation worse is that even after the speaker completes the translation operation, with the directions she gets from the Fregean proposal, the hearer may still need to do another translation, for example, to translate the speaker's first person expression carrying the public sense into the

hearer's second person expression with the term 'you'. Because of this second translation, the Fregean proposal makes the translation operation unnecessarily complex and redundant. Another famous criticism of Frege's proposal comes from John Perry. This criticism focuses on the feature of the alleged 'limited accessibility' of the private sense in Frege's proposal. Perry argues that

If one finds plausible reasons elsewhere for believing in a universe that has, in addition to our common world, myriad of private perspectives, the idea of propositions of limited accessibility will fit right in. I have no knockdown argument against propositions, or the metaphysical schemes that find room for them. But I believe only in a common actual world. (Perry, 2000, p. 38)

This criticism is in a sense powerful because most philosophers today believe only in a common actual world, just as Perry does. But I think his criticism missed a difference, which is not trivial, between Frege's proposal and Perry's own proposal. If we compare those two proposals, it is not hard to discern that the scopes of their proposals are different: Perry takes the indexical attitude in general as the target of his discussion, but Frege picks out the first person indexical from various kind of indexicals and limits his discussion to de se communication only. One possible (and also reasonable as I understand it) explanation for Frege's choice is that the distinction between the private sense and the public sense can be applied properly only to the first person indexical, but not other types of indexicals. If this were true, perhaps Frege already had some awareness about the difference between the modality of applicability of translation transmission model to de se communication (which is necessity) and the modality of applicability of translation transmission model to other types of indexical belief communication (which is possibility). It is also noticeable that the difference between the scope of Perry's proposal and the scope of Frege's proposal is similar to the difference between Recanati's view on the applicability of translation transmission model and my view. This is not contingent. At the end of section 1.3.2, I will try to elucidate an important difference between the first person indexical and other kinds of indexical by appealing to Perry's

proposal.

1.3.2 Perry on Structured Proposition

In Section 1.2.4, I mentioned a scenario, in which Peter said to John 'I feel good'. It is obvious that if John wants to show that he understands Peter's utterance, he should say 'You feel good' rather than repeating 'I feel good'. But 'you feel good' is apparently different from 'I feel good'. How could Peter and John reach a consensus while communicating with each other using different sentences? Or, to put it in a more specific way, how can those two sentences reach co-aboutness when they are not same-saying? John Perry offers a possible answer to this question in his 1979 article 'The Problem of the Essential Indexical', explaining how two sentences containing different indexical elements can be used to express the same belief.

For Perry, the content of belief attitude in general is a standard proposition *simpliciter*, whose truth value is absolute, i.e. it does not vary from one context to another. But the indexical belief is of a special type. The content of the indexical belief attitude, according to Perry, is the structured proposition which is composed of two elements. The first element is what Perry calls the belief state, and the second element is what Perry calls the object of belief. The two elements are relatively independent of each other: we can have indexical beliefs about different objects of belief but share the same belief state, and we can also have indexical beliefs which are in different belief states but about the same object of belief. Perry asserts that:

The proposal, then, is that there is not an identity, or even an isomorphic correspondence, but only a systematic relationship between the belief states one is in and what one thereby believes. (Perry, 2000, p. 40)

So, the next question is: what are the belief state and object of belief? For Perry,

the object of belief is the standard proposition (i.e. proposition with a determined truth value) as the tradition picture of belief shows. The belief state, according to Perry, is the way that a believer in a context is situated to handle the relevant object of belief. Belief state is individuated according to the indexical expression that the believer uses when expressing her indexical belief. Perry writes that:

We use sentences with indexicals or related propositions to individuate belief states, for the purposes of classifying believers in ways useful for explanation and prediction. (Perry, 2000, p. 39)

This is a crucial point which is worthy of emphasis. Readers might have already noticed that, unlike Frege, Perry does not embed his discussion of indexical attitudes into a context of (verbal) communication. So, the connection between the indexical belief and the expression of indexical belief can hardly be constituted by appealing to the context of communication. This apparently differentiates Perry's approach from that of Frege. However, this does not imply that in Perry's proposal there can be no connection between the indexical belief and its linguistic expressions. Perry provides another way to build up this connection: if the belief state of an indexical belief is individuated according to the indexical expression that the speaker uses to express that indexical belief, as Perry claims in the quotation, the individuation relation then connects belief to its linguistic expression. For example, the standard proposition 'It is cloudy on November 14, 2015' can be an object of belief under the belief state 'It is sunny *today*' for a person who holds that belief on November 14, 2015. Accordingly, in that conversation scenario involving Peter and John, both Peter and John share a common object of belief, which is 'Peter feels good', but the ways by means of which they hold this object of belief are different: Peter keeps it under the belief state that 'I feel good', using the indexical expression 'I', but John holds it under the belief state that 'you feel good', using the indexical expression 'you'. At this moment, we encounter with Perry's answer to the question mentioned at the beginning of this section: how can those two sentences reach co-aboutness when they are not same-saying? Perry's answer is that the object of belief is what is

responsible for reaching co-aboutness in an indexical communication, and the belief state, with its relevant indexical expression, is what is responsible for reaching same-saying, therefore, two beliefs can reach co-aboutness without reaching same-saying as long as they keep the same objects of belief under different belief state.

It should be noted that no problem of limited accessibility is involved here. Unlike Frege's proposal, Perry's proposal does not rely on a 'myriad of private perspective' that there is something which is specifically private to someone. The *de se* belief state 'I feel good' is not a belief state that can only be obtained by Peter. On the contrary, that belief state can also be possessed by John. But of course, when John possesses the belief state 'I feel good', his object of belief is 'John feels good' rather than 'Peter feels good'. A similar situation takes place when the object of belief is concerned. In Perry's proposal, neither belief state nor object of belief is 'private', therefore neither of them is bothered by the alleged 'problem of limited accessibility'.

Now let us suppose the conversation described above leads to a successful communication, and John forms a new belief based on his acceptance of Peter's utterance that 'I feel good', then it is clear what John did when he acquired the belief is that he abandoned Peter's belief state (the *de se* belief state) but at the same time preserved Peter's object of belief (the standard proposition 'Peter feels good'), and then subsumed this object of belief into his own belief state (the second person belief state). I think this operation is equivalent to the translation operation in my Translation Transmission Model. So here we find answers for the two questions I raised at the beginning of this section in Perry's proposal:

Question 1: what is translated in an indexical communication?

Perry's answer: the belief state is what has been translated, not the object of belief.

Question 2: where and when does the translating operation take place?

Perry's Answer: the operation takes place in the hearer's mind when she tries to acquire a belief based on her acceptance of the speaker's expression.

Question 3: how is the translating operation done?

Perry's Answer: the operation is done by preserving the object of belief while substituting the hearer's belief state for the speaker's belief state.

It is easy to discern that unlike Frege's answer to the second question, Perry's answer to the second question conforms to our intuition on his question. But it may be unclear whether the translation operation is necessary for the hearer or not, because in his proposal Perry talks about the indexical communication in general, not a specific type of communication. As I argued in previous sections, not all types of indexical communication necessarily involve the Translation Transmission Model. At the end of the section 1.3.1, I mentioned a difference between Perry's and Frege's proposal, in that Perry's proposal of the structured proposition can be applied to various kinds of indexical beliefs, while Frege's proposal only applies to *de se* belief. Now since we have some knowledge of Perry's proposal, it is time for me to explain the difference in detail. Let us imagine another scenario in which Peter made a phone call to John and on the phone, he told John 'It is cloudy today' and John replied 'Oh, yes, it is cloudy today'. In this phone call scenario, if both Peter and John were in Melbourne, which implies that they were on the same date (suppose it was November 14, 2015), both Peter's and John's belief are the same: they share not only the same object of belief ('It is cloudy on November 14, 2015') but also the same the belief state ('It is cloudy today'). Therefore, it should be noted that in the phone call scenario, what takes place is an indexical communication, but no operation of translation is involved in it, and what gets involved is only replication: Peter and John use the same linguistic expression to express the belief, therefore rather than translation, John just replicates Peter's belief. However, if, when making the phone call, Peter was in Melbourne, and John was in New York, which implies that they were on different date (suppose it was November 14, 2015, for Peter, and

it was November 13, 2015, for John), their beliefs share the same belief state ('It is cloudy today') but they do not share the same object of belief. Therefore, in their conversation, Peter and John simply talk past each other, and there is no real communication in that conversation. To avoid talking past each other, Peter and John must adjust their belief states to make sure their objects of belief are the same. The adjustment of belief state amounts to a translation operation.

In section 1.2.4, I mention a difference between the de se communication and other types of indexical communication. The difference is that the translation transmission model necessarily applies to de se communication, but it only possibly applies to other types of indexical communication. Now the phone call scenario provides evidence for the second clause of this assertion: it is possible for non-de-se indexical communications to adopt the replication transmission model for the transmission of attitudes. In Perry's terms, the difference between de se belief communication and other types of indexical belief communication can be formulated in this way. For any attitude other than the de se attitude, it is possible for two agents to hold the same object of belief under the same belief state and it is also possible for them to hold different objects of belief under the same belief state. But if two agents both adopt the first person belief state, it is impossible for them to share the same objects of belief.

I argued in Section A of the Introduction that Perry's proposal does not offer a satisfactory response to the challenge raised by indexical belief communication because according to Perry's proposal what is communicated is the non-perspectival object of belief, which is a standard proposition, but the perspectival belief state is not included in what is communicated. Perry's proposal destroys the perspectivity of indexical communication, so it can hardly explain how we can acquire a perspectival, indexical belief from an indexical belief communication. But this does not imply that Perry's proposal is without merits. One of its merits, as I articulate in this section, is that it can help us distinguish the de se communication from other types of indexical communication. The distinction between de se communication and other types of

indexical communication is important for my subsequent argument.

1.3.3 Recanati on Systematic Transformation

Between all the three proposals discussed in section 1.3, Recanati's proposal is the only one that explicitly proposes the idea that in conducting a de se communication, we perform a translation operation, rather than a replication operation (see section 1.1.3). I agree with Recanati on this point. However, I am not satisfied with Recanati's detailed characterizations of the translation in de se communication with regard to a number of aspects. In this section, I will discuss the three reservations I have with Recanati's proposal.

First, I am not satisfied with Recanati's argument on the rejection of 'thought expressed by utterance'. He argues that in order to reject the replication transmission model, the idea of 'thought expressed by utterance' must first be rejected. In section 1.2.3, when characterizing the replication transmission model, I introduced two principles based on a quotation from Stephan Torre. Those two principles are the mind to speech principle and the speech to mind principle:

We can take the Simple Transfer Model of communication to involve commitment to the following two principles:

The Mind-to-Speech Principle: the content of the utterance is identical to the content of a belief the speaker expresses.

The Speech-to-Mind Principle: the content of the utterance is identical to the content of a belief the hearer acquires. (Torre, 2015, pp. 12-13)

It is clear that two principles must be either modified or rejected if we want to replace the replication transmission model with the translation transmission model. Recanati's proposal is to reject the idea of 'the content of the utterance', which occurs in both

principles and acts as a bridge which connects the speaker's belief to the hearer's belief. He claims that:

We have to give up the very notion of 'the thought expressed by the utterance'. There is no such thing: there is the speaker's thought and the thought formed by the hearer if communication is successful (Bezuidenhout 1997). That is all there is as far as thoughts are concerned. We don't need an additional entity, the thought expressed by the utterance. (Recanati, 2016, p. 154)

This seems to be a simple and easy solution, because once the 'content of the utterance' was rejected, there would be nothing to bridge the speaker's belief and the hearer's belief, so that the replication transmission model fails. I think this solution is problematic, because it is in contradiction with a basic intuition which is always at work in our verbal communication. This intuition is that I acquire certain pieces of information from whom I communicate with, based on my understanding of her words. As I mentioned in section 1.2.3, I, as a human being have no direct access to those mental states in other people's minds. The only way I can get access to mental states in other people's minds is to appeal to some medium. In verbal communication, the medium is linguistic expression. If the idea of 'the content expressed by the utterance' as a whole was rejected, it would be very tricky for us to explain the role that linguistic expressions play in communication in general, let alone the de se communication in particular. This point is closely related to my second disagreement with Recanati's proposal.

My second dissatisfaction is that I think the status of language in Recanati's proposal on the explanation of de se communication is puzzling: it seems that he rejects the role which language plays when he claims that the 'content of the utterance' must be abandoned, but then he introduces language back into his picture when he argues for the co-ordination role of linguistic mode of presentation. Recanati first makes a distinction between two kinds of 'mode of presentation': the linguistic mode of presentation and the psychological mode of presentation:

Linguistic modes of presentation are whatever information is linguistically encoded about the reference. ‘I’ encodes the information that the referent is the speaker, ‘you’ that it is the hearer, and so on and so forth. Psychological modes of presentation are the ways the subject thinks of the reference. (...) Psychological modes of presentation, in my framework, are mental files. (Recanati, 2016, p. 153)

In a de se communication, according to Recanati’s proposal, when a speaker expresses one of her de se attitudes using the first person linguistic expression, the linguistic meaning of that expression, or to put it in Recanati’s terms, the linguistic mode of presentation associated with that expression (I think Recanati has the Fregean sense in mind when he deploys the phrase ‘mode of presentation’) is the same for both the speaker and the hearer. However, the psychological mode of presentation associated with the speaker’s de se attitude is different from the psychological mode of presentation associated with the hearer’s relevant attitude:

When the speaker says ‘I’, she thinks of herself in the first person way, but the hearer who understands her thinks of her in a different way. The interpersonal variability of psychological modes of presentation establishes that linguistic and psychological modes of presentation are distinct, for the linguistic mode of presentation is constant: ‘I’ means the same thing for the speaker and the hearer. (...) So the linguistic mode of presentation corresponds to a piece of information that is common to both the speaker’s and the hearer’s files about the speaker. (Recanati, 2016, pp. 153-154)

Now it is clear that since the speaker’s psychological mode of presentation is different from the hearer’s psychological mode of presentation, the first person term’s linguistic mode of presentation cannot be identical with both of them at the same time. However, this does not automatically lead to Recanati’s conclusion that there is not ‘content of the utterance’, or even if there is ‘content of the utterance’, it does not play a role in the explanation of de se communication. It should be noted that in de se

communication the co-aboutness requirement is still at work, which means that participants of a de se communication are still required to talk about the same subject matter, otherwise they simply talk past each other and their communication fails. Therefore, although the speaker's psychological mode of presentation is different from the hearer's psychological mode of presentation, what is presented through their modes of presentation must still be identical. To put it in Recanati's terms, although the speaker and the hearer adopt different mental files in a de se communication, the object which their mental files are about must be the same. How is this possible? Recanati appeals to the co-ordination role that linguistic mode of presentation plays:

So the linguistic mode of presentation corresponds to a piece of information that is common to both the speaker's and the hearer's files about the speaker. These files are distinct, and they are of different types, but they are constrained by the linguistic mode of presentation. The linguistic constraint on files is a means for coordinating them. On this interpretation the linguistic modes of presentation are not ersatz ingredients of thought. They are not thought constituents, but they are suitably related to the relevant thought constituents. The thought constituents are the files—the psychological modes of presentation—which the linguistic mode of presentation constrains. (Recanati, 2016, p. 154)

In this paragraph, Recanati claims that it is by means of the linguistic mode of presentation that the hearer's thought is properly related to the speaker's thought. However, it should be noted that when rejecting the 'thought expressed by the utterance', Recanati argues that in a communication there are only the speaker's thought and the hearer's thought, and 'we do not need an additional entity', which is, according to Recanati, the thought expressed by the utterance. If no additional entity is needed, why does Recanati himself introduce a new additional entity called the linguistic mode of presentation? I think this is a question that Recanati has to face if he wants to keep his theory coherent. What is more, in the discussion of my third dissatisfaction, it can be discerned that the new additional entity introduced by

Recanati is not unproblematic.

My third dissatisfaction with Recanati's proposal is that he does not provide a clear explanation of how the role which the linguistic mode of presentation plays in his conception of de se communication should be cashed out. In the case of the traditional picture of communication, the role played by the 'thought expressed by the utterance' is pretty clear: the speaker encodes her thought into the linguistic apparatus (utterance), and the hearer extracts the thought from the linguistic apparatus that she hears. The content of utterance, or the thought expressed by the utterance, plays a bridging role between the speaker and the hearer by means of replication. However, based on the discussions in previous sections, it is clear that in de se communication, the language does not play the same type of role. According to Recanati, the role that language (or in his term, the linguistic mode of presentation) plays in de se communication is called 'coordination':

The thought constituents are the files—the psychological modes of presentation—which the linguistic mode of presentation constrains. It follows that communication involves, not replication, but coordination of thoughts. The coordination aspect is obvious in cases of joint attention accompanying the use of demonstratives. The speaker and the hearer do not have the same perspective on the object of attention but their attentional acts are coordinated and converge on the object. Similarly, the mental acts of reference performed by the speaker and the hearer by means of their respective files about the speaker are coordinated through, *inter alia*, the linguistic mode of presentation encoded by the word 'I'. (Recanati, 2016, p. 154)

In this paragraph, Recanati describes 'coordination' as the constrain that linguistic mode of presentation sets on both the speaker's thought and hearer's relevant thought. A de se communication, based on this description, is a coordination between the speaker's thought and the hearer's thought and the coordination is determined by the linguistic expression that the speaker uses. However, it should be noted that in

another paragraph which I quoted in section 1.1.3, Recanati says something a bit different:

The Naïve Conception of Communication rests on the idea that communication is the replication of thoughts: the thought the hearer entertains when he understands what the speaker is saying is the very thought which the speaker expressed. But in the indexical case, there is no replication, but some kind of systematic transformation. The truth-conditions can be preserved across contexts only if the internal content of the thought is modified or adjusted so as to compensate for the shift in point of view from the speaker to the hearer. It follows that the Naïve Conception is false. (Recanati, 2016, pp. 149-150)

Here Recanati provides a different description of how an indexical communication is conducted: it is conducted through a ‘systematic transformation’, or to put it in my terminology, a translation. But systematic transformation is not co-ordination. It seems that either Recanati himself does not notice this problem in his argument, or he does not take it to be a problem. But I argue here that it is a real problem: a *de se* communication cannot be a co-ordination and a systematic transformation at the same time, because there is a non-trivial difference between those two: co-ordination does not involve the process of acquisition, but transformation involves the process of acquisition. This difference can be illuminated by an example. Suppose Oliver told Smith ‘I am hungry’ and Smith replied to Oliver ‘Oh, you are hungry. Let us have some snacks.’ I think it can be widely accepted that in this conversation a typical *de se* communication takes place. If we adopt Recanati’s co-ordination view to explain this communication, the explanation is like this: before the conversation started, Oliver had a belief ‘I am hungry’ in his mind and Smith had a belief ‘You are hungry’ in his mind, but those two beliefs were not co-ordinated with each other; then the conversation started, and Smith’s belief was then co-ordinated with Oliver’s relevant belief by means of the linguistic mode of presentation associated with Oliver’s utterance ‘I am hungry’. This explanation of *de se* (belief) communication, however, is not in

accordance with our intuition that the hearer acquires a belief from the speaker through the hearer's understanding of the speaker's utterance, because in this explanation Smith the hearer did not acquire a belief from Oliver the speaker (Smith had already had the belief 'you are hungry' in mind before the conversation started), but what Smith acquired from Oliver through his understanding of Oliver's utterance is merely an co-ordination of beliefs. On the other hand, if we adopt Recanati's systematic transformation view, we get a different explanation of the communication in that conversation: Oliver had a belief 'I am hungry' in mind and he expressed his belief using the utterance 'I am hungry'; then Smith, who heard that utterance, transformed what he received from the utterance in order to understand it, and he expressed his understanding by using second person sentence 'You are hungry' in order to meet the co-aboutness requirement (to be co-about with Oliver's utterance/belief). This explanation is in accordance with our intuition that communication involves a process of acquisition. Recanati includes both the coordination view and the systematic transformation view in his proposal, and this causes unnecessary perplexity about how to cash out the role that linguistic mode of presentation plays in de se communication. A possible solution to this problem is to abandon the co-ordination view and only obtain the transformation view, which I think is a better choice.

1.3.4 Preliminary Characterization of Translation Transmission Model

Now I have discussed three proposals on how to characterize the translation transmission model. Each of them has its own advantages, but I am not satisfied with any of them. Therefore, in this section, I try to provide a preliminary outline of my own characterization of the translation transmission model. First, the aim of my adoption of the translation transmission model in de se communication is to meet the co-aboutness requirement. In de se communication, if participants insist on replicating

each other's first person linguistic expressions and first person mental attitudes, they necessarily talk past each other, and therefore no successful communication is conducted. To avoid the situation of talking past each other, I introduce the translation transmission model. Second, the translation operation involved in the translation transmission model takes place in the hearer's mind. It is the hearer, not the speaker, who is responsible for conducting the translation. Third, what is translated in de se communication is not only the de se thought expressed by the first person linguistic expression, but also the linguistic expression itself. This is because in verbal communication, the hearer has access to what the speaker expresses by means of understanding the meaning of the utterance that the speaker uses to express her attitude.

The fourth feature of the translation transmission model is by far the most important feature. I argue that in the translation transmission model, not every type of translation is evaluated as being equally useful or valid, especially when we use that model to explain de se communication. There is a certain type of translation which can better secure the achievement of the co-aboutness condition in de se communication, compared with other types of translation. In fact, a similar situation can be discovered in the replication transmission model as well. Not every type of same-saying is evaluated as being equally useful or valid: the same-saying of a certain word is evaluated as being a better guarantee for the co-aboutness in one type of communication, but the same-saying of a different word is evaluated as being a better guarantee for the co-aboutness in another type of communication. Let us introduce an example here to illustrate my point. At the end of Section 1.2.4, I argue that the replication transmission model is followed by both the non-indexical communications and some of the non-de-se indexical communications. Now let us compare a non-indexical communication with a relevant non-de-se indexical communication: Suppose that one day, Peter and Smith meet on the street and they start a conversation:

Conversation 1: non-indexical communication

Smith: 'Melbourne is a beautiful city.' Peter: 'Yes, Melbourne is a beautiful city.'

Conversation 2: non-de-se indexical communication

Smith: 'It is a beautiful city here.' Peter: 'Yes, it is a beautiful city here.'

In both scenarios, participants of those communications follow the replication transmission model, i.e. they use the same-saying strategy to make sure they are talking about the same subject matter. However, the difference between them is that, in Conversation 1, Smith and Peter achieve the co-aboutness condition by using the same-saying of 'Melbourne', but in Conversation 2, they achieve the co-aboutness condition by adopting the same-saying of 'here'. Can we say that the same-saying of proper name 'Melbourne' and the same-saying of indexical 'here' are equally effective at guaranteeing the achievement of co-aboutness condition? I do not think so. I argue that in fact the same-saying of indexical 'here' is a better choice for securing the co-aboutness in this case compared with the same-saying of proper name 'Melbourne'. This is because, the same-saying of 'Melbourne' is prone to the error of misidentification, but the same-saying of 'here' is not prone to that error. We know there are at least two cities whose names are both Melbourne in the world. One Melbourne city is in Victoria, Australia, and the other Melbourne is in Florida, USA. So, in Conversation 1, it is fairly possible that Smith's 'Melbourne' refers to the Melbourne city in Victoria, Australia, but Peter's 'Melbourne' refers to the Melbourne city in Florida, USA. In this sense, it is quite reasonable for Smith and Peter to doubt 'Are we talking about the same Melbourne?' in Conversation 1, because it is possible that they make an error of misidentification. However, it is not reasonable for Smith and Peter to doubt 'Are we talking about the same here?' in Conversation 2. This is because, given the context of their conversation (they meet each other on the street), it is contextually guaranteed that their tokens of 'here' both refer to the same place. The same-saying of 'here', in Conversation 2, is immune to the error of misidentification. The satisfaction of co-aboutness condition in Conversation 2 is better secured compared with the satisfaction of co-aboutness condition in Conversation 1. For the

type of communication which is similar to the Conversation 1 and 2 in that scenario, same-saying of spatial indexical is a better choice for securing the co-aboutness than same-saying of proper name of a certain location. I argue that we can find similar situations in translation transmission model as well. For each type of communication which conforms to the translation transmission model, for example de se communication, there is a certain type of translation which can better secure the achievement of the co-aboutness requirement, compared with other types of translation. What is more, the co-reference associated with that translation is also more prominent than the co-reference associated with other types of translation. Let us call the translation which successfully secures the co-aboutness of de se communication the 'X translation'. In the third chapter of my dissertation, I will specify what 'X translation' is.

My current characterization of the translation transmission model is far from being comprehensive, but it outlines the basic features of the translation transmission model in my conception. A comprehensive characterization of those features is the task of the following chapters of this dissertation. It should be clear from my discussion that among the four features that I discussed in this section, the fourth feature is the most important. In Chapter 3, I will discuss this feature in detail. But before launching my further characterization of the translation transmission model, I first need to introduce what is transmitted in accordance with that model. In the next chapter, I will discuss what is communicated in de se communication, i.e. the de se thought. I will clarify what it is to think a de se thought by discussing two important theses concerning de se thought. They are the essential indexical thesis and the immunity to error through misidentification thesis.

Chapter II De Se Thought and Identification

In Chapter 1, I introduced the communication framework, proposing the co-aboutness requirement, and discussed the necessary application of the translation transmission model as an idiosyncratic feature of de se communication. But this is not where the story ends: besides the form of communication, there is also the content of communication. To put it in other words, besides the information transmission model, there is also what is communicated in the communication. In de se communication, what is communicated is the de se thought. In order to express my de se thoughts and communicate them to someone else, I must first be able to produce de se thoughts about myself. Here we encounter a new question about what it means to produce and think de se thought. This is a deep question which is worthy of a book length discussion on its own merits, and it can hardly be properly answered in a single chapter. Therefore, in this chapter I will not try to provide a comprehensive answer to this question, but rather, I will attempt to articulate one aspect of the production of de se thought, which is whether or not our production of de se thought involves identification (and/or mis-identification). I call this the identification problem of de se thought. Although the identification problem covers only one aspect of what is communicated in de se communication, it is by no means trivial. On the contrary, as my analysis in this chapter shows, the answer to the question of whether our producing and entertaining of de se thought involves identification places an important constraint on our understanding of what it means to produce and think de se thought. Once the identification problem is explored, we make significant progress towards a comprehensive understanding of the production of de se thought.

In order to clarify whether or not a process of identification is involved in our production of de se thought, I introduce two theses as the background of my clarification. Those two theses are the essential indexical thesis and the immunity to error through misidentification thesis (IEM). Each of them is a significant

philosophical thesis on its own merit, and I will discuss both of them in the following sections. However, I notice that the connection between them has rarely been discussed in the current literature. In this chapter, I argue that the philosophical significance of the connection between those two theses is worthy of exploration, and it can be properly understood if we take a closer look at both of them in relation to de se thought. In other words, the binary relation between the two theses can be better revealed by a scrutiny of the triple relation between the essential indexical, the immunity to error through misidentification and the de se thought. On the one hand, if we take the essential indexical thesis to be a thesis about the connection between the first person perspective and the performance of intentional action, according to a certain interpretation of the thesis, the essential indexical thesis implies an identification thesis concerning de se thought. On the other hand, if we attribute the epistemic property of IEM to de se thought, according to some interpretations of the immunity thesis, this attribution results in another identification thesis concerning de se thought. What I discover here is that the identification thesis implied by the essential indexical thesis may be incompatible with the identification thesis caused by the attribution of the property of IEM. This incompatibility, in turn, implies that it may not be possible for de se thought to be both essentially indexical and immune to error through misidentification at the same time. I think this would be a very problematic claim. However, I argue that the alleged incompatibility is only apparent. De se thoughts are both essentially indexical as well as immune to error through misidentification. In order to justify this argument and reject the alleged incompatibility between the essential indexical thesis and IEM, a substantial exploration of the essential indexical thesis, the immunity thesis, and their relations to de se thought must be developed.

In this chapter, I first introduce some general characterizations of the essential indexical thesis. Then I proceed to introduce two skeptical response to the thesis. In recent years, there has been a strong wave of de se skepticism, in which some philosophers make sophisticated arguments to cast doubt on the truth of the

essential indexical thesis. Cappelen & Dever's criticism of the essentiality of indexicality per se and Magidor's criticism of the property account of de se content are prominent examples. I discuss each of them in turn. I will argue that Cappelen & Dever's criticism is not comprehensive enough to cover all aspects of essential indexical, and Magidor's criticism is based on an inaccurate reading of David Lewis' property account of de se content. My discussion shows that the essential indexical thesis is closely related to an identification thesis, which is that in entertaining a de se thought, I do not need to appeal to any representation of myself for the purpose of identifying myself. Although there is plenty of room for debate about what the identification thesis exactly means, it reveals a close connection between de se thought and the identification. In order to adequately discuss the implication of that connection, I introduce two defenses of the essential indexical thesis against the de se skepticism. Those are Matthew Babb's indexical intention defense and Juliana Faccio Lima's action-explanation defense. Both of them clearly argue that an identification thesis implied by the essential indexical thesis is crucial for establishing the connection between first person perspective and the performance of intentional action. However, this argument may seem to be incompatible with another thesis concerning de se thought. Philosophers seem to agree that some de se thoughts, if not all de se thoughts, have the epistemic property of immunity to error through misidentification.²⁸ According to the Simple Account/Simple Explanation of IEM, de se thought is immune to the error of misidentification because entertaining de se thought does not involve identification. This seems to be incompatible with the identification thesis which is implied by the essential indexical thesis as I mentioned above. Hereby we are facing a dilemma: if we insist on arguing that de se thought is essential for the performance of intentional action because of an identification role it plays, we have to reject a popular explanation of why de se thought is IEM, which claims that de se thought does not involve any identification; but if we accept that

²⁸ See Cappelen, Herman., Dever, Josh., Chapter 7 Indexicality and Immunity to Error, pp. 127-129, in *The Inessential Indexical*, Oxford University Press, 2013.

explanation of the immunity thesis, we then have to look for a different explanation of why *de se* thought is essential for the performance of intentional action.

To avoid this dilemma, I argue that the relation between *de se* thought and the immunity thesis should be cautiously re-examined. What does ‘being immune to the error of misidentification’ really mean for *de se* thought? To answer this question, I need to clarify first what identification (and also, misidentification) means here. I introduce Wittgenstein’s distinction between the use of ‘I’ as subject and the use of ‘I’ as object, Gareth Evans’ distinction between ‘identification-dependent’ and ‘demonstrative identification’, James Pryor’s distinction between *de re* identification and which-object identification, and Manuel García-Carpintero’s distinction between identification_P and identification_C . All those distinctions are devices philosophers use to solve the tension between the essential indexical thesis and the Simple Account of IEM. I am not satisfied with their solutions, because those distinctions on which their solutions are based are heavy theoretical commitments for their solutions, and those distinctions face philosophical challenges. Because of those theoretical commitments, the Simple Account is not that simple. My own plan is to solve the apparent incompatibility between the essential indexical thesis and the IEM thesis by rejecting the Simple Account and replacing it with the Definition Account. The basic idea of the Definition Account is that *de se* thought has the property of IEM simply because of its definition. Since there is no ‘identification free’ condition in the Definition Account, there is no incompatibility between the essential indexical thesis and the Definition Account of IEM. This account may be a bit boring, but I argue that it works, and it is even simpler than the Simple Account.

2.1 Essential Indexical

First, I introduce the essential indexical thesis. The term ‘essential indexical’ was

coined by John Perry²⁹. According to Perry's initial argument, the problem of the essential index is a challenge to the traditional doctrine of proposition, which claims that the content of belief is a proposition that 'has a truth-value in an absolute sense, as opposed to merely being true for a person or at a time.' (Perry, 1979, p. 6) . He says:

The problem of the essential indexical reveals that something is badly wrong with the traditional doctrine of propositions. (Perry, John., 1979, p.8)

For a belief that is expressed by a sentence like 'I am an international student', there seems to be no way for its content to have a truth value in the absolute sense, as the indexical term 'I' in that sentence shifts its referent in accordance with different contexts. When I hold this belief, this is a true belief as I am indeed an international student, but when my colleague Adam holds this belief, this is a false belief as he is a domestic student. We have so many beliefs that are similar to the belief 'I am an international student' in our everyday lives. Perry argues that any attempt to substitute the indexical terms in the linguistic expressions of those beliefs with non-indexical terms fails, because first, propositions expressed by those non-indexical sentences we get after the substitution have lost the self-locating function that the original sentences have, and second, non-indexical sentences cannot account for the fact that a single sentence can be true in context A, while the same sentence can be false in context B, which is to say, non-indexical sentences cannot account for the variation of their truth value in different contexts. That's why Perry argues that the traditional doctrine is not enough to deal with the essential indexical problem. It can be easily discovered that the essential indexical is a thesis initially motivated by a concern about the role of indexical terms in the linguistic expressions of belief states.

²⁹ Please refer to John Perry, The Problem of the Essential Indexical, *Noûs*, Vol. 13, No. 1 (Mar. 1979), pp. 3-21, Blackwell Publishing.

However, because of the thought provoking ‘messy shopper’ case³⁰ that Perry introduced in his paper, the subsequent development of discussions on the essential indexical thesis deviates from Perry’s original conception in a philosophically interesting way: now the most popular reading of essential indexical thesis no longer takes it to be a linguistic thesis, but rather a thesis about the connection between intentional action and indexical thought. Most philosophers currently are more interested in discussing the connection between the change of the messy shopper’s beliefs and the change of his behavior. Here their focus is no longer about the true value of linguistic sentence (or proposition), but rather about the relation between our performance of certain intentional actions and our capacity to use certain indexical elements to organize our reasons to perform those actions. For example, Cappelen & Dever discussed various ways of understanding what the philosophical significance of first person perspective could be about, but they acknowledged that the most important way to understand its significance is to understand that as a revelation of a connection between indexical and agency:

We suspect that Chapter 3 is the most important chapter in this book. In that chapter we argue that there are no interesting or distinctive explanatory connections between indexicality and agency. In talking to philosophers about these issues over many years, we find that attempts to articulate the essence of Perry and Lewis’ insight tend to come back to something or other about a

³⁰ Perry’s original narrative of the messy shopper is:

I once followed a trail of sugar on a supermarket floor, pushing my cart down the aisle on one side of a tall counter and back the aisle on the other, seeking the shopper with the torn sack to tell him he was making a mess. With each trip around the counter, the trail became thicker. But I seemed unable to catch up. Finally, it dawned on me. I was the shopper I was trying to catch. I believed at the outset that the shopper with a torn sack was making a mess. And I was right. But I didn't believe that I was making a mess. That seems to be something I came to believe. And when I came to believe that, I stopped following the trail around the counter, and rearranged the torn sack in my cart. My change in beliefs seems to explain my change in behavior. (Perry, John., 1979, p. 3)

distinctive connection to agency. (Cappelen and Dever, 2013, p. 18)

Simon Prosser formulates the essentiality indexical thesis in a similar way:

The Doctrine of the Essential Indexical: There is a class of actions, C_A , such that no action in C_A can be performed by an agent who lacks the relevant egocentric mental state, the presence of which is normally indicated by the use of an egocentric indexical. (Prosser, Simon., 2015, p. 6)

In this chapter, I generally take the essential indexical thesis to be a thesis about the relation between agency and thought. For me, the essential indexical thesis is that the performance of intentional action requires the agent to have a first person perspective with respect to at least some of her thought contents. This is only a general description of the essential indexical thesis, not a strict definition, but it is sufficient for my current purpose. When I occasionally use the essential indexical in its old-fashioned sense, i.e. a thesis about linguistic proposition, I will make special clarifications. However, I think it is worthy of being noted that in Prosser's formulation, the indexicality per se does not play any essential role. The deployment of indexical term is simply an indication of the presence of a relevant mental state, which Prosser calls the egocentric mental state. By formulating the thesis in this way, Prosser has already suggested a view that what is really essential in the essential indexical thesis is neither the indexicality per se nor any other types of indexical, but rather, it is the first person perspective.

2.2 Two Criticisms of Essential Indexical

Although essential indexical thesis has already been widely discussed for many years, it has not been accepted by all philosophers. The controversy around it has never ended. Recently there has been a wave of de se skepticism. In this section I introduce two outstanding criticisms of the essential indexical thesis emerging from that wave. They are Herman Cappelen and Josh Dever's denial of the essentiality of indexicality

and Ofra Magidor's criticism of David Lewis' property account (or centred world account) of de se content. The two criticisms in a broad sense share a common target, but they criticize the target from different angles. Cappelen and Dever take the indexicality per se as the Achilles' heel of the essential indexical thesis. In contrast, Magidor focuses on a certain explanation of the essential indexical thesis, arguing that David Lewis' centred world account of de se content does not do justice to the de dicto/de se distinction. She complains that Lewis' proposal may lead to a failure in establishing a real distinction between de se belief and de dicto belief. However, as my analysis will show, both Cappelen and Dever's criticism of indexicality, and Magidor's criticism of the property account, are problematic. I argue that the former fails to establish what it is expected to establish, and the latter is based on a misunderstanding of Lewis' proposal. In the subsequent sections, I analyze Cappelen & Dever's criticism of essential indexicality first, and then analyze Magidor's criticism of the centred world account.

2.2.1 Generic Opacity instead of Essential Indexicality

When two referential expressions refer to the same referent, and we know that they refer to the same referent, we say their co-referential relation is transparent to us. If two referential expressions refer to the same referent, but we do not know that they refer to the same referent, in this case their co-referential relation is opaque to us. After scrutinizing Perry's messy shopper case carefully, Herman Cappelen and Josh Dever draw a conclusion is that it was the co-referential opacity, rather than the alleged 'indexicality', in that it was the shopper's epistemic status that prohibited the shopper from performing the right action (which is to clean the sugar on the floor), and it was the removal of the co-referential opacity and the realization of the co-referential transparency in the shopper's epistemic status that enabled the shopper to realize which action he should perform in that scenario. In the case that Perry

deliberately conceived, the relevant opacity is an indexical opacity, because that opacity involves the first person pronoun 'I', which is generally considered to be an indexical term. However, Cappelen and Dever argue that indexicality per se is not essential to the opacity, because indexical opacity is merely an instantiation of the more general generic opacity:

Here is one way to summarize Perry's claim:

Indexical Opacity. There's a set of indexicals, I-SET, that cannot be substituted *salva veritate* in action-explanation contexts by any other expressions.

It should be clear that this is an instantiation of the more general thesis:

Generic Opacity. Co-referential referring expressions cannot be substituted *salva veritate* in action-explanation contexts. (Cappelen and Dever, 2013, p. 33)

For Cappelen and Dever, the generic opacity does not involve indexicality essentially, because it can take place when no information about indexical term per se is involved in an agent's epistemic status. To show this point more clearly, they construct a thought experiment as below:

Superman/Clark Kent

Pushing my cart down the aisle I was looking for CK to tell him he was making a mess. I kept passing by Superman, but couldn't find CK. Finally, I realized, Superman was CK. I believed at the outset that CK was making a mess. And I was right. But I didn't believe that Superman was making a mess. That seems to be something that I came to believe. And when I came to believe that, I stopped looking around and I told Superman to clean up after himself. My change in beliefs seems to explain my change in behavior.

We take the S/CK case as showing that in general action explanations don't have their explanatory force preserved by substitution of co-referential

singular terms. Seeing Indexical Opacity as an instance of Generic Opacity suggests that there's nothing deeply central about indexicals here. (Cappelen and Dever, 2013, p. 33)

The generic opacity here simply means that the co-referentiality of two (or more) referring expressions (or mental files) is not available to the relevant agents. In order to reveal the co-referentiality, an identification belief must be introduced into the context. In the messy shopper case, the identification belief is 'I am the shopper with a torn sack', and in the Superman/Clark Kent case, the identification belief is 'Superman is Clark Kent'. Since the identification belief does not have to contain an indexical element, Cappelen and Dever insist that indexicality is not essential for the generic opacity. What is more, they argue that opacity involving indexical expressions, which they call 'indexical opacity', is not distinctive compared to opacity which does not involve indexical expressions. Both indexical opacity and non-indexical opacity pattern with the classic Fregean opacity:

Our goal in what follows is to consider whether substitution failures involving indexicals present problems of a distinctive kind. We first present six of the classic puzzle cases (or thought experiments or examples) involving indexicals. For each of them we show that they pattern in obvious ways with classic Frege puzzle cases. (Cappelen and Dever, 2013, pp. 60-61)

This is an overall picture of Cappelen and Dever's argument of eliminating essential indexicality and defending generic opacity in their explanation of the messy shopper case. It should be noted that the generic opacity thesis is a thesis concerning an agent's epistemic status about the co-referentiality of terms, and the co-referentiality of terms is an instance of the more general co-aboutness that I discussed in previous chapters. In this respect, I agree with Cappelen and Dever that the generic opacity as an epistemic thesis concerning co-aboutness is a crucial point in understanding the messy shopper case and in explaining the essential indexical thesis.

However, I do not think that their argument can do what it is expected to do, which is to correctly reveal the connection between the co-aboutness (and the relevant opacity/transparency) and the explanation of intentional action, because their argument is not comprehensive enough. To show the non-exhaustiveness of Cappelen and Dever's argument, it should be noticed that there are at least two things which Cappelen and Dever neglect in their argument.

The first thing is that Cappelen and Dever discuss merely the opacity of co-reference, but they neglect the transparency of co-reference. In Perry's original version of the messy shopper case, the shopper changed his behaviour only after the co-aboutness relation between the belief 'the shopper with a torn sack was making a mess' and the *de se* belief 'I was making a mess' became transparent to him. Analogously, in Cappelen and Dever's 'Superman / Clark Kent' case, the shopper 'stopped looking around and told Superman to clean up' only after the co-aboutness relation between the belief 'Superman was making a mess' and the belief 'Clark Kent was making a mess' became transparent to him. Now we know the transparency of co-aboutness is presented in both cases, but are the transparency in the messy shopper case and the transparency in the 'Superman/Clark Kent' case homogeneous? Are they both instances of a generic transparency type, just like indexical opacity and non-indexical opacity? If not, are they heterogeneous, or at least different from each other at an important aspect?

I argue that the co-aboutness transparency which involves *de se* thought, for example the transparency of co-referential relation between 'the shopper with a torn sack' and 'I', and the co-aboutness transparency which does not involve *de se* thought, for example the transparency of co-referential relation between 'Superman' and 'Clark Kent', are different in an important respect, which is that they posit different epistemic thresholds for agents. In the 'Superman/ Clark Kent' case, it is likely that an agent is ignorant about which person is Superman and it is also likely that the agent is ignorant about which person is Clark Kent. I call this 'symmetric ignorance', which means if some agents fail to recognize the co-referential relation between 'X' and 'Y', it is likely that their failure is caused by their ignorance about

what 'X' refers to, and it is also likely that their failure is caused by their ignorance about what 'Y' refers to. However, in the 'shopper with a torn sack / I' case, there is no symmetric ignorance. On the contrary, there is asymmetric ignorance. It is likely that the agent is ignorant about which person the shopper with a torn sack is, but it is highly unlikely that an agent is ignorant about which person he is. I call this the asymmetric ignorance, which means if some agents fail to recognize the co-referential relation between 'X' and 'Y', the possibility that their failure is caused by their ignorance about what 'X' refers to is much greater than the possibility that their failure is caused by their ignorance about what 'Y' refers to.

To achieve the co-referential transparency in non-de-se cases like the 'Superman / Clark Kent' case, agents need to overcome symmetric ignorance, but to achieve the co-referential transparency in de se cases like the 'shopper with a torn sack / I' case, agents need to overcome asymmetric ignorance. I call the difference between overcoming symmetric ignorance and overcoming asymmetric ignorance 'the difference of epistemic threshold'. This difference indicates a possible way of distinguishing de se cases from non-de-se cases. I believe this strategy can be applied to distinguish other types of indexical cases from non-indexical cases as well. For example, suppose that Emily lives in Melbourne. It is likely that she does not know how to locate Melbourne and it is also likely that she does not know how to locate the capital city of Victoria, but it is highly unlikely that she can be ignorant about how to locate her current *here*. So, for Emily, there is also a difference of epistemic threshold between recognizing the co-reference between 'Melbourne' and 'the capital city of Victoria' and recognizing the co-reference between 'Melbourne' and 'here'. Cappelen and Dever fail to realize the epistemic difference between asymmetric ignorance and symmetric ignorance, perhaps because they pay attention to opacity only, but ignore transparency.

The second thing Cappelen and Dever neglect in their argument is the co-referentiality of tokens of the same word type. When discussing the co-referential opacity (either the indexical opacity or the generic opacity), Cappelen and Dever's argument covers only the co-referential relation between different words, for

example, between 'Clark Kent' and 'Superman', but their argument does not cover the co-referential relation between different tokens of the same word, for example, between one token 'I' and another token 'I'. This neglect leads to a further neglect, which is the neglect of the opacity and transparency associated with the co-referential relation between tokens of the same word. I think the transparency/opacity of co-reference between tokens of the same word is obviously different from the genetic transparency/opacity that Cappelen and Dever discuss in their argument. For example, it would not be irrational of me to challenge the co-reference between 'Superman' and 'Clark Kent', because it is fairly possible that their co-reference is opaque to me. However, if I challenge the co-reference between two tokens of 'I' which I myself produced, there must be something highly unusual going on, because in most cases their co-reference is transparent to me. Here we find another difference of epistemic threshold. It should be noted that the co-reference between tokens of 'I' can be different from the co-reference between tokens of 'Clark Kent', and the transparency/opacity associated with them can also be different. If I know many people whose names are 'Clark Kent', for me it is easy to challenge the co-reference of tokens of 'Clark Kent' I produced. Maybe one token 'Clark Kent' refers to this Clark Kent, but another token 'Clark Kent' refers to that Clark Kent. However, it is not easy for me to challenge the co-reference of tokens of 'I' I myself produced. This difference is closely related to a point I made at the end of Chapter 1. That point is: not every type of same-saying is evaluated as being equally useful. The same-saying of a certain word is a better guarantee for the co-aboutness in a certain type of communication than the same-saying of other words.

The two points that are not covered by Cappelen and Dever's argument are not trivial. As I will argue in Chapter 3, the co-referential relation of tokens of 'I' used for a single person is idiosyncratic. The co-reference of tokens of 'I' is, in a sense, independent of the reference of each token of 'I'. What is more, there is a special type of transparency associated with the co-referentiality of tokens of 'I': their co-reference must be transparent to the person who uses those tokens, otherwise there will be an

infinite regress associated with eliminating the opacity of co-reference between tokens of 'I'. The explanation of co-reference of 'I-I' same-saying is an important clue for my explanation of the co-reference of 'X-translation' in a de se communication.

2.2.2 De Dicto/De Re and De Se

Besides the co-referentiality and its opacity/transparency, there is another aspect of de se thought which has been touched on, but not elaborated, in Cappelen and Dever's criticism of the essential indexical thesis. This aspect is identification. According to a Fregean framework, which is generally accepted by Cappelen and Dever when they discuss co-referential opacity³¹, the alleged generic opacity takes place within certain opaque contexts (for example, the action explanation context) when two or more referential expressions are used to refer to the same referent by means of different senses. To put it in another way, the same object has been identified multiple times through different criteria of identification. Therefore, the co-referential opacity is closely related to the co-identificational opacity, which means when identifying an object in a certain way, I do not know that it can also be identified in other ways, and when I identify an object that I once identified a while ago in a different way, I do not realize that it is still the same object.

Although we did not find an elaboration of the identification aspect of de se thought in either Perry's original account of the essential indexical thesis or Cappelen and Dever's criticism of it, there is another early proponent of the essential indexical thesis,

³¹ For example, when discussing whether there is a diagnostic distinctiveness of indexical opacity or not, Cappelen and Dever say:

Since the question of whether Diagnostic Distinctiveness is true is potentially enormously broad (requiring us to look at every proposed account of opacity), we limit our discussion to how contemporary followers of the German philosopher and mathematician Gottlob Frege (so-called Fregeans) deal with indexicals. (Cappelen and Dever, 2013, p. 59)

whose theory contains an important insight concerning the identification aspect of de se thought. The proponent is David Lewis. While Perry argues that the special feature of indexical belief is exhibited by how we get access to the content of indexical belief (to put it into Perry's terms, the belief state), Lewis argues that the special feature of de se belief lies in its very idiosyncratic type of content. As I said in Section A of the Introduction, the content of an ordinary propositional belief is a proposition with definite truth value, but the content of a de se belief, Lewis argues, is a property:

In response to these difficulties Lewis proposes an account that, characteristically, is at once simple and radical. Do not think of de se thoughts as attitudes to propositions at all. Think of them rather as self-ascriptions of properties. When Perry realizes that he is following his own trail, he self-ascribes the property of spilling the sugar; when Kaplan realizes that it is his own reflection that he can see, he self-ascribes the property of having flaming pants. Moreover, the de dicto can then be thought of as a special case of the de se: the case in which what is self-ascribed is the property of membership of a world. (Holton, 2015, p. 400)

When believing a de se belief (or making a de se judgement), I self-ascribe a property, which is the content of the belief. A property itself is not a proper bearer of truth value. Only after it has been ascribed to an individual can it be true or false. Therefore, a de se belief by itself has no definite truth value, and only after it has been believed by a person can it be properly evaluated as being true or false. So far so good.

The problem begins when Lewis continues to characterize de se belief using his possible world semantic framework. The ordinary propositional belief has a set of possible worlds as its content, according to the possible world semantics. However, a de se belief, according to Lewis, has a set of centred worlds as its content. A centred world is a possible world paired with a center, which is usually an individual in that

world.³² Therefore, for Lewis, the essential indexical thesis is at the end of the day a thesis about the essentiality of centred worlds. Ofra Magidor summarizes Lewis' centred world semantics as below:

Let a centred-world be a pair $\langle w, x \rangle$, such that w is a possible world, and x is an individual. Consider the attitude of belief. On the centred-worlds account, an agent's total (centred) belief state is taken to be a set of centred-worlds (intuitively: those centred-worlds $\langle w, x \rangle$ such that is compatible with the agent's beliefs that w obtains, and she is x). Let a centred-proposition be a set of centred-worlds. An agent whose total centred belief state is S believes the centred-proposition p , if and only if S includes p . For example, suppose Jill believes (de se) that she is tall. This will be construed as a belief attitude towards the centred-proposition $\{\langle w, x \rangle: x \text{ is tall in } w\}$. Jill believes this centred-proposition just in case all the centred-worlds in her belief state are included in the proposition, i.e. just in case she believes that whichever individual she might be and whatever the world might be like, the world is such that she is tall. (Magidor, 2015, pp. 250-251)

By tracing the indexicality back to the type of content associated with de se belief, which is the set of centred worlds, Lewis provides an explanation for essential indexicality: indexicality is essential because the beliefs which motivate the

³² Richard Holton says,

In response to worries of this kind, Lewis suggests that we should take centres to be, not spatio-temporal locations, but, rather, individuals. (Holton, 2015, p. 405)

So, both Holton and Magidor believes that the Lewisian centers are individuals. I take this point.

It is suggested that Lewis's realist conception of possible world may be incompatible with the idea that one individual can be the center for a set of centred worlds, because according to the realist conception of possible world, no individual can exist in multiple worlds. To overcome this problem, I think we can conceive individuals which are centers of a set of centred worlds as being in the counter-part relations to one another.

performance of intentional actions have centred worlds as their contents, which cannot be replaced by ordinary propositional attitudes in which contents are worlds without centers.

However, Magidor argues that the Lewisian distinction between the de dicto belief.³³ and the de se belief is problematic, because, she argues, the boundary between de se belief and de dicto belief is not clear enough. There seems to be a certain type of belief, according to Magidor, which is both de dicto belief and de se belief at the same time. This is because the centred world account can be adopted not only for the explanation of de se content, but also for the explanation of de dicto content. As Magidor says:

Lewisians have a common response to this worry. Call a centred proposition p ‘boring’ just in case, for all worlds w , and individuals x , if p is true in $\langle w, x \rangle$, then for all individuals y , p is true in $\langle w, y \rangle$. Call a centred-proposition ‘interesting’ just in case it is not boring. The proposal is that distinctively de se attitudes are attitudes towards interesting centred-propositions, while ordinary de dicto attitudes are attitudes towards boring centred-propositions. (Magidor, 2015, p. 264)

Therefore, the content of de dicto belief and the content of de se belief can both be formulated in the centred way. Their difference, Magidor claims, is that in the case of de dicto content, if one individual X in a possible world W can be a center, other individuals in that world can be centers as well, but in the case of de se content, only a certain individual, namely the believer herself, can be the center. We have a pair of equations at hand according to Magidor’s interpretation³⁴ of Lewis’ centred world

³³ Holton notices that Lewis treats propositional belief and de dicto belief as being equivalent, see Holton, 2015, p. 401.

³⁴ Thanks to a reminder from one of my examiners, I should note here that the distinction between the interesting centred proposition and the boring centred proposition is not an idiosyncratic interpretation of Lewis by Magidor, but rather, this distinction is widely accepted and deployed by several philosophers, including Andy Egan (2019) and Richard Holton (2015).

account:

1. de dicto belief = belief which has a boring set of centred worlds as content

and

2. de se belief = belief which has an interesting set of centred worlds as content

This pair of equations is supposed to be exhaustive and unambiguous. If a belief is found that belongs to neither side of this pair, or belongs to both sides of this pair, the definitions fail. But Magidor claims that she indeed discovers a belief as such. She says:

Suppose for example, that in each possible world either there is a heaven, in which case everyone is such that if they are good they go to heaven or there is no heaven, in which case no one is such that if they are good they go to heaven. Now suppose that I hope that I will go to heaven if I am good. This is the sort of attitude that would typically be classified as de se: the content of my hope concerns what might happen to me, and I can have this attitude even if I have suffered severe amnesia and do not remember who I am. Yet, this is an attitude towards the centred-proposition $\{ \langle w, x \rangle : \text{in } w, \text{ if } x \text{ is good } x \text{ goes to heaven} \}$, which is a boring centred-proposition (for all worlds w and individuals x, y , if $\langle w, x \rangle$ is a member of the proposition, then so is $\langle w, y \rangle$). So by the above proposal it would be classified as a de dicto rather than a de se attitude. (Magidor, Ofra., 2015, pp. 264-265)

Magidor argues that there is at least one belief, which cannot be properly accommodated by the presumably exhaustive and unambiguous pair of equations. It is the belief 'I will go to heaven if I am good'. I think people, including me, share the intuition in general that this is a de se belief. But Magidor's analysis tries to show that this belief does not have interesting centred worlds in Lewisian sense as its content. This is because, if 'I will go to heaven if I am good' is true for one individual (which may be the believer herself) in a certain world, it is also true for many other individuals in that same world (other people). Therefore, as Magidor argues, this belief either

belongs to both sides of the equation pair, or it belongs to neither side of the equation pair: it belongs to the first equation because its content is a boring set of centred worlds, and it also belongs to the second equation because it is widely recognized as a *de se* belief; but at the same time, it does not belong to the first equation because it is not widely recognized as a *de dicto* belief, and it does not belong to the second equation because its content is not an interesting set of centred worlds. Magidor suggests that the very existence of the proposition 'if I am good, I will go to heaven' implies that the equation pair is not as exhaustive and unambiguous as we expect it to be, and the inexhaustive-ness and ambiguity turn out to be the fatal defect of the Lewisian distinction between *de dicto* belief and *de se* belief³⁵.

However, I find Magidor's argument, which is based on her analysis of Lewis'

³⁵ It should be noted that Magidor is not the only philosopher who criticizes the Lewisian explanation of *de se* attitude by appealing to the traditional *de dicto/de re* distinction. Steven E. Boer and William G. Lycan, in their joint authored paper 'Who, Me?', also claim that:

A growing number of theorists now maintain that the ascription of first-person attitudes or attitudes "de se" cannot be reduced to the ascription of attitudes *de re* or *de dicto*. (...) We are not persuaded. We believe the Irreducibility Thesis should be repudiated. In what follows, we shall set out our case for rejecting the Thesis and for maintaining that a version of the traditional *de re/de dicto* distinction is adequate in itself to accommodate Castañeda's data. (Boer and Lycan, 1980, p. 427 and p. 430)

Boer and Lycan admit that there are some special features associated with the *de se* attitudes, but they believe that the differences between *de se* attitudes and the traditional *de dicto/de re* distinction should be explained from the angle of pragmatics, not from the angle of semantics:

Our denial that attitudes (and references) *de se* are a surd or kink in semantics is not a denial that there is something special about the sorts of reflexives that Castañeda has singled out and abbreviated by starred pronouns. Certainly, there is something distinctive about them. Our contention is that they are pragmatically distinctive, not semantically *sui generis* in Castañeda's way. (Boer and Lycan, 1980, p.432)

It could be easily seen from this paragraph that Boer and Lycan believe the *de se* attitude is not semantically distinctive. This is in contradiction to Lewis' claim that *de se* attitude is semantically distinctive, because the semantic content of a *de se* attitude is a set of centred worlds, while the semantic content of other propositional attitudes is a set of ordinary possible worlds without centers. According to Boer and Lycan, the semantic content of a *de se* attitude and the semantic content of a *de re* attitude belong to the same type:

Bluntly, what we contend is that attitudes *de se* are simply attitudes *de* their owner. (Boer and Lycan, 1980, p. 432)

Therefore,

In particular, contra Castañeda, Perry, Lewis, and the rest, we need not admit that the content of an attitude *de se* is inexpressible by any nonperspectival, third-person sentence. (Boer and Lycan, 1980, p. 433)

centred world account, to be unconvincing. This is because, as I will show, she misunderstands the relation between de se attitude and the traditional de dicto/de re attitude in Lewis' account. This misunderstanding is so central to her analysis that it undermines her entire interpretation of Lewis. Magidor assumes that there must be a clear and sharp distinction between de se attitude and de dicto/de re attitude. But this assumption is problematic. As I will show, the relation between de se attitude and de dicto/de re attitude is more complicated than what Magidor assumes.

Lewis argues that the centred world account (i.e. the property account) can be adopted to explain the content of the traditional de dicto/de re attitude. For example, when I hold a de dicto belief that ' $1+1=2$ ', I ascribe to myself the property of 'being inhabiting in a world which belongs to the set of worlds in which " $1+1=2$ " is true', according to the centred world account. This property, or centred proposition, is boring in Magidor's definition, because if it is true for one center in a possible world, it is also true for other centers in that world, and if it false for one center in a possible world, it is also false for other centers in that world.

Here one important question emerges: does the boringness of its content prevent this belief to be a de se belief? Magidor's answer to this question is: yes, it does. But I argue that Lewis' answer to this question should be: no, it does not. Magidor assumes that a de se attitude must have an interesting centred proposition as its content. In other words, she assumes that there is only interesting de se content. A boring de se content, for her, is a categorical mistake, just like a round square. But for Lewis, a boring de se content is acceptable, as it is just a species of de se content. In the example I mentioned above, the property 'being inhabiting in a world which belongs to the set of worlds in which " $1+1=2$ " is true' is a typical boring de se content. When generalizing his centred possible world framework to explain the de dicto/de re attitude, Lewis uncovers a possibility that de se attitude can be an all-embracing attitude. It is all-embracing in the sense that the traditional de dicto/de re attitude can be regarded as a special species of de se attitude. This special species of de se attitude is marked by its distinctive feature that it has boring centred proposition as its content,

in contrast with other de se attitudes which have interesting centred propositions as their contents. Therefore, according to Lewis, a sharp and clear distinction between de se attitude and de dicto/de re attitude is simply unnecessary.

Once we understand that a sharp distinction between de se attitude and de dicto/de re attitude is unnecessary for Lewis, we can then overturn Magidor's criticism of Lewis. Magidor argues that the belief 'if I am good, I will go to heaven' is a belief which 'that would typically be classified as de se' (Magidor, 2015, p. 265), and she also argues that this belief has a boring centred proposition as its content. So this belief becomes a challenge to the Lewisian centred world framework, as according to Magidor, a de se belief must not have a boring centred proposition as its content under the Lewisian framework. But now we know that Lewisian framework can well tolerate the existence of de se attitude with boring de se content, so the belief 'if I am good, I will go to heaven' is no longer a challenge to the Lewisian framework.

However, if the distinctive feature of de se attitude under the Lewisian framework is not the interesting centred proposition as its content, then what is the real distinctive feature of de se attitude? Manuel Garcia-Carpintero provides a condensed and precise answer to this question:

This provides a nice solution to the initial problem of de se thought. If no conception of the subject (including those allowing for de re thought on Millian strictures) is sufficient for de se thought, and none appears to be needed, this is because the subject is not represented as part of the content but is brought to bear for purposes of evaluation by the act of judging itself, not by its content. (Garcia-Carpintero, 2015, section 1)

and also,

Lewis's proposal is that the content of Perry's epiphanic belief is the property of making a mess, which, in judging it, he self-ascribes (at the time of the believing) the way ordinary propositions are supposed to be ascribed to the

actual world when they are believed. The other shopper does not believe (i.e., self-ascribe) this property; she instead believes a traditional de re proposition: that the property making a mess applies to Perry. (Garcia-Carpintero, 2015, section 1)

Garcia-Carpintero here argues that the distinctive feature of de se attitude is that, in the case of de se attitude, the individual which the attitude is about, i.e. the producer/holder of this attitude, is not represented in the content of the attitude, while in the case of the traditional de dicto/de re attitude, the subject matter which the attitude is about is indeed represented in the content of the attitude. The representation of the individual which a de se attitude is about is absent in the content of that de se attitude, or in Lewis' terms, the center of the set of centred worlds is not represented. This is the distinctive feature of de se attitude, according to Lewis.³⁶

³⁶ The idea of 'absence of representation of self' in the de se mental state can also be found in other philosophers' writings. For example, when discussing his notion of the 'first person redundancy', Simon Prosser also suggests something similar to Lewis's idea. Prosser introduces the idea of 'first person redundancy' in this way:

Suppose that C did in fact perceive C due to a reflection or refraction of some kind. This would make no difference to C's interactions with Os. C would still need to respond to only one parameter, represented directly in experience, in order to interact with the Os. Any experiential representation of C would therefore be redundant in C's interactions with Os. This would be equally true of C's beliefs. C might believe that there was an O at a certain distance, and react accordingly, even if C did not perceive the O. Again, only one parameter need be involved in this. C would need only to believe that the O was near; C would not need to believe that the O was near to C. Since C's representations of C would be redundant, I shall say that C's representation of the spatial relation between C and an O is first-person redundant. (Prosser, 2015, p. 15)

Prosser continues to argue that it is the first person redundancy in the representations of the alleged 's-e relation' that grants those representations the power to motivate intentional actions:

This can be shown by a regress argument. Suppose that I am S and I stand in s-e relation *R* to *o*, and I wish to perform an action on *o* that depends on my relation to it. My mental representation of this relational state will typically be first-person redundant. For example, from my point of view, *o* might appear nearby and to the left. This might be because I see it

However, based on my analysis of Magidor's criticism of Lewis, I find an important insight concerning the identification aspect of de se thought in Lewis' account. This insight is captured by a combination of two doctrines in Lewis' property account of de se content. Lewis argues that the content of a de se belief is a property, which is not an eligible truth-bearer and does not have a truth value. To make the de se belief a proper truth-bearer (so that it can have a truth value), the believer must assign the property to a certain individual, which is the believer herself. Lewis named this operation the 'self-ascription'³⁷. So, we get the first doctrine we need for capturing the insight:

Doctrine 1: In believing a de se belief, the believer self-ascribes a property to herself.

The second doctrine has already been introduced during my analysis of Magidor's criticism of Lewis:

Doctrine 2: The believer who believes a de se belief is not represented in the

as nearby and to the left, or it might be because I merely believe this to be the case (perhaps it is dark, or *o* is invisible). This representational state gives me what I need in order to perform the action. But the equivalent non-redundant state (i.e. a state which is not first-person redundant) would not. (Prosser, 2015, p. 17)

³⁷ Holton characterizes the notion of 'self-ascription' as:

As I have already stressed, taking self-ascription as primitive is crucial to Lewis's account. We normally think of ascription as a two-place relation: one ascribes a property to a thing. Self-ascription would then be the special case where the thing is the self. But that won't do the work here. If the self is just thought of extensionally, then we would have no way to distinguish the belief that one's pants were on fire from the belief that the pants of someone, who is you though you don't realize it, are on fire. (...) So we have to think of self-ascription as a one-place relation: one simply self-ascribes a property. There is a real question whether we can make sense of this, but I'll postpone that till later. (Holton, 2015, p. 403)

Please note that Holton's characterization of self-ascription is compatible with the insight that I introduce later.

content of that de se belief.

Once we combine those two doctrines together, we capture an insight contained in Lewis' property account of de se content:

Insight: In believing a de se belief, we do not need to appeal to any representation of self to complete the self-ascription.

This insight is valuable, because it is one of the few detailed characterizations of 'self-ascription' that I can find in Lewis' theory. Cappelen and Dever also notice the importance of the concept 'self-ascription', and they correctly point out that David Lewis does not provide an adequate characterization of what 'self-ascription' is:

What we argue now is that even if one accepts that there is a special problem to be solved here, it is not the revisionary theory of content that does the solving, but the unexplained notion of self-ascription of a property. (...) But we can say more than this. In fact, it is self-ascription that is doing all of the work. The switch from propositions to properties as semantic values is completely inert. (Cappelen and Dever, 2013, p. 103 and p. 105)

Since self-ascription is an 'unexplained notion' in Lewis' own account³⁸, any

³⁸ Richard Holton's paper 'Primitive Self-Ascription: Lewis on the De Se' is a comprehensive study on the notion of 'self-ascription'. Holton also complains that not only other philosophers, but also (even) Lewis himself, do not provide adequate explanations on the notion of 'self-ascription':

The second part, which is much less to the fore in Lewis's presentation and in the subsequent discussion, involves treating our attitude to these properties as that of self-ascription. I hyphenate the term to stress that this is a primitive notion for Lewis. (...) I will argue that, as a result of underplaying its second part, the radical nature of Lewis's account has been largely missed. In particular, much recent literature has tried to incorporate his account simply by treating the objects of the attitudes as centered worlds, where a centered world is an ordered pair of a possible world together with a spatiotemporal location. But centered propositions just correspond to properties. So whilst they give us the right objects for the Lewis account, they go no way to providing the right attitude. (...) The idea of primitive self-ascription is an obscure one. (Holton, 2015, p. 400)

detailed characterization of ‘self-ascription’ is very valuable. The insight I have gained is not only valuable, but also important for my account of what it is to produce and think a de se thought, because it is a characterization of the identification aspect of de se thought. I think it can be generally acknowledged that in order to ascribe a property to a certain object, we must identify that object first³⁹, and it is not unusual that we appeal to certain representations of that object to identify it. However, here Lewis informs us that we ourselves are not represented in the content of our de se beliefs, which are by definition beliefs about ourselves, but we still self-ascribe certain properties to ourselves while believing de se beliefs. This leaves open two possibilities:

(1) Self-ascription does not involve identification (of the self)⁴⁰,

³⁹ Philosophers sometimes call this the Russell Principle, for example, Annalisa Coliva says

Indexicals and demonstratives allow us to think about particulars directly, by satisfying what is known in the literature on the topic as Russell Principle. Russell Principle: In order to think of an object one must know which object one’s thought is about. (Coliva, 2017, p.242)

⁴⁰ Recanati accepts the possibility (1). He distinguishes between two types of de se thought: the implicit de se thought and the explicit de se thought (Recanati, 2007, p. 260). The implicit de se thought is ‘identification-free (they ‘concern’ the subject without being about him or her)’ (Recanati, 2007, p. 260). This means, when entertaining a Recanatian implicit de se thought, I self-ascribe a property (or properties) without appealing to any representation of myself, but this self-ascription does not enable me to identify and locate myself. Besides the implicit de se thought, Recanati argues, there is also the explicit de se thought. When entertaining a Recanatian explicit de se thought, I ascribe a property (or properties) to myself in virtue of some representations of myself. This ‘ascription by means of representation’ grants the explicit de se thought an identificational function. So, Recanati argues that only by entertaining an explicit de se thought, I can identify and locate myself.

Recanati’s distinction between the implicit de se thought and the explicit de se thought has been criticized by García-Carpintero, who argues that Recanati fails to recognize the distinction between identification_P and identification_C (for an explanation of the distinction between identification_P and identification_C, please refer to Section 2.4.2). Recanati’s implicit de se thought is both identification_P-free and identification_C-free, and his explicit de se thought is both identification_P-involving and identification_C-involving. However, García-Carpintero argues that some de se thoughts are identification_P-free but identification_C-involving. See García-Carpintero, 2015, Section 1.

or

(2) Self-ascription does involve a certain type of identification (of the self), but identification of that type does not depend on any representation of self.

At this stage, it seems that we do not have sufficient reason to decide which possibility we should choose. In Section 2.3, I will argue that we should choose the second possibility.

2.3 Defenses of the Essential Indexical

In the previous section, I discussed Cappelen and Dever's criticism of the essential indexical thesis and Magidor's criticism of the centred world account of de se content. My discussion has uncovered two aspects of entertaining de se thought: the co-referential aspect and the identification aspect. I will discuss the co-referential aspect in the third chapter of this dissertation. With regard to the identification aspect, there is a question left unanswered:

Is it the case that (1) entertaining a de se thought does not involve an identification (of the self), or is it the case that, (2) entertaining a de se thought involves a special type of identification (of the self) which does not depend on any representation of the self?

In this section, I suggest that it is much easier to choose the second option, rather than the first option, because it is much easier to explain the connection between de se thought and the performance of intentional action if we accept the identification role of de se thought. To make this point manifest, I introduce Matthew Babb's and Juliana Lima's defenses of the essential indexical thesis. My analysis of their defenses shows

In Section 2.3, I will provide my own reasons for rejecting possibility (1) and accepting possibility (2).

that we should say 'Yes' to the second question above.

2.3.1 Matthew Babb on Indexical Intention

Matthew Babb argues that the indexicality is essential for the performance of intentional action, because the intention, based on which the intentional action is driven, is essentially indexical:

This is the Argument from Intentions: the Essential Indexicality Thesis holds because (1) intentions are essential to intentional action and agency and (2) intentions are essentially indexical. (Babb, 2016, pp. 455-456)

So, the next question to ask is why indexicality is essential for those intentions which drive our intentional actions. Babb argues that the key to answering this question is the referential role he associates with indexicality. According to his reading, in Perry's messy shopper case, the crucial realization which leads the messy shopper to change his behavior from pursuing the alleged messy shopper to rearrange the torn sack in his own cart, is that the shopper eventually succeeds in referring to the correct individual:

According to Perry, my realizing that I am the messy shopper is essential to explaining why I eventually acted the way I did, readjusting my sugar sack: my behaviour cannot be completely explained without indexical reference to myself. Not even a non-indexical, third-personal thought about me being the messy shopper (e.g., the thought that the messy shopper is Chester Fox, where 'Chester Fox' refers to me) seems to suffice for fully explaining my behaviour, since I may be unaware that the object of that thought (the thing referred to by 'Chester Fox') is me. That is the basics of Perry's argument for the essential indexicality of intentional action; indexicals are essential to intentional action because in order to fully explain such actions indexical reference must be made to the agent whose action it is. (Babb, 2016, pp. 441-442)

The indexical reference that Babb mentions in this paragraph is very similar to what I mean by the phrase ‘identification aspect of de se thought’. Here we may meet a potential objection: Babb does not use the phrase ‘de se’ or ‘first person’, but simply uses the term ‘indexical’. So, does this imply that he does not limit the scope of his discussion to the first person indexicals notion only, but rather includes all types of indexical and demonstrative notions into the scope of his discussion?

My reply to this objection is that, although he does not explicitly state that by ‘indexical’ he means the first person indexical, Babb explicitly takes ‘non-indexical’ to be the same with ‘third-personal’ in this paragraph when he says ‘not even a non-indexical, third-personal thought about me being the messy shopper’. So, it is highly likely that by ‘indexical’ Babb simply means something non-third-personal. It should be noted that the distinction which matters for Babb here is between the non-indexical and the indexical, or to put it in another way, the distinction between the third personal and the non-third-personal. Other distinctions within the range of indexical, for example, the distinction between the first person indexical and the second person indexical, does not matter for Babb. Babb continues to argue that even in the ‘Superman/Clark Kent case’, which Cappelen and Dever claim to be absolutely not indexical-involving, there is still an indexical identification that is implicitly involved in the shopper’s mind when he changes his behavior from looking for Clark Kent to chatting with Superman. In order to make this implicit identification explicit, we need to appeal to indexical notions. Babb says that:

My thought is more along the lines of that Clark Kent *as in that guy I just saw* is Superman. And it is this indexico-demonstrative reference to Clark that leads me to chase him down and tell him he is spilling sugar. Without this reference, I simply have a thought of the form *x is identical to y*, but this alone would not lead me to chase down an individual. I need to first relate this person to myself and my surroundings, and I do this, in thought, through an indexico-demonstrative use of ‘himself’. Thus indexical reference is involved in the Messy Superman case, albeit implicitly. (Babb, 2016, pp. 443-444)

I agree with Babb that a pure identity thought of the form, 'x is identical to y', by itself does not give us any clue how to identify either individual. If I have no idea about which individual Superman is when I am looking for Clark Kent, simply informing me that 'Clark Kent is Superman' does not give me any substantial help. However, if I know Superman as 'that guy I just saw' or 'the guy I met two years ago', informing me that 'Clark Kent is Superman' is then very useful for helping me in looking for Clark Kent. It is the indexico-demonstrative cluster that plays the identification role in this case, not the proper name itself. Therefore, in order to form an intention that can effectively drive an intentional action, Babb argues, we need to include some indexical components in that intention.

2.3.2 Juliana Lima on Principles of Action Explanation

Juliana Faccio Lima defends the essential indexical from a different angle. She argues that a plausible explanation of intentional action should follow at least two principles. The first principle is the following:

P3) If a set of (true) beliefs and desire explain a particular action X, then the desire is not satisfied by chance.

This principle states that in a correct explanation of an intentional action the desire cannot be satisfied by chance or as a byproduct of the action. (Lima, 2018, p. 7)

The second principle is:

P4) If a set of beliefs and desire explain an action X, then it cannot also explain a significantly different and available action Y.

The sense of 'significantly different' I am appealing to here entails that acting on different objects makes the action significantly different. (Lima, 2018, p. 8)

Based on those two principles, Lima argues that all non-indexical explanations of intentional actions fail, because they violate at least one of those two principles. She says:

In the remainder of this paper, I will argue that AIM, or any indexical-free model for that matter, does not offer correct explanations of intentional bodily actions because they violate at least one of the two principles I will propose here. (Lima, 2018, p. 7)

Why does an indexical-free explanation of intentional action fail to follow at least one of the two principles? Lima's answer to this question is that, if an explanation is indexical-free, it is also identificational-free:

An identificational element found only in indexical beliefs is necessary to motivate intentional action (Lima, 2018, p.1)

and

A different way of singling out indexical beliefs is to characterize them in terms of an identificational element: indexical beliefs are beliefs about an object in the believer's immediate environment (at least in non-controversial cases) which she cannot fail to take it to be about the correct object (a.k.a Immunity to Error through Misidentification – IEM). (Lima, 2018, p.3)

Here she establishes a close connection between indexicality and identificational elements, and she continues to argue as to why an explanation of intentional action without an indexical component (therefore without identificational element) violates the first principle of action explanation:

After all, when a person has only identification-free beliefs and desires, she is subject to misidentification, that is, her beliefs are about an object, but she might take it to be about a different object. In this case, two outcomes are possible: either Herman's action ends up targeting the right object (because he takes his beliefs and desire to be about the right object in his immediate

environment) or it does not (because he takes them to be about a different object in his immediate environment). (Lima, 2018, p. 10)

Lima here suggests that when conducting an object-orientated intentional action, an agent who lacks identificational information is threatened by the risk of misidentification. The crucial point is that according to an identification-free action explanation, the agent not only has to face the risk of misidentification, but also has no resources to overcome that risk. This leads to an unacceptable result: it is purely accidental that the agent's action ends up targeting the correct object. The agent's intentions and/or desires which drive her actions are satisfied only by chance, if at all. This result is in contradiction with the first principle that Lima proposes.

Now let us think about the misidentification case in another way. Suppose there are two scenarios. In both scenarios, Oliver wants to cuddle his daughter Emily. However, in the first scenario, Oliver successfully cuddles Emily, but in the second scenario, Oliver mistakenly cuddles Emily's friend Laura, because Oliver has no identificational information. I believe it can be generally acknowledged that in those two scenarios, the intentional actions that Oliver ends up performing are significantly different, because in the first scenario, Oliver cuddles Emily, but in the second scenario, Oliver cuddles Laura. However, Lima points out, those two significantly different actions can be explained by the exact same non-identificational explanation. This is because, what can differentiate the explanation of the second action from the explanation of the first action is that the explanation of the second action involves a misidentification. But this differentiation is impossible in a non-identificational explanation. So, if we stick to the non-identificational explanation, we have no way to distinguish between the explanation of the second action and the explanation of the first action. This result violates the second principle that Lima proposes.

I am not going to discuss all the details of Lima's argument here, but it is worth noting that Lima explicitly argues that in order to follow the two principles she

proposes, the indexical-identificational component must be included in any successful explanation of intentional action. One potential objection here is to doubt whether Lima's argument is a complete response to the question I raise at the end of Section 2.2.2 or not. We may get the impression that Lima's argument can be applied to object-orientated intentional action only, because in all the examples she uses, she argues that the indexical-identificational element is needed to locate the object of an intentional action properly. If an action is not object-orientated, for example, an action of jumping driven by the *de se* intention 'I want to jump', we may tend to think that this action is not included in the scope of Lima's argument.

But this impression is incorrect. According to contemporary linguistics, the sentence 'I want to jump' contains a pronominal placeholder PRO. So the content of the sentence 'I want to jump' is equal to 'I want PRO to jump', in which the 'PRO' can be roughly understood as the first person indexical 'myself'. I think the same applies to the agent's intention 'I want to jump': when I produce an intention 'I want to jump', I want myself to jump. The pronominal placeholder PRO can be taken as the indexical-identificational element required by Lima's argument. If we replace the pronominal placeholder PRO with a non-indexical-identificational element, for example, if we replace the PRO in the intention 'I want PRO (=myself) to jump' with 'Oliver', we get a new intention 'I want Oliver to jump'. Then we again will face those difficulties in following the two principles that I discussed above. However, this response essentially relies on the assumption that the pronominal placeholder PRO is a constituent of intention, and therefore it cannot convince people who believe the pronominal placeholder PRO appears only in linguistic expressions, but not in intentions. For those people, I can convince them in a different way. It is possible for an action to be non-object-oriented, but it is impossible for an action to be performed without a subject. So, an intention which drives the performance of an intentional action, be it object-oriented or not, must identify a subject as the performer of that action. In order to identify the subject, an indexical-identificational component must be included in the intention. For example, the 'I' in the intention 'I

want to jump'. If we replace it with a non-indexical-identificational component, for example, if we replace the 'I' in the intention 'I want to jump' with 'Oliver', we get a new intention 'Oliver wants to jump'. Then again, we have to face those difficulties in following the two principles that I discussed above. So, in a nutshell, the scope of Lima's argument covers not only object-oriented intentional actions, but also intentional actions driven by non-object-oriented de se intentions.

After examining Matthew Babb's and Juliana Lima's defenses of the essential indexical thesis, I think it is clear that philosophers have a consensus that if we reject the cognitive identification function of indexical thought, we have to face some difficulties in explaining our intentional actions. Instead, if we accept the cognitive identification function of indexical thought, it would be much easier for us to explain the intentional action. De se thought, of course, is a kind of indexical thought. This consensus gives us an important clue for answering the question I raised at the end of Section 2.2.2. The question is, which option shall we choose? (1) Self-ascription does not involve identification (of self), or (2) Self-ascription does involve a certain type of identification (of self), but identification of that type does not depend on any representation of self?

My analysis of Babb's and Lima's arguments enables me to choose the second option:

(2) Self-ascription does involve a certain type of identification (of self), but identification of that type does not depend on any representation of self.⁴¹

The consequence of choosing the second option is that, I agree I do not need to appeal to any representation of myself to ascribe properties to myself in entertaining

⁴¹ There may be a question about why this thesis is needed for my argument in this dissertation, and how this thesis is related to other parts of my argument. As one of my examiners points out, she/he has difficulty in understanding what exactly the proposal or the relevant distinctions were. However, once we realize that there is a connection between non-representative identification thesis in the second chapter and the co-reference independence thesis in third chapter, the question itself dissolves. The connection has been addressed when I discuss the analogy between the co-reference of intrapersonal de se thoughts and the co-reference of interpersonal 'I-You' communication in the third chapter. Please refer to Section 3.2.3 and Section 3.2.4.

a de se thought, and this self-ascription does involve a certain type of identification (of myself), but identification of that type does not depend on any representation of myself. Therefore, in a nutshell, the identification aspect of de se thought can be characterized in this way:

Entertaining a de se thought involves a special type of identification of the self which does not depend on any representation of the self.

2.4 De Se Thought and IEM

Besides the essential indexical thesis, there is another thesis concerning a special aspect of entertaining de se thought. This thesis is the immunity to error through misidentification thesis. Philosophers in general agree that at least a large portion of de se thoughts, if not all de se thoughts, have the epistemic property of immunity to error through misidentification (IEM). However, this consensus may be incompatible with the conclusion of Section 2.3.2, if the immunity thesis is interpreted in a specific way. What does it mean for a de se thought to be immune to the error of misidentification? One answer might be that (some) de se thoughts have that immunity property because entertaining those de se thoughts do not involve identification. This answer seems to be in contradiction with the claim that entertaining a de se thought involves a special type of identification which does not appeal to any representation of self, which the conclusion of Section 2.3.2. So how can we deal with this apparent contradiction? Does this apparent contradiction imply that for the sake of preserving the immunity property of some de se thoughts, we must reject the identification function of de se thought, and therefore break the connection between de se thought and the essential indexical thesis? My answer to this question is 'No'. In this section, I will show that de se thought's immunity property and its identification function are compatible with each other, and their apparent

contradiction can be resolved if we have a proper understanding of why *de se* thought is immune to the error of misidentification.

2.4.1 Immunity to Error through Misidentification

At the end of Chapter 1, I briefly mentioned the immunity to error through misidentification (IEM). There IEM was introduced as a semantic feature of the spatial indexical 'here'. But in this chapter, the immunity to error through misidentification is an epistemic thesis concerning the aboutness of our thought, which is different, although not totally separated, from the immunity feature I mentioned in the first chapter. So, I need to introduce the immunity thesis again. What is immunity to error through misidentification? An easy way to answer this question is first to define what the error of misidentification is. Sydney Shoemaker, who is believed to be the inventor of the terminology 'immunity to error through misidentification', claims that

To say that a statement "a is F" is subject to error through misidentification relative to the term "a" means that the following is possible: the speaker knows some particular thing to be F, but makes the mistake of asserting " a is F" because, and only because, he mistakenly thinks that the thing he knows to be F is what "a" refers to. (Shoemaker, 1984, pp. 7-8)

James Pryor argues the type of misidentification that Shoemaker describes in this paragraph is the so-called 'de re misidentification', because it is a mistaken equation between two different individuals, one is whatever the term 'a' refers to, and the other is the thing which is known to be F.⁴² James Pryor has a condensed but precise summary of what the de re misidentification is:

⁴² 'Here it seems that Shoemaker is envisaging an error of the sort we discussed in Examples 1 through 6: that is, an error of *de re* misidentification.' (Pryor, 1999, p. 286)

Say that we have a case of de re misidentification whenever the following three conditions obtain: (i) There is some singular proposition about x , to the effect that it is F , that a subject believes or attempts to express. (ii) The subject's justification for believing this singular proposition rests on his justification for believing, of some y , and of x , that y is F and that y is identical to x . (iii) However, unbeknownst to the subject, $y \neq x$. (Pryor, 1999, pp. 274-275)

If a singular proposition/thought/judgment is immune to the error of de re misidentification, it implies that the individual, which the proposition/judgement/thought is about, cannot be mistakenly identified for another individual. However, Pryor does not believe that 'immunity to error through misidentification' has been exhausted by 'immunity to error through de re misidentification', because he believes 'misidentification' has not been exhausted by 'de re misidentification'. Pryor argues that there is another type of misidentification, which he calls the 'which-object misidentification' or 'wh-misidentification' for short. He says:

Now, with this notion in hand, we can say that wh-misidentification occurs when: (i) A subject has some grounds G that offer him knowledge of the existential generalization $\exists xFx$. (ii) Partly on the basis of G , the subject is also justified, or takes himself to be justified, in believing of some object a that it is F . (iii) But in fact, a is not F . Some distinct object (or objects) y is F , and it's because the grounds G "derive" in the right way from this fact about y that they offer the subject knowledge that $\exists xFx$. (Pryor, 1999, p. 282)

According to Pryor, the possibility of 'which-object misidentification' lies in the difference (or split) between the evidence which supports the claim ' $\exists xFx$ ' and the evidence which supports the claim ' Fa '. If there is a situation in which the evidence that sufficiently supports ' $\exists xFx$ ' does not sufficiently support ' Fa ', and some additional evidence (together with the evidence that sufficiently supports ' $\exists xFx$ ') is needed to sufficiently support ' Fa ', then in that situation the claim ' Fa ' is prone to error through

wh-misidentification. However, if there is a situation in which the evidence that sufficiently supports $\exists xFx$ also sufficiently supports 'Fa', and no additional evidence is needed, then in that situation the claim 'Fa' is immune to the error through wh-misidentification. Pryor's official definition of immunity to error of wh-misidentification is

When that proposition is believed on grounds G, it is possible for those grounds to be defeated by undercutting evidence in such a way that the following two conditions hold: (i) the combination of G and that undercutting evidence no longer justifies you in believing of a that it is F; but (ii) the combination of G and that undercutting evidence could, by itself, offer you knowledge that $\exists xFx$. If it is not ever possible for G to be undercut in that way, then the belief is immune to wh-misidentification when justified by grounds G. (Pryor, 1999, pp. 284-285)

The distinction between de re (mis)identification and which-object (mis)identification is coined by Pryor. But this distinction is not unprecedented. I think it is anticipated in Gareth Evans' distinction between 'demonstrative identification' and 'identification-dependence'. Evans' distinction can be found in the following paragraphs:

When knowledge of the truth of a singular proposition, 'a is F', can be seen as the result of knowledge of the truth of a pair of propositions, 'b is F' (for some distinct Idea, b) and 'a = b', I shall say that the knowledge is identification-dependent: it depends (in part) on the second basis proposition, which I shall call the identification component. (Evans, 1982, p. 180)

and

The word 'identify' can do us a disservice here. In one sense, anyone who thinks about an object identifies that object (in thought): this is the sense involved in the use I have just made of the phrase 'demonstrative identification'.

It is quite another matter, as we saw, in effect, in 6.6, for the thought to involve an identification component for the thought to be identification-dependent.
(Evans, 1982, p. 218)

It may seem that Evans' distinction is a bit superficial compared with Pryor's distinction, because Evans merely appeals to the difference in the form of identification to draw the distinction (identification-dependence is of the equation form of 'a=b', while demonstrative identification is of the form of demonstration), but Pryor appeals to the difference of evidence which grounds the two types of identification to draw his distinction. Nevertheless, both Evans and Pryor are aware of the difference between two types of identification/misidentification, which is important for my subsequent discussions.

2.4.2 IEM and First Person Thought

In the previous section, I discuss two different types of immunity to error through misidentification. Both of them are epistemic properties of singular thought. In this section, I try to articulate philosophers' ideas about the relation between IEM and the first person thought. By 'first person thought', I mean the thought which is reported via the (singular) first person linguistic expression.

Although he did not use the terminology 'immunity to error through misidentification', Wittgenstein is believed to be the first philosopher in the history of modern philosophy to explore the relation between IEM and first person thought. In a famous paragraph from *The Blue Book*, Wittgenstein says:

There are two different cases in the use of the word "I" (or "my") which I might call "the use as object" and "the use as subject"....One can point to the difference between these two categories by saying: The cases of the first category involve the recognition of a particular person, and there is in these

cases the possibility of an error...It is possible that, say in an accident, I should feel a pain in my arm, see a broken arm at my side, and think it is mine, when really it is my neighbor's....On the other hand, there is no question of recognizing a person when I say I have toothache. To ask "are you sure it's you who have pains" would be nonsensical. (Wittgenstein, 1958, pp. 66–67)

Wittgenstein distinguishes the use of the term 'I' as object from the use of the term 'I' as subject, and he argues that the first person thoughts which involve the use of 'I' as object are prone to the error of misidentification (for example, the broken arm case), while other first person thoughts which involve the use of 'I' as subject are immune to the error of misidentification (for example, the toothache case). I think it can be generally agreed that when Wittgenstein talks about some first person thoughts 'being prone to the error of misidentification', his notion of misidentification is similar to the *de re* misidentification, because the example he uses in that paragraph involves a *de re* misidentification between a person whose arm is broken and me. But it is unclear which type of misidentification Wittgenstein is talking about when he talks about some first person thoughts 'being immune to the error of misidentification'.

After Wittgenstein, other philosophers also propose their own ideas concerning the relation between IEM and the first person thought. Let us look at a few of them:

To say that a term is immune to error through misidentification is to say that this sort of mistake is not possible. It is recognized nowadays that this needs to be supplemented by relativity to grounds. Almost any judgment that is IEM when based on one kind of ground is subject to error when based on another. Coming to know via introspection that someone has a headache doesn't leave open the room for the possibility that I may have misidentified the sufferer. But coming to know by looking at a CAT scan of a brain evidently undergoing a migraine does, because I may be wrong about whom the CAT scan belongs to. And IEM doesn't extend to other ways of referring to oneself. If I know via introspection that someone has a headache, then I know that I have a headache.

But I may not know that JI has a headache, even though I am JI. If I have amnesia, for example, or suffer from delusions about my identity. (Ismael, 2012, pp. 62-63)

A judgment “I am F” with the IEM property in virtue of the information on which it is based is not susceptible to a particular type of error. The person making the judgment cannot be mistaken about whom or what it is that is being judged to be F. Alternatively put, there is no gap between learning that someone is F and learning that one is oneself F. There is no such gap because these judgments are made in ways that do not involve identifying a particular person as oneself. They are, in Evans’s phrase, identification-free.

(Bermúdez, 2017, p. 53)

Nonetheless, some past-tense judgments are susceptible to error through misidentification, even when derived from autobiographical memories. We can illustrate this through a past-tense version of one of Wittgenstein’s examples from the Blue Book. Suppose that I was involved in an accident and on waking up in hospital I have a vivid memory of feeling a pain in my arm and seeing an injured arm by my side with a large bruise. This is plainly an autobiographical memory. It is suitably episodic and self-specifying. On the basis of that memory I form the judgment “My arm was bruised.” As it happens, however, the bruised arm I saw was not my own and in fact the pain in my arm came from a fracture rather than a bruise. This is clearly an error of misidentification. (Bermúdez, 2017, p. 137)

From those quotations, it can be easily discerned that Ismael seems to argue that the first person judgement based upon introspection is immune to the error through misidentification, although it is unclear which type of misidentification she has in mind, and Bermúdez seems to argue that judgements of the form ‘I am F’ is immune to the error of wh-misidentification, as long as there is no epistemic gap between knowing that ‘I am F’, and knowing that ‘there is someone who is F’. Therefore, based

on my analysis of Wittgenstein's, Ismael's and Bermúdez's writings, I think it is safe to say that many philosophers believe that there is a certain type of first person thought, although not all types of first person thought, which is immune to error through a certain type of misidentification.

The next question that needs to be answered is: why does that type of first person thought have the epistemic property of IEM? There is a very intuitive answer to this question: there is a certain type of first person thought, which is immune to error through a certain type of misidentification, because entertaining that type of first person thought does not involve a certain type of identification. To put it in a simpler way, some first person thoughts have the property of IEM, because they are identification-free. This answer is the alleged 'Simple Explanation/Simple Account' of the relation between first person thought and IEM, and it is a popular theory accepted by many philosophers. To demonstrate how popular this explanation is, we just need to look at a few quotations:

I do so by arguing that the best explanation of immunity to error through misidentification, whenever it arises and not just in the *de se* case, is what I call the Simple Explanation. The Simple Explanation just consists in the observation that a judgment will be immune to error through misidentification when it is not based on an identification. The Simple Explanation may not appear very exciting – or perhaps even very explanatory. However, what it says does at least seem true.' (Morgan, 2012, p. 105)

The Simple Account helps us to distinguish two senses for a thought to have an identification component. In the first sense ('identification_p'), the epistemic grounds for the thought include an identity-premise. In the second sense ('identification_c'), the thought includes a concept that identifies what it is about. On the Simple Account, having an identification_c component is compatible with a thought being IEM; for being IEM is lacking an

identification_P component, and thoughts having identification_C components may well be identification_P free — they might even be epistemically basic. (García-Carpintero, 2013, p. 202)

When philosophers claim that a given form of self-knowledge issues in first-person judgments with immunity to error through misidentification, it is normally in the service of showing that there is something special about that way of coming to know about ourselves. What is special is not merely that it keeps us safe from making this particular kind of mistake — that doesn't seem to be of any very great interest in itself. What is special about the forms of self-knowledge marked out by this epistemic feature, rather, is that it reveals them to be a way of knowing about oneself that is unmediated through any kind of identification. An identification, after all, would have brought with it the possibility of a misidentification. (Lea Salje, 2016, p. 129)

Morgan, García-Carpintero and Lea Salje all agree that some first person thoughts have the property of IEM because they are identification-free. Morgan and Salje do not specify which type of identification they have in mind when they assert that 'it is not based on an identification' and 'that is unmediated through any kind of identification'. But García-Carpintero clearly states he distinguishes two senses of 'being identification-free', which are identification_P-free and identification_C-free. Identification_P is similar to the de re identification, and identification_C is similar to which-object identification. García-Carpintero suggests that the two types of identification are not tightly connected, as he admits that a first person thought can be identification_P-free without being identification_C-free. I will come back to this point later.

So far everything seems to be fine. However, if we put the conclusion of Section 2.3.2 and the Simple Account together, we seem to face a contradiction:

(1). *Conclusion of Section 2.3.2*: entertaining a de se thought involves a special type of identification which does not appeal to any representation of self⁴³.

(2). *Simple Account*: there is a certain type of first person thought, which is immune to error through a certain type of misidentification, because entertaining that type of first person thought does not involve a certain type of identification⁴⁴.

It may seem like that we have to choose only one of the two claims, because they seem to be incompatible with each other: the first asserts that de se thought involves identification, but the second asserts that some first person thoughts do not require it. If we take de se thought to be identical with first person thought, then Simple Account is incompatible with the conclusion of Section 2.3.2. This apparent incompatibility in turn leads to more significant incompatibility, which is the incompatibility between the essential indexical thesis and the IEM thesis: it may seem that a de se thought cannot be essential for the performance of intentional action and also be immune to the error through misidentification at the same time, because the former presupposes an identification function, but the latter assumes an identification-free status. This is an unsatisfactory result which I cannot accept. So how can we avoid being troubled by this unsatisfactory result?

I argue that the alleged incompatibility between the essential indexical thesis and the IEM thesis is merely apparent. On the contrary, the two theses are perfectly compatible with each other. In fact, I am not alone in believing that those theses are compatible. Other philosophers also believe in their compatibility. Their strategy for justifying the compatibility is to appeal to a distinction that I mentioned above,

⁴³ Here the 'special type of identification' is not specified, but later in this chapter, I will show that many philosophers argue that the 'special type of identification' here is the wh-identification.

⁴⁴ Here the 'certain type of identification/misidentification' is not specified, but later in this chapter, I will show that many philosophers argue that the 'certain type of identification/misidentification' here is the de re identification/misidentification.

namely, the distinction between two senses of 'identification'. For example, Evans distinguishes between 'identification-dependence' (which is similar to *de re* identification and identification_p) and 'demonstrative identification' (which is similar to *wh*-identification and identification_c), and he reminds us of the danger of moving from the fact that there is no identification in the first sense (i.e. no identification-dependence) to a conclusion that there is no identification in the second sense (i.e. no demonstrative identification):

When knowledge of the truth of a singular proposition, "a is F", can be seen as the result of knowledge of the truth of a pair of propositions, "b is F" (for some distinct Idea, b) and "a = b", I shall say that the knowledge is identification-dependent: it depends (in part) on the second basis proposition, which I shall call the identification component. We might say that knowledge of the truth of a singular proposition is identification-free if it is not identification-dependent. (...) For our purposes, it will be more useful to define a narrower notion of identification-freedom: knowledge of the truth of a singular proposition is identification-free in the narrow sense if (i) it is not identification-dependent and (ii) it is based on a way of gaining information from objects. (Evans, 1982, pp. 180-181)

and

The word "identify" can do us a disservice here. In one sense, anyone who thinks about an object identifies that object (in thought): this is the sense involved in the use I have just made of the phrase 'demonstrative identification'. It is quite another matter, as we saw, in effect, in 6.6, for the thought to involve an identification component for the thought to be identification-dependent. There is a danger of moving from the fact that there is no identification in the latter sense (that no criteria of recognition are brought to bear, and so forth) to the conclusion that there is no identification in the former sense. (Evans, 1982, p. 218)

Besides Evans, García-Carpintero also adopts a similar strategy. He argues that a first person judgement can be identification_P-free without being identification_C-free. As my analysis in Section 2.3 shows, entertaining a de se thought enables me to pick out a certain individual, which is myself. So, I think it is safe to say the ‘special type of identification’ in the conclusion of Section 2.3.2 is the wh-identification/identification_C. García-Carpintero (and also Evans) argues that this is how his distinction between two senses of ‘identification’ provides a justification for the compatibility between the Simple Account and the conclusion of Section 2.3.2: the identification in the conclusion of Section 2.3.2 is of the identification_C/wh-identification type, but the identification (and misidentification) in the Simple Account is of the identification_P/de re identification type. A first person thought can be identification_P-free but identification_C-involving at the same time, so the conclusion of Section 2.3.2 and the Simple Account can be compatible with each other. This in turn guarantees the compatibility between the essential indexical thesis and the IEM thesis.

This strategy may seem to work well, but it is by no means perfect, as it faces a potentially fatal risk: because of its essential reliance on the ‘de re identification/wh-identification’ distinction (or other distinctions of the similar type), this strategy may simply collapse if the ‘de re identification/wh-identification’ distinction is challenged. Now, we know there are indeed some philosophically sophisticated challenges to the validity of ‘de re identification/wh-identification’ distinction.⁴⁵ If the distinction is proved to be invalid, this strategy loses its essential foundation. So, it is worthwhile exploring a new strategy which is not subject to that risk, i.e. does not rely on the ‘de

⁴⁵ For example, Annalisa Coliva challenges the ‘de re misidentification/wh-misidentification’ distinction in her 2006 and 2017 papers. She argues that the alleged ‘wh-misidentification’ can be reduced to the ‘de re misidentification’. For an extensive articulation of Coliva’s idea on this issue, please refer to the second section of her 2006 paper ‘Error through misidentification: some varieties’ (Coliva, 2006, pp. 409-415) and the fourth section of her 2017 paper ‘Stopping points: “I”, immunity and the real guarantee’ (Coliva, 2017, pp. 246-250).

re identification/wh-identification' distinction. In the next section, I will outline my own strategy for justifying the compatibility between those two theses.

2.4.3 IEM and De Se Thought

In the previous section, I highlighted the apparent incompatibility between the essential indexical thesis and the IEM thesis. In this section, I try to explain this apparent incompatibility away. The first step is to find out where the apparent incompatibility comes from. As I understand it, the incompatibility is caused by introducing the Simple Account to explain why some first person thoughts are immune to error through misidentification. To be more precise, it is the 'identification-free' condition in the Simple Account that causes the apparent incompatibility. If we can find a different explanation of the connection between de se thought and the IEM thesis, which, in contrast to the Simple Account, does not contain an identification-free condition, the apparent tension between the essential indexical thesis and the IEM thesis can be easily explained away. In fact, I have already found an explanation as such.

I call my explanation of the connection between de se thought and the IEM thesis the Definition Account. The crucial difference between the Simple Account and my Definition Account is that the former contains an identification-free condition, while the latter does not:

Definition Account: there is a certain type of first person thought, i.e. de se thought, which is immune to error through misidentification.⁴⁶, simply in virtue

⁴⁶ I do not specify which type of misidentification is involved here, the de re misidentification or the wh-misidentification. This is because the Definition Account does not rely on the distinction between de re (mis)identification and wh-(mis)identification. Later in this chapter, I will show that the Definition Account can cover both the two types of misidentification.

of the definition of de se thought.

In the Introduction of this dissertation, a de se thought was initially defined as ‘a thought I hold about myself as myself’. In Section 2.2.2, I characterized de se belief as: by believing a de se belief, a believer self-ascribes a property (or properties) to herself, but the believer herself is not represented in the content of the de se belief. In Section 2.3.2, I introduced another characterization of de se thought: entertaining a de se thought involves a special type of identification of the self that does not depend on any representation of the self. Based on those discussions, I define de se thought as

De se thought =_{df} a first person thought which is about its thinker and is only about its thinker. By entertaining this thought, its thinker is aware that this thought is about herself. In other words, the thought alone enables its thinker to self-ascribe a property (or properties).

In Section 2.4.2, I introduced my definition of the first person thought:

First person thought =_{df} a thought which is reported via a (singular) first person linguistic expression.

In the previous sections, I have shown that philosophers generally agree that not all types of first person thought are immune to the error of misidentification: some first person thoughts (i.e. some thoughts reported via first person linguistic expressions) have the epistemic property of IEM, but others do not have that property. However, philosophers have come to no agreement on how to draw the distinction between the first person thoughts which have the IEM property and the first person thoughts which do not have IEM property. According to my Definition Account, de se thought is the type of first person thought which has the IEM property, and other types of first person thought do not have the IEM property. This is because my definition of de se thought implies that it makes no sense to challenge whether a de se thought is about its producer/holder or not: as long as it is a de se

thought, it must be about, and only about, its producer/holder. If it is not about its thinker or not only about its thinker, it cannot be fully qualified to be a de se thought.

This definition alone guarantees that de se thought can be about nothing other than its thinker, and this in turn explains why de se thought is immune to error through misidentification. Since there is no 'identification-free' condition in the Definition Account, the Definition Account is entirely compatible with the conclusion of Section 2.3.2.⁴⁷ This compatibility in turn resolves the apparent incompatibility between the essential indexical thesis and the IEM thesis.⁴⁸ Both of them can be assigned to de se thought at the same time without any tension. Compared with the Simple Account, the Definition Account may seem to be boring, but it does not have the theoretical commitments that the Simple Account has.⁴⁹ Therefore, it is not exposed to the risk that the Simple Account is exposed to.

However, it can be anticipated that there are at least two potential challenges to the Definition Account. The first challenge focuses on my unusual distinction between first person thought and de se thought. It can be discerned that the validity of the Definition Account essentially depends on the 'first person thought/de se thought' distinction, but this distinction may seem to be suspicious: for many philosophers, de se thought and first person thought are identical. Philosophers rarely discuss the possibility that there is a distinction between de se thought and

⁴⁷ The conclusion of Section 2.3.2 is

Entertaining de se thought involves a special type of identification of the self which does not depend on any representation of the self.

⁴⁸ This is because the essential indexical thesis implies the conclusion of Section 2.3.2, and the Definition Account is an explanation of the IEM thesis. For a general description of the essential indexical thesis, please refer to Section 2.1.

⁴⁹ The Simple Account appeal to the distinction between de re (mis)identification and wh-(mis)identification to account for the compatibility between the essential indexical thesis and the IEM thesis. For a detailed argument on this point, please refer to Section 2.4.2. The Definition Account does not need to appeal to that distinction to account for that compatibility.

first person thought. But I argue that although all de se thoughts are first person thoughts, not all first person thoughts are de se thought. If there were no distinction between first person thought and de se thought, the Definition Account would simply fail.

The second challenge is connected to the first challenge, but it has a different focus. It focuses on what is the bearer of the IEM property. In the previous sections, I have shown that philosophers use various expressions to designate what they believe to be the bearer of the IEM property: first person judgement, I-thought, first person thought, de se thought, and even self-knowledge. So, I need to explain why I choose de se thought, rather than first person thought, to be the bearer of the IEM property.

The two challenges are closely related. So, my response will cover both of them. I argue that according to my definition of de se thought, there are at least four ways in which a thought can fail to be a de se thought:

First, a thought can fail to be a de se thought if it is not 'about oneself', i.e. it is not about its thinker.

Second, even if a thought is in fact about its thinker, it can still fail to be a de se thought if it is not reported via the first person linguistic expression. In other words, even if a thought is in fact about its thinker, it can still fail to be a de se thought if it is not a first person thought.

Third, even if a thought is about its thinker and is reported via first person linguistic expression, it can still fail to be a de se thought if its thinker fails to recognize that the thought is about herself.

Fourth, even if a thought is about its thinker, and it is reported via first person linguistic expression, and its thinker succeeds in recognizing that the thought is about herself, it can still fail to be a de se thought if it is also about other individuals at the same time, e.g. if it is ambiguous or it partially refers to two or

more objects.

The first and the second ways of failing to be a de se thought are plausible and familiar, but the third and the fourth ways of failing to be de se thought may seem to be very obscure: how can someone fail to recognize that a first person thought in her mind, which is in fact about her, is about herself? how can a singular first person thought be about its current thinker and other individuals at the same time? In the next section I will show that those are indeed possible.

It is clearly discernible from the four ways of failing to be de se thought that de se thought is a subset of first person thought. As the third and the fourth ways of failing to be de se thought show, a thought can be first person thought without being de se thought. I call a thought like this a *non-de-se first person thought*. A non-de-se first person thought is reported via the first person linguistic expression, but it is either about more than one individual, or taken to be about an individual who is not its thinker. The existence of 'non-de-se first person thought' justifies my distinction between de se thought and first person thought. This is my response to the first challenge.

A possible misunderstanding of the Definition Account is that it only explains why de se thought is immune to the error of wh-misidentification, but does not specify whether de se thought is immune to the error of de re misidentification. This is incorrect. In the next section, I will argue that the production of non-de-se first person thought can involve both the de re misidentification and the wh-misidentification concerning the thinker of the thought. Once a first person thought is affected by the error of misidentification, be it the de re misidentification or the wh-misidentification, it will no longer be fully qualified to be a de se thought. In this way, the Definition Accounts explains how we can exclude first person thoughts that are affected by the error of misidentification from the domain of de se thought. Since de se thought is immune to errors of misidentification, and non-de-se first person thought is not immune to those errors, I choose de se thought to be the bearer of

the epistemic property of IEM, rather than first person thought. This is my response to the second challenge.

2.4.4 The Defense of the Definition Account

My responses to the two challenges can be properly vindicated only after I clearly demonstrate the existence of non-de-se first person thought and the connection between non-de-se first person thought and the error of misidentification. Therefore, in this section, I will first show that some first person thoughts are not qualified to be de se thoughts because of the cognitive impact of wh-misidentification, and then I will proceed to show that some first person thoughts are not qualified to be de se thoughts because of the cognitive impact of de re misidentification. Those non-de-se first person thoughts associated with wh-misidentification are related to the third way of failing to be de se thought that I described in Section 2.4.3, and the non-de-se first person thoughts associated de re misidentification are related to the fourth way of failing to be de se thought.

Can a first person thought, i.e. a thought reported via the (singular) first person linguistic expression, be cognitively impacted by a wh-misidentification concerning which individual the thought is about? Some philosophers' answer is 'No'. For example, proponents of the reflexive account of first person thought claim that a first person thought is immune to the error of wh-misidentification concerning its subject. This is because, according to the reflexive account, the individual about which a first person thought is held is solely determined through a reflexive rule. The rule is: a first person thought is about the individual who produces and thinks that thought. Daniel Morgan describes James Higginbotham's reflexive account as

Higginbotham's reflexivist account of de se thoughts is a distinctive version of the familiar Fregean idea that to think a de se thought is to think about

oneself using a special first-person mode of presentation. The distinctiveness lies in the account Higginbotham gives of the special mode of presentation. According to reflexivism, being in a state that involves thinking about oneself using the first-person mode of presentation is being in a reflexive state, i.e. a state that involves thinking about oneself as the subject of that very state. (Morgan, 2012, pp. 106-107)

And Lucy O'Brien describes the reflexive account as

We have supposed that a subject's awareness of her use of 'I', with mastery of SRR, immediately warrants knowledge that she is using 'I', and in so doing warrants knowledge that she is referring to herself. The explanation of such a warrant is something we will come back to. (O'Brien, 2007, pp. 80-81)

According to the reflexive account, a first person thought is solely about the person who produces and entertains the thought. It seems that the reflexive characterization of first person thought does not leave any room for possible cognitive impact from a wh-misidentification concerning about which individual that thought is held. But I argue that this is not the case.

The central point of my argument is that, even if I am entertaining a first person thought characterized by the reflexive account, I do not necessarily recognize myself as the thinker of that thought. As long as I do not recognize myself as the thinker of that first person thought, that thought is not a *de se* thought for me. This point, according to García-Carpintero, has been accepted by several philosophers, including O'Brien and Recanati. García-Carpintero says

(O'Brien and Recanati) Both consider the case of a schizophrenic patient who believes that some mental states he is conscious of are not really his but are the thoughts of some other person that have somehow been implanted in him. In Recanati's example, the patient is conscious of the judgment that one

would express by accepting “I am good and omnipotent”; he thereby comes to accept that the owner/judger of this state is good and omnipotent, but he does not accept that he is the owner of that thought—it is rather “the other” who has implanted it. (García-Carpintero, 2015, Section 1: De Se Thoughts: Two Conflicting Accounts)

The schizophrenic patient succeeds in recognizing that there is a first person thought ‘I am good and omnipotent’ in his mind, and he also agrees that the first person thought is about its producer/thinker. So, it is safe to say that the schizophrenic patient endorses the reflexive account of first person thought. However, the schizophrenic patient neither recognizes himself as the producer of the first person thought, nor takes himself to be willingly entertaining that thought. He insists that the first person thought is produced by ‘the other’⁵⁰. ‘The other’ implants

⁵⁰ The thought insertion that I discuss here is the insertion of first person thought. Lisa Bortolotti and Matthew Broome describe the first person thought insertion as

(...)we want to explore this relationship within psychopathology and provide an analysis of thought insertion according to which subjects affected by this condition lack self knowledge. We argue that this failure of self knowledge consists in failing to ascribe to oneself a thought that is accessed first-personally and in failing to either endorse the content of that thought with reasons or manifest commitment towards it in behaviour. (Bortolotti and Broome, 2009, pp. 205-206)

and they provide an example of the report of first person thought insertion:

2. “One evening one thought was given to me electrically that I should murder Lissi.” (Jaspers 1963, page 580) (Bortolotti and Broome, 2009, p. 215)

Michael Sollberger also discusses the first person thought insertion and argues

In particular, it will be argued that there is nothing wrong with the idea of being first-personally aware of a thought which is presented in consciousness as being someone else’s. (Sollberger, 2014, p. 590)

However, besides the first person thought insertion, there is another type of thought insertion, i.e. the insertion of non-first-person thought. Analisa Coliva describes the non-first-person thought insertion as

the first person thought in to his mind and he cannot help thinking about it. Therefore, the patient believes that the first person thought ‘I am good and omnipotent’ is about ‘the other’, not about himself. This belief alone, argues García-Carpintero, makes the thought ‘I am good and omnipotent’ not qualified to be a de se thought for the patient. García-Carpintero says,

We should distinguish between two different mental states. There is, in the first place, the judgment about oneself—to the effect that one is good and omnipotent—made by being in a state whose significance is captured by the token-reflexive interpreted form the owner (judger) of this very state is good and omnipotent. There also is the impression about oneself—to the effect that one is judging oneself to be good and omnipotent—had by being in a state whose significance is rather captured by the token-reflexive form the owner of this very state appears to be judging herself to be good and omnipotent. Normally, when one is in the latter belief-state, one is thereby also in the former, but this is not so in the case of the schizophrenic patient. Recanati is right that to be in a state with a reflexive content is not sufficient to ground a self-ascription’ (García-Carpintero, 2015, Section 1: De Se Thoughts: Two Conflicting Accounts)

I agree with García-Carpintero on his claim that ‘Recanati is right that to be in a state with a reflexive content is not sufficient to ground a self-ascription’. In the

Notice, however, that for this report to be a counterexample to the thesis that present-tense introspectively based self-ascriptions of mental states are logically IEM, it should be expressed in the first person. But, interestingly, the subject is not saying “I am thinking ‘Kill God.’” Rather, she is saying that Chris is thinking the thought, which is occurring in her mind. So, if anything, this report would show that it is conceivable to make third-person psychological ascriptions on the basis of introspective knowledge of someone else’s mental states. (Coliva, 2002, p. 30)

In this chapter, I focus on the discussion of first person thought insertion. The non-first-person thought insertion is not in the scope of my discussion.

schizophrenic patient case, the patient is aware that there is a first person thought in his mind, and he accepts the reflexive account of first person thought, but nevertheless, by entertaining the thought 'I am good and omnipotent' he does not self-ascribe the property of *being good and omnipotent*. Instead, he ascribes the property of *being good and omnipotent* to 'the other', because he believes 'the other' produced that thought and implanted it into his mind. In this sense, the schizophrenic patient makes an error of wh-misidentification: he misidentifies the producer/thinker of the first person thought. We all know that in fact he is the producer/thinker of the first person thought, but he himself does not think so. The cognitive impact of this wh-misidentification is that the first person thought fails to be a de se thought for the patient in the third way that I described in Section 2.4.3. For this patient, the thought 'I am good and omnipotent' is a typical non-de-se first person thought.

The next step is to show some first person thoughts that are cognitively impacted by the de re misidentification also fail to be de se thought in the fourth way that I described in Section 2.4.3. In other words, I need to demonstrate the existence of this possibility: some first person thoughts are not only about their thinkers, but also about other individuals at the same time. Let us imagine a scenario like this:

I was sitting in a room with many people. There was a CCTV on the ceiling of that room, and a large screen in front of us displaying what was being filmed by the CCTV. So, people in the room could see everyone's head on that screen. My colleague Oliver, sitting next to me, suddenly told me, 'Look at the screen. There is a pigeon feather on Smith's head.' I did not know how to locate Smith in that room, and I did not know how to locate Smith on that screen. But I was so curious that I turned to look at the screen. I noticed that there was a person, on whose head there was a small piece of feather. So, I thought, 'That person has a feather on his head, so that person must be Smith.' But as my observation continued, I suddenly had an epiphany that the person on whose head there was a feather was me! (Suppose that the feather is too light to be felt through bodily

sensation) So, I thought, ‘That person is me, so I have a feather on my head.’
 Because of my understanding of this thought, I tried to get rid of the feather on my head using my hands. Then it turned out that my hands did not touch anything, so I found out there was no feather on my head. After this discovery, I returned to my second thought ‘That person has a feather on his head.’

I call this scenario the ‘Screen Scenario’. In the Screen Scenario, I have three different thoughts and three updates of information, which can be listed as below:

Time 1	First Thought: ‘There is a pigeon feather on Smith’s head.’
Update of Information 1	Watching the screen
Time 2	Second Thought: ‘That person has a feather on his head, therefore that person must be Smith.’
Update of Information 2	Watching the screen and epiphany
Time 3	Third Thought: ‘That person is me, so I have a feather on my head.’
Update of Information 3	Touching the top of my head by hands
Time 4	Second Thought*: ‘That person has a feather on his head.’

Let us compare my second thought at Time 2 with my third thought at Time 3. At Time 2, I draw an equation between a name-like content (Smith) and a demonstrative content (that person), but at Time 3, I draw an equation between a demonstrative content (that person) and a, in Coliva’s terminology, first person content (I).⁵¹ This is not a trivial difference.

⁵¹ If we stick to the Lewisian account of de se content which I introduced in Section 2.2.2, the thought

According to Annalisa Coliva's theory of 'stopping point of inquiry', the two equations have different cognitive impacts: Coliva argues that an indexical content is not a further clarification of a demonstrative content (and *vice versa*), but a demonstrative/indexical content is indeed a further clarification of a name-like content. She says:

I do think there are irreducibly first-personal contents. They are not entirely unique in what makes them special, though. They share what makes them distinctive with other indexical contents – particularly with here and now contents – as well as with genuinely demonstrative contents. The aim of this paper is to bring out exactly what makes them special, by showing that they perform a distinctive cognitive function. Namely, they are stopping points of inquiry. (Coliva, 2017, p. 234)

For Coliva, indexical content and demonstrative content are the stopping points of inquiry. This is their 'distinctive cognitive function', which implies that other types of content, for example the name-like content, do not have this function. So, what does the 'stopping point of inquiry' mean? Coliva argues that the stopping point of inquiry marks the end of a certain type of inquiry. When doing that type of inquiry, we can conduct multiple layers of clarification concerning the subject matter of inquiry. As long as I still have some ignorance regarding the subject matter, I can continue to do the relevant clarification. But once I arrive at the stopping point of inquiry, I have no room for further clarification, so the inquiry must be stopped. This is because the certain type of ignorance concerning the subject matter of the inquiry has been blocked by the stopping point. In a nutshell, the indexical/demonstrative content blocks a certain type of ignorance concerning the subject matter of the inquiry,

'that person is me' at Time 3 should be understood as ascribing the relational property of *being identical to each other* to the two individuals, me and that person. The only difference is that I self-ascribe the relational property without appeal to any representation of myself, but I ascribe the relation property to that person by means of the representation that I obtain by observing his image on the screen.

allowing no room for any further clarification. Coliva provides a detailed articulation of this point:

Nevertheless, here is the simple idea: once you entertain an I-content like ‘I am F’ (and I will focus on propositional contents for convenience in the rest of this paper, if not explicitly indicated otherwise), there is no room for a certain kind of question. Namely, assuming the proposition ascribes a certain property to the individual, it makes no sense to ask ‘Who does (or seems to) have that property?’; or, equivalently, ‘Which person does (or seems to) have that property?’. To put it otherwise, once you entertain a first-person content, you have reached a stopping point of inquiry. If you are entertaining a genuinely first-personal content, you automatically know that it is you that have (or seem to have) the property in question and there is no room left for an inquiry concerning who has that property. (Coliva, 2017, p. 235)

In this paragraph, Coliva specifies the type of inquiry for which the indexical/demonstrative content can be the stopping points. This type of inquiry is the inquiry of locating an individual, or more precisely, it is the inquiry of locating an individual which is the bearer of a certain property (or properties). For example, if I was asked the question, ‘Which one is Prof. Tsai’s tallest student?’, a question about the bearer (which one) of a property (being Prof. Tsai’s tallest student), I could answer the question by saying ‘Emily is Prof. Tsai’s tallest student.’ It is highly likely that this answer does not clear up all the relevant ignorance in the questioner’s mind. This answer may not enable the questioner to locate Prof. Tsai’s tallest student in her egocentric spatial-temporal system. So, this answer allows room for further clarifications. The questioner can continue to ask, ‘If so, which one is Emily?’ I can answer this question by pointing to a person over there and saying, ‘That person is Emily’ or ‘She is Emily.’ This answer is a clarification to the previous answer, but this answer itself does not need further clarification, because it enables the questioner to locate ‘Emily’ (or ‘Prof. Tsai’s tallest student’) in her egocentric spatial-temporal system. In fact, Coliva’s position here is even stronger: this answer allows no room for

further clarification. This is because, once the egocentric locating is done, it does not make sense for the questioner to continue to ask, 'If so, which one is that person/which one is she?'. An indexical/demonstrative content thoroughly cures the questioner's ignorance about the egocentric spatial-temporal location of an individual which is the bearer of a certain property, blocking the possibility of further clarification. That is why Coliva argues that the indexical/demonstrative content is the 'stopping point of inquiry':

To repeat, what these samples collectively show, if we focus on the passages in italics, is that homing in a first-person thought (or indeed a demonstrative one as well, like in Perry-2) is the stopping point of one's inquiry. For it resolves, once and for all, one's previous ignorance regarding who was supposed to have the properties ascribed by a given (list of) predicative concept(s) in the relevant (list of) proposition(s). If that is right, then no wonder that first-personal, or more generally indexical and genuinely demonstrative contents, have been taken to play a fundamental role in our cognitive lives. For they actually allow us to locate ourselves spatially, temporally, and with respect to other objects, even when we cannot locate ourselves and our spatio-temporal location in objective – that is to say, non-perspectival – ways.¹ Conversely, the passages reported show that, just by itself, the objective individuation of ourselves and our whereabouts is not sufficient to secure that we know which person, place and time we are thinking about. (Coliva, 2017, p. 236)

In this paragraph, Coliva mentions 'locating ourselves and our spatio-temporal location in objective ways' and 'objective individuation of ourselves and our whereabouts'. I believe those are the cognitive roles performed by the non-egocentric name-like content. If so, the non-egocentric name-like content is exactly what needs to be clarified by the egocentric indexical/demonstrative content. The purpose of clarification is, in Coliva's terms, 'to secure that we know which person, place and time we are thinking about'.

There is one important point that I would like to emphasize: for Coliva, an indexical/demonstrative content should not be understood as a further clarification of another indexical/demonstrative content. Coliva's arguments in previous quotations show that the indexical content and the demonstrative content are equally effective at being stopping point of inquiry. When talking about the stopping point of inquiry, Coliva always mentions both the indexical content and the demonstrative content at the same time, which implies that she does not make further distinctions between indexical and demonstrative contents. It should be noted that Coliva never picks out any specific type of indexical (or demonstrative) content and claims that it is a more radical stopping point of inquiry. So according to Coliva, indexicals and demonstratives are equal in the function of being the stopping point of inquiry:

They share what makes them distinctive with other indexical contents – particularly with here and now contents – as well as with genuinely demonstrative contents. (Coliva, 2017, p. 234)

and

Homing in a first-person thought (or indeed a demonstrative one as well, like in Perry-2) is the stopping point of one's inquiry. (Coliva, 2017, p. 236)

and

Only I-contents (and more generally indexical and demonstrative ones) can perform the function of stopping points of inquiry. (Coliva, 2017, p. 237)

and

The real guarantee, enjoyed by first-person contents, or indeed, *mutatis mutandis*, by indexical and genuinely demonstrative ones (...) (Coliva, 2017, p. 238)

I summarize Coliva's doctrines of 'stopping point of inquiry' like this:

Doctrine 1: the name-like content can be further clarified by the indexical/demonstrative content with regard to which individual they are about, but not *vice versa*.

Doctrine 2: an indexical/demonstrative content cannot be further clarified by another indexical/demonstrative content with regard to which individual they are about.

Those two doctrines are epistemic claims about the clarification as being veridical. Coliva clearly argues that the clarifying demonstrative/indexical content is a better guide to the fact about the reference of the name-like content that is clarified. So, from these two epistemic doctrines, we can gain two semantic doctrines:

Doctrine 3: if a content A is further clarified by a content B, the aboutness/reference of content B overrides the aboutness/reference of content A.

Doctrine 4: If a content A is not clarified by a content B, the aboutness/reference of content B does not override the aboutness/reference of content A.

So, being informed by the four doctrines, we are now able to examine the Screen Scenario. Before the first update of information, I have a thought about a person called Smith. After the first update of information, I produce a new thought about the individual to whom I refer as 'that person' and I construct a de re identification between that person and Smith. According to the Doctrine 1, the cognitive impact of this de re identification is that my thought 'That person has a feather on his head' at Time 2 becomes a further clarification of my thought 'There is a pigeon feather on Smith's head' at Time 1. Then, According to Doctrine 3, the aboutness/reference of the thought 'That person has a feather on his head' overrides the aboutness/reference of the thought 'There is a pigeon feather on Smith's head'.

However, this characterization of the cognitive impact of de re identification does not work for the de re (mis)identification between that person and me, which takes

place at Time 3. This is because, according to Doctrine 2, the de re (mis)identification at Time 3 (that person is me) does not indicate any clarification relation. Therefore, the thought 'I have a feather on my head' at Time 3 is not a further clarification of the thought 'that person has a feather on his head' at Time 2, and *vice versa*. According to Doctrine 4, the aboutness/reference of the thought 'I have a feather on my head' does not override the aboutness/reference of the thought 'That person has a feather on his head'.

If so, what is the real cognitive impact of the (mis)identification 'that person is me' at Time 3? According to Doctrine 2, this (mis)identification does not indicate any clarification relation. I (mis)identify that person as myself, just as I (mis)identify myself as that person. To put it in Lewis' terms, I ascribe the relational property of *being identical to each other* to both the two individuals, that person and me, at the same time, although the property is ascribed to them in different ways (I self-ascribe the relational property without appealing to any representation of myself, but I ascribe the relational property to that person by means of my representation of him). As I said, the aboutness of the thought 'I have a feather on my head' does not override the aboutness of the thought 'that person has a feather on his head', and *vice versa*. To put it in Lewis' terms, my self-ascription of the property *having a feather on the head* does not override my ascription of that property to that person, and *vice versa*. Consequently, at Time 3, the thought 'I have a feather on my head' bears more than one aboutness: by entertaining this thought, I not only self-ascribe the property of *having a feather on the head*, but also ascribe the property of *having a feather on the head* to that person at the same time. But at Time 3, I am not aware of this double-aboutness: if I was asked at Time 3, 'which individual is your thought about? Yourself, or that person whose head you see on the screen?', my answer to this question would be, 'both, because they are the same individual'. In a nutshell, because of the de re misidentification between that person and me, or the mis-ascription of the relational property of *being identical* to that person and me, my thought 'I have a feather on my head' is about two individuals at the same time,

although I do not realize this at Time 3.

Therefore, the real cognitive consequence of introducing the de re (mis)identification between that person and myself is that I produce a non-de-se first person thought. At Time 3, thought I express with 'I have a feather on my head' is both about me and that person at the same time. This indicates that this first person thought fails to be a de se thought in the fourth way that I described in Section 2.4.3. So, it is a non-de-se first person thought. The Screen Scenario demonstrates the existence of non-de-se first person thought caused by de re misidentification.

Now I have shown that first person thoughts affected by the wh-misidentification fail to be de se thoughts in the third way that I described in Section 2.4.3, and first person thoughts affected by the de re misidentification fail to be de se thoughts in the fourth way. The existence of non-de-se first person thought vindicates my distinction between de se thought and first person thought, and as I said at the beginning of this section, this distinction is exactly what I need to defend my Definition Account against the two potential challenges. Since my responses to those challenges are properly supported by the distinction, my Definition Account of the connection between de se thought and immunity to error through misidentification is vindicated.

2.4.5 A Possible Objection

In this section, I reply to a possible objection to my Definition Account of the relation between de se thought and immunity to error through misidentification.

This possible objection argues that according to the Definition Account, the connection between IEM thesis and de se thought is simply a corollary of the definition of de se thought: a de se thought is immune to the error of misidentification, as long as it is a de se thought. To challenge whether a de se

thought is about and is only about its thinker is to challenge whether it is a de se thought at all. In other words, we cannot challenge whether a de se thought is about and is only about its thinker without challenging whether it is a de se thought. Therefore, according to the Definition Account, attributing the IEM property to de se thought does not impose any additional constraint on the account of de se thought. In this sense, the connection between de se thought and IEM thesis is philosophically boring. But many philosophers have a much higher expectation: they expect that a proper explanation of the connection between de se thought and IEM thesis should impose some additional constraints on the account of de se thought.⁵² In other words, they expect the connection between de se thought and IEM thesis to be philosophically interesting. The objection against my Definition Account accuses the Definition Account of failing to fulfill those philosopher's expectation.

My response to this objection is that we should first evaluate whether those philosophers' expectation is reasonable or not. Daniel Morgan is deeply skeptical about those philosophers' expectation. He says,

It may be worth stepping back a little and asking whether there is any good reason to expect an explanation of fp-immunity, or of any other kind of immunity to error through misidentification, that goes beyond the Simple Explanation. What is it about immunity to error through misidentification that makes it look as though it is crying out for some nonobvious explanation? Very often, one's sense of what needs to be further explained, and what can be taken for granted, will partly depend on one's sense of what it might be possible to further explain. Is there something about fp-immunity that makes it look as though it might be possible to provide a deeper explanation of it than the Simple Explanation does? (Morgan, 2012, p. 123)

Morgan here expresses his skepticism towards the plausibility of the expectation

⁵² For more discussions on this point, please refer to Daniel Morgan's paper 'Immunity to error through misidentification: what does it tell us about the de se?' (Morgan, 2012, pp. 104-123)

of *deeper explanation*. By ‘expectation of deeper explanation’, he means the expectation that an explanation of the connection between de se thought and IEM thesis can impose tight constraints on the account of de se thought. He asks: where does this expectation come from? Is it plausible to hold this expectation? But unfortunately, he does not provide answers to these questions. Morgan argues that instead of a deeper explanation, the connection between de se thought and the IEM thesis can be properly explained by a shallow explanation. The shallow explanation that he has in mind is the Simple Explanation, which I introduced in Section 2.4.2:

Explaining fp-immunity does not require us to accept that de se thoughts are reflexive (as Higginbotham claims). Neither does it require us to accept that some de se thoughts are implicit (as Recanati claims). The most plausible explanation of fp-immunity – the Simple Explanation – imposes no tight constraints on an account of the de se. Constraints on such an account will need to be gathered from somewhere else. (Morgan, 2012, pp. 122-123)

The Simple Explanation is philosophically shallow because according to the Simple Explanation, attributing the IEM property to de se thought does not impose tight constraints on the account of de se thought. However, as my analysis in Section 2.4.2 shows, the Simple Explanation contains an *identification-free* condition, and it relies on the distinction between de re (mis)identification and wh-(mis)identification. Those theoretical commitments can still be counted as *constraints* on the account of de se thought, although they may not be tight constraints.

I am sympathetic to Morgan’s skepticism concerning the expectation of deeper explanation, and I agree with Morgan that attributing the IEM property to de se thought does not impose tight constraints on the account of de se thought. However, I think Morgan’s Simple Explanation is still too deep, or in other words, it is not shallow enough. The theoretical commitments that the Simple Explanation carries

make it vulnerable to certain philosophical challenges⁵³. Furthermore, the Simple Explanation does not provide answers to Morgan's questions about where the expectation of a deeper explanation comes from and whether the expectation is plausible or not.

My Definition Account does not carry the theoretical commitments that the Simple Explanation carries, so it is not vulnerable to those challenges that the Simple Explanation is vulnerable to. This is an advantage of my Definition Account. According to my Definition Account, the connection between IEM thesis and de se thought is simply a corollary of the definition of de se thought. So, the Definition Account does not impose any additional constraint on the account of de se thought, let alone *tight* constraints. In this sense, it is even shallower than the Simple Account. This is the second advantage of my Definition Account. The third advantage of my Definition Account is that it answers Morgan's questions about where the expectation of a deeper explanation comes from and whether or not the expectation is plausible. My observation is that in general, philosophers who have the high expectations of a 'deeper explanation' do not recognize the existence of non-de-se first person thought. So, they do not recognize my distinction between de se thought and first person thought. For them, there is no difference between de se thought and first person thought. This is not a trivial point.

From Section 2.4.2 to Section 2.4.4, my analysis shows that although the connection between de se thought and IEM thesis is philosophically boring, the connection between first person thought and IEM thesis is philosophically interesting. To be more precise, the question of why some first person thoughts are immune to the error of misidentification and some first person thoughts are not immune to that error, is philosophically interesting. I have introduced some philosophers' answers to this question in Section 2.4.1 and Section 2.4.2⁵⁴. My own

⁵³ For more discussions on those possible challenges, please refer to the end of Section 2.4.2.

⁵⁴ Please refer to Section 2.4.1 and Section 2.4.2. There I introduce Wittgenstein's distinction between

Definition Account can be seen as a response to this question⁵⁵, too. Therefore, we can reasonably expect an explanation of the connection between first person thought and IEM thesis to impose additional (or even tight) constraints on the account of first person thought. In other words, it is reasonable to expect an explanation of the connection between first person thought and IEM thesis to be philosophically interesting. If those philosophers limit the scope of their expectation of deeper explanation to the connection between first person thought and IEM thesis, their expectation is quite reasonable. However, since those philosophers do not recognize the distinction between de se thought and first person thought, they extend the scope of their expectation to the connection between de se thought and IEM thesis. That is why they seek for a deeper explanation on this issue. In a nutshell, those philosophers' expectations of a deeper explanation on the connection between de se thought and the IEM thesis comes from a combination of two points:

(1) they correctly diagnose that an explanation of the connection between the first person thought and the IEM thesis can be philosophically interesting,

But

(2) they mistakenly identify the de se thought with the first person thought.

My Definition Account emphasizes the distinction between first person thought and de se thought, therefore it can dissolve those philosophers' expectation by dissolving

the use of I as subject and the use of I as object, Gareth Evans' distinction between 'identification-dependent' and 'demonstrative identification', James Pryor's distinction between de re identification and which-object identification, and Manuel García-Carpintero's distinction between identification_p and identification_c.

⁵⁵ Basically, my answer to this question is that, a de se thought is immune to the error of misidentification because of its definition, and a non-de-se first person thought is prone to the error of misidentification because, as Section 2.4.4 shows, the error of misidentification can be constitutive of the non-de-se first person thought, i.e. the production of non-de-se first person thought can be based on the cognitive impact of the error of misidentification.

the second point. This is my answer to Morgan's first question.

Since the expectation of deeper explanation is based on a confusion between de se thought and first person thought, I do not think it is a plausible expectation. So, I do not think a proper explanation of the connection between de se thought and IEM thesis is obligated to fulfill that expectation. This is my answer to Morgan's second question. So much for my reply to the objection.

Now, I have demonstrated the possibility that in certain circumstances I can produce and think a non-de-se first person thought. I think this is enough to defend my Definition Account against the two challenges. The distinction presupposed by the Definition Account, which is the distinction between the de se thought and the first person thought in general, is a valid distinction. In addition, the non-de-se first person thought can be affected by the error through misidentification, so the bearer of the property 'being immune to error through misidentification' should be the de se thought, not the first person thought. Since the Definition Account of IEM is properly defended, I now have explained why the two theses below both can be attached to de se thought without any tension:

1. Entertaining a de se thought involves a special type of wh-identification of the self which does not appeal to any representation of self.
2. De se thought is immune to the error of misidentification because of its definition.

Those two theses characterize the identification aspect of de se thought. In the next chapter, I will make use of those two theses in developing an answer for the co-aboutness problem of de se communication.

Chapter III Interpersonal Co-aboutness in De Se Communication

In the first chapter, I have discussed the central question of my dissertation, which is the co-aboutness problem of de se communication. An essential requirement of a successful communication is co-aboutness: participants of a communication are reasonably expected to talk about the same subject matter. If they talk past each other, they fail to conduct a successful communication. In non-indexical communication, the co-aboutness requirement can be satisfied by the application of the replication transmission model, according to which the co-aboutness condition is achieved by replicating the participant's linguistic expressions. I call the replication of linguistic expression same-saying. However, in the case of de se communication, the same-saying strategy is trapped in a confusing situation: if I say 'I am tired now' and you reply 'I am tired now', we are same-saying but we are not talking about the same subject matter; if I say 'I am tired now' and you reply 'You are tired', we are talking about the same subject matter but no same-saying has been performed. The guaranteed relation between same-saying and co-aboutness is broken in de se communication. So, a new model of communication, which is different from the replication transmission model, is needed for correctly explaining de se communication. I call the new model the translation transmission model.

In this chapter I present my articulation of the translation transmission model, especially the fourth feature of the translation transmission model that I introduced at the end of Chapter 1. A crucial part of this articulation is to introduce a specific type of translation into the translation transmission model. The function and significance of translation in the translation transmission model are equivalent to the function and significance of same-saying in the replication transmission model. Just as not every type of same-saying is equally good at securing the satisfaction of the co-aboutness requirement, not every type of translation is equally good at securing the achievement of the co-aboutness condition. I argue that there is a specific type

of translation, which guarantees the co-aboutness for de se communication better than other types of translation. In Chapter I, I temporarily call that type translation the 'X translation'. In this chapter, I will reveal what 'X translation' actually is.

However, it is not a simple task to reveal what 'X translation' is. First, I introduce three normative theses concerning the connection between the translation and the co-aboutness condition in communications that conforms to the translation transmission model. They are the modified version of Laura Schroeter's theses that I introduced in Section B of the Introduction. Schroeter argues that a proper connection between same-saying and co-aboutness should neither be too loose nor too tight, and it should be transparent to participants of the communication. Based on Schroeter's argument, I argue that a proper connection between translation and co-aboutness should neither be too loose nor too tight, and it should be transparent to participants of communication. For me, they are criteria for determining whether or not the 'X translation' properly secures the fulfillment of the co-aboutness requirement in de se communication or not. Second, I introduce Alex Worsnip's analogy strategy as a background for my subsequent discussions. Worsnip argues that there is an analogy between the interpersonal disagreement and the intrapersonal incoherence. He encourages us to conceive disagreement as the 'interpersonal incoherence'. I think the analogy between the interpersonal case and the relevant intrapersonal case is a seminal approach, and I adopt this approach to draw an analogy between the intrapersonal de se inference and the interpersonal de se communication. Let us call the co-referentiality of two tokens of 'I' in de se inference the 'I-I' co-reference. I argue that there is a certain type of translation in de se communication. The co-referentiality of that translation is analogous to the 'I-I' co-reference. This is the 'I-You' translation. 'I-You' translation is the 'X translation' that I am looking for. Let us call the co-referentiality of 'I-You' translation the 'I-You' co-reference.

The analogy between 'I-You' co-reference and 'I-I' co-reference is plausible, because there are some non-trivial similarities between them. The first similarity is that both of them are established without appealing to any specific representation of the

relevant subjects, and therefore are not affected by any misrepresentation of those subjects. The second similarity is the co-referentiality of 'I-I' same-saying is transparent to the producer of de se inference, just as the co-referentiality of 'I-You' translation is transparent to the participants of de se communication. The third similarity between 'I-I' coreference and 'I-You' coreference is the most important one, because it explains why they have the first and the second similarities. The third similarity is that the 'I-I' co-reference and the 'I-You' co-reference are both explained by explanations of the same approach, which I call the externalist approach. I will introduce Jenann Ismael's externalist explanation of the co-reference of I-thoughts held by a single person. Ismael's externalist approach not only explains why the co-reference of I-thoughts can be established without appealing to any specific representation of their thinker, but also explains why the co-reference of I-thoughts is transparent to their thinker. This is a paradigm for my discussion on 'I-You' translation's fulfillment of the co-aboutness condition. I argue that based on Ismael's externalist approach, a similar externalist explanation of 'I-You' co-reference can be developed. Why does my 'I' co-refer with your 'you' in a de se communication? According to my externalist explanation, the 'I-You' co-reference is explained and secured by factors of the external aspect of communication. Those factors are the mutual recognition between participants of a communication and the 'I-You' reciprocity as a norm which governs the use of 'I' and 'You'. The match between the mutual recognition and the 'I-You' reciprocity not only explains why the co-reference of the 'I-You' translation can be secured without appealing to specific representations of either 'I' or 'You', but also explains why the 'I-You' co-reference is recognizable for participants of de se communication. In this sense, it explains the transparency of the co-referentiality of 'I-You' translation.

Once the externalist explanation of 'I-You' co-reference is articulated, I continue to explain why other possible translations for de se communication do not surpass the 'I-You' translation. This is because, unlike 'I-You' coreference, the co-referential relations of other translations do not have the three similarities with 'I-I' co-reference that I discuss above. At the end of this chapter, I will examine whether the 'I-You' translation

follows the three criteria that I proposed. The result of my examination shows that 'I-You' translation does follow those criteria, so I think it is safe to say that 'I-You' is the 'X translation' that I am searching for. Since 'X translation' is a crucial part of the translation transmission model, by clarifying the 'X translation', we make substantial progress in articulating the translation transmission model.

3.1 Co-aboutness and 'I-You' Translation

At the end of Chapter 1, I listed my four preliminary characterizations of the translation transmission model of de se communication. Those characterizations can be summarized in this way:

De Se Communication = Translation Transmission Model (translation) → Common Ground (co-aboutness) [The '→' stands for a guarantee relation]

Let us compare this summary with my characterization of the traditional picture of communication:

Traditional Picture of Communication = Replication Transmission Model (same-saying) → Common Ground (co-aboutness) [The '→' stands for a guarantee relation]

At first glance, it may seem that those two pictures of communication are very much alike: both of them require the co-aboutness condition to be satisfied and to be included in the common ground. However, the crucial difference between the two pictures is that, in the traditional picture of communication, it is the same-saying that plays the role of guaranteeing that the satisfaction of co-aboutness condition is included in the common ground, while in de se communication, it is the translation that plays the same role. As I have argued in Section 1.3.4, not every type of translation is evaluated as being equally effective at securing the co-aboutness in translation transmission model. For de se communication, there is a certain type of translation that can play the role of securing co-aboutness better than other types of

translation. In Chapter 1, I temporarily call it 'X translation'. So, we have a second version of the translation transmission model of de se communication:

De Se Communication* = Translation Transmission Model (X translation) → Common Ground (co-aboutness) [The '→' stands for a guarantee relation]

In this section, I will specify the 'X translation'.

3.1.1 Criteria of Co-aboutness

An important preparation for the specification of 'X translation' is to find out the criteria for determining whether or not in communication the co-aboutness condition is properly satisfied. In Section B of the Introduction, I discuss Laura Schroeter's criteria of what a proper connection between same-saying and co-aboutness should be like. Schroeter reminds us that

On an acceptable account of meaning identity, sameness of meaning must guarantee sameness of subject matter, while allowing for fallibility about the nature of that subject matter. (Schroeter, 2012, p. 178)

This implies that the connection between same-saying and co-aboutness should not be so loose that same-saying fails to play the role of 'ensuring that all parties to a conversation pick out the very same subject matter', but on the other hand, the connection should not be so tight that it blocks 'legitimate inquiry or debate into the nature of the subject matter'. I agree with Schroeter on her characterization that we should adopt those criteria which are neither too loose nor too tight. This characterization of criteria, which can be used to determine the satisfaction of the co-aboutness condition in communications that conform to the replication transmission model, gives me a clue of what the criteria for determining the connection between translation and co-aboutness in de se communication should be like. If I modify Schroeter's characterization by replacing 'sameness of meaning/same-saying' in her

characterization with 'translation', I get a characterization of the criteria I am looking for, which is:

A semantic function of translation is to ensure that participants of de se communication pick out the very same subject matter with these words. Translation must somehow guarantee sameness of subject matter – and it must do so in a way that is accessible to ordinary speakers. At the same time, however, the translation must allow for speakers' fallibility and disagreement about the nature and properties of the subject matter. Translation must not close off legitimate inquiry or debate into the nature and properties of the subject matter. (Please refer to Section B of the Introduction to see the original quote from Schroeter.)

Three criteria for determining the connection between translation and co-aboutness can be deduced from this characterization. The first criterion is that, for de se communication, the connection between translation and co-aboutness should be tight enough to ensure that participants of the communication concentrate on the same subject matter if they use proper translation. But at the same time, the connection should also be loose enough to leave open the possibility that participants of the communication can have disagreements over the nature and properties of the subject matter. This is the second criterion. The third criterion is that 'it must do so in a way that is accessible to ordinary speakers', which means that for competent participants of de se communication, the connection between translation and co-aboutness should be transparent to them, and it is transparent to them because of the translation which they adopt to do the communication. At the end of this chapter, I will examine the 'X translation' that I find for de se communication by those three criteria.

3.1.2 Analogy between Intrapersonal and Interpersonal

From this section, I begin my specification of ‘X translation’. The strategy that I adopt is to establish an analogy between the intrapersonal case and the interpersonal case. This strategy is derived from the work of Alex Worsnip, who claims that there is an analogy between interpersonal disagreement and intrapersonal incoherence:

Thus, the final view is:

Disagreement in the wide sense – final pass. You and I disagree in the wide sense if and only if I hold some (positive²⁶) attitude or attitudes A1, and you hold some (positive) attitude or attitudes A2, such that the combination of A1 and A2 in a single individual would produce an incoherence (that was not already present in either A1 or A2 taken individually). (Worsnip, 2019, p. 259)

and

That is: disagreement is to the interpersonal case as incoherence is to the intrapersonal case. (Worsnip, 2019, p. 252)

I am convinced by Worsnip’s argument. I believe the interpersonal disagreement is analogous to the intrapersonal incoherence. What is more, I argue that If we look deeper into the analogy, it can be discovered that the analogy is not only important for illuminating a wider conception of disagreement, but also important for elaborating the co-aboutness in the communication, because the analogy itself presupposes that there is a co-aboutness condition in the intrapersonal case, which is analogous to the co-aboutness condition in the interpersonal case. To make this point explicit, we can look at an example like this:

Intrapersonal incoherence	Interpersonal disagreement
Smith: ‘Moon is a satellite, but Moon is not a satellite.’	Smith: ‘Moon is a satellite.’ Peter: ‘No, Moon is not a satellite.’

This is an analogy between Smith’s intrapersonal incoherence and Smith and Peter’s

interpersonal disagreement. For Smith’s two ideas to be incoherent with each other, they must be about the same subject matter, otherwise there will be nothing incoherent between them. This is what I call the co-aboutness condition in intrapersonal inference. Analogously, for Smith and Peter to disagree with each other, Smith’s idea and Peter’s idea must be about the same subject matter, otherwise they will simply talk past each other. The co-aboutness in those two occasions is guaranteed and made transparent by (intrapersonal or interpersonal) same-saying. Since the interpersonal disagreement is analogous to the intrapersonal incoherence, the co-aboutness of interpersonal disagreement is also analogous to the co-aboutness of intrapersonal incoherence. This, in turns, assumes what guarantees the co-aboutness in interpersonal disagreement is analogous to what guarantees the co-aboutness in intrapersonal incoherence. So, the intrapersonal the same-saying is analogous to interpersonal same-saying. This is an important insight, which is a premise for my subsequent discussions on de se communication.

Now we can develop a further analogy based on Worsnip’s analogy: if the interpersonal disagreement is analogous to the intrapersonal incoherence, the disagreement in interpersonal de se communication must also be analogous to the incoherence in intrapersonal de se inference⁵⁶. I can use an example to illustrate this analogy:

De se inference (intrapersonal)	De se communication (interpersonal)
Smith: ‘I have a meeting today, but I have no meeting today.’	Smith: ‘I have a meeting today.’ Peter: ‘No, ___ [Translation] have no

⁵⁶ The idea of ‘the analogy between the intrapersonal case and the interpersonal case’ is supported by Simon Prosser, who says,

I first argue that, by parallel with the synchronic intrapersonal case, the sharing of a MOP should involve epistemic transparency between the token thoughts of the two thinkers. (Prosser, 2019, p. 465)

	meeting today.'
Result: Incoherence	Result: Disagreement

In this example, the de se inference is an instance of intrapersonal incoherence, and the de se communication is an instance of interpersonal disagreement. So, the analogy between de se inference and de se communication in this example is simply an instance of Worsnip's more general analogy that I have analyzed above, and what I have argued above about the co-aboutness relation in Worsnip's analogy can also be applied to the co-aboutness relation in this example. In order to make the de se disagreement analogous to the de se incoherence, the co-aboutness between the first person sentence 'I have a meeting today' and its relevant translation must be analogous to the co-aboutness between the two first person sentences of the de se inference. But which term shall I choose to be a better translation of 'I' in de se communications? Or in other words, which term shall I pick to fill the blank in the example discussed above? The analogy between intrapersonal de se inference and interpersonal de se communication gives us an important clue for answering those two questions: the term, whose co-reference with 'I' in a de se communication best imitates the co-reference between tokens of 'I' in de se inference, is a good translation of 'I'. I argue that the co-referential relation between 'You' and 'I' in interpersonal communication best imitates the co-referential relation between two tokens of 'I' in intrapersonal inference, so 'You' is a good translation of 'I' for the purpose of conducting de se communication, and 'I-You' translation is the 'X translation' that I am looking for. Now I can rewrite the example discussed above by filling the blank in it with 'I-You' translation:

De se inference (intrapersonal)	De se communication (interpersonal)
Smith: 'I have a meeting today, but I have no meeting today.'	Smith: 'I have a meeting today.' Peter: 'No, <u>you</u> have no meeting today.'

Result: Incoherence	Result: Disagreement
---------------------	----------------------

3.1.3 'I-You' Translation

What is 'I-You' translation? By this terminology, I mean first and foremost the linguistic translation between sentences containing first person singular indexical terms (e.g. I, me, my, mine and so on) and sentences containing second person singular indexical terms (e.g. you, your, yours and so on). For example, the translation between 'I have a meeting today' and 'You have a meeting today' is a typical 'I-You' translation. However, people may continue to ask: does what I mean by 'I-You' translation include a cognitive translation between de se thought and de tu thought⁵⁷? At first glance, the answer to this question seems to be very straightforward. However, it is not actually the case, because there are philosophers who deny the very existence of de tu thought. For example, Richard G. Heck, who claims that

Consider the indexical "you". As a matter of its standing meaning, an utterance of "you" refers to the person addressed in that utterance. But in the sense that there is such a thing as a self-conscious, first-person belief, there is no such thing as a second-person belief, or so it seems to me. Of course, I can identify someone descriptively, as the person to whom I am now speaking, and may have beliefs whose contents involve that descriptive identification. But that is not what I mean to deny: I mean to deny that there is any such thing as an essentially indexical second-person belief. The phenomenon of the second-person is a linguistic one, bound up with the fact that utterances, as we make them, are typically directed to people, not just made to the cosmos. (If there

⁵⁷ By the phrase 'de tu thought', I mean roughly the second person thought about the addressee as you.

were speakers of a language who never directed their utterances to their fellows, they would have no use for the second-person.) The word 'you' has no correlate at the level of thought: if not, then the contents of the beliefs we express using the word 'you' have very little to do with its standing meaning. (Richard G. Heck, Jr., 2002, p. 12)

This is a very famous and thought-provoking paragraph. Heck's claim in this paragraph is that while there are first person thoughts associated with first person linguistic expressions, there are no second person thoughts associated with second person linguistic expressions. There are second person sentences, but they 'have no correlate at the level of thought'. If Heck is correct, there is no 'cognitive translation between de se thought and de tu thought', because, according to Heck, de tu thought does not exist. Although Heck himself says that 'I don't really know how to argue for this claim: it just seems right to me, even obviously so.' (Richard G. Heck, Jr., 2002, p.12), other philosophers seem to have many ideas on how to argue for (or argue against) this claim.⁵⁸ In this chapter, I am not going to discuss whether there is an irreducible second person thought or not. My reason for this decision is that the focus of this dissertation is on the connection between the fulfillment of co-aboutness condition and our use of linguistic expression. In Section 1.3, I argue that what is translated in the translation transmission model is, first and foremost, the linguistic

⁵⁸ Léa Salje provides a good summary of philosophers' ideas on this question:

According to one family of responses to the first question—taken up recently by Sebastian Rödl (2007), Michael Thompson (2012), Jose Luis Bermudez (2005) and Guy Longworth (2013, 2014), between others—the kind of thought canonically expressed by our uses of “you” is not distinct from the kind of thought canonically expressed by uses of “I”. For Longworth, for instance, these two ways of thinking are two faces of the same cognitive capacity. For Rödl, the words “I” and “you” are two possible modes of expression of one and the same thought; he writes, “‘You...’ said by me to you and ‘I...’ said by you in taking up the address, express the same act of thinking, they express the same thought. Therefore it is wrong to oppose second-person thought to first-person thought. This is a difference in the means of expression, not in the thought expressed.” (Rödl 2007, p. 197). (Salje, 2017, p. 817-818)

expressions that we use to make communications. This is one of the crucial differences between my characterization of translation transmission model and Recanati's characterization of that model. In addition, in Section 1.2.3, by citing Schroeter's characterization of the fulfillment of co-aboutness condition, I emphasize that it is the 'semantic function' of linguistic expressions, not their cognitive function, which guarantees that participants in communication pick out the same subject matter. Also, in the replication transmission model, it is the 'sameness of meaning', not the sameness of thought, which guarantees the sameness of subject matter. Therefore, with regard to the translation transmission model, I insist on a similar standpoint: it is the translation of language, not the translation of thought, which first and foremost guarantees the achievement of co-aboutness requirement in de se communication. Even if there is no 'second person thought' (therefore no translation between de se thought and de tu thought), there is still the translation between first person language and second person language. In this chapter, I take the 'I-You' translation as a linguistic translation to be the primary target of my discussion. Whether or not there is irreducible de tu thought may be a serious philosophical problem in itself, but it is not within the scope of my dissertation.

However, there are two real problems that I must resolve in this chapter. The first problem is why the analogy between the interpersonal 'I-You' translation and the intrapersonal 'I-I' same-saying is a legitimate analogy, and the second problem is why I claim that the 'I-You' translation best imitates the 'I-I' same-saying, and therefore is the alleged X translation that I am looking for. Or in other words, why I claim 'I-You' translation is a better guarantee for the achievement of the co-aboutness requirement in de se communication, compared with other types of translation. For me, resolutions to those two problems are closely related. In subsequent sections of this chapter, I argue that there are three important similarities between 'I-You' translation and 'I-I' same-saying, which make 'I-You' translation the best imitation of 'I-I' same-saying. The analogy between 'I-You' translation and 'I-I' same-saying is also sustained by the three similarities, and that is why this analogy is legitimate. The first similarity is that the 'I-

'I' co-reference in de se inference and the 'I-You' co-reference in de se communication are both explained by explanations of the same approach, which I call the externalist approach, and their externalist explanations are analogous to one another. The second similarity is that the 'I-I' co-reference is transparent to the relevant producer of de se inference, just as the 'I-You' co-reference is transparent to the relevant participants of de se communication. The third similarity is that both the co-reference of 'I-I' same-saying and the co-reference of 'I-You' translation are not affected by specific misrepresentations of the relevant subjects. When I have misrepresentations about myself, the 'I-I' same-saying in my de se inference still guarantees the fulfillment of the co-aboutness condition, and analogously, when I have misrepresentations about my addressee, or when my addressee has misrepresentations about me, the 'I-You' translation still fulfills the co-aboutness condition in our de se communication. This 'immunity to effect of misrepresentation' is somehow similar to the immunity to error of de re misidentification that I discussed in Chapter 2. The three similarities cannot be found in comparisons between 'I-I' same-saying and other types of translation, but can only be found in the analogy between 'I-I' same-saying and 'I-You' translation. Therefore, 'I-You' translation is the best interpersonal analogue of the intrapersonal 'I-I' same-saying. By discussing the three similarities, I also point out three features of 'I-You' translation, which are (1) the co-referentiality of 'I-You' translation is externally secured by a factor of the communication context, (2) the co-referentiality of 'I-You' translation is transparent to participants of de se communication, and (3) 'I-You' translation is immune to the effect of misrepresentation. The three features are unique to 'I-You' translation, and they make 'I-You' translation a better option for guaranteeing the satisfaction of co-aboutness condition in de se communication than other translations. In the next section, I will discuss the three similarities in detail.

3.2 'I-You' Translation and 'I-I' Same-saying

In previous sections, I argue that 'I-You' translation is a good guarantee for the co-aboutness of de se communication, because 'I-You' translation is the best interpersonal analogue of the intrapersonal 'I-I' same-saying⁵⁹. I also argue that the analogy between 'I-You' translation and 'I-I' same-saying is sustained by three similarities between them. In this section, I discuss the three similarities. First, I introduce Jenann Ismael's externalist explanation of the co-referentiality of 'I-I' same-saying, which relies on an insight concerning the identificational aspect of de se thought. This insight is closely related to the conclusion I got at the end of Section 2.3.2. In her explanation, Ismael argues that the co-referentiality of 'I-I' same-saying is secured by external factors. I think this is an illuminating approach: given that 'I-You' translation is analogous to 'I-I' same-saying, can 'I-You' translation be explained by an externalist approach which is similar to the externalist explanation of 'I-I' same-saying? My answer to this question is yes. I make a distinction between the internal aspect and the external aspect of communication and argue the co-referentiality of 'I-You' translation is secured by a special factor of the external aspect of communication, which I call the mutual recognition. Now, let us take a closer look at the first similarity between 'I-I' same-saying and 'I-You' translation.

3.2.1 Co-referentiality of Tokens of 'I'

In section 3.1.2, I introduced an example of intrapersonal de se inference:

Smith: 'I have a meeting today, but I have no meeting today.'

⁵⁹ It should be noted that the co-reference of 'I-I' same-saying which I discussed in this chapter is the co-reference of 'I-I' same-saying in de se inference, and the co-reference of 'I-You' translation which I discuss in this chapter is the co-reference of 'I-You' translation in de se communication. The 'I-You' co-reference outside of de se communication and the 'I-I' co-reference outside of de se inference are not in the scope of this chapter.

I think it can be generally agreed that this is a typical example of intrapersonal incoherence. For Smith's two sentences in this example to be incoherent with each other, they must be about the same subject matter, which implies the 'I' in Smith's first sentence must be co-referential with the 'I' in his second sentence. But why do the two tokens of 'I' co-refer? How can we explain this co-referentiality of 'I-I' same-saying? How can we tell whether they are indeed co-referential or not? A usual strategy to explain the co-referentiality of two referential expressions is to appeal to the criteria of identity of their referent (or referents). Jenann Ismael describes the strategy of appealing to criteria of identity as:

Reidentification and intersubstitution require knowing what type of object the expression picks out and one has to apply criteria of identity associated with the type to determine when reidentification has been achieved or when intersubstitution is licensed. (Ismael, 2010, p. 211)

If we adopt this strategy to explain the co-referentiality of tokens of 'I' in the example, we need to find out for whom Smith's first token 'I' is used, and also find out for whom his second token 'I' is used. It is easy to see in the example that Smith uses both tokens for himself. We then adopt the relevant criteria of identity to determine whether the two individuals (i.e. two selves) are identical or not. If they are identical, the two tokens of 'I' are co-referential. Ismael describes this procedure as applying criteria of identity for selves in judging whether one token 'I' is intersubstitutable with another token 'I':

To fit "I" into this mold, we would have to say that occurrences of "I" refer to the selves that produce them, and intersubstitution of "I" in temporally separated thoughts involves application of criteria of identity for selves. We would have to say that each of us applies criteria of identity for selves in judging that some past or future thought is ours, and that the "I" in that thought can be intersubstituted with the "I" in this one, as for example we apprehend and apply criteria of identity for days in judging that a 'today'

occurrence on one occasion is intersubstitutable with a 'today'-occurrence on another. (Ismael, 2010, p. 212)

In a nutshell, according to the usual strategy, tokens of 'I' are co-referential because they are used for the same self, and we determine whether a self is identical with another self by means of the criteria of identity of self. Ismael calls this 'the view of man on the street':

This view is probably closest to that of the man on the street. Ask a nonphilosopher what makes two occurrences of 'I' coreferential and he will tell you they belong to the same self, end of story. (Ismael, 2010, p. 213)

However, Ismael argues that 'the view of man on the street' is incorrect, because this view assumes that we can use tokens of 'I' in such a way that they are co-referential only after we know what the relevant criteria of identity of self are. If a man on the street says that two tokens of "I" are co-referential because they are about the same self, we can continue to ask him what it is for two selves to be identical, or what it is for a self to be identical with itself. I think we can rationally expect that he will be confused by those questions. Anyone who supports 'the view of man on the street' must be able to answer these questions, because if she does not know what the criteria of identity of self are, she then is not in a proper position to make co-referential use of tokens of 'I'. Ismael argues this is a deeply problematic assumption:

It makes possession of the notion of a self prior to the ability to reidentify oneself over time, and that doesn't seem phenomenologically right. It seems too much thought to attribute to the child, for example, who says I enjoyed the ice cream yesterday and I want to have some more. The ability to reidentify oneself over time seems to precede any grasp of the concept of a self. (Ismael, 2010, p. 213)

I think the scenario that Ismael mentions in this paragraph is very plausible. If a six-

year-old child says, 'I enjoyed the ice cream yesterday and I want to have some more today', it is very clear that this child succeeds in using tokens of 'I' in such a way that her first token 'I' is co-referential with her second token 'I'. But does this child know what the criteria of identity of self are? Ismael's answer is no, and I agree with Ismael. It is obvious to me that my capacity of using tokens of 'I' in such a way that they are co-referential does not depend on my grasp of the criteria of identity of self (in Ismael's words, 'the ability to reidentify oneself over time seems to precede any grasp of the concept of a self.'). Therefore, it does not make sense to appeal to the criteria of identity of self to account for the co-referentiality of tokens of 'I'.

Next, Ismael continues to argue that besides the co-reference of tokens of 'I', the reference of a token 'I' also does not depend on criteria of identity of self:

Reference – even reference to a continuant - can in this way precede apprehension of criteria of identity. Apprehension of such criteria will be unnecessary for purposes of coordination, and if it is ever made, it will present itself as a discovery. (Ismael, 2010, p. 220)

This view, as Ismael argues, has already been accepted by many philosophers:

There is a general recognition in the literature on the self that the self isn't picked out by an individuating conception of the kind that could serve as a Fregean sense. (...) It was discussed also by Strawson, Evans, Anscombe, and others, and often expressed by saying that no criterion of identity is employed in picking out the self as the object of a referential act. (...) It's the lack of employment of criteria of identity in judgments of same-self relatedness across contexts that distinguishes selfidentification from other kinds of indexical and demonstrative reference. (Ismael, 2010, p. 212)

I agree with Ismael's view mentioned in this paragraph that 'no criterion of identity is employed in picking out the self as the object of a referential act'. This view is closely related to the conclusion I come to at the end of Section 2.3.2, which is

‘entertaining de se thought involves a special type of identification of self which does not depend on any representation of self’. Although the conclusion is not the same as Ismael’s view (because a criterion of identity of self is not the same as a representation of a specific self), they still share an important consensus: I pick out myself without mediation. In Chapter 2, I argue that I pick out myself without the mediation of representation of myself when entertaining a de se thought; and in this chapter, I agree with Ismael that I pick out myself not only without the mediation of criteria of identity of self when producing a token ‘I’ for myself.

Since the reference of a token ‘I’ and the co-reference of tokens of ‘I’ both do not depend on criteria of identity (of self), the co-referentiality of tokens of ‘I’ should not be explained by the apprehension of criteria of identity (of self). Therefore, the usual strategy and the view of man on the street do not work in this case. A new explanation of co-reference of tokens of ‘I’, which does not depend on the ‘apprehension of criteria of identity’, is needed. In the next section, I will introduce Ismael’s externalist explanation of co-referentiality of tokens of ‘I’, which is exactly the explanation we need.

3.2.2 Externalist Explanation of ‘I-I’ Co-reference

The primary target of Ismael’s explanation is the co-aboutness of I-thoughts⁶⁰, not the co-referentiality of ‘I-I’ same-saying, which means that her primary focus is on

⁶⁰ Here I should explain the difference between Ismael’s terminology and my terminology. In Chapter 2, I make an unusual distinction between de se thought and non-de-se first person thought. So, my use of the phrase ‘de se thought’ in this dissertation is a bit idiosyncratic. I do not think Jenann Ismael’s ‘I-thought’ perfectly matches my ‘de se thought’, because she does not make any distinction between de se thought and non-de-se first person thought. However, I still believe that by ‘I-thought’, Ismael means roughly ‘thought about oneself as oneself’, which is also a definition of de se thought that I accept.

the level of thought, not on the level of language. She explains why we can take it for granted that an I-thought of mine is about the same subject matter as another I-thought of mine from an externalist perspective. However, I argue that a crucial thought experiment that Ismael uses in her argument can be interpreted as a thought experiment concerning the co-reference of linguistic expressions. Therefore, with some modifications, Ismael's explanation on the co-aboutness of I-thought can also be applied to explain the co-reference of 'I-I' same-saying. I call Ismael's explanation the 'externalist explanation', because it appeals only to external factors which can play the explanatory role even if we do not apprehend or understand them. In Ismael's own words, those factors are 'external to, and need not be represented in, thought.' (Ismael, 2010, p. 223) In contrast, the usual strategy, or the view of man on the street, which I discussed in the previous section, can be seen as an internalist explanation, because in this explanation, the criteria of identity of self can be adopted to explain the co-reference of 'I-I' same-saying only after it has been apprehended or grasped by the relevant language user. If language users do not apprehend or grasp those criteria, they cannot play the explanatory role that they are expected to play in the relevant language games. But in Ismael's explanation, the external factors can always play the role of explaining and securing the co-aboutness of I-thoughts, whether the relevant thinker apprehends those factors or not.

Ismael begins her enquiry on the co-aboutness problem of I-thoughts by proposing a creative reading of a classical thought experiment designed by John Perry and Simon Blackburn. Perry and Blackburn invite us to imagine a scenario, which I call the 'Z-land scenario':

Consider a small isolated group, living in a place we call Z-land. Z-landers do not travel to or communicate with residents of other places, and they have no name for Z-land. When a Z-lander sees rain, he will say to others not in a position to look outdoors, 'It is raining'. His listeners then act appropriately to there being rain in Z-land: they close the windows in Z-land, cancel plans for Z-land picnics, and grab umbrellas before going into the Z-land out-of-doors.

They have no other use for 'It is raining'. They do not call their sons in far-off places, or listen to the weather news, or read newspapers with national weather reports. (Perry and Blackburn, 1986, p.144)

For Perry and Blackburn, the purpose of designing this fictional scenario is to show that in some propositions there is a kind of semantic component called the 'unarticulated constituent'. The unarticulated constituent makes a contribution to the semantic content of that proposition, but is not articulated in that proposition, which means no words in sentences which express that proposition indicate the existence of that constituent. For example, in the Z-land scenario, when a Z-lander uses the proposition 'It is raining', it is semantically equal to the proposition 'It is raining here'. It is obvious that the term 'here' refers to Z-land, but in Z-lander's original sentence, there is no 'here'. Therefore, in this scenario, Z-land is an unarticulated constituent for the proposition 'It is raining'. Perry and Blackburn use this thought experiment to promote an idea concerning semantics, not an idea concerning thought. However, Ismael discovers that according to Perry and Blackburn's own comment about the Z-land scenario, it is possible to interpret this thought experiment in a different direction. Perry and Blackburn claim that

Z-land is a constituent of the practice, or language game, in which the sentence It is raining plays a role. There is no need to postulate a concept or idea of Z-land as a component of their thought, to secure the connection to Z-land. (Perry and Blackburn, 1986, p. 145)

Perry and Blackburn mention 'a concept or idea of Z-land as a component of their thought'. This comment suggests the possibility of reading the 'Z-land scenario' in a different way, which is not only about semantics, but also about thought. Let us imagine that one day, Smith and Peter, who are both Z-landers, meet on the street. Smith says: 'It is raining today' and Peter replies: 'Yes, it is raining'. It is obvious that Smith's words and the thought expressed by his words are about the weather of Z-land. It is also obvious that Peter's words and the thought expressed by his words are

also about the weather of Z-land. Smith's weather talk/thought and Peter's weather talk/thought are both connected to Z-land. In addition, Smith and Peter know their weather talk/thought are about the same subject matter, which means the co-aboutness of their weather talk/thought is transparent to them. However, nothing in their words indicates they have any concept or idea of Z-land: they do not mention any description or name of the place, and they do not use the locational indexical 'here' or 'there'. Perry and Blackburn tell us 'there is no need to postulate a concept or idea of Z-land as a component of their thought, to secure the connection to Z-land.' This implies that the connection between Z-land and their weather talk/thought is not based on or mediated by their apprehension of the concept of Z-land. This in turn implies that the co-aboutness between Smith's weather talk/thought and Peter's weather talk/thought is not based on or mediated by their apprehension of the concept of Z-land. So, we should not appeal to any concept or idea of Z-land to account for the co-aboutness of their weather talk/thought.⁶¹ But then, how should we explain the co-aboutness of Peter's and Smith's weather talk/thought, and how should we explain the connection between Z-land and their weather talk/thoughts?

Ismael argues that the co-aboutness of Peter's and Smith's weather talk/thought, and the connection between Z-land and their weather thoughts, should both be explained by a series of external facts, which include but are not limited to: (1) Peter and Smith are both Z-landers who live in Z-land, and (2) Z-landers do not travel to other places, and (3) Z-landers do not communicate with residents of other places, and (4) Smith and Peter's weather conversation takes place in Z-land. Those external facts about Z-land and Z-landers by themselves explain the co-aboutness of Peter's and Smith's weather talk/thought and secure the connection between their weather talk/thought and Z-land. As long as those external facts persist, Peter's and Smith's

⁶¹ This is very similar to the conclusion I get at the end of Section 3.2.1, which is 'since the reference of a token 'I' and the co-reference of tokens of 'I' both precede the apprehension of criteria of identity, the co-referentiality of tokens of 'I' should not be explained by the apprehension of criteria of identity.'

weather talk/thought will continue to be about the same subject matter, which is the weather of Z-land. The Z-landers do not need to know anything about those facts, and their ignorance has no impact on the explanatory role of those facts. I think this externalist approach is very illuminating. It explains the co-aboutness of thoughts by factors which are external to thought. Those factors do not need to be apprehended or represented in thought in order to play their explanatory role. Even if Z-landers have no apprehension of them, those external factors can still explain the co-aboutness of Z-landers' weather talk/thought.

In the previous section, I introduced Ismael's argument that a competent language user can use tokens of 'I' in such a way that they are co-referential, even if the language user does not apprehend the relevant criteria of identity of self. This is very similar to the situation that I discuss in this section, which is that Z-landers can make their weather talk/ thought about the same subject matter, which is the weather of Z-land, even if Z-landers have no idea of Z-land. Ismael constructs an analogy between the intrapersonal co-reference of I-thoughts and the interpersonal co-reference of Z-landers' weather thoughts: if the interpersonal co-reference of Z-landers' weather thoughts can be explained by an externalist explanation, can the intrapersonal co-reference of I-thoughts be explained by an analogous externalist explanation? Ismael's answer to this question is 'Yes'. She argues that:

Explicit definition does indeed require prior conceptualization of the referent, but if we can learn how to use 'I' by a purely mechanical procedure, one that can be described without any explicit reference to or conceptualization of its target, we can let the de facto relations between tokens that in fact get intersubstituted when 'I' is used in accordance with the procedure implicitly define the referent. (...) We don't need semantic links between terms (or their mental analogues, ideas) and their objects where there are architectural links in place. It is only when we break architectural links by moving to different contexts that we require mediating concepts. (Ismael, 2010, p.220)

In this paragraph, Ismael mentions the ‘purely mechanical procedure’, the ‘de facto relations between tokens (of I-thought)’ and the ‘architectural links’. They are Ismael’s descriptions of the external factors which secure the co-aboutness of I-thoughts. Those factors are external in the sense that they do not need to be represented in thought, or in other words, they do not need to be apprehended by the person who holds I-thoughts. Ismael describes the effect of those factors as ‘processing that is invisible to the subject’ and ‘complicated action behind the scenes’:

There are immediate connections between the mental states of a single subject, but immediacy in this sense is shorn of the traditional encumbrances of transparency and certainty, and compatible with the kind of complicated action behind the scenes (processing that is invisible to the subject, processing that doesn’t get represented in thought) that Dennett emphasizes in his critiques of the Cartesian Theater. (Ismael, 2010, p. 223)

It should be noted that Ismael should not be understood as claiming that the co-referentiality of I-thoughts is shorn of transparency in this paragraph. On the contrary, what Ismael actually claims is that the external factors which explain the co-referentiality of I-thoughts are shorn of transparency, because they are not represented in our thought. But what are those external factors that have been described as ‘purely mechanical procedure’, ‘de facto relations between tokens (of I-thought)’ and ‘architectural links’? Ismael’s answer is:

What is special about “I”, wholly unique to it as a term in thought, is that the facts that govern intersubstitutability of ‘I’-occurrences are external to, and need not be represented in, thought. (...) Whatever ideas I have about my future self play no role in determining which of the world’s future inhabitants is me. If I think that I will be the first woman president of the United States, but Marla Maples beats me to it, that doesn’t make my present thoughts about my future self thoughts about Miss Maples. Whoever gathered these memories and sent them to me along the discriminating internal channels that Nature has

provided inside the head, with whatever internal embellishments and emotional resonances that past self added, that was me. And whoever receives these thoughts and experiences along the same channels will be me as well. (Ismael, 2010, p.224)

In this paragraph, Ismael mentions ‘the discriminating internal channels that Nature has provided inside the head’ twice. This is the most detailed description of the external factor that I can find in Ismael’s paper. Here Ismael emphasizes, again, that ‘the facts that govern intersubstitutability of “I”-occurrences are external to, and need not be represented in, thought.’ I think ‘the facts that govern intersubstitutability of “I”-occurrences’ must include ‘the discriminating internal channels that Nature has provided inside the head’. So, in a nutshell, according to Ismael, the co-reference of I-thoughts is secured and explained by a series of external factors, for example the ‘discriminating internal channels that Nature has provided inside the head’, which do not need to be represented in thought or apprehended by the holder of I-thoughts.⁶².

⁶² It should be noted that Jenann Ismael is not the only philosopher who is sympathetic to the idea that what explains the co-reference of de se thoughts does not need to be represented in thought. I think Recanati and Peacocke are also sympathetic to this idea. For example, Recanati argues:

But the idea that various pieces of information cluster into a single file when they are ‘taken to concern the same object’ can be understood in a way that does not presuppose a prior identity judgement. It may be entirely a matter of subpersonal binding of information. Thus in the case of proto-files at least it is the cognitive system, not the subject, that takes the pieces of information to concern the same object and cluster them within a file. The subject does not judge that the pieces of information concern the same object. Identity is presupposed, it is built into the way the information is (subpersonally) packaged. Bochner’s twofold objection can therefore be met by appealing to the subpersonal mechanisms underlying informational clustering: the cognitive system tracks the source of information by distributing information into files based on distinct ER relations, and that is what accounts for informational clustering. (Recanati, 2012, p.98)

In this paragraph, Recanati tries to explain why various pieces of information clusters are taken to be about the same object, which is a typical co-reference (or co-aboutness) problem. To resolve this problem, Recanati appeals to the ‘sub-personal packaging’ and ‘sub-personal binding’ of information.

Ismael's externalist explanation of co-reference of I-thoughts is based on an analogy between the co-reference of I-thoughts and the co-reference of Z-landers' weather thoughts, and the externalist explanation of co-reference of Z-landers' weather thoughts is extracted from Perry and Blackburn's Z-land scenario by means of Ismael's creative reading. But as I have shown in this section, we can use the externalist explanation to explain the co-reference of Z-lander's weather talk, too. So, analogously, we should also be able to use the externalist explanation of co-reference of I-thought to explain the co-reference of 'I-I' same-saying in de se inference: I can use tokens of 'I' in such a way that they are co-referential, because this co-referentiality is secured and explained by a series of external factors. Those factors are the external factors which explain the co-aboutness of I-thoughts.

If the packaging or binding process is really sub-personal, it is not represented in thought, and it does not need to be apprehended at the personal level. Peacocke also has similar ideas on this issue. He argues that:

There can be an integrating apparatus in an individual that is not conscious at all, an integrating apparatus that forwards representations to the individual's file on itself. There can be nonconscious states with de se contents, even in entirely nonconscious individuals. The presence of an integrating apparatus and a self-file are merely necessary constitutive subpersonal conditions for the existence of conscious subjects. They are not sufficient conditions. (Peacocke, 2014, p. 20)

and

This more general notion of a subject must involve some kind of integration of states and events, and a control system for action linked to that integrating apparatus. The difference from the case of conscious subjects is that the integration does not result in the occurrence of a conscious state or event. The integration has rather informational and explanatory significance. (Peacocke, 2014, p. 44)

Peacocke's unconscious sub-personal integrating apparatus of information is also not represented in thought and does not need to be apprehended at the personal level.

3.2.3 The 'External/Internal' Distinction

In Section 3.1.2, I introduce an analogy between the intrapersonal de se inference and the interpersonal de se communication. Based on this analogy, in Section 3.1.3, I introduce a further analogy between the co-reference of 'I-I' same-saying and the co-reference of 'I-You' translation, which is sustained by three similarities between them. One of the similarities is that the co-reference of 'I-I' same-saying in de se inference and the co-reference of 'I-You' translation in de se communication are both explained by explanations of the same approach, which I call the externalist approach. In Section 3.2.2, I discuss Ismael's externalist explanation of the 'I-I' coreference. Inspired by Ismael's explanation, I will develop a similar (although not identical) externalist explanation of the 'I-You' coreference in de se communication. However, before the explanation, there are two questions which need to be addressed. First, what does the 'external/internal' distinction mean for de se communication? What is the external and the internal in de se communication? Second, why do proposals of the internalist approach fail to explain the 'I-You' coreference in de se communication? In this section, I answer the first question, and in the next section, I will answer the second.

As my discussions in Section 3.2.1 and Section 3.2.2 show, for Ismael, an explanation of (semantic or cognitive) intrapersonal co-reference is an internal explanation if it appeals to factors that need to be apprehended or represented in thought in order to play the explanatory role, and an explanation of intrapersonal co-reference is an external explanation if it appeals to factors that do not need to be apprehended or represented in thought in order to play the explanatory role. Therefore, for Ismael, the intrapersonal 'internal/external' distinction is between 'being represented in thought' and 'not being represented in thought'. Can I draw an interpersonal 'internal/external' distinction based on this intrapersonal analogue? Now, if we take the analogy between intrapersonal de se inference and interpersonal de se communication seriously, we shall then accept that a producer of a de se

inference is analogous to participants of a de se communication, and what is apprehended or represented in that producer's thought is analogous to what is communicated between participants of that de se communication. So, the interpersonal analogue of the intrapersonal 'internal/external' distinction can be described like this: for de se communication, the 'internal/external' distinction is a distinction between what is communicated within de se communication and what is not communicated within de se communication. It should be noted that 'what is not communicated within de se communication' is a very broad characterization. In fact, it is so broad that anything that is not communicated within de se communication, whether it is relevant to the de se communication or not, can be included as 'what is not communicated within de se communication'. So, this characterization is not precise enough and should be refined. I think a better characterization of the 'internal/external' distinction is: for a de se communication, its internal aspect is what is communicated, and its external aspect is what is not communicated but has been presupposed by that communication for the sake of making it successful. This new characterization of the external aspect of de se communication is more accurate than the previous one, because it contains a new constraint on what should be included as the external aspect. The new constraint is 'being presupposed by that communication for the sake of making the communication successful'. It is closely related to a notion that I discussed in Section 1.2.2, which is the notion of 'common ground' (or common presupposition). Generally speaking, the common ground of a communication is whatever participants of that communication take for granted as the background of that communication. The common ground can be updated, challenged, enlarged or reduced during the communication, but at the very beginning of the communication, participants must share some essential common presuppositions as the common ground of their communication, otherwise the communication can never begin. So, it is obvious that what is included in the initial common ground cannot be what is communicated in the communication. After I add the new constraint into the characterization of 'external aspect', it is clear that the common ground that I discussed in Section 1.2.2 is a part of the external aspect of

communication.

3.2.4 Internalist Explanation of 'I-You' Co-reference

I have characterized what is internal and what is external for interpersonal communication. The next step is to answer the central question of this chapter:

Why is there 'I-You' co-reference? Or, how does my 'I' (or 'you') co-refer with your 'you' (or 'I') in a communication?

Based on the 'internal/external' distinction, there are two possible answers to the central question: the internalist explanation of 'I-You' co-reference and the externalist explanation of 'I-You' co-reference. In this section, I introduce the internalist explanation. In Section 3.2.5, I will argue that the internalist explanation cannot explain some cases of 'I-You' co-reference, and therefore it is deeply flawed. In Section 3.2.6 and Section 3.2.7, I will argue that the externalist explanation is a better option for explaining why my 'I' and your 'you' co-refer in a de se communication.

For me, a typical internalist explanation is that which appeals to the epistemology of meaning of 'I' and 'You' to account for the 'I-You' co-reference. When participating in a (verbal) communication, people communicate with one another using linguistic expressions, and they understand what is communicated by understanding the linguistic expressions that they use to express what they want to communicate. So, accordingly, in de se communication, people communicate de se thoughts by using sentences containing, for example, words like 'I', 'me', 'my' and 'mine'. People find out which person the de se communication is about by means of their understanding of the meaning of first person terms. I call the operation of discovering whatever a term means the 'semantic consumption' of that term. Therefore, for a singular referential term like 'I' or 'you', the semantic consumption of it is to find out the

individual to which the term refers⁶³. So, semantic consumption of 'I' (and also 'you') is the way participants of a de se communication gain access to the individual about which the de se communication is. Proponents of the internalist explanation argue that the co-reference of 'I-You' translation can only be explained after the semantic consumption of 'I' and 'You' is done. For example,

One day, Smith met Oliver. Smith told Oliver: 'I feel sick today.' Oliver replied: 'What happened to you?'

Proponents of internalist explanation explain this conversation in this way: both Smith and Oliver are talking about the same subject matter, which is Smith's feeling, because Smith's 'I' refers to himself, and Oliver's 'you' refers to Smith. The semantic consumption of 'I' and 'you' in this conversation shows that the two terms refer to the same individual. This is how the co-reference of 'I-You' translation guarantees the co-aboutness of de se communication⁶⁴. In this example, the internalist explanation

⁶³ It should be noted here that the meaning of a singular referential term is its referent, not the description of the connection between the (token) term and its referent. For example, a token 'now' refers to the moment in which the token 'now' is used. But 'the moment in which the token "now" is used' is not the meaning of that token 'now'. It is merely a description of connection between the token 'now' and its meaning. The real meaning of the token now is the actual moment. Therefore, knowing that this token 'I' refers to its current user does not amount to knowing the meaning/referent of this token 'I'. This argument is supported by Heck, who says,

So, if we are to defend thesis (5)—if we are to defend Frege's identification of what is said with a Thought—we must deny that what is said by an utterance of, say, 'You are a philosopher' involves its standing meaning, as the identification of ways of thinking with standing meaning would have it. The sense of an utterance of 'You are a philosopher' cannot be anything like "The addressee is a philosopher." Kaplan, of course, has familiarized us with one way of denying that standing meaning is any part of what is said. On his account, standing meaning is a determinant of what is said but is no part of what is said: character, as he puts it, determines but is no part of content. I have argued, on grounds different from his, that Fregeans must accept this component of Kaplan's view. (Richard G. Heck, Jr., 2002, p. 12)

⁶⁴ This internalist explanation of 'I-You' co-reference is somehow similar to the internalist explanation of 'I-I' co-reference, which also explains the co-reference of tokens of 'I' by appealing to their

seems to be useful. However, I argue that the internalist explanation cannot handle some special cases. Later in this chapter I will show that semantic consumption is not sufficient for explaining the co-reference of 'I-You' translation. In the next section, I will argue that a factor of the communication context is very important for the externalist explanation of 'I-You' co-reference.

3.2.5 The Defect of the Internalist Explanation

Now I have articulated what the internalist explanation of 'I-You' co-reference in de se communication is. In this section, I will explain why I do not adopt the internalist explanation. A central claim of the internalist explanation is that the co-reference of 'I-You' translation in de se communication is achieved by participants' semantic consumption of 'I' and 'You'. I believe this claim is incorrect. I argue that it is possible for us to establish the co-reference of 'I-You' translation without appealing to the semantic consumption of 'I' and 'You'. The 'I-You' co-reference is independent of the semantic consumption of 'I' and 'You'.⁶⁵ Here I introduce an example to

referent's being identical.

⁶⁵ The idea that the 'I-You' co-reference is independent of the references of 'I' and 'You' is anticipated by Simon Prosser, who argues that the semantic co-ordination role of 'I' and 'You' facilitates the interpersonal trading on identity of the (co-)reference of 'I' and 'You'. He says:

The second obstacle to the extension of the account of shared MOPs to the indexical case concerns the fact that speaker and hearer utter different words if they are in different contexts. If I say "I" and you say "you" when both referring to S, there is no doubt that we use different words. This obstacle is easier to overcome than the first. The solution appeals to the fact that anyone who properly understands the use of an indexical term must understand the way in which its reference varies with context. Consequently, anyone who understands the word "here," and encounters a token of it, must understand that the token has the same reference as a token of "there" in an appropriately related context, and they can rely on this fact in trading on identity. Strictly speaking it is not shared words per se that facilitate interpersonal trading on identity; it is their coordinating role that matters. Provided speaker

demonstrate the existence of this possibility. I call this example the ‘loud neighbor’ case:

Imagine that you are staying in a hotel room. There are some people staying in the room next to yours, and they are talking very loudly. Their talk is so loud that you can hear it clearly. If one of them says 'I need a cup' and another one of them replies 'There is one on your left', you can well understand the co-reference of 'I' and 'You' in their conversation. But you are not in a proper position to consume the semantic value of either the token 'I' or the token 'You'. You cannot see through the wall, and you have never met those people before, so you have no idea of which individuals their 'I' and 'You' refer to. But nevertheless, you know there is a co-reference of 'I-You' translation in their de se communication.

The ‘loud neighbor’ case shows that the co-reference of ‘I-You’ translation can be established without the semantic consumption of ‘I’ and ‘You’. A crucial point here is that no complete semantic consumption of either ‘I’ or ‘You’ can be achieved in the ‘loud neighbor’ case. In Section 3.2.4, I propose that for a singular referential term like ‘I’ or ‘you’, the semantic consumption of it is to find out the individual to which the term refers. So, a complete semantic consumption of ‘I’ (or ‘you’) must include picking out the individual to which the token ‘I’ refers. and a complete semantic

and audience both understand the way in which a word functions, be it indexical or otherwise, the coordinating role can be achieved. (Prosser, 2019, pp. 479-480)

In this paragraph, Prosser argues that it is the semantic co-ordination role of indexicals like ‘I’ and ‘You’, not the actual reference of ‘I’ or ‘You’, that ensures the interpersonal trading on identity of the (co-)reference of ‘I’ and ‘You’. His notion of ‘trading on the interpersonal identity’ is similar to my notion of ‘co-aboutness’. So Prosser is sympathetic to my idea that the ‘I-You’ co-reference is independent of the references of ‘I’ and ‘You’. But unfortunately, Prosser neither tells us how to cash out the semantic co-ordination role of ‘I’ and ‘You’, nor does he tell us what the ‘appropriately related context’ means in his theory. In this sense, Prosser’s explanation of the ‘I-You’ co-reference is incomplete.

consumption of ‘you’ must include picking out the individual who is the addressee of the user of the token ‘you’. However, in the ‘loud neighbor’ case, you cannot see your loud next-door neighbor at the time of hearing their talk, and you have never met them before. You can by no means pick out the individual who is the referent of the token ‘I’ (or ‘you’) that you hear. What you can understand is nothing more than that ‘this token “I” refers to the owner of this voice’, or ‘this token “you” refers to the addressee of the owner of that voice’. This is not a complete semantic consumption, because no individual is picked out. But it does not prevent you from knowing the co-reference of ‘I’ and ‘you’ in the ‘loud neighbor’ case. We can establish the co-reference of ‘I-You’ translation (and therefore guarantees the relevant co-aboutness) without doing semantic consumption of ‘I’ and ‘You’. The ‘I-You’ co-reference is independent of the semantic consumption of ‘I’ and ‘You’. So, the internalist explanation is not the correct explanation for ‘I-You’ co-reference in de se communication.

3.2.6 Externalist Explanation of ‘I-You’ Co-reference

From my discussion on the ‘loud neighbor’ case, I conclude that the co-reference of ‘I-You’ translation in de se communication is independent of the semantic consumption of ‘I’ and ‘you’. The next step is to reveal the real guarantee which explains and secures the ‘I-You’ co-reference. In this section, I introduce an example which includes a comparison between two different interpretations of four sentences, to help reveal what the real guarantee is.

Here is a group of four sentences, which can be interpreted in two different ways:

Interpretation 1	Interpretation 2
‘I have a meeting today.’	[‘I have a meeting today.’

'No, you have no meeting today.'	'No, you have no meeting today.'
'How do you know that?'	'How do you know that?'
'Because I know your working schedule.'	'Because I know your working schedule.']

According to Interpretation 1, this group of sentences consist of four unconnected sentences. Each of those sentences is a separate unit, which has no connection with other members in the group. Therefore, if we adopt Interpretation 1, we will not find out any co-referential relation between tokens of 'I' and tokens of 'you' in that group of sentences. However, according to Interpretation 2, this group of sentences is interpreted as a record of a communication. The four sentences are connected in a meaningful way: the second sentence is a negation of the first sentence, and the third sentence is a question towards the second sentence, and the fourth sentence is a reply to the third sentence. I use the square brackets to indicate a group of sentences as a whole. We can then extract at least two cases of 'I-You' co-reference from this 'communication reading': the 'I' in the first sentence is co-referential with the 'you' in the second sentence, and the 'you' in the third sentence co-refers to the 'I' in the fourth sentence. This is the crucial difference between Interpretation 1 and Interpretation 2: according to the former, there is no 'I-You' coreference in those sentences, but according to the latter, there are at least two cases of 'I-You' co-reference in this communication. This crucial difference cannot be explained by the internalist explanation, because the semantic consumption of 'I' and 'You' in Interpretation 1 is exactly the same as the semantic consumption of 'I' and 'You' in Interpretation 2. The internalist explanation, which appeals to the semantic consumption of 'I' and 'You', is totally insensitive to the difference that according to Interpretation 1 there is no 'I-You' co-reference in those four sentences but according to Interpretation 2 there is an 'I-You' co-reference in that communication. So how can this difference be explained?

I argue that the difference can be explained if we find what causes that difference,

and we can find the cause of that difference by conceiving Interpretation 1 and Interpretation 2 in chronological order. At the very beginning, following Interpretation 1, we take the four sentences we encountered to be unconnected and independent. At the moment, we do not find 'I-You' co-reference in those sentences, and I think it is safe to claim that at the moment there is no 'I-You' co-reference in those sentences. Then we shift to Interpretation 2 by embedding those four sentences into a communication context. After this shift, we take that group of sentences to be a record of communication. At this time, we find cases of 'I-You' co-reference in that group of sentences, and I think it is safe to say there is 'I-You' co-reference in this communication. The chronological conception of Interpretation 1 and 2 clearly shows that the 'I-You' co-reference appears after we embed the group of sentences into a communication context.⁶⁶ This makes me curious about the relation between the communication context per se⁶⁷ and the co-reference of 'I-You' translation in de se communication. This relation can also be discovered in the 'loud neighbor' case of Section 3.2.5. When listening to the loud talk of people staying next door, you take it for granted that they are communicating with one another. So, when you hear the two sentences, 'I need a cup' and 'There is one on your left', you take it for granted that those sentences are parts of a communication. It is under this presupposition that you realize the 'I-You' co-reference in those two sentences. The connection between the communication context per se and the 'I-You' co-reference in de se communication has been discovered again and again. But how should this

⁶⁶ It should be noted that the 'communication context' in question here is a communication context of the bare minimum. It is a very thin conception of communication, not a rich one. When introducing the communication context into the interpretation of those four sentences, I do not specify which type of communication does it belong to, nor do I specify when and where it takes place, let alone how many people participate in the communication. This is because I do not need to. A basic, thin conception of communication can very well play the role that I expect it to play in explaining the co-reference of 'I-You' translation.

⁶⁷ It should be noted here I simply use the phrase 'communication context per se', but do not use the further specified phrase 'de se communication context per se'.

connection be characterized? My answer is that: an essential factor of communication context per se guarantees the co-reference of 'I-You' translation in de se communication by its match with the 'I-You' reciprocity.

This characterization of the connection between communication context and 'I-You' co-reference is very important for my explanation of 'I-You' co-reference. But to understand this characterization, first we need to understand what the 'I-You' reciprocity is. The 'I-You' reciprocity is a very idiosyncratic norm which governs both the use of the term 'I' and the use of the term 'You'. In fact, according to 'I-You' reciprocity, the use of 'I' and the use of 'you' are intrinsically connected to the extent that we cannot understand how to properly use one of them without the understanding of how to properly use the other. Léa Salje has a vivid description of the 'I-You' reciprocity:

In order for it to be appropriate to express a thought about you using 'you', but not 'that person', it seems that there also has to be awareness going in the other direction: you also have to be aware of me. This kind of reciprocity was part of what was involved in the notion of addressing drawn on in the last section—you cannot notice that my utterance is directed towards you without also being aware of me. If the appropriateness of using 'you' (but not 'that person') is bound up with the notion of addressing, then we can at least say that the felicity conditions for 'you' (but not 'that person') minimally involve the hearer's awareness of the speaker, as well as the speaker's awareness of the hearer. (Salje, 2017, p. 821)

In a nutshell, 'I-You' reciprocity means that the use of 'I' and the use of 'you' are reciprocally connected.

'I-You' Reciprocity:

When X is in a proper position to accept Y's referring to Y-self as 'I', X must also be in a proper position to refer to Y as 'you'; and when X is in a proper position to be

referred by Y as 'you', X must also be in a proper position to refer to X-self as 'I' for Y. When X is not in a proper position to be referred by Y as 'you', X is also not in a proper position to refer to X-self as 'I' for Y. If X is not in a proper position to refer to Y as 'you', X is also not in a proper position to accept Y's referring to Y-self as 'I'.

It should be noted that the 'I-You' reciprocity is not a norm of semantic consumption of 'I' and 'You'. If we compare the 'I-You' reciprocity with the semantic consumption that I introduced in Section 3.2.5, we can easily discern their difference: the semantic consumption results in locating an individual as the referent of a token 'I' or 'you' given a certain context, but the 'I-You' reciprocity as a norm does not function in that way. 'I-You' reciprocity is not a norm of semantic consumption. 'I-You' reciprocity is a norm of the proper use of 'I' and 'You'. It outlines an essential precondition for establishing the 'I-You' co-reference: in order to use their token 'I' and token 'you' in such a way that the token 'I' co-refers with the token 'you', the user of token 'I' and the user of token 'you' must be interlocked in the way prescribed by 'I-You' reciprocity. What's more, if those language users are interlocked in that way, the 'I' used by one of them will be co-referential with the 'you' used by another one of them, regardless of their semantic consumptions of 'I' and 'You' ⁶⁸. Then, the next question: how can we satisfy

⁶⁸ Léa Salje's view on this point is even stronger. She argues that once the two language users are interlocked in the way prescribed by 'I-You' reciprocity, the only proper way they can plausibly refer to each other is to refer to each other using 'I' and 'You'. In other words, being interlocked in this way prevents those language users from using other referential expressions (third personal indexical, for example). Salje says,

Once the two are interlocked in this reciprocal conscious interaction, her thought is no longer appropriately expressible with a third-personal utterance, even if that utterance is too soft to be audible. (Salje, 2017, p. 825)

For my current purpose, I do not need to accept Salje's stronger view. In fact, I do not think Salje's view is correct on this point. I believe that what guarantees and explains the co-reference of 'I-You' translation in de se communication does not prevent us from using other possible translations in de se communication, although other translations may not be as good at guaranteeing the satisfaction of co-aboutness condition as 'I-You' translation.

this requirement of 'being interlocked in the way prescribed by 'I-You' reciprocity'?

My answer to this question is that simply by being in a communication context with one another, we satisfy the requirement of being interlocked in the way prescribed by 'I-You' reciprocity. More precisely, it is one of the constitutive factors of the communication context that enables the satisfaction. This factor is what I call the 'mutual recognition relation'⁶⁹ between participants of the communication.

⁶⁹ The term 'mutual recognition' is also used by Léa Salje (Salje, 2017) in her discussion on the proper condition of entertaining a second person thought about someone else. Salje argues that properly entertaining a second person thought about a person presupposes the 'mutual recognition of mutual awareness' between that person and the thinker of the second person thought. She says,

Imagine, for instance, that she reveals herself to Leontes by a wink of the eye. Leontes' crashing realization, we may imagine, will bring with it a wave of shame and horror. But the change brought about is not only in him. Introduced onto the scene is now a mutual recognition of mutual awareness, and with it a new restriction on the felicitous modes of expression available to Hermione for the articulation of her thoughts about her husband. (Salje, 2017, p. 825)

However, there are at least four differences between my conception of 'mutual recognition relation' and Salje's conception of 'mutual recognition'.

First, what is recognized in my 'mutual recognition' is different from what is recognized in Salje's mutual recognition. According to Salje's conception, what is mutually recognized is the mutual awareness. By the term 'mutual awareness', Salje means something like a conjunction of X's awareness of Y's awareness of X's awareness and Y's awareness of X's awareness of Y's awareness. However, what is recognized in my characterization of 'mutual recognition' is an interpersonal relation between prospective participants of a communication. There is no recognition of (other people's) recognition of (other people's) recognition in my mutual recognition. A prospective participant of a communication must be ready to be recognized by other prospective participants, but she does not need to recognize another prospective participant's recognition.

Second, the number of people involved in Salje's mutual recognition is different from the number of people involved in my mutual recognition. Salje's 'mutual recognition' is limited to a recognition between two people, but my 'mutual recognition' can be a recognition between multiple people (so more than two people).

Third, Salje's 'mutual recognition' is related to reference, but my 'mutual recognition' is related to

Mutual Recognition Relation:

In order to start a communication with someone, prospective participants must be in the mutual recognition relation with one another. In a two-person communication, X must first recognize Y as whom X-self is going to communicate with. But at the same time, X must also recognize X-self as Y's prospective communicator. In addition, X must be ready to be recognized by Y as Y's prospective communicator. Otherwise there will not be any communication between X and Y. The situation is the same for Y. Y must recognize X as Y-self's communicator, recognize Y-self as X's prospective communicator, and be ready to be recognized by X as X's prospective communicator, at the same time.

This is a typical description of mutual recognition in communications involving only two participants. But mutual recognition is by no means confined to binary communications. In communications involving more than two participants, there must also be mutual recognitions between their participants, but it would be more complex to describe the structure of mutual recognition in those cases. It should be noted that the constitution of the mutual recognition relation does not presuppose the wh-

co-reference. For Salje, the mutual recognition of mutual awareness is a presupposition for properly entertaining a second person thought about someone and properly using the second person pronoun to refer to someone. In other words, her 'mutual recognition' is only related to the referential use of 'You'. But my 'mutual recognition' is a contextual factor which explains the 'I-You' co-reference in communications.

Fourth, it seems that Salje's 'mutual recognition' presupposes the which-object identification of people who are involved in entertaining a second person thought. In order to constitute Salje's mutual recognition, participants must be able to identify and locate each other individually. But the constitution of my 'mutual recognition' relation does not presuppose the which-object identification of the (prospective) participants of a communication. Participants of a communication, who are in my mutual recognition relation, does not necessarily need to identify and locate each other individually. I will explain more on this point in Section 3.2.7.

Because of those four differences, my 'mutual recognition' should not be read as a synonym of Salje's 'mutual recognition'.

identification of the (prospective) participants of a communication. I will explain this point in detail in Section 3.2.7.

The mutual recognition relation is a constitutive factor of the communication context. Without it, a communication context will simply lapse. It is clear from my description of 'I-You' reciprocity and my analysis of the structure of mutual recognition that the mutual recognition relation satisfies the requirement set by the 'I-You' reciprocity. 'Being in a mutual recognition relation' matches 'being interlocked in the way prescribed by "I-You" reciprocity'⁷⁰. Therefore, the fact that some people are in a communication context per se makes it guaranteed that in a communication, the 'I' used by one of them is co-referential with the 'you' used by another one of them, as long as they are in the same communication.

Now we know that, when participating in a de se communication, one participant can use the term 'you' to translate the term 'I' produced by another communication to achieve co-referentiality, which in turn guarantees that they are communicating about the same subject matter, which is the alleged co-aboutness condition of communication. The co-reference of 'I-You' translation is secured by the match between the mutual recognition as a constitutive factor of communication per se and the 'I-You' reciprocity which is an intrinsic norm governing the use of 'I' and 'you'. According to my distinction of the internal and external aspect of communication, the internal aspect is what is communicated in the communication, and the external aspect is what is not communicated but has been presupposed by the communication for the sake of making it successful. It is obvious that both the mutual recognition and the norm of 'I-You' reciprocity belong to the external aspect of communication. So, my

⁷⁰ It should be noted that the norm of 'I-You' reciprocity prescribes a relation between two individuals, and the mutual recognition relation can be between more than two individuals. But this does no harm to the match of 'being in a mutual recognition relation' and 'being interlocked in the way prescribed by "I-You" reciprocity'. This is because the mutual recognition relation between more than two individuals always implies the mutual recognition between any two individuals in that relation.

explanation of 'I-You' co-reference in de se communication appeals only to factors of the external aspect of communication. In this sense, it is an externalist explanation of 'I-You' co-reference, and the 'I-You' co-reference in de se communication is externally secured.

It should be noted that mutual recognition is not a special constitutive factor presupposed only by de se communication. Instead, it is a universal constitutive factor presupposed by communication in general, although it does play a special role for de se communication which it does not play in other types of communication. This implies that the externalist explanation is not a very demanding explanation. It does not take a lot of effort to secure the 'I-You' co-reference in the external way. On the contrary, the 'I-You' co-reference can be easily secured in the external way because the external explanation appeals to a factor, i.e. the mutual recognition, that is found in every type of communication. The low threshold of securing 'I-You' co-reference by means of external factors, which is an advantage of 'I-You' translation, is an important reason for choosing 'I-You' translation to guarantee the co-aboutness of de se communication. Other types of possible translation for guaranteeing the co-aboutness of de se communication requires much more than what 'I-You' translation requires. For example, if you translate my sentence 'I will do a presentation tomorrow' as 'Smith's best friend will do a presentation tomorrow', then, in order to secure the co-referential relation between 'I' and 'Smith's best friend', you need to know who Smith's best friend is and who the token 'I' refers to. In other words, you have to secure the co-referential relation between 'I' and 'Smith's best friend' through an internalist approach, which involves the semantic consumptions of 'I' and 'Smith's best friend'. That semantic consumption is not required by the explanation of 'I-You' co-reference. So, 'I-You' translation is a less demanding way of satisfying the co-aboutness condition in de se communication, compared with other possible translations.

Based on my discussions in this section, I think it is safe to say that 'I-You' translation, whose co-referentiality is externally secured by factors in the external aspect of communication, fulfils the first criterion for regulating the connection between

translation and co-aboutness that I proposed in Section 3.1.1: the connection between translation and co-aboutness should be tight enough to ensure that participants of the communication concentrate on the same subject matter.

3.2.7 Four Possible Objections

In this section, I will discuss four possible objections against my externalist explanation of 'I-You' co-reference in de se communication. By responding to those possible objections, I also make some further clarifications on the externalist explanation.

The first objection challenges the coherence of my externalist explanation. The objection claims that the constitution of mutual recognition relation between participants of a communication presupposes the wh-identification⁷¹ of those participants. But the wh-identification of the participants is identical with the result of the semantic consumptions of 'I' and 'You'. In Section 3.2.5 and 3.2.6, I argue the 'I-You' co-reference can be established with appealing to semantic consumptions of 'I' and 'You', and the externalist explanation of 'I-You' co-reference does not involve semantic consumptions of 'I' and 'You'. If the constitution of mutual recognition relation presupposes the wh-identification of those participants, it seems that this is equivalent to say the constitution of mutual recognition relation presupposes the result of the semantic consumptions of 'I' and 'You'. From the first objection's point of view, at the beginning I claim that the externalist explanation does not involve semantic consumption, but then I re-introduce the result of semantic consumption back into the externalist explanation under the name of 'mutual recognition'. So, according to this objection, my characterization of externalist explanation is

⁷¹ Please refer to Section 2.4.1 for a definition of which-object identification (wh-identification).

incoherent.

My reply to the first objection is that it misunderstands the concept of ‘mutual recognition relation’. I never argue that the mutual recognition between participants of a communication presupposes the wh-identification of those participants. On the contrary, I argue that the mutual recognition between participants of a communication can be independent of wh-identification of those participants. This argument consists of two sub-theses: the mutual recognition can be constructed without appealing to wh-identification, and the mutual recognition can be discovered without appealing to wh-identification. I will discuss those two sub-theses.

First, the mutual recognition relation can be constructed without appealing to wh-identification of participants of communication. Let us imagine a scenario called the ‘phone scam case’:

I received a phone call from an unknown number. I answered it and found out that it was a phone scam. I did not know who the scammer was. But nevertheless, during my conversation with the scammer over the phone, my ‘I’ and the scammer’s ‘You’ succeeded in co-referring with each other, and so did my ‘You’ and the scammer’s ‘I’.

This scenario can be viewed as a typical example of showing how the mutual recognition can be constructed without appealing to wh-identification. I think it can be generally agreed that when I say ‘I’ and the scammer say ‘you’ during the conversation on the phone, my ‘I’ co-refers with the scammer’s ‘you’. According to the externalist explanation, this ‘I-You’ co-reference presupposes a mutual recognition between the scammer and me. But in that scenario, I have no idea of who the scammer is, therefore I do not make wh-identification about the scammer. On the other hand, it is highly likely that the scammer also has no idea of who I am, therefore the scammer does not make wh-identification about me. This clearly shows that the construction of mutual recognition can be independent of wh-identification of participants of communication. So, how is the mutual recognition between the scammer and myself

being constructed? My answer is that the mutual recognition is constructed by both the scammer's phone call and my answering that phone call. By calling someone on the phone, a scammer recognizes whomever she is going to talk with as the prospective communicator and recognizes herself as the prospective communicator for whomever she is going to talk with, and she is ready to be recognized by whomever she is going to talk with as the prospective communicator. By answering a call from an unknown number, I recognize whomever I am going to talk with as a prospective communicator and recognize myself as a prospective communicator for whomever I am going to talk with, and I am ready to be recognized by whomever I am going to talk with as her prospective communicator. The repeated occurrence of the term 'whomever' clearly indicates that I do not need to (wh-)identify and locate the individual who is on the other end of the phonecall. Therefore, the wh-identification does not need to be involved in the construction of mutual recognition relation.

Here I am aware of the possibility that my argument can be challenged by a more complicated version of the phone scam case:

I received a phone call from an unknown number. I answered it and found out that it was a phone scam. I did not know who the scammer was. In fact, there were two scammers speaking on the other end of the call, whose voices were modified by an audio editor to the extent that their modified voices were too similar to be distinguished from each other. But I did not know this, and I thought there was only one scammer. During my conversation with those scammers, we exchanged some 'I-You' talk. But this time, it was unclear whether my 'I' (or 'you') always succeeded in being co-referential with those scammers' 'you' (or 'I').

Proponents of the first objection may argue that in this scenario, the anonymous mutual recognition relation between participants of a de se communication fails to always secure the 'I-You' co-reference in that communication. However, I argue that their diagnosis of the problem in this more complicated scenario is incorrect. The real problem in this scenario is that I take the mutual recognition relation to be between

two participants of the communication, but those scammers take the mutual recognition relation to be between three participants of the communication. In short, I have no agreement with those scammers on how many participants are there in the communication. This lack of agreement leads to the result that there is no real mutual recognition between the two scammers and me. It is because there is no real mutual recognition between participants of that de se communication, that it is unclear whether my 'I' (or 'you') succeeds in being co-referential with those scammers' 'you' (or 'I'). Therefore, the more complicated version of the phone scam case is not a real challenge to the function of mutual recognition in my externalist explanation of 'I-You' co-reference.

Second, the mutual recognition relation can be discovered without appealing to wh-identification of participants. In the 'loud neighbour' case that I introduced in Section 3.2.5, you have no wh-identification information concerning those loud neighbours who stay in the room next to yours: you have never met them before, and you cannot see through the wall when hearing them talking. When you hear them talking to one another using sentences containing 'I' and 'you', you do not know to whom the token 'I' refers and to whom the token 'you' refers. But all those mentioned above do not prevent you from discovering that those people are communicating with one another, and do not prevent you from discovering the mutual recognition relation between them. You do not need to (wh-)identify and locate the referents of the tokens of 'I' and 'You' that you heard in order to realize the 'I-You' co-reference. You only need to realize the mutual recognition relation between participants of that talk. That is how you can realize the 'I-You' co-reference in their talk ('I need a cup.' 'There is one on your left.'). So, wh-identification does not have to be involved in the discovery of mutual recognition relation.

Since the mutual recognition can be constructed and discovered without appealing to wh-identification, the first objection against my externalist explanation fails.

The second objection challenges the sufficiency of the externalist explanation. It

claims that the match between the mutual recognition and 'I-You' reciprocity is not sufficient to guarantee or explain the co-reference of 'I-You' translation in de se communication. To show how this is possible, let us imagine a scenario like this,

Smith, Oliver and Peter all mutually recognize one another as prospective communicators. Smith says, 'I need a cup.' and Oliver replies, 'There is one on your left.' But suddenly Peter looks at the sky and says, 'I just saw a UFO!'

In this scenario, it is obvious that Smith 'I' co-refers with Oliver's 'you' (or, 'your', to be precise). According to the externalist explanation, it is because of the match between 'I-You' reciprocity and the relevant mutual recognition, i.e. the mutual recognition between Smith and Oliver. However, if the externalist explanation is a sufficient explanation or guarantee for 'I-You' co-reference, Oliver's 'you' should also be co-referential with Peter's 'I', because there is also a mutual recognition between Oliver and Peter. However, it is obvious that Oliver's 'you' does not co-refer with Peter's 'I'. So, we are tempted to draw the conclusion that the match between the mutual recognition and 'I-You' reciprocity is not a sufficient guarantee or explanation for the co-reference of 'I-You' translation in a de se communication.

My reply to this objection is that it highlights the importance of the criteria of individuating communication. It should be noted that what the externalist explanation explains is the co-reference of 'I-You' translation *in a de se communication*. The phrase 'in a de se communication' is not trivial. If a token 'I' and a token 'you' do not appear in the same communication, their relation (be it co-referential or not) is not the target of my externalist explanation. In the scenario described above, I think it is intuitively clear that Smith's utterance and Oliver's utterance belong to one communication, but Peter's utterance is not a part of that communication. Since Peter's 'I' and Oliver's 'you' are not in the same communication, their relation is not in the scope of my externalist explanation. Although there are mutual recognitions between Peter, Oliver and Smith, this does not imply that Smith, Oliver and Peter are all in the same communication. The mutual recognition is just one of the constitutive factors of communication

context, but it is not the communication context itself. In fact, what the second objection shows is that the mutual recognition relation is not a sufficient condition for individuating a communication...⁷²

The third objection challenges the compatibility between my conception of ‘mutual recognition’ and Stalnaker’s definition of ‘common ground’. I argue that the mutual recognition relation is a part of the external aspect of communication. By ‘external aspect of communication’ I mean ‘what is not communicated in the communication but has been presupposed by the communication for the sake of making the communication successful’. This definition is similar to Stalnaker’s definition of ‘common ground’, which I introduced in Section 1.2.2. However, according to Stalnaker’s definition, only propositions can be included in common ground. In other words, only propositions can be commonly presupposed by participants of communication. But based on my analysis of the structure of mutual recognition in Section 3.2.6, I do not think that the mutual recognition can be taken as a type of propositional attitude with propositional content. In other words, I do not think the mutual recognition is propositional. If the mutual recognition is non-propositional, or at least can be non-propositional, it then can no longer be included in Stalnaker’s common ground. There are, obviously, two ways to resolve this incompatibility: the first way is to modify my conception of mutual recognition, and the second way is to modify the definition of common ground. I think it is better to choose the second option. The reason for my choice is that, as Salje suggests, even non-verbal communications presuppose the mutual recognition between their participants. She says,

A natural way to proceed here is to ask what we care about when we care

⁷² However, I must admit here that the assertion ‘Peter’s utterance does not belong to the communication to which Smith’s and Oliver’s utterances belong’ is merely based on my intuition. I do not have well-formulated criteria for individuating a communication. The question of ‘what the criteria for individuating a communication are’ is by itself a very important question, but it is beyond the scope of this dissertation.

whether or not we have succeeded in addressing another. A plausible-looking answer is that we care whether or not the other is receptive to us, whether they are suitably sensitive to our attempts to engage their attention, whether we have made our attempts at contact sufficiently salient to them. In the standard case of verbal communication, we care whether they have noticed that our utterance is directed towards them, but the phenomenon is just as familiar in cases of nonverbal interpersonal contact. In non-verbal cases, it's just that we care whether we have succeeded in reaching out to them in other ways. (Salje, 2017, p.831)

It is easy to discover that by 'reaching out to them in other ways', Salje means 'reaching out to them in ways other than directing our utterances towards them'. Jane Heal also holds a similar idea on this issue,

It is not true; however, that one person's cognitive stances are all private, hidden from others, unless and until they get linguistic expression. Something may be common knowledge between agents, where the thought of each is open to the other, not in virtue of their speech but in virtue of their situation and/or their non-linguistic actions. And also it is not obviously true that, in such a situation of common knowledge, one agent cannot have a thought which is 'addressed' to the other, in some sense of 'address' appropriate to the thought having a second person character. So, for anything we have yet seen, perhaps one person can 'address' another without uttering the word 'you', and perhaps without engaging in speech at all. We need to reflect more on what phenomenon, or range of phenomena, might fall under the umbrella of this word 'addressing'. (Heal, 2014, p. 320-321)

It seems that Jane Heal's standpoint about the existence of 'non-propositional' mutual recognition is even stronger, because she explicitly claims that mutual recognition (an instance of the 'the thought of each is open to the other') between participants of non-verbal communication can be achieved 'in virtue of their situation and/or their

non-linguistic actions'. The idea that mutual recognition can be non-propositional seems to be supported by philosophers. Actually, I believe that the 'phone scam case' in my reply to the first objection is a typical example of constructing mutual recognition in virtue of non-linguistic actions: in that case, the mutual recognition between the scammer and me is constructed by the scammer's action of calling me on the phone and my action of answering that call. Therefore, my reply to the third objection is that: there is an incompatibility between my conception of mutual recognition and Stalnaker's definition of common ground. But in order to resolve this incompatibility, we should modify the definition of common ground, rather than modifying my conception of mutual recognition. The purpose of the modification is to make the common ground compatible with the non-propositional conception of mutual recognition. But I will not discuss the details of the modification in this chapter.

The fourth objection is related to the grammatical gender of pronouns. In English, there is no gender-neutral third person singular pronoun. It is either 'he/him/his' or 'she/her/her'. In recent years, people begin to use 'they' as a gender-neutral third person singular pronoun. Now let us imagine there is a language called LANG. In LANG, there is no gender-neutral second person singular pronoun. It is either the 'masculine You' or the 'feminine You'. Therefore, if people use LANG to make de se (verbal) communication, it seems that they need to know which second person singular pronoun they should use in order to secure the co-reference of their 'I-You' talk. However, my externalist explanation by itself does not provide any direction for which second person singular pronoun they should use. It seems that when making de se communications, participants who speak LANG have to appeal to the internal aspect of their communications to decide which second person singular pronoun they should use, the 'masculine You' or the 'feminine You'.

I take this objection to be a serious challenge. My response to this objection consists of two parts. First, if in LANG there is also no gender-neutral first person singular pronoun, the solution to this problem is relatively straightforward. In a de se communication, a participant who would like to convey her de se thought in LANG has

to choose either the 'masculine I' or the 'feminine I', and other participants can use the corresponding second person singular pronoun to satisfy the co-aboutness condition. In this case, they do not need to appeal to the internal aspect of communication, and my externalist explanation of 'I-You' co-reference still works. However, if in LANG there is a gender-neutral first person singular pronoun, and if in a de se communication, a participant who speaks LANG, uses this gender-neutral first person singular pronoun to convey her de se thought, the situation becomes very perplexing. It seems that other participants of that communication have to appeal to the internal aspect of their communication to decide which second person singular pronoun they should use. I think the best way to avoid this perplexing situation is to create a corresponding gender-neutral second person singular pronoun for LANG. We should be clear that grammatical gender is not an essential component of the first or second person pronoun. In English, the first person pronoun 'I' and the second person pronoun 'you' has no grammatical gender, but people speaking English has no problem in making de se communication.

3.2.8 Transparency and Immunity

In Section 3.1.3, I introduce three similarities between 'I-I' same saying and 'I-You' translation. The first similarity is that 'I-I' co-reference in de se inference and 'I-You' co-reference in de se communication are both explained by explanations of the externalist approach, and their externalist explanations are analogous to one another. I have articulated this similarity in Section 3.2.2 and 3.2.6. The second similarity is that the 'I-I' co-reference is transparent to the relevant producer of de se inference, just as the 'I-You' co-reference is transparent to the relevant participants of de se communication. In this section, I will articulate this similarity.

The co-reference of 'I-You' translation in de se communication is externally secured by the match between 'I-You' reciprocity and mutual recognition. Therefore,

the 'I-You' co-reference will be transparent to participants of de se communication once the match between mutual recognition and 'I-You' reciprocity become transparent to participants of de se communication. As I have shown in Section 3.2.6, when I shifted from Interpretation 1 to Interpretation 2, I embedded those four sentences into a communication context, regarding them as a record of de se communication. Once I did that, the co-reference between 'I' and 'You' in those sentences suddenly became transparent to me. If I am a participant of de se communication, the mutual recognitions between participants of that communication are definitely transparent to me, otherwise I cannot be legitimately counted as a participant of that communication. But 'I-You' reciprocity is not transparent to participants of communication in that way. It is a piece of linguistic knowledge concerning the norm of using 'I' and 'You', which we acquire by learning. If I never learn this piece of knowledge, 'I-You' reciprocity then is opaque to me, so are the 'I-You' co-reference. But fortunately, we can reasonably expect an average language user to know the norm of 'I-You' reciprocity. Therefore, as long as participants of de se communication are average language users, we can reasonably expect the match between 'I-You' reciprocity and mutual recognition to be transparent to them, which in turn guarantees that co-reference of 'I-You' translation is transparent to them. This is how the 'I-You' translation fulfils the third criterion that I proposed in Section 3.1.1, which is for participants of de se communication, the connection between translation and co-aboutness should be transparent to them, and it is transparent to them because of the translation they adopt to do the communication

The transparency of 'I-I' same-saying, or same-saying in general, is more philosophically interesting than the transparency of 'I-You' same-saying. The basic argument to support the transparency of same-saying is that

1. The co-reference of same-saying must be transparent, because
2. If it is not transparent, it needs to be articulated. But

3. Articulating the co-reference of same-saying will result in the infinite regress.

So,

4. The co-reference of same-saying must be transparent from the very beginning.

Recanati and Schroeter both support this argument. Recanati argues,

Argument A

Cicero is bald

Tully is well-read

Cicero = Tully

Someone is bald and well-read

Argument B

Cicero is bald

Cicero is well-read

[implicit premise : Cicero = Cicero]

Someone is bald and well-read

The suppressed premise in argument B is meant to ensure that the two occurrences of 'Cicero' in the explicit premises of the argument actually corefer (if they did not corefer, the argument would be invalid, indeed). As Campbell and many others have shown, however, this view of argument B as enthymematic and resting on a suppressed premise is indefensible. In general, the attempt to reduce presumptions of identity to implicit identity judgments launches an infinite regress. (Recanati, 2011, p. 10)

And Schroeter argues,

Moreover, even subjects who are able to make explicit identity judgments must rely on some more basic mechanism for recognizing sameness of reference between their own thoughts. To insist that the subject must make an explicit identity judgment before she can recognize that two thoughts are about the same thing would be to invite a vicious regress - for even the simplest inference from 'P' to 'P' would then require infinitely many explicit identity judgments to establish the co-reference of premise and conclusion.

The moral here is much the same as the one Lewis Carroll drew in the case of modus ponens: we must have some basic way of taking two thoughts to be co-referential which does not require an explicit identity judgment (Carroll 1895). (Schroeter, 2008, p. 115)

Both Recanati and Schroeter mention the ‘infinite regress/ vicious regress’. Let me use Recanati’s example to show why articulating the co-reference of same-saying leads to an infinite regress. Suppose we are reading a record of an inference,

P1. Cicero is bald

P2. Cicero is well-read

Someone is bald and well-read

Then suppose the co-reference of same-saying of ‘Cicero’ is not transparent to me, and this makes it too hard for me to understand this inference. Then an equation must be introduced to articulate this co-referential relation for me. So,

P1. Cicero is bald

P2. Cicero is well-read

P3. Cicero is Cicero

Someone is bald and well-read

Now, in order to understand Premise 3 as an articulation of the co-referential relation between ‘Cicero’ in Premise 1 and ‘Cicero’ in Premise 2, I must understand the first ‘Cicero’ in Premise 3 co-refers with the ‘Cicero’ in Premise 1 and the second ‘Cicero’ in Premise 3 co-refers with the ‘Cicero’ in Premise 2. But since the co-reference of same-saying is opaque to me, I do not understand that. So, another equation needs

to be introduced to articulate the co-referential relation between the first 'Cicero' in Premise 3 and the 'Cicero' in Premise 1, and the co-referential relation between the second 'Cicero' in Premise 3 and the 'Cicero' in Premise 2, for me. So,

P1. Cicero is bald

P2. Cicero is well-read

P3. Cicero is Cicero

P4 Cicero is Cicero, and Cicero is Cicero

(infinite regress)

Someone is bald and well-read

Operations above can be repeated infinitely. That is how the articulation of co-reference of same-saying leads to an infinite regress. So, the co-reference of same-saying must be transparent at the very beginning.

A possible objection is that, here we can utilize the 'object language/metalanguage' distinction and introduce a type of metalanguage to talk about the object language. By this strategy, we can articulate the co-reference of same-saying without being troubled by infinite regress. However, I argue that this strategy cannot solve the problem of infinite regress, because the relation between object language and metalanguage needs to be articulated, too. So, we have

P1. Cicero is bald

P2. Cicero is well-read

P3. 'Cicero' in P1 is 'Cicero' in P2

P4. "'Cicero' in P1' in P3 is 'Cicero' in P1 and "'Cicero' in P2' in P3 is 'Cicero' in P2

P5. "'Cicero' in P1' in P3' in P4 is "'Cicero' in P1' in P3 and "'Cicero' in P2' in P3'

in P4 is “Cicero’ in P2’ in P3

(infinite regress)

Someone is bald and well-read

Therefore, the best strategy to avoid infinite regress is to accept the transparency of co-reference of same-saying at the very beginning. So much for the second similarity between ‘I-You’ translation and ‘I-I’ same-saying.

The third similarity between ‘I-I’ same-saying and ‘I-You’ translation is that both of them are not affected by any specific misrepresentations of the relevant subject. I call this feature the ‘immunity to the effect of misrepresentation’. This immunity is closely related to the immunity to error through misidentification that I discussed in Chapter 2, but they are not identical. By ‘misidentification’, philosophers mean both the *de re* misidentification and which-object identification. I have discussed the two types of misidentification in Chapter 2. But by ‘misrepresentation’, I mean the mis-ascription of properties to an individual. It is clear that my ‘I-I’ co-reference is not affected by my misrepresentation of myself. If I correctly represent myself as ‘Smith’s best friend’, but at the same time, I misrepresent myself as ‘the author of the paper *Attitudes de dicto and de se*’, my utterance ‘I am Smith’s best friend’ still shares co-aboutness with my utterance ‘I am the author of the paper *Attitudes de dicto and de se*’, as both of them are about me. This is because, according to Ismael’s externalist explanation, the co-reference of ‘I-I’ same-saying is secured externally by a series of external factors. But the misrepresentation of the referent of ‘I’ takes place at the level of thought, which, according to Ismael’s ‘external/internal’ distinction, belong to the internal range. The internal factors cannot affect what is secured and explained by external factors, therefore the externally secured ‘I-I’ co-reference is immune to the effect of misrepresentation of the referent of ‘I’.

The co-reference of ‘I-You’ translation is also immune to the effect of

misrepresentation. But this immunity is more philosophically interesting than the immunity of 'I-I' co-reference. To reveal its philosophical interest, let us compare two scenarios, one of which involves 'I-You' translation, the other of which involves a different translation between 'I' and another term:

Imagine that Smith mistakenly believes that Oliver is a library thief who steals books from libraries (but in fact Oliver is not). Oliver meets Smith and wants to borrow Smith's textbook. Here are two possible conversations between Smith and Oliver

Scenario 1	Scenario 2
Oliver: 'Can I borrow your textbook?' Smith: 'I won't lend it to you.'	Oliver: 'Can I borrow your textbook?' Smith: 'I won't lend it to a library thief.'

In both scenarios, Smith has misrepresented Oliver in the same way: Oliver has been misrepresented as a library thief. In Scenario 2, this misrepresentation prevents Smith from being co-referential with Oliver (because the phrase 'a library thief' in Smith's words does not co-refer with the term 'I' in Oliver's words), and therefore they fail to fulfill the co-aboutness condition of communication. However, in Scenario 1, the same misrepresentation does not prevent Smith from being co-referential with Oliver (since Smith's 'you' co-refers with Oliver's 'I'), and therefore they succeed in fulfilling the co-aboutness condition. It is a good example which shows that the 'I-You' co-reference is immune to the effect of misrepresentation. As I have argued in previous sections, the establishment of co-reference of 'I-You' translation does not depend on the semantic consumptions of 'I' and 'you'. This in turn implies that the establishment of 'I-You' co-reference does not depend on any specific representation (and/or misrepresentation) of the referent of 'I' and 'You'. Therefore, the co-reference of 'I-You' translation is not affected by misrepresentations about any participants of the de se communication.

The 'immunity to effect of misrepresentation' implies that co-reference of 'I-You' translation can be compatible with various different representations of the referent of

'I' (and 'you'). This enables the participants of de se communication to have different ideas about which property can be ascribed to the referent of 'I' (and 'You'), but at the same time they still focus on the same referent of 'I' (and 'you'), i.e. the co-aboutness condition is still satisfied. This how the 'I-You' translation fulfils the second criterion I proposed in Section 3.1.1, which is that the connection between translation and co-aboutness should be loose enough to leave open the possibility that participants of the communication can have disagreements over the nature and properties of the subject matter. Other translations, for example the translation between 'I' and 'a library thief', cannot fulfil this criterion, because when a participant of de se communication adopts other translations, their different ideas about the referent of 'I' make it highly likely that they fail to focus on the same subject matter.

After examining the 'I-You' translation by all the three criteria that I proposed in Section 3.1.1, I conclude that the 'I-You' translation passes all the examinations. 'I-You' translation is the 'X translation' that I am looking for. The translation transmission model of de se communication can now be further specified by introducing the 'I-You' translation into its formulation. So, we have:

De Se Communication** = Translation Transmission Model ('I-You' translation) →
Common Ground (co-aboutness) [The '→' stands for a guarantee relation]

It is clear that the central question of this dissertation has been answered:

Participants of de se communication achieve the co-aboutness condition by using the 'I-You' translation.

3.3 Summary

An essential requirement of successful communication is that participants of communication must talk about the same subject matter. I call this requirement the co-aboutness condition of communication. According to the traditional picture of

communication, what is communicated in a communication is transmitted between participants of that communication in accordance with the replication transmission model. In that case, the co-aboutness condition is satisfied by same-saying, which is the replication of one another's linguistic expressions. The basic idea behind the same-saying strategy is that we talk about the same subject matter if we use the same words. However, for an idiosyncratic type of communication, which is the de se communication, same-saying simply does not work. De se communication is the communication of de se thought. In de se communication, we will definitely talk past each other if we just replicate one another's words. For example, if I say, 'I am happy' and you reply, 'I am happy', we are not communicating with each other because I am talking about my feelings and you are talking about yours. So, how can we satisfy the co-aboutness condition in de se communication? This question is the central question of this dissertation. I argue that in de se communication, what is communicated is transmitted in accordance with a new model, which is the translation transmission model. In order to satisfy the co-aboutness condition of communications which conform to the translation transmission model, a new strategy must be developed. I call the new strategy 'translation'. Participants of a de se communication talk about the same subject matter by means of translating one another's words. The most important part of this translation strategy is the translation of first person terms. The singular first person indexical 'I' is a representative of first person terms. So, in this dissertation I focus on the translation of 'I'. For the purpose of fulfilling the co-aboutness condition in de se communication, which term is a good translation of 'I'? In order to answer this question, I introduce Worsnip's analogy between the intrapersonal incoherence and interpersonal disagreement. Based on this analogy, I develop a further analogy between the intrapersonal de se inference and the interpersonal de se communication. This analogy gives us an important clue for looking for a good translation of 'I': if 'X' is a good translation of 'I', then the co-referential relation between 'X' and 'I' must well imitate the co-referential relation between two tokens of 'I'. So, the term whose co-referential relation with 'I' best imitates the co-

referential relation between two tokens of 'I' is the good translation of 'I' that I am looking for. This term, I argue, is 'you'. The analogy between the 'I-I' same-saying and 'I-You' translation is sustained by three similarities between them. The first similarity is that the 'I-I' co-reference in de se inference and the 'I-You' co-reference in de se communication are both explained by explanations of the externalist approach, and their externalist explanations are analogous to one another. This similarity is the most important one of the three similarities. Why is there 'I-You' co-reference in de se communication? How does my 'I' co-refer with your 'you'? I argue that the co-reference of 'I-You' translation is established by the match between the mutual recognition relation and the 'I-You' reciprocity in de se communication. The second similarity is that the 'I-I' co-reference is transparent to the producer of de se inference, just as the 'I-You' co-reference is transparent to the participants of de se communication. The third similarity is that both the 'I-I' co-reference and the 'I-You' co-reference are immune to the effect of misrepresentation. I articulate those three similarities. This justifies the legitimacy of the analogy between 'I-I' same-saying and 'I-You' translation. During my articulation, I examine the 'I-You' translation using three criteria for determining the proper connection between translation and co-aboutness. After the examinations, I conclude that the 'I-You' translation satisfies all three criteria. At the end of the dissertation, I provide my answer to the central question of this dissertation: when participating in a de se communication, we fulfill the co-aboutness requirement by using 'I-You' translation.

Chapter IV Conclusion

In this chapter, I present the main conclusions of this dissertation.

I begin the first chapter of my dissertation by asking what the real problem of *de se* attitudes is. Philosophers have already provided many answers to this question. I introduce Dilip Ninan's, Matheus Valente's and François Recanati's answers to this question. Dilip Ninan argues that the real problem concerning the *de se* attitude is that there is a conflict between multiple people's agreement on a certain doxastic attitude and an explanation of their actions based on that attitude, when the attitude in question is a *de se* attitude. Matheus Valente argues that the real problem associated with the *de se* attitude takes place when we try to share our *de se* attitudes with other people. He argues that there are normally three principles to which an attitude together with its content normally conforms. However, when I try to share a *de se* attitude with others, Valente argues, I have to reject at least one of those principles. François Recanati argues that the communication of indexical attitudes raises a challenge for the traditional understanding of how a (verbal) communication works, due to what he calls the limited accessibility. I somehow agree with Recanati on his general strategy for answering this question, but I do not agree with the specific answer that he provides.

For me, one of the real problems of *de se* attitude is how to satisfy the co-aboutness requirement of a *de se* communication. For communication that conforms to the replication transmission model, the co-aboutness condition is achieved by utilizing the same-saying strategy. But in my view, *de se* communication is distinguished from other types of communication, because the replication transmission model is incompatible with *de se* communication. Instead, *de se* communication necessarily conforms to the translation transmission model.

I then introduce Frege's, Perry's and Recanati's characterizations of the translation transmission model. According to Frege's characterization of the translation

transmission model, what is translated in a de se communication is the sense of first person sentence. The translation operation takes place in the speaker's mind when she tries to communicate her de se belief to someone else, and the translation operation is to substitute a public sense of a first person sentence for its private sense. Perry's characterization of the translation transmission model is different from Frege's characterization. For Perry, what is translated in an indexical communication is the belief state, not the object of belief, and the translation operation takes place in the hearer's mind when she tries to acquire a belief based on her acceptance of the speaker's utterance. The translation operation is to substitute the hearer's belief state for the speaker's belief state but preserve the object of belief. Recanati's characterization of the translation transmission model is more sophisticated compared with Frege's and Perry's characterizations. But I am not satisfied by Recanati's characterization for a couple of reasons. First, I am not satisfied with Recanati's argument because of its rejection of *thought expressed by utterance*. He argues that in order to reject the replication transmission model, the idea of *thought expressed by utterance* must first be rejected. I think this solution is problematic, because it is in contradiction with a basic intuition that is prevalent in our verbal communications. The intuition is that I acquire information from the person whom I communicate with based on my understanding of her words. My second dissatisfaction is that the status of language in Recanati's explanation of de se communication is perplexing: it seems that at first he rejects the role that language usually plays in communications when he claims that the content of the utterance must be abandoned, but then he introduces the role of language back into his picture when he argues for the coordination role of linguistic mode of presentation. My third dissatisfaction with Recanati's proposal is that he does not provide a clear explanation of how the role that the linguistic mode of presentation plays in his conception of de se communication should be cashed out.

Based on my analysis of the three proposals of the translation transmission model, I propose my own characterization of that model. In my opinion, the translation transmission model has four important features. First, the aim of my adoption of the

translation transmission model in de se communication is to meet the co-aboutness requirement. Second, the translation operation involved in the translation transmission model takes place in the hearer's mind. It is the hearer, not the speaker, who is responsible for conducting the translation. Third, what is translated in de se communication is not only the de se thought expressed by the first person linguistic expression, but also the linguistic expression itself. This is because in verbal communication, the hearer gets access to what the speaker expresses by understanding the meaning of the utterance that the speaker uses to express her thought. The fourth feature is the most important one. I argue that in translation transmission model, not every type of translation is evaluated as being equally effective, especially when we use that model to explain the de se communication. There is a certain type of translation which can better secure the achievement of the co-aboutness condition in de se communication, compared with other types of translation. I call this type of translation the X translation. In Chapter 3, I reveal what the X translation actually is.

Although the second chapter is not about the de se communication *per se*, it is still important to my dissertation, because it is about what is communicated in the de se communication, i.e. de se thought. In order to express my de se thoughts and communicate them to someone else, I must first be able to produce de se thoughts about myself and think of them in mind. In Chapter 2, I try to articulate one aspect of the production of de se thought, which concerns whether or not our production of de se thought involves identification (and/or mis-identification). I call this the 'identification problem of de se thought'. In order to clarify whether or not a process of identification is involved in our production of de se thought, I introduce two theses as the background of my clarification. They are the essential indexical thesis and the immunity to error through misidentification thesis (IEM). If we take the essential indexical thesis to be a thesis about the connection between the first person perspective and the performance of intentional action, according to a certain interpretation of the thesis, the essential indexical thesis implies an identification

thesis concerning de se thought, which is that I identify myself by entertaining a de se thought. But if we attribute the epistemic property of IEM to de se thought, according to some interpretations of the immunity thesis, this attribution results in another identification thesis concerning de se thought, which is that de se thought is identification-free. It seems that the identification thesis implied by the essential indexical thesis may be incompatible with the identification thesis implied by the immunity thesis. If so, this would be a very problematic result.

However, I argue that the alleged incompatibility is only apparent. De se thoughts are both essentially indexical and immune to error through misidentification. There is no tension between those two theses. First, I introduce some general characterizations of the essential indexical thesis. Then I proceed to introduce two skeptical studies of the thesis, which are Cappelen & Dever's criticism of the essentiality of indexicality *per se* and Magidor's criticism of the property account of de se content. I argue that Cappelen & Dever's criticism is not comprehensive enough to cover all aspects of essential indexical, and Magidor's criticism is based on an inaccurate reading of David Lewis' property account of de se content. Cappelen & Dever's de se skepticism, as I understand it, leaves open the possibility that an idiosyncratic co-aboutness of de se thought is essential for our performance of intentional actions. My discussion on Magidor's criticism of David Lewis shows that the essential indexical thesis is closely related to an identification thesis, which is that in producing and holding a de se thought, I do not need to appeal to any representation of myself for the purpose of identifying myself. Although there is plenty of room for debates about what the identification thesis exactly means, it reveals a close connection between de se thought and the identification. In order to discuss the implication of that connection, I introduce Matthew Babb's indexical intention defense and Juliana Faccio Lima's action-explanation defense. Both of them argue that an identification thesis implied by the essential indexical thesis is crucial for establishing the connection between the first person perspective and the performance of intentional actions. However, this argument may seem to be incompatible with another thesis concerning de se thought,

which is the immunity thesis. Philosophers seem to agree that some de se thoughts, if not all de se thoughts, have the epistemic property of immunity to error through misidentification. According to the Simple Account/Simple Explanation of IEM, de se thought is immune to the error of misidentification, because producing and entertaining de se thought does not involve identification. This seems to be incompatible with the identification thesis implied by the essential indexical thesis. But I argue that this incompatibility is apparent.

What does 'being immune to the error of misidentification' really mean for de se thought? To answer this question, we need to clarify first what identification (and subsequently, misidentification) means here. I introduce Wittgenstein's distinction between the use of I as subject and the use of I as object, Gareth Evans' distinction between 'identification-dependent' and 'demonstrative identification', James Pryor's distinction between de re identification and which-object identification, and Manuel García-Carpintero's distinction between identification_P and identification_C. But I am not satisfied by those philosophers' solutions to the problem that I mention above, because their solutions are based on those distinctions they draw, and those distinctions are heavy theoretical commitments for their solutions. What is more, those distinctions face philosophical challenges. Because of those theoretical commitments, the Simple Account is not that simple. My own plan is to solve the apparent incompatibility between the essential indexical thesis and the IEM thesis by rejecting the Simple Account and replacing it with the Definition Account. The basic idea of Definition Account is that de se thought has the property of IEM simply because of its definition. Since there is no 'identification free' condition contained in the Definition Account, there is no incompatibility between the essential indexical thesis and the Definition Account of IEM. A crucial step for establishing my Definition Account is to vindicate my distinction between de se thought and first person thought. To vindicate this distinction, I need to demonstrate the existence of non-de-se first person thought. I appeal to some deliberately designed scenarios to discuss the possibility that I do not regard a first person thought in my mind as being about myself, and the

possibility that a first person thought in my mind is not only about myself, but also about other individuals. The basic idea here is that because of the cognitive impact of misidentification, including both *wh*-misidentification and *de re* misidentification, some of our first person thoughts are not *de se* thoughts for us.

After addressing my discussion on the identification aspect of *de se* thought, I return to the topic that I discussed in the first chapter, which is the co-aboutness problem of *de se* communication. In Chapter 3, I present my articulation of the fourth feature of the translation transmission model that I introduced at the end of Chapter 1. A crucial part of this articulation is to introduce a specific type of translation into the translation transmission model. I call this specific type of translation the 'X translation'. Just as not every type of same-saying is equally effective at securing the satisfaction of co-aboutness requirement, not every type of translation is equally effective at securing the achievement of the co-aboutness condition in *de se* communication. I argue there is a specific type of translation, which guarantees the co-aboutness for *de se* communication better than other types of translation. This is the 'X translation'. I reveal what 'X translation' is in Chapter 3. First, I introduce three normative theses concerning the connection between the translation and the co-aboutness condition in communications that conform to the translation transmission model. Those theses are the modified version of Laura Schroeter's theses that I introduced in Section B of the Introduction. Schroeter argues that a proper connection between the same-saying and the co-aboutness should neither be too loose nor be too tight, and it should be transparent to participants of communication. She uses those theses as criteria for determining whether the co-aboutness condition is properly achieved or not in a communication. Based on Schroeter's argument, I argue that a proper connection between the translation and the co-aboutness should neither be too loose nor be too tight, and it should be transparent to participants of communication. For me, those are criteria for determining whether or not the 'X translation' properly securing the fulfillment of the co-aboutness requirement in *de se* communication. Second, I introduce Alex Worsnip's analogy strategy for articulating a wider conception of

disagreement as a background for my subsequent discussion. Worsnip argues that there is an analogy between the interpersonal disagreement and the intrapersonal incoherence, so he encourages us to conceive disagreement as the ‘interpersonal incoherence’. I think the analogy between the interpersonal case and the relevant intrapersonal case is a seminal approach, and I adopt this approach to draw an analogy between the intrapersonal de se inference and the interpersonal de se communication. Let us call the co-referentiality of tokens of ‘I’ in de se inference the ‘I-I’ co-reference. I argue there is a certain type of translation, namely, the ‘I-You’ translation in de se communication, whose co-referentiality is analogous to the ‘I-I’ co-referentiality in de se inference. ‘I-You’ translation is the ‘X translation’ that I am looking for. Let us call the co-referentiality of ‘I-You’ translation the ‘I-You’ co-reference. The analogy between ‘I-You’ co-reference and ‘I-I’ co-reference is plausible, because there are some non-trivial similarities between them. The first similarity is that both of them are established without appealing to any specific representation of their subjects. Therefore, they are not affected by any misrepresentation of their subjects. The second similarity is that the co-referentiality of ‘I-I’ same-saying is transparent to the producer of de se inference, just as the co-referentiality of ‘I-You’ translation is transparent to the participants of de se communication. The third similarity between ‘I-I’ coreference and ‘I-You’ coreference is the most important one, because it explains why they have the first and the second similarities. The third similarity is that the ‘I-I’ co-reference and the ‘I-You’ co-reference are both explained by the same approach, which I call the externalist approach. Here I introduce Jenann Ismael’s externalist explanation of the co-reference of I-thoughts in a single person’s mind. Ismael’s externalist approach not only explains why the co-reference of I-thoughts can be established without appealing to any specific representation of the self, but also explains why the co-reference of I-thoughts is transparent to their thinker. This is a paradigm for my subsequent discussion on ‘I-You’ translation’s fulfillment of the co-aboutness condition. I argue that based on Ismael’s externalist approach, a similar externalist explanation of ‘I-You’ co-reference can be developed. Why does my ‘I’ co-refer with your ‘you’ in a de se communication? According to my externalist explanation, the ‘I-You’ co-reference is

explained and secured by factors of the external aspect of communication. But what are those factors? A crucial task of my externalist explanation is to answer this question. I argue that those factors are the mutual recognition relation between participants of a communication and the 'I-You' reciprocity as a norm which governs the use of 'I' and 'You'. The match between the mutual recognition and the 'I-You' reciprocity not only explains why the co-reference of 'I-You' translation can be secured without appealing to specific representations of either the referent of 'I' or the referent of 'You', but also explains why the 'I-You' co-reference is recognizable for participants of de se communication. In this sense, it explains the transparency of the co-referentiality of 'I-You' translation. Once the externalist explanation of 'I-You' co-reference is addressed, I continue to explain why other possible translations for de se communication do not surpass the 'I-You' translation. This is because, unlike 'I-You' co-reference, the co-referential relations of other translations do not have the three similarities with 'I-I' coreference that I discuss above. At the end of Chapter 3, I examine whether the 'I-You' translation follows the three criteria that I propose at the beginning of this chapter. The result of my examination shows that 'I-You' translation does follow those criteria, so I think it is safe to say that 'I-You' is the 'X translation' that I am searching for. Since 'X translation' is a crucial part of the translation transmission model, by clarifying the 'X translation', we make substantial progress in articulating the translation transmission model.

Bibliography

Anscombe, G.E.M., *The Collected Philosophical Papers of G.E.M. Anscombe, Volume II: Metaphysics and the Philosophy of Mind*, Oxford: Basil Blackwell, 1981.

Babb, Matthew., 'The Essential Indexicality of Intentional Action', in *The Philosophical Quarterly*, Volume 66, Issue 264, pp. 439–457, 2016.

Barth, Christian., *Objectivity and the Language-Dependence of Thought*, New York: Routledge, 2011.

Bermúdez, José Luis., *Understanding 'I': Language and Thought*, Oxford University Press, 2017.

Block, Ned., 1990, 'Inverted Earth', in *Philosophical Perspectives* (Volumes 4), J. Tomberlin (ed.), Atascadero, CA: Ridgeview Publishing Company.

Boer, Steven E., and Lycan, William G., 'Who, Me?', in *Philosophical Review*, Vol. 89, Issue. 3, pp. 427-466, 1980.

Bortolotti, Lisa. and Broome, Matthew., 'A role for ownership and authorship in the analysis of thought insertion', in *Phenomenology and the Cognitive Sciences*, Vol 8, Issue 2, pp. 205-224, 2009

Brinck, Ingar., *The Indexical 'I': the First Person in Thought and Language*, Kluwer Academic Publishers, 1997.

Cappelen, Harman., and Dever, Josh., *The Inessential Indexical: On the Philosophical Insignificance of Perspective and the First Person*, Oxford University Press, 2013.

Coliva, Annalisa., Thought Insertion and Immunity to Error Through Misidentification, in *Philosophy, Psychiatry, & Psychology*, Vol 9, No. 1, pp. 27-34,

2002.

Coliva, Annalisa., 'The First Person: Error through Misidentification, the Split between Speaker's and Semantic Reference, and the Real Guarantee', in *The Journal of Philosophy*, Vol. 100, No. 8, pp. 416-431. 2003.

Coliva, Annalisa., 'Error through misidentification: some varieties', in *Journal of Philosophy*, Vol. 103, No.8, pp. 403-425, 2006.

Coliva, Annalisa., 'Stopping points: "I", immunity and the real guarantee', in *Inquiry*, Vol 60, Issue 3, pp. 233-252, 2017.

De Gaynesford, Maximilian., *I: The Meaning of the First Person Term*, Oxford University Press, 2006.

Egan, Andy., 'De Se pragmatics', in *Philosophical Perspectives*, Volume 32, Issue 1, pp. 144-164, 2018.

Evans, Gareth., *The Varieties of Reference*, Oxford: Oxford University Press, 1982.

Feit, Neil., *Belief About the Self*, Oxford: Oxford University Press, 2008.

Feit, Neil., and Capone, Alessandro., ed. *Attitudes De Se: Linguistics, Epistemology, Metaphysics*, CSLI Publications, 2013.

Frege, Gottlob., 'Thought', in Beaney, Michael. Ed. *The Frege Reader*, pp. 325-345, Oxford: Blackwell, 1997.

Garcia-Carpintero, Manuel., 'The Self File and Immunity to Error Through Misidentification', in *Disputatio*, Vol. V, No. 36, pp. 191-206, 2013.

Garcia-Carpintero, Manuel., 'De Se Thought', Oxford Handbook Online, <https://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199935314.001.0001/oxfordhb-9780199935314-e-61>, 2015.

Garcia-Carpintero, Manuel., and Torre, Stephen. Ed. *About Oneself: De Se thought*

and communication, Oxford University Press, 2016.

Garcia-Carpintero, Manuel., 'The philosophical significance of the De Se', in *Inquiry*, Vol. 60, No. 3, pp. 253 - 276, 2017.

Gertler, Brie., *Self-Knowledge*, London: Routledge, 2011.

Gjelsvik, Olav. 'Indexicals: what they are essential for', in *Inquiry*, Vol. 60, No. 3, pp. 295-314, 2017.

Glock, Hans-Johann, and Hacker, P.M.S., 'Reference and the First Person Pronoun', in *Language & Communication*, Vol. 16, No. 2, pp. 95 105, 1996.

Harman, G., 'The Intrinsic Quality of Experience', in *Philosophical Perspectives (Volume 4)*, J. Tomberlin (ed.), Atascadero, CA: Ridgeview Publishing Company, 1990.

Heal, Jane., 'Second Person Thought', in *Philosophical Explorations*, Vol 17, Issue 3, pp. 317-331, 2014.

Heck, Richard G. (Heck, Richard Kimberly), 'Do Demonstratives Have Senses?', in *Philosophers' Imprint*, Vol. 2, No. 2, <http://www.philosophersimprint.org/002002/>, 2002.

Holton, Richard., 'Primitive Self-Ascription: Lewis on the *De Se*', in *A Companion to David Lewis*, Loewer, Barry. and Schaffer, Jonathan. (eds), pp. 399-410, Oxford: Wiley-Blackwell, 2015.

Ismael, Jenann., 'Me, Again', In *Time and Identity: Topics in Contemporary Philosophy, Volume 6*, Keirn-Campbell, O'Rourke and Shier (eds.), pp. 209-228, Cambridge: MIT Press, 2010.

Ismael, Jenann., 'Reflexivity, Fixed Points, and Semantic Descent: How I Learned to Stop Worrying and Love Reflexivity', in *Acta Analytica*, Vol. 26, Issue. 4, pp 295–310, 2011.

Ismael, Jenann., 'Immunity to Error as an Artefact of Transition Between

Representational Media', in Prosser, Simon., Recanati, Francois. Ed. *Immunity to Error through Misidentification: New Essays*, pp. 62-80, Cambridge University Press, 2012.

Jaszczolt, Kasia M., 'First-Person reference in Discourse: Aims and Strategies', in *Journal of Pragmatics*, special issue 'Focus on the Speaker', Volume 48, Issue 1, 2013.

Kapitan, Tomis., Review of *The Inessential Indexical: On the Philosophical Insignificance of Perspective and the First Person*, on Notre Dame Philosophical Reviews, 19/09/2014, <https://ndpr.nd.edu/news/the-inessential-indexical-on-the-philosophical-insignificance-of-perspective-and-the-first-person/>

Kaplan, David, 'Demonstratives', in Almog, Perry, and Wettstein Ed. *Themes from Kaplan*, pp. 481-563, 1989.

Kindermann, Dirk., 'Varieties of Centering and De Se Communication', in Garcia-Carpintero, Manuel., Torre, Stephen. Ed. *About Oneself: De Se thought and communication*, pp. 307-339, Oxford University Press, 2016.

Lewis, David., 'Attitudes De Dicto and De Se', in *The Philosophical Review*, Vol. 88, No. 4, pp. 513-543, 1979.

Lima, Juliana Faccio, 'Indexicality and action: why we need indexical beliefs to motivate intentional actions', in *Inquiry*, 2018. DOI: 10.1080/0020174X.2018.1487881

Loewer, Barry. and Schaffer, Jonathan. (eds), *A Companion to David Lewis*, Oxford: Wiley-Blackwell, 2015.

Longworth, Guy., 'You and me', in *Philosophical Explorations*, Vol 17, Issue 3, pp. 289-303, 2014.

Magidor, Ofra., 'The Myth of the De Se', in *Philosophical Perspectives*, Vol. 29, No. 1, pp. 249-283, 2015.

Maier, Emar., 'Why My I Is Your You: On the Communication of De Se Attitudes', in Garcia-Carpintero, Manuel., Torre, Stephen. Ed. *About Oneself: De Se thought and*

communication, pp. 220-245, Oxford University Press, 2016.

May, Robert., 'The Invariance of Sense', in *The Journal of Philosophy*, Vol. 103, No. 3, pp. 111-144, 2006

Morgan, Daniel., 'Can you think my "I"-thoughts?', in *Philosophical Quarterly*, Vol. 59, No. 234, pp. 68-85, 2009.

Morgan, Daniel., 'Immunity to error through misidentification: what does it tell us about the de se?', in Prosser, Simon., Recanati, Francois. Ed. *Immunity to Error through Misidentification: New Essays*, pp. 104-123, Cambridge University Press, 2012.

Morgan, Daniel., 'The Demonstrative Model of First-Person Thought', in *Philosophical Studies*, Vol. 172, Issue. 7, pp. 1795–1811, 2015

Moss, Sarah., 'Updating as Communication', in *Philosophy and Phenomenological Research*, Vol. 85, pp. 225–248, 2012

Ninan, Dilip, 'De Se Attitudes: Ascription and Communication', in *Philosophy Compass*, Vol. 5, Issue. 7, pp. 551-567, 2010.

Ninan, Dilip., 'What is the Problem of De Se Attitudes', in Garcia-Carpintero, Manuel., Torre, Stephen. Ed. *About Oneself: De Se thought and communication*, pp. 86 -110, Oxford University Press, 2016.

O'Brien, Lucy., *Self-Knowing Agents*, Oxford University Press, 2007.

Pagin, Peter., 'De Se Communication: Centred or Uncentred?', in Garcia-Carpintero, Manuel., Torre, Stephen. Ed. *About Oneself: De Se thought and communication*, pp. 272-306, Oxford University Press, 2016.

Peacocke, Christopher., *The Mirror of the World: Subjects, Consciousness, and Self-Consciousness*, Oxford University Press, 2014.

Perry, John., 'The Problem of the Essential Indexical', in *Noûs*, Vol. 13, pp. 3-21, 1979.

Perry, John., Blackburn, Simon., 'Thought without Representation', in *Aristotelian Society Supplementary Volume*, Vol. 60, Issue 1, pp. 137–166, 1986.

Perry, John., *The Problem of The Essential Indexical and Other Essays*, Stanford, California: CSLI Publications, 2000.

Predelli, Stefano., 'Intentions, indexicals, and communication', in *Analysis*, Vol. 62, pp. 310-316, 2002.

Prosser, Simon., Recanati, Francois. Ed. *Immunity to Error through Misidentification: New Essays*, Cambridge University Press, 2012.

Prosser, Simon., 'Why Are Indexicals Essential', in *Proceedings of the Aristotelian Society*, Vol. 115 (Issue 3), pp. 211-233, 2015.

Prosser, Simon., 'Shared modes of presentation', in *Mind and Language* Vol. 34, Issue 4, pp. 465-482, 2019.

Pryor, James, 'Immunity to Error through Misidentification', in *Philosophical Topics* (Volume 26): 271-304, 1999.

Recanati, François., *Perspectival Thought: A Plea for (Moderate) Relativism*, Oxford University Press, 2007.

Recanati, François., 'Mental Files and Identity' (Gareth Evans Memorial Lecture, Oxford, 25 January 2011), <http://www.unige.ch/lettres/philo/mulligan/festschrift/Recanati-paper.pdf>

Recanati, François., *Mental Files*, Oxford University Press, 2012.

Recanati, François., *Mental Files in Flux*, Oxford University Press, 2016.

Recanati, François., 'Indexical Thought: The Communication Problem', in Garcia-Carpintero, Manuel., Torre, Stephen. Ed. *About Oneself: De Se thought and communication*, pp. 141-178, Oxford University Press, 2016.

Romdenh-Romluc, Komarine., 'First-Person Thought and the Use of "I"', in *Synthese*, Vol. 163, No. 2, pp. 145-156, 2008

Rovne, Carol, A., 'The Epistemology of First-Person Reference', in *The Journal of Philosophy*, Vol. 84, No. 3, pp. 147-167, 1987.

Salje, Lea., 'The Subjective Perspective in Introspection', in *Journal of Consciousness Studies*, 23: pp. 3-4, 2016.

Salje, Léa., 'Thinking About You', in *Mind*, Vol. 126, Issue 503, pp. 817-840, 2017.

Schroeter, Laura., 'Why Be an Anti-Individualist?', in *Philosophy and Phenomenological Research*, Vol. 77, No. 1, pp. 105-141, 2008.

Schroeter, Laura., 'Bootstrapping our way to samesaying', in *Synthese*, Vol. 189, Issue 1, pp. 177-197, 2012.

Schroeter, Laura., Schroeter, François., 'Rationalizing Self-Interpretation: the epistemology of reference in Rationalizing Self-Interpretation', in Daly C. (eds) *The Palgrave Handbook of Philosophical Methods*, pp. 419-447, Palgrave Macmillan, London, 2015.

Schroeter, Laura. and Schroeter, François., 'Semantic Deference versus Semantic Coordination', in *American Philosophical Quarterly*, Vol. 53, Issue 2, pp. 193-210, 2016.

Seeger, Max., Authorship of thoughts in thought insertion: What is it for a thought to be one's own?, in *Philosophical Psychology*, Vol 28, Issue 6, pp. 837-855, 2015.

Shoemaker, Sydney., *Identity, Cause, and Mind*, Cambridge: Cambridge University Press, 1984.

Sollberger, Michael., Making Sense of an Endorsement Model of Thought-Insertion, in *Mind & Language*, Vol. 29, No. 5, pp. 590–612, 2014.

Stalnaker, R.C., 'Assertion', in *Context and Content: Essays on Intentionality in Speech and Thought*, pp. 78-95, Oxford: Oxford University Press, 1999.

Stalnaker, R.C., 'Indexical Belief', in *Context and Content: Essays on Intentionality in Speech and Thought*, pp. 130-149, Oxford: Oxford University Press, 1999.

Stalnaker, R.C., 'Common ground', in *Linguistics and Philosophy*, Vol. 25, Issue 5-6, pp. 701–721, 2002.

Stanley, Jason., *Know How*, Oxford: Oxford University Press, 2011.

Torre, Stephen., 'De Se Thought and Communication: An Introduction', in Garcia-Carpintero, Manuel., Torre, Stephen. Ed. *About Oneself: De Se thought and communication*, pp. 1-24, Oxford University Press, 2016.

Valente, Matheus., 'What is special about indexical attitudes?', in *Inquiry*, Vol. 61, No. 7, pp. 692-712, 2017, <https://doi.org/10.1080/0020174X.2017.1402701>.

Wittgenstein, Ludwig., *The Blue and Brown Books*, Oxford: Blackwell, 1958.

Wittgenstein, Ludwig., *Philosophical Investigations*, 4th edition, 2009, P.M.S. Hacker and Joachim Schulte (eds. and trans.), Oxford: Wiley-Blackwell.

Worsnip, Alex., 'Disagreement as Interpersonal Incoherence', in *Res Philosophica*, Vol. 96, No. 2, pp. 245-268, 2019.