Pedro Henrique Gomes Muniz

The dynamics of *de se* thoughts

PUC-Rio - Certificação Digital Nº 1313515/CA

Tese de Doutorado

Thesis presented to the Programa de Pós-graduação em Filosofia of PUC-Rio in partial fulfillment of the requirements for the degree of Doutor em Filosofia.

Advisor: Prof. Ludovic Soutif

Rio de Janeiro
June 2017

Pontifícia Universidade Católica
DO RIO DE JANEIRO

Pedro Henrique Gomes Muniz

"The dynamics of *de se* thoughts"

Tese apresentada como requisito parcial para a obtenção do grau de Doutor pelo Programa de Pós-Graduação em Filosofia do Departamento de Filosofia do Centro de Teologia e Ciências Humanas da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada:

Prof. Ludovic Soutif
Orientador
Departamento de Filosofia – PUC-Rio

Prof. Carlos Mario Marquez Sosa
Departamento de Filosofia – PUC-Rio

Prof. Marco Antonio Caron Ruffino
Universidade Estadual de Campinas – UNICAMP

Profª. Karla Chediak
Universidade do Estado do Rio de Janeiro – UERJ

Prof. André Leclerc
Universidade de Brasília – UnB

Profª. Monah Winograd
Coordenadora Setorial de Pós-Graduação e Pesquisa do Centro de Teologia e Ciências Humanas – PUC-Rio

Rio de Janeiro, 5 de junho de 2017.

Pedro Henrique Gomes Muniz

After his graduation at Universidade Federal do Ceará in 2011, Pedro Muniz obtained a Master's degree at PUC-Rio in 2013 with the support of CNPq. His main area of interest has been, from the outset, philosophy of language and of mind. Between 2010-2011, he took part in a research program for undergraduate students supported by CNPq to study spontaneous linguistic understanding. Between 2014-2015, he was in an exchange program in Paris to participate in a research entitled "Proofs, demonstrations and representation", financed by CAPES, at the Institut d'Histoire et de Philosophie des Sciences et des Techniques. He was a visiting student at Institut Jean-Nicod during the same period.

Acknowledgements

I also want to thank the teachers and everyone who works hard at our department at PUC-Rio to help us become competent graduate students. I thank you all, for all the seminars and all the help given directly or indirectly, all of which has been certainly very important to me and to my fellow MA and PhD students.

I count myself lucky for having been able to count on true friends during these years I've lived in Rio. I would not have made it alone. There are a few names I must mention. Chris, for letting me live at her place and for so much more she's done for me; Max, for the time I spent at his house and for his friendship and support; Felippe, Mel and Fred, for all the moments of simple and sincere friendship we've shared; Daniel, for those times when we just sat and talked and the messages that gave me strength; Roberta, who has been always been there for me; all the OrginAs and all my friends from Toca RJ: what my life in Rio would be without you, I woudn't dare to imagine!

At last (though they are not last in my thought), I thank my parents, who are and will always be with me.

Abstract

Muniz, Pedro Henrique Gomes; Soutif, Ludovic (Advisor). **The dynamics of *de se* thoughts.** Rio de Janeiro, 2017. 182 p. Tese de Doutorado - Departamento de Filosofia, Pontifícia Universidade Católica do Rio de Janeiro.

The overall aim of this work is to show that we need to account for the dynamics of *de se* or I-thoughts (or else, thoughts usually expressed by the use of the first-person pronoun) within the broader picture of a dynamist theory of mental acts – of the kind advocated by Dokic (2001). I argue that the move is necessary if we are to deal with the so-called "problem of cognitive dynamics" in such a way that the account be able to capture the distinguishing features of *de se* thoughts and related attitudes. The doctoral dissertation is divided into four main parts. In Chapter One I review the arguments in support of the claim that *de se* thoughts are special and irreducible to other kinds of thoughts – in particular, *de re* and *de dicto*. Chapter Two tackles the issue of cognitive dynamics as originally put by Kaplan (1989) and discusses the extent to which it applies to *de se* thoughts *qua* indexical thoughts. Chapter Three considers some of the objections raised by *de se* skeptics (notably, Cappelen & Dever 2013) against the view that *de se* and essential indexicality on the whole are deep and interesting phenomena and tries to answer them. Chapter 4 shows that *de se* thoughts exhibit features not captured by a theory of indexicals. However, given that an account of their dynamics is still needed, I point toward the possibility of broadening the scope of the problem formulated by Kaplan for indexical thoughts. If the problem of cognitive dynamics, as I argue, outstrips indexical thoughts and concerns the unfolding of all intentional states over time, one might want to reformulate it as a problem of internal (mental) continuity. I conclude that a study of the conditions under which *de se* thinkings are entertained over time amounts to elaborating a (neo-Lockean) theory of personal identity.

Keywords

Cognitive dynamics; essential indexicality; *de se* thoughts.

Resumo

Muniz, Pedro Henrique Gomes; Soutif, Ludovic. **A dinâmica dos pensamentos *de se*.** Rio de Janeiro, 2017. 182 p. Tese de Doutorado - Departamento de Filosofia, Pontifícia Universidade Católica do Rio de Janeiro.

O objetivo global deste trabalho é mostrar que precisamos explicar a dinâmica de pensamentos *de se* ou pensamentos em primeira pessoa (ou ainda, pensamentos normalmente expressos com o uso do pronome da primeira pessoa) dentro de um quadro mais amplo de uma teoria dinamista dos atos mentais – nos moldes da teoria defendida por Dokic (2001). Argumento que esse movimento é necessário se quisermos lidar com o assim chamado "problema da dinâmica cognitiva" de tal forma que a explicação seja capaz de capturar as características distintivas dos pensamentos *de se* e atitudes relacionadas. A tese está dividida em quatro partes principais. No Capítulo 1, faço uma revisão dos argumentos a favor da afirmação de que pensamentos *de se* são especiais e irredutíveis a outros tipos de pensamentos – *de re* e *de dicto* em particular. O Capítulo 2 lida com o problema da dinâmica cognitiva como este foi originalmente formulado por Kaplan (1989), e discute até que ponto ele se aplica a pensamentos *de se* enquanto pensamentos indexicais. No Capítulo 3, considero algumas das objeções levantadas por céticos com relação ao *de se* (notadamente, Cappelen & Dever 2013) contra a ideia de que o *de se* e a indexicalidade essencial como um todo são fenômenos profundos e interessantes e que precisamos dar respostas a eles. O Capítulo 4 mostra que os pensamentos *de se* apresentam características que não são capturadas por uma teoria dos indexicais. No entanto, dado que uma explicação de sua dinâmica ainda é necessária, aponto para a possibilidade de expandir o escopo do problema que foi formulado por Kaplan para pensamentos indexicais. Uma vez que, como argumento, o problema da dinâmica cognitiva está para além de pensamentos indexicais e tem a ver com o desenrolar de todos os estados intencionais no decorrer do tempo, podemos reformulá-lo como um problema da continuidade interna (mental). Concluo que um estudo das condições nas quais pensamentos *de se* são apreendidos com o passar do tempo equivale a elaboração de uma teoria (neo-Lockeana) da identidade pessoal.

Palavras-chave

Dinâmica cognitiva; indexicalidade essencial; pensamentos *de se*.

Table of contents

Éste, no lo olvidemos, era casi incapaz de ideas generales, platónicas. No sólo le costaba comprender que el símbolo genérico perro *abarcara tantos individuos dispares de diversos tamaños y diversa forma; le molestaba que el perro de las tres y catorce (visto de perfil) tuviera el mismo nombre que el perro de las tres y cuarto (visto de frente). Su propia cara en el espejo, sus propias manos, lo sorprendían cada vez. Refiere Swift que el emperador de Lilliput discernía el movimiento del minutero; Funes discernía continuamente los tranquilos avances de la corrupción, de las caries, de la fatiga. Notaba los progresos de la muerte, de la humedad. Era el solitario y lúcido espectador de un mundo multiforme, instantáneo y casi intolerablemente preciso.*

Jorge Luis Borges, "Funes El Memorioso"

Un homme qui dort tient en cercle autour de lui le fil des heures, l'ordre des années et des mondes. Il les consulte d'instinct en s'éveillant, et y lit en une seconde le point de la terre qu'il occupe, le temps qui s'est écoulé jusqu'à son réveil; mais leurs rangs peuvent se mêler, se rompre.

Marcel Proust, *Du côté de chez Swann*

Introduction

Our apprehension of reality and of ourselves, as well as the formation and subsequent preservation of all the concepts that we use in our everyday life (a large part of which we employ unconsciously) are processes that necessarily happen over time. In that sense, and by the very fact that they are processes, they are dynamic – no concept can be acquired in an immediate way and in a determinate and eternal form. Even the supposedly unmistakable "I" (which in an important way reflects the conception we have of ourselves) is not static, being also subject to constant evolution and change. Still, in spite of these inevitable changes, the very formation and apprehension of concepts in general must depend on something being preserved. Apprehension in that sense must mean conservation of something, which simply means this: a part of that concept must *somehow* remain static – at least for certain periods of time, short though they may be. I suspect that those processes (that is, apprehension and preservation of concepts in general) depend on (among other things) a certain capacity of generalization. Borges's character Funes, a little like Heraclitus as depicted by Plato, ends up devoid of language, for he perceives absolutely all the changes in the world that surrounds him. For him, every single moment and absolutely every element of reality is unique, there is no connection between them, so he loses his capacity of generalizing. He is no longer able to cognitively preserve anything, since he is no longer capable of grasping any concept. On the other hand, there are those who on occasion lose themselves in time and space, those to whom changes and the flow of time can get cloudy or mixed up, as it is the case for the main character of Proust's famous novel *Du côté de chez Swann*. While Funes, on the one hand, is *only* able to perceive change, one may *not* realize that something important changed in one's surroundings. I think that in the latter case, at least in some circumstances, missing some of the changes around us does not necessarily entail failure in retaining some possible object of our thought (like ourselves, a certain place or a certain object or moment in time), some epistemic attitude we have, or that very thought. Part of what I want to enquire in this work is this: given that thoughts, concepts and ideas change or evolve through

time, and given that our relation to those thoughts, concepts and ideas also change, how do those processes occur? What is preserved and what is lost? And what determines that preservation and loss? What mechanisms must be in place so that we can say that the same thought or concept is preserved? What kind of *activity* must a thinking subject perform in order to compensate for the changing reality which consistitues and surronds the object of her thought? Are those activities or mechanisms conscious? In short (and in slightly more technical terms), how does internal continuity in relation to conceptual contents occur – how do we retain thoughts and concepts through time?[1] The way I understand it, those are some of the questions that are raised by what we may call "cognitive dynamics" (though as we will see, there may be more than one way to understand what cognitive dynamics really is, or at least of understanding what its business or scope really is).

The first time attention was drawn to what came to be called cognitive dynamics, it was in connection with indexicality,[2] through the work of David Kaplan.[3] Such as it has been presented by him, the problem of cognitive dynamics

---

[1] Throughout this work, I generally prefer to employ the term "thought", as opposed to the equally common "proposition". The term "thought" is an allusion to Gottlob Frege's ideas on language, having been used by him in the course of the development of his own semantic theory. In his writings, Frege deliberately chose to employ the German word *gedanke*, usually translated as "thought", even though Frege himself wanted to avoid any form of "subjectivism" or "psychologism" in his main philosophical project, *i.e.,* the foundation of arithmetic on logic. He held that thoughts were completely objective entities, being a fundamental part of what constitutes the meaning of complex linguistic expressions, such as sentences. That meaning should be accessible to anyone (that is, any competent speaker of the language). Though the present work keeps some distance from the Fregean anti-psychologist spirit, I hope to be able to give answers to the problems raised by the dynamics of *de se* thoughts taking as a starting point the general framework of a semantic theory inspired by Frege's ideas. I am not against Frege in thinking that it is necessary to clearly distinguish the level of content of what *he* called a thought (that is, the part of the meaning of a sentence which is objective and can be properly said to be true or false, being thus the object of investigation of the sciences), from other levels of content that are rather the object of inquiry of epistemology, psychology, a theory of pragmatics in language or a theory of action. Still, if the scope of our investigation is larger than that of Frege (who was particularly concerned with logic and mathematics), we have to take into account all those possible levels of meaning. In that sense, the word "thought" carries a certain ambiguity which, far from being a disadvantage, can rather be of interest to a research such as this one. Just as we can take a thought to be an objective entity, part of the meaning of complex expressions of language (at least partially in accordance with Frege's ideas), we can also employ the term to refer to *real* mental episodes belonging to a particular individual. What is it that we intuitively call a "thought", if not a certain kind of mental act that carries some sort of content?

[2] Indexicality is understood here in a very simple way, as the study of indexical expressions in language. Roughly speaking, indexicals are terms whose meaning is sensitive to contextual changes. Part of their meaning (the referent) varies according to the context in which they are used, because they can be used to refer to different things in different contexts. Common examples of indexicals are: adverbs of time and place, pronouns, verb tenses, etc. I say more on indexicals in the following paragraphs.

[3] Kaplan has coined the term "cognitive dynamics" in KAPLAN, D. 1989.

is that of explaining the preserving of a certain subject's epistemic attitude towards indexical thoughts in face of contextual changes (such as the passage of time) and given that in order to re-express the same indexical thoughts in different contexts, their syntactic form must be changed.[4] If we understand it that way, the problem which we are presented with is, at least to some extent, derived from a more general problem, *i.e.,* the explanation of the semantics of indexical terms and the thoughts that contain those terms. So, in a way, we would be facing an essentially linguistic problem, which concerns the expression and re-expression of indexical thoughts taking into account changes in the contexts in which they are expressed.

It can be said that Frege is responsible for discovering this particular facet of the issue, writing briefly about it in his famous paper "The Thought".[5] The following passage of Frege's paper gained the reputation of initiating the discussions on what came to be known as cognitive dynamics:

> If someone wants to say the same today as he expressed yesterday using the word "today", he must replace this word with "yesterday". Although the thought is the same its verbal expression must be different so that the sense, which would otherwise be affected by the differing times of utterance, is readjusted. The case is the same with words like "here" and "there".[6]

Frege goes on to sketch that which could be taken to be his view on indexicality. In short, Frege's opinion seems to be that indexicals such as "yesterday", "today", "there", "here", among others, only acquire a complete meaning in each specific context of use. Elements of the context (the time and place, gestures, etc.) supply the incomplete sense with what it lacks. Indexicals thus differ from proper names, for example, in possessing "incomplete senses", sensitive to contextual changes.[7]

---

[4] The most common example of an epistemic attitude would be the attitude of belief, and that is the example with which I shall be dealing most of the time, for simplicity. This kind of attitude is of course also commonly called "propositional attitude". A propositional attitude is basically any kind of attitude a subject can have towards a thought or a proposition, such as belief, hope, doubt, etc. The thought/proposition is that which is represented after the word "that" in sentences such as "I believe that Rio de Janeiro has an amazingly beautiful geography", and the epistemic attitude is obviously showed by that which comes before.

[5] I use the term "facet" because, as it will soon be clear for the reader, I think what we are actually facing is a broader phenomenon (or at least potentially broader), and if so we should not lose sight of its real scope.

[6] FREGE, G. 1918, p. 296.

[7] Of course, several basic notions of the semantic theory inspired by Frege are being presupposed here, like the notion of *sense* itself, as well as that of *thought*. A little more space will be given to it in the first chapter. Fregean semantics is ternary (if we count the linguistic sign as an element): we have 1. a sign, that is, a word or some other linguistic term; 2. a reference (the *Bedeutung*), or the entity denoted/presented by the sign; and 3. the sense (*Sinn*), which is the way that entity is denoted/presented. There is an idea present in the background of Frege's theory according to which

We can justifiably choose to restrict our investigation to this way of understanding the cognitive dynamics phenomenon, which mainly comes from David Kaplan's interpretation of Frege's text just quoted above. Upon reading Frege's own text, we can clearly see that what Kaplan later called "cognitive dynamics" is indeed presented in close connection to the phenomenon of indexicality. So, understanding cognitive dynamics under this light, our *main* task seems to be that of explaining which linguistic adjustments are necessary for a given subject to express retention or change of a propositional attitude she or he has in face of a change in the context.

Although it had never been Frege's goal to deal with this issue, in the passage quoted above he does sketch what would be an intuitive account of this linguistic phenomenon. Such an account is retaken by Kaplan in his own text and adapted to his own theory. According to Kaplan's interpretation, which remains as faithful to Frege's as it can be, in order for a subject to retain the same attitude in relation to an indexical thought in spite of the passage of time, it is necessary for him to re-express such a thought while adequately modifying the indexical terms (such as the adverb and the verbal tense, for example). In order to preserve the attitude so that the same thought can really be re-expressed, the substitution of the indexical must be made within one and the same "family" of indexicals. One example of such a family would be the set: [today, yesterday, the day before yesterday] when referring to the past, or the set: [today, tomorrow, the day after tomorrow] when expressing thoughts about the future. Thus, just like in Frege's example, in order to re-express a thought expressed the day before with a sentence such as "It is a beautiful day today", it is necessary to adjust that sentence to "It was a beautiful day *yesterday*".[8]

---

we only access the world (and refer to entities in it) through a kind of conceptual apparatus, which would in Frege's theory be represented by the notion of sense. What I mean by this is basically the theory of what we can call "descriptivism" (such as it is presented, for example, in RECANATI, F. 2012, p. 3 and what follows). "Thought" is the name Frege gives to the senses of complex linguistic expressions, *i.e.,* the senses of sentences. It is for that reason that we can say that a given thought is expressed by a given sentence. There is much that can be said about the whole semantic theory inspired by Frege's work, but this is not the place to do it. So, I shall not to dwell upon the question of it being acceptable or not in a Fregean theory to defend the existence of incomplete senses that could be completed by potentially non-conceptual entities. Perhaps that is indeed what Frege argues for in "The Thought": that indexical senses are incomplete and could be completed by entities which are not strictly speaking *senses* (*Sinne*), not being of the same nature as the entities that they are supposed to complete. Of course, one could elaborate on that idea of incomplete senses so as to make it compatible with the rest of Frege's theory. I address the issue of indexicality in the second chapter of this work, but overall I am favorable to Frege's ideas at this point.

[8] It is important to keep it in mind that, although Frege sketched an explanation for the phenomenon, the truth is that, in principle, there should not exist within Frege's system any problems related to

Such as by Kaplan interpreted it, the Fregean intuitive answer faces some problems, and I will explain and deal with them in detail in the second chapter of this work. For now, I just want to give the reader a taste of the ways in which the phenomenon of cognitive dynamics might relate to *de se* thoughts, which is, after all, the main topic of this work.[9] I want to apply cognitive dynamics to that kind of thought both in the way the phenomenon seems to have been understood by Kaplan himself and in the way it has been dealt with in the literature after his "Demonstratives". The

indexicals. Frege's main philosophical project was to ground arithmetic firmly on logic, which he believed was the most basic science of all, in the sense of potentially serving as the ultimate ground not only to mathematics, but to all other sciences. To be successful in this life-long project, Frege had to formulate a new logical language, for the tools of logic then available to him were not sophisticated enough for the accomplishment of his desired goal. That new logical system is developed in the *Begriffsschrift* as Frege's "conceptual notation". (FREGE, G. 1879.) That innovative notation should function more or less like a "perfect language", particularly suited for science and free from all the inaccuracies and general "imperfections" of ordinary language. To give just one example, the Begriffsschrift (which, besides being the title of Frege's work, is how this notation is called) would not contain indexical signs, or indeed any term that could admit different referents depending on the context in which it is used. There is no place for any kind of ambiguity in Frege's ideal language, so for every object there is one name, and every sense is eternal and unchangeable, always presenting a referent. Taking that into account, we can easily see why indexicals would be eradicated in that system of logic. And that is why strictly speaking Frege is not concerned with the semantic of those expressions. The idea that indexicals and in a way cognitive dynamics do not present a problem for Frege is to be found, for example, in RUFFINO, M. 2007. But I do not think a strict Fregean would really be completely exempt from the need of investigating cognitive dynamics. Supposing that we agree with Frege and take thoughts to be eternal and unchangeable, to explain cases of a subject's changing or retaining her attitude towards a given thought would be easy enough in what concerns the linguistic condition of being disposed to re-express (or not) the (very) same thought, which would be expressed the same way. If that was all there was to cognitive dynamics, there would not be a problem indeed. But as it will be clear later in this work, I think that even in the case of eternal thoughts, there are other aspects of cognitive mechanism of retention/change that still need an explanation, since, to begin with, the content of those thoughts is not given to us effortlessly and all at once. That is still the case, even if one believes that thoughts are unchangeable entities somewhere in a Platonic realm, just waiting to be grasped by us, humans. I believe that, no matter if they are eternal or not, thoughts are always subject to revision, refinement, enlargement – maybe not in themselves, if you are a strict Fregean, but at least cognitively speaking, that is, considering how individuals *apprehend* those thoughts. Of course, the original Fregean project does not comprise this kind of investigation, which would probably be too psychological for Frege's own taste, but if Frege were to develop a truly complete theory of thought, this kind of issue would have to be addressed, or so I believe. In short, even a strict Fregean whose interest lies in Frege's main life-long project would need an explanation of how an individual who apprehends a given thought/concept does just that, though he could choose to ignore the issue, leaving it for the cognitive dynamics theorist. I will of course come back to the discussion of the scope of cognitive dynamics in chapers 2 and 4.

[9] Throughout this work I employ "*de se* thoughts" to talk about that which we can also more simply call "thoughts about oneself", "first-person thoughts"/"thoughts in the first person", "egological thoughts" or "I-thoughts". In employing "*de se* thoughts" I am clearly hinting at the notion originally put forward by David Lewis, which is supposed to be in contrast with Quine's previously stablished notions of *de re* and *de dicto* propositions/thoughts. I am also alluding to the discussions on the special kind of information carried by a way of thinking about oneself that is *de se* as when a person thinks of herself as herself in contrast to thinking about herself in an impersonal way. The most recurrent example of that kind of thought has been put forward by Perry in PERRY, J. 1979. This point is discussed in what follows, in chapter 1. We may also use "*de se*" to talk about (essential) indexicality more generally, but my use of the term has a narrower connotation.

dynamics of *de se* thoughts is also a topic I will tackle in chapter 2, particularly in the case of *de se* thoughts whose expression involves the use of tokens of the first-person pronoun "I".

If one understands cognitive dynamics the way Kaplan did, it may seem strange to try to explain the dynamics of first-person thoughts. In order to see how strange this may seem, we only have to take under consideration Kaplan's main goal when writing "Demonstratives": he was trying to come up with an appropriate semantic theory for indexicals and demonstratives (which can be regarded as a special case of indexical expressions) and, based on that, develop a logic to deal with those expressions. Explaining how indexicals behave should naturally also involve explaining the fact that the same thought can be re-expressed in different contexts through the use of different linguistic terms. And that is of course related to how persons retain or change their epistemic relations to those thoughts. This is what Kaplan took as being the business of cognitive dynamics.

Again, according to Kaplan's view, the most fundamental problem of cognitive dynamics is that of accounting for the continuation of the same attitude connected to the appropriate re-expression in different contexts of the thought that expresses that attitude. In the case of indexical thoughts, that involves the correct substitution of the indexical term(s). But we know that in the case of first-person thoughts – thoughts a person has about herself – the first-person pronoun "I" never needs to be substituted by any other indexical in order for the same thought to be successfully re-expressed in another context. Normally someone trying to re-express her previous I-thought has to adapt the verb tense previously used, or maybe other indexicals used along with the first-person pronoun (such as adverbs of time). But the first-person pronoun itself does not seem to belong to a "family" of indexicals similar to the one "today" belongs to, at least not in the same sense. That is why Kaplan took the problem of cognitive dynamics in the case of I-thoughts to be the problem of accounting for the communication of one's I-thoughts to others. In that kind of situation, the pronoun "I" can indeed be said to belong to certain sets of indexicals, such as the set: [I, you], when one's interlocutor is supposed to understand and re-express one's I-thoughts by substituting the first-person pronoun "I" for the second-person singular pronoun "you".

So, strictly speaking, and strictly according to how Kaplan formulated the problem, cognitive dynamics does not seem to apply to first-person thoughts held

by the same subject through time. So, we have already departed from Kaplan if we say that the problem of cognitive dynamics should apply to *all kinds of indexical thoughts*, including I-thoughts, and not only to thoughts whose syntactical forms evolve depending on context. If *de se* thoughts are cases of indexical thoughts, an account of cognitive dynamics should apply to the dynamics of this specific kind of thought as well.

Be it as it may, it seems clear to me that the problem of cognitive dynamics such as it has been formulated by Kaplan does *not* apply to the case of *de se* thoughts in a straightforward way. I think that when dealing with cognitive dynamics, we can start from the same problems as Kaplan did, but they will have to be restated in different ways to apply for the case of I-thoughts. An explanation of the dynamics of first-person thoughts will have to take into account, for example, the fact that in order to re-express the same I-thought in different contexts, I do not have to substitute the indexical "I". This is of course only one of the characteristic features of *de se* thoughts which have to be taken under consideration when trying to account for their dynamics. I discuss other characteristics of *de se* thoughts in the first chapter, imeadiately following this introduction. My intention at that point is to investigate in more detail the very nature of that kind of thought, which, along with cognitive dynamics, is to be one of the main subjects of this work. My goal is to establish what distinguishes *de se* thoughts from other kinds of thoughts, such as *de dicto* and *de re*.

The existence of the *de se* has been famously defended by David Lewis, John Perry, as well as a few other authors (though each has a distinct line of argumentation).[10] Perry's idea can be formulated roughly thus: there must exist a level of content expressed by any particular indexical term that is irreducible and cannot be explained by any other means, such as another indexical indicator or a description. This is particularly evident in the case with the first-person pronoun – there supposedly exists a first-person perspective from which we can never escape. As it has been originally demonstrated by Perry, the attempt to substitute a pronoun such as "I" by some other expression changes the very content of the thought expressed. When an individual acts on the basis of some thought he has about

---

[10] It can already be seen in Frege, for example, such as in his famous "The Thought" from 1918. As for Perry, *cf.* PERRY, J. 1979, but also LEWIS, D. 1979. These two authors claim to be indebted to Hector Neri Castañeda. An important text by the latter on this topic is CASTAÑEDA, H. N. 1967.

himself in a first-person way, that thought carries a content (or even a kind of knowledge, we might say) that cannot be captured if we substitute the first-person pronoun for another expression (such as a description). The resulting attempt of re-expression of the I-thought could not be used in an explanation of the subject's action, for it does not play the same role in the explanation as the original indexical thought. In Perry's example, he is in a supermarket following a shopper around a counter, because that shopper has a torn sack of sugar and is making a mess, leaving a white trail of sugar on the floor. After a while trying to reach the shopper with the torn sack and not succeeding, Perry realizes *he himself* is the shopper who is making a mess, he then thinks a *de se* thought (while before he was thinking of himself without realizing it, through a description – "the man with the torn sack"), and that prompts him to bend down and fix the torn sack. Of course, this explanation of the issue is overly simplified, but for now it suffices to say that such is the sort of observation that is supposed to give us reason to think that indexicality and the *de se* are *essential*.

However, the reader who is already familiar with those issues will know that Herman Cappelen and Josh Dever argued against the interest in studying indexicality in the sense of an irreducible, essential phenomenon. They regarded with skepticism the relatively widespread belief among philosophers that indexicality and the *de se* are real and philosophically interesting phenomenona. They write:

> We argue that [...] there are no such things as essential indexicality, irreducibly *de se* attitudes or self-locating attitudes. Our goal is not to show that we need to rethink those phenomena – that they should be explained in ways different from how, e.g. Perry and Lewis explained them. Our goal is to show that the entire topic is an illusion – there's nothing there.[11]

More specifically, they argue that the arguments advanced by Perry and Lewis in favor of the existence of irreducibly *de se* thoughts (and essential indexicality) are not conclusive, though over time they have been taken for granted by many authors. But here I want to point out that even if Cappelen and Dever are right, if those arguments in favor of an irreducible indexicality are not good enough, that does not entail the non-existence of essential indexicality or irreducibly *de se* thoughts. In any case, they themselves admit that if we were to find good arguments to replace those advanced by Perry in favor of his irreducibility thesis, for example,

---

[11] CAPPELEN, H.; DEVER, J. 2013, p. 3.

then that would represent a real advancement in the debate in which they engage over the very the existence of indexicality.[12]

I believe that (essential) indexicality is a real phenomenon, though I think Cappelen and Dever are actually right in pointing out that the "classical" arguments in favor of it are not strong enough (assuming that they are really arguments in favor of it at all, which according to the two authors is debatable). I think it is important to address Cappelen and Dever's criticisms in a work such as this one, because I presuppose the object of my study, that is, *de se* thinking, to really exist and be something worth investigating. So, I dedicate the third chapter of this work to try and defend, in my own way, Perry's basic idea. There is an explanatory value in *de se* thoughts that has to be accounted for, I think there are indeed special ways of thinking about oneself that have to be explained. My main goal in chapter 3 is to present counter-arguments to Cappelen and Dever's criticisms. For the sake of brevity, I give special attention to the attacks they make on those features of *de se* thoughts which are particularly relevant to a study of their dynamics (such as they have been discussed in chapters 1 and 2). If there is one thing that is undeniably good in their attacks on the ideas most of us have on indexicality and the *de se*, it is this: they have prompted us to formulate arguments that are more solid.

Now, even if the reader is not convinced by the arguments I present in the third chapter against what may be called *de se* skepticism, in chapter 4 I argue that, independently of one's being able to come up with good arguments in favor of the essential indexical thesis or not, *de se* thoughts are not *only* instances of indexical thoughts. It is fair to say that from a semantic point of view, they are indeed paradigmatic cases of indexical thoughts. The first-person pronoun is a standard example of an indexical. But once we broaden the scope of our analysis and start taking into account epistemic and cognitive facts more generally, we realize that the specific way of thinking about oneself, via the first-person pronoun "I", does not always play a role in the identification of the referent, as it happens with other indexicals. If we follow the line of thought put forward by Stéphane Chauvier,[13]

---

[12] *Ibid.*, p. 41. In this particular case, this is said in connection to a claim attributed by Cappelen and Dever to Perry, which they call the "Impersonal Incompleteness Claim" and which reads "In some action explanations/rationalizations, indexicals occur ineliminably". *Ibid.*, p. 39. But again, one of the main criticisms made by the two philosophers in their own book is that the arguments on which the essential indexicality thesis is based are two weak.

[13] To cite just one example from which I draw inspiration to put forward my own arguments: CHAUVIER, S. 2001.

what he calls "egological thoughts" register the event of *being* the conscious instance of certain properties. Indeed, Chauvier defends that we should understand that kind of thought less as something that an individual *has*, or *grasps* than as something an individual *is*, for they are a constitutive part of the individual's *personhood*. I discuss some of Chauvier's ideas and suggest we go beyond indexicality both in what concerns *de se* thoughts and cognitive dynamics. What I mean by this is the following: there are certain peculiarities of *de se* thoughts which cannot be captured by any theory of indexicality.

As I show in chapter 2, it is meaningful and worthwhile to study the dynamics of *de se* thoughts, even if that means going a little beyond what Kaplan himself thought of as being the business of cognitive dynamics. Here I suggest we go further still from Kaplan's original ideas. We have seen that, such as it has been understood by Kaplan (under the influence of Frege's remark in "The Thought"), the problem of cognitive dynamics concerns first and foremost the (re-)expression of indexical thoughts. It is the business of the congnitve dynamics theorist to explain the retention and change of a subject's attitude towards this specific kind of thought: indexical thought. That is the original scope of the problem. But in order to fully explain the dynamics of first-person thoughts, I suggest we have to go beyond that original formulation of the problem, since first-person thoughts present certain features which are not explained by appealing to their indexical character. And if that is indeed the case, we are allowed to at least ask if cognitive dynamics itself, as a phenomenon that needs explanation, does not extend beyond indexical thoughts – to all kinds of thought.

In short, here is a list of the main ideas I put forward in this work:

1. *De se* thoughts are indexical thoughts, but they constitute a special kind of indexical thought, presenting special characteristics. Among those, in order for the same person to re-express the same de se thought in a different context, no change of the indexical "I" is required, as it is often the case with other indexicals.

2. The problem of cognitive dynamics was originally presented by Kaplan as a problem involving indexical thoughts, particularly the retention of attitudes towards indexical thoughts whose re-expression involves some sort of substitution of the indexical terms employed. In this sense, and as we shall see it in more detail in chapter 2, cognitive dynamics is the dynamics of the cognitive significance which

is (in Kaplan's account) always attached to the characters of indexicals, which in turn must change with each new context of use.

3. The problem should in principle be raised to the case of *de se* thoughts, since they are indexical thoughts. However, in a way the problem should not apply to *de se* thoughts, since they are always expressed via the same indexical: the first-person pronoun "I". A substitution of indexicals is not needed for the re-expression of this kind of thought – the same character ("I") is always used, and consequently the same cognitive significance is always present.

4. But we can show that cognitive dynamics actually has a slightly broader scope than it has been originally thought by Kaplan. We can show that the task of the cognitive dynamics theorist is not *only* explaning what happens in cases of attitude retention towards thoughts whose re-expression involves a change in the indexicals employed. We can show that there are other dynamic processes involved in attitude retention which must be accounted for. In short, the dynamics is not only linguistic: it is not only a dynamics of the characters used. In view of this, it makes sense to account for the dynamics of *de se* thoughts. This dynamics must be explained in view of the "peculiarities" of this kind of thought which, in spite of those peculiarities, continues to be indexical.

5. I believe *de se* and indexicality skepticism is wrong, and counter-arguments can be formulated against the skeptical ones put forward by Cappelen and Dever, even if no knock-out argument can be formulated *in favor of* essential indexicality and the *de se*.

6. If the reader is still not convinced by what I have to say against skepticism concerning indexicality and the *de se*, I argue that some of the "peculiar" characteristics of *de se* thoughts are not entirely captured by any available theory of indexicals. Given that, as I show in chapter 2, it is still possible to study the dynamics of those thoughts, in order for their dynamics to be fully explained, the problem has to be reformulated beyond its original scope (as presented in point 2).

7. If that is the case, we can at least ask if cognitive dynamics as a whole does not extend beyond its original scope – beyond the study of indexical thoughts – to a dynamist investigation of all kinds of thought.

Perhaps a truly complete theory of cognitive dynamics can only be developed by taking into account and analyzing more deeply all the particular dynamic processes of human thinking – not just "*de se* thinking", but other sorts of thinking

mechanisms as well. I believe all of them have that feature in common of being dynamic processes, but each presents different characteristics that render them unique, deserving an analysis in their own right. In the end, I hope I will be able to spell out at least the general lines of an account of the (dynamic) mechanisms involved in *de se* thinking – a form of thinking which is at once so mysterious and so fundamental in our lives.

1
The specialness of *de se* thoughts

1.1.
Introduction

In this chapter my main objective is to investigate the distinguishing features
of *de se* thoughts and attitudes. Intuitively, *de se* thought constitutes a special kind
of mental state: the kind of thought I have about myself in a self-conscious way is
arguably different from other kinds of thoughts I may have either about myself or
about other things in the world. In particular, *de se* thoughts seem to be irreducible
to what we call *de dicto* and *de re* thoughts. Many have put forward arguments in
favor of that difference, trying to show the specialness of *de se* thinking – I also
want to explore that in this chapter.

De se thoughts are standardly (though not always) expressed and ascribed
using the first-person pronoun "I". By looking at some cases of failure in the

---

[14] Translation by Charles Martin. New York: W. W. Norton & Company, 2004, verses 594-599.

substitution of that pronoun by coreferential expressions, we can reveal certain mechanisms of self-reference which seem to provide at least *prima facie* evidence for the specialness of *de se* thoughts. The substitution of expressions such as a proper name like "Pedro" or a description like "the person who is writing this thesis" by the first-person pronoun "I" in constructions of the type "ξ is worried", for example, fail to preserve meaning and may fail to preserve truth as well. This seems to indicate that "I" has a kind of meaning which is neither reducible nor fully explained by coreferential terms.

Another way of showing the same thing is by looking at the modal status of some sentences containing the first-person pronoun before and after we substitute that pronoun by a corefering singular term. "I am uttering nothing", for instance, is *contingently* false. By uttering that sentence I am obviously uttering something, but while that is true in this world, I am indeed uttering nothing in other possible worlds. On the other hand, "The person uttering this token is uttering nothing" is false in all possible worlds (in which the sentence is tokened), being self-contradictory.[15] That is also due to the ways in which the first-person pronoun and a description refer.

Substitutions of singular terms in oblique contexts (e.g. in *de se* belief reports) in which the *quasi-indicator* cannot be eliminated "without loss of content and change of truth-value"[16] also point to the specialness of *de se* attitudes. Consider, for instance, the contruction "The Editor-in-Chief of *Manuscrito* believes *he (himself)* is a lucky guy" used to report Manuscrito's Editor-in-Chief's own utterance ("I am a lucky guy", uttered by Marco Ruffino) and, presumably, his own *de se* belief.[17] Note that here it is the quasi-indicator that seems to be *essential* to the expression of the *de se* belief, not the first-person pronoun.

Section 1.2 is dedicated to establishing what differentiates *de se* thoughts from other kinds of thoughts, particularly *de re* and *de dicto* ones. In section 1.3 I discuss well-known examples of cases available in the philosophical literature that lend support to the claim that *de se* thoughts have a special status. As I go through those examples, I also go through a critical review of the main accounts of *de se* thoughts that we can find – the theories developed to make sense of the example cases. I discuss the ideas put forward by Frege, Recanati, Perry, Kaplan, Lewis and

---

[15] KAPITAN, T. 1999, p. 3
[16] *Ibid.*
[17] *Ibid.*, p. 4-5.

Evans, among some other authors. I conclude with a (provisional) list of at least some of the unique features of *de se* thoughts. I leave out all the controversies surrounding the notion of *de se* attitudes and indexicality, which are to be dealt with in chapter 3. The particular characteristics of *de se* thoughts I identify in this chapter will serve as a point of depart for attacking, in chapter 3, the skeptic arguments *against* the existence of irreducible indexicality and *de se* thoughts put forward by adherents of some forms of *de se* skepticism.

1.2.
De re, de dicto and *de se* thoughts

Traditionally, philosophers of language endorse a view on thoughts/propositions seeing them as being either some sort of conceptual apparatus through which we represent things in reality (and which may or may not be descriptive) or as being sorts of structured *ensembles* of the things we refer to (if we consider that we refer to them directly).[18] In both cases, the "thought" is an important part of the meaning of a complex linguistic expression, and it is a unit of information.[19] Even though these conceptions of what "thoughts" are may seem strange at first, the ideas underlying them are actually simple and rather intuitive for most of us.

When someone thinks of Mont Blanc in France, for example, given that such a person has never been to Charmonix (the French town from where we can climb le Mont Blanc) nor anywhere near the mountain, she will only have knowledge of the highest mountain in Europe through descriptions read somewhere or heard from other people and pictures seen in books and on Google, for example. She will only have this sort of "second hand" knowledge of the thing in question (le Mont Blanc). In such a case, we could say that one is thinking about the object *descriptively*.

But one can also think of something when one is in *direct contact* with it. Right here and now, while I am looking at this computer with which I am writing, for example, I can entertain such a thought. In this case I have an informative relation with the object in question on the grounds of which I can (potentially)

---

[18] This is of course a very rough picture, and such as it is presented, it probably sounds a little mysterious. But we will have a chance to explore the notion of direct reference in a little more detail in what follows.

[19] Though not just *any* unit of information, of course. I shall clarify this point in the next few paragraphs.

receive all kinds of information through my sense organs. The idea here is that the relation itself is responsible for the *singularity* of my thought (to employ a kind of terminology that is common coin). I can think about a certain *singular* object because of that relation I have to it. And more importantly, both because of that relation and because of the object itself, my very thought "acquires" singularity. The kind of knowledge I have in this sort of case can be regarded as being descriptive or not, though saying that it is descriptive can, at least in some cases, be problematic. This is the case of indexical thoughts, for example.

There exists still other ways of thinking about the objects we encounter in our lives, of course. Among them, I think one of the most relevant ones is through the *memory* of those objects. And here we are dealing with a phenomenon that is present in our day-to-day lives: we constantly use information on objects drawn from our memory of them. In this sort of case, it seems that a sort of mental representation of the object is at work. Maybe one could ask if there is not already a mental representation of the object when it is directly presented to my sense organs and I think about it. If one were to believe there was, that would probably involve the assumption (traditionally attributed to a kind of Fregeanism) that thought is always somehow "detached" of its object – in the sense that *Sinn* is that which intermedeates the thinker and whatever she thinks about. But although this may be the case sometimes, I think there must be situations in which thought is a direct consequence of our contact with the object of our thought, as in some cases of indexical thinking. In such a case there is no intermediate representation between my thought and its object. "Direct cases" like these must be assumed, I believe, if we do not want to be forever caught in a world of representation – in a Fregean world of *Sinn*.

Be it as it may, we think of different things all the time, and there seems to be a great variety of things which can become the objects of our thoughts, as well as a great variety of ways we can think about those objects (and then express our thoughts in sentences). We can (and we often choose to) say that a certain thought is "indexical", "singular", "*de re*", "descriptive", "*de dicto*" or "*de se*" (among other possible labels) grounding our choice on various factors. Among the most recurrent ones in the philosophical literature are: 1. The nature of the object thought of and 2. The nature of the relation the subject has with that object, or both. But is it really right to say that the nature of our thoughts differ according to the nature of those

different objects and/or the way we think about them? Thoughts are not only conceptual apparatus or *ensembles* of things we refer to, they also *convey information*, as it has been pointed out before.[20] It is one of the characteristics of what we, philosophers of language, call "thoughts" that they are epistemically rich, they are informative – a thought has, as we say, an informative value. So, if information is part of what constitutes a thought, the kind of information that is incorporated in a given thought will have an influence on the kind of thought that we have as a result. If the quality of the information we receive from objects depends on the object's being present or not, on the nature of the object and on other factors, then we start seeing more clearly why philosophers put thoughts into categories depending on their objects and the thinker-object relation.

In what follows I want to explore some of the characteristics that are supposed to account for a few examples among the great variety of types of thought we may have. I want discuss some of the arguments in favor of the idea that there are possibly various kinds of thought which can be entertained by a subject *at* a time and *over* time. I also want to investigate some of the characteristics of what we call *de se* thoughts – those which make them similar but at the same time irreducible to other kinds of thought, particularly those we identify as *de dicto* and *de re*.

Some of those arguments can be found in the literature on propositional attitudes and attitude reports. It was W. V. Quine who kindled the contemporary discussions on *de re* and *de dicto* thoughts after his 1956 paper, "Quantifiers and propositional attitudes". He wanted to offer us a solution to a kind of ambiguity that we find in sentences expressing propositional attitudes, such as expressions of belief.[21] Quine shows the ambiguity with the help of a few examples, such as:

"Ralph believes that someone is a spy" and

"I am looking for a dog that talks".

Quine argues that the ambiguity arises because those attitude verbs have two *senses*, which he calls the *relational sense* and the *notional sense*. The difference between these two senses is explicitly made in the way the sentences containing

---

[20] As I have also pointed out in the beginning of this section, that is not just *any kind* of information. It is not the place to enter into the intricate details of this discussion, but one could, in a certain sense at least, say that DNA, for example, also conveys information. The difference between the kind of information conveyed by our thoughts and the one we may say that we find in our DNA is probably in what we call *intentionality*. I dedicate part of section 4.3 in chapter 4 to discuss that notion.
[21] Just check QUINE, W. V. 1956.

those verbs are expressed in certain Romance languages, such as Spanish and Portuguese. In these languages, we can sometimes express those different senses through different verb modes. Both in Spanish and in Portuguese, the sentence "I am looking for a dog that talks" can be expressed in two different ways.

So, in the two following instances: "Procuro un perro que habla" and "Procuro um cachorro que fala", the propositional attitude verb has a relational sense, and the sentences are interpreted as meaning that I am looking for *a particular dog* that can talk. In more formal writing, it can be rendered thus:

($\exists x$) (x is a dog $\land$ x talks $\land$ I seek x)

On the other hand, when the subjunctive mode is used, as in: "Procuro un perro que hable" and "Procuro um cachorro que fale", the verb has a notional sense, and the sentences are interpreted as meaning that I am just looking for any dog that is capable of talking. In Quine's formal rendering:

I strive that ($\exists x$) (x is a dog $\land$ x talks $\land$ I find x)

Quine interprets the act of looking for something into a propositional attitude as *striving that* one finds something.

In the other of Quine's examples that I have presented above, the ambiguity is always present, be it in English, Spanish or Portuguese. In that case, the sentence "Ralph believes that someone is a spy" can be interpreted as meaning either that: 1. *Someone is such that Ralph believes that* she *is a spy*, that is, ($\exists x$) (Ralph believes that x is a spy) – this interpretation showing the relational sense of the verb; or that 2. *Ralph believes that there are spies*, that is, Ralph believes that ($\exists x$) (x is a spy). In other words, Ralph believes that there is at least one spy in the world. This second interpretation shows the notional sense of the verb. Note that Quine does not interpret the ambiguity as being purely syntactical or grammatical. And that in spite of the fact that he takes the relational interpretation of a sentence containing a propositional attitude verb to be problematic, because it involves quantifying into a referentially opaque context. As Quine himself puts it, the referential interpretation involves "quantifying into a propositional-attitude idiom from outside", and that is a "dubious business".[22] Contexts of attitude report are common examples of *referentially opaque contexts*.

---

[22] *Ibid.*, p. 178-179.

For Quine, in such opaque contexts, any singular term appearing after the propositional attitude verb simply does not refer as it does in transparent contexts. According to him, the "that" that appears in propositional attitude reports must be viewed as sealing the following sentences off in such a way that the terms we find in those sentences no longer have any relation to an exterior realiy (at least not the way they would normally have in other circunstances).[23] So, as Quine's argument goes, it is *meaningless* to *quantify into* opaque contexts, because quantifying only makes sense if the terms concerned actually refer. The *x* immeadiatly following the quantifier must be of the same kind as in whatever context the quantifier is being applied (referential).

It is not my aim to discuss here the minutiae of Quine's ideas, nor problems which may arise from them. I think I have written enough to get to my objective, which is just pointing out the fact that we can call these transparent interpretations of belief reports "*de re*", because they involve a sort of direct reference to an object. This stems from the theoretical background of Quine's theory, because for him quantification is always on particular object-positions. Quantification is always *de re*, always transparent.

Semantically, in *de re* or transparent contexts, an exchange between co-referential terms can be made without any loss in the truth-value of the sentence or anything else – what matters is the very object referred. The same does not happen in *de dicto* contexts. A sentence expresses a *de re* proposition just in case there can be a substitution of co-referential terms with preservation of the the proposition's truth value. If such a substitution cannot be made *salva veritate*, the sentence expresses a *de dicto* proposition.

Looking at syntax only, if we follow the first interpretation of the ambiguous sentence "Ralph believes someone is a spy", (that is, the interpretation according to which the sentence means something like "Someone is such that Ralph believes that he is a spy"), the *quantifier's scope* is *large*, so it quantifies over a variable (x) that occurs inside the scope of the verb "believes": ($\exists x$) (Ralph believes that x is a spy). Following the second interpretation (according to which the sentence means something like "Ralph believes that there are spies"), the existential quantifier has a narrow scope, quantifying over "believes": Ralph believes that ($\exists x$) (x is a spy).

---

[23] *Ibid.*, p. 179.

We can say that a sentence expressing a propositional attitude is syntactically *de re* if there is in it a pronoun or a free variable within the scope of a verb typically used to express propositional attitudes (such as "to believe") – such a variable must be anaphorically connected to a singular term or a quantifier outside the scope of the verb. If that is not the case, if there are no such variables and no anaphora, the sentence is said to be syntactically *de dicto*.

Of course, from Quine's ideas on quantification and opaque/transparent contexts, we should not draw the conclusion that belief can never be characterized as being *de re* – as being about something directly. In reality, the distinction is neither so intuitive nor so clear-cut as one might at first suppose. For example, as Tyler Burge points out,[24] even though the distinction is commonly characterized appealing to cases of substitution of coreferential terms, there are cases in which the substitutivity of the terms cannot be made without loss, but we would nevertheless be inclined to call the sentence in question *de re*. As an example, we can suppose that someone is at a bar, sees a man in the corner and then formulates a thought about him (as the man in the corner). The thinker in this case may not be inclined to think of the man he sees in the corner as "the firstborn in Kiev in 1942", for example, even if that description would fit him perfectly. He might be inclined to think of him solely on the basis of the description "the man in the corner", so substitutivity is unwarranted. However, intention may exist that the ascription of attitude towards the man in the corner still be *de re*.[25]

The *de re*/*de dicto* distinction is very often presented in terms of opacity and transparency in certain contexts, and thus in terms of the possibility of substitution of coreferential terms in those contexts. But we must keep it in mind that 1. Such presentation is not free of controversy (as shown in the previous paragraph), and 2. There are actually many ways in which the difference between *de re* and *de dicto* thoughts may be presented.

In any case, some of the ideas put foward by Quine, Burge and other authors lead us to think that a distinction must be drawn between two sorts of cases: the expression of thoughts in so-called "transparent" contexts, also called *oratio recta*; and the attribution of thoughts in "opaque" or "oblique" contexts, also called *oratio obliqua*. Being transparent or opaque, *de re* or *de dicto* are clearly properties of

---

[24] In BURGE, T. 1977.
[25] *Ibid.*, p. 341.

linguisitic contructions. These seem to be categories of different ways we can *talk* about things in the world – different ways in which thoughts are expressed. That is, I take it, pretty obvious. But perhaps they can also be taken to be different ways of *thinking* about things in the world, as I suggested in the beginning of this section. These different ways of thinking about objects, when expressed in language (in constructions that we would call *de re* or *de dicto*) would end up influencing the truth-value of the linguistic construction in question. So, the question is: what reason do we have to believe that natural language (or some regimented one) provides us with the means to detect different *kinds* of thoughts, as opposed to different *ways of expressing or reporting* the same kind of thought?

Bach, for instance, is skeptical to the view that the *de dicto*/*de re* distinction as applied to the expression of *thoughts* would mirror a distinction between *kinds of thought*. He thinks that the *de re*/*de dicto* distinction such as it has been originally employed should be restricted to modalities, and not be extended to attitudes reports and attitudes themselves.[26] Bach defends that, on the one hand, "acquaintance (familiarity) with an individual is not necessary for thinking of it [in a singular way]",[27] and on the other, expressing a thought which has the syntactical form of a *de re* thought does not guarantee a truly *de re* thought. For Bach, thinking of an object in a *de re* manner is still to think a mediated thought – *de re* modes of presentation, as he puts it, connects us to objects in a causal-historical way. The relation can be more or less remote, but the object is always determined relationally (as opposed to satisfactionally).

Quine, Burge and other authors are *not* skeptical about the possibility of applying the *de re*/*de dicto* distinction to kinds of thoughts. I think they do provide us with good reasons to go beyond the view that the *de dicto-de re* and *de re-de se* distinctions are purely syntactical or grammatical. Even for Quine, who (as we have seen) is skeptical to the consistency of the semantic interpretation of quantification *into* belief contexts, it still makes sense to speak of a *relational* as opposed to a *notional sense* of "believes".

Recanati, in his turn, provides some interesting epistemological grounds for the *de re-de se* distinction. The ways we think about things in the world seem to depend on our epistemic relations to those things (on how we get information from

---

[26] BACH, K. 2010, p. 44.

[27] *Ibid*., p. 57.

them). So, the disctinction between *de re* and *de dicto* thoughts should also concern our theory of knowledge. Again, the distinction is clear enough in the level of language, but we can also present it as an epistemic distinction between ways of thinking about things and gathering knowledge from them. We can think about things *descriptively*, and be related not to whatever we are thinking in a direct way, but via a *dictum*. In this case, the kind of thought we entertain could be labeled *de dicto*. And we can think about things as we are (somehow) directly related to them – and this kind of thought could be characterized as *de re*.

Recanati distinguishes *de dicto* and *de re* thoughts by appealing to the difference between descriptive and nondescriptive thoughts respectively. He argues that the distinction should not be spelled out by appealing to any notion similar to direct reference or direct thinking, because even in cases in which we would be more inclined to say that there is direct reference, as in demonstrative thinking, there is actually *a way in which we think of the object* – in the example just given, there is a *demonstrative* mode of presentation of the object.[28] As I have been trying to say, this way of thinking about the object, or its mode of presentation, supposes a certain kind of informational link with it – which is direct in the case of *de re* thoughts, but not so in the case of *de dicto* thoughts.

As Recanati argues, the mode of presentation will be responsible for the truth-conditions of the thought one is expressing. Remember that, just a few paragraphs above, we have seen how the *de re*/*de dicto* distinction in the case of constructions containing modal operators (such as "it is necessary that" or "it is possible that") also affect the truth-value of the sentence and the thought expressed. But in that case other things were in play (such as the scope of the operator or a discussion on the status of essential properties and necessity, for example). Here, on the other hand, To give an example similar to his, if I see a woman, say, Juliette, and think "*she* is French" – thus, an indexical (demonstrative) thought –, my thought will be true if, and only if, *Juliette* is French. As Recanati puts it, the truth-condition of my thought is singular, because it has to do with that French girl directly. If, on the other hand, I happen to think of Juliette descriptively, say, through the description "the wealthiest girl in the world", thus thinking "The wealthiest girl in the world is French", my thought will not involve Juliette directly. It can be true or false

---

28 RECANATI, F. 2009, p. 249. Incidentally, this does not seem very far from Bach's view in his paper quoted above.

depending on the following conditions: if some girl, be it any girl, is the wealthiest girl in the world, and if that girl is also French, than my thought will be true. To explain this using a slightly more formal terminology, my thought will be true if, and only if, there is a girl x such that for every girl y distinct from x, x is wealthier than y and x is French.[29] So, if we follow Recanati, what distinguishes *de re* from *de dicto* constructions is really the role played by the mode of presentation in establishing the the thought's truth-conditions – and modes of presentation are always present, be them "direct", as in demonstrative thinking, or "indirect", as in descriptive thoughts.

We could sum up the kind of view defended by Recanati by saying that *de dicto* thoughts are descriptive, while *de re* thoughts are nondescriptive. Frege presented senses as modes of presentation of the reference. Following this idea, when one has a *de dicto* thought about a given object, such object is presented via a descriptive mode of presentation – a *dictum*, which may include some property attributed to the object (the wealthiest person in the world). On the other hand, when one has a *de re* thought about something, this thing, whatever it is, is presented in a nondescriptive way – the access to the *res* is not given via its properties. It is worth emphasizing that this is no longer Frege's original view, for he would not have accepted the idea that we can have a direct access to objects we think about. But it is a neo-Fregean thesis defended by some authors, including Recanati. In the kind of theory put forward by Recanati, what determines the reference in the case of a *de re* thought is something external to the thought itself: it is the very contextual relation in which the *res* stands to the representation in the thinker's mind – such a relation is not represented in the thought, even though it constitutes its ground. Here Recanati prepares the terrain to his *mental files theory* (which we shall discuss in detail later on), by arguing that "*de re* thoughts are based upon relations in virtue of which the subject can gain information from the object".[30] I think that for now we can just continue our investigation on *de re* and *de dicto* thoughts (until we get to *de se* variety) without worrying about the mental files metaphor. For Recanati, mental files play the role of modes of presentation (or Fregean thoughts, in a Fregean theoretical framework) and as the name indicates, they are units for storage

---

[29] *Ibid*., p. 249-250.
[30] *Ibid*., p. 251.

of information about the referent. I will just go on talking about thoughts and modes of presentation without adding one more metaphor.

If we try to think of examples of thoughts in which objects are not presented through any specific property, but through concepts of the objects in question which just stand for the objects themselves, the easiest examples would be of indexicals and demonstratives. These do not need to be the only examples, though. One can arguably have a *de re* thought about a certain man, calling him by a certain name, and still not think of him via any particular property. Still, token-reflexive words, such as the first-person pronoun "I", are also *de re* in the sense of having their meaning determined by certain contextual relations that happen in each particular context of utterance. Again, their token-reflexive modes of presentation make indexicals and demonstratives into examples of *de re* expressions. And following Recanati's line of reasoning, other expressions such as proper names, when employed in a *de re* fashion, exploit contextual relations, and are in this aspect similar to indexicals.[31] Even though I may have a thought about a certain individual referring to him by his name and ascribing to him a certain property (of being the wealthiest girl in the world, for instance), the idea is that the name itself is not represented in my thought through any kind of description, at least not necessarily: the name is used "in *de re* mode" when it does not present the referent through any of the referent's features. Such features or properties the referent may present, they play no role in determining what the representation of the referent is about, and they do not fix the referent (in this case, one may even ascribe false properties to the referent – it will make no difference).[32] In the case of *de re* thinking, what determines the reference is not any property it may possess, but the relation the thinker's mental representation has with the object in question.

If we understand *de re* thoughts in those terms, we can envisage the possibility of taking *de se* thoughts to be a special case of *de re* thoughts, at least in some specific cases. That is a possibility put forward by Recanati: "A *de se* thought is a

---

[31] I will not go into the details of how exactly thoughts in whose expression we find terms other than indexicals and demonstratives (such as a proper name) can be *de re* in the sense of containing a conceptual representation of the object in question which is not determined via any property of the object. The *de re* feature is rather grounded on some relation which is not itself represented in the thought. In the case of historical names such as Cicero (the Roman orator), Recanati's suggestion is that a causal chain of communication allows us to entertain *de re* thoughts about these historical characters.

[32] RECANATI, F. 2009, p. 253.

de re thought about oneself, that involves a particular mode of presentation, namely a first person mode of presentation".[33] There is a variety of possible modes of presentation of the first person, but there are arguably some which are specific of the first person – which only the first person has. Following Recanati, these are grounded on particular kinds of perception, such as proprioception.[34] Proprioception and other non-traditional senses which are particular to the first person are examples of the special relation I can have with myself, and which give rise to *de se* thoughts. There is arguably special modes of presentation of the first person which are grounded on those special sensory relations.

In the next section I introduce the main examples we can find in the philosophical literature in favor of the existance of *de se* thoughts. As we will see, these are usually presented in a way that *de se* thoughts are seen in opposition to *de re* thoughts. But the basic idea is this: there exist certain situations in which it is possible for a subject to attribute certain properties to *x* while ignoring the fact that he is himself *x*. When it happens that the subject in question is indeed *x*, what we observe is an instance of a "Frege case" in which the same referent in presented under two distinct modes of presentation. One happens to be demonstrative and *de re* (when the subject thinks about some individual *x*), the other can be *de se* if the thinker realizes the individual *x* is actually himself. After such realization, it may be that the thinker switches from a demonstrative relation to the object (which happens to be himself) to a properly first-person relation (through proprioception,

---

[33] *Ibid.*, p. 255.

[34] Roughly speaking, proprioception, sometimes also called the kinesthetic sense, is the capacity we have, without the help of other traditional senses such as vision, of recognizing and locating our own bodies in space, their position and orientation, as well as the strength necessary to move our muscles and the position of our body parts relative to each other. The sense organs directly responsible for proprioception are sensory receptors usually referred to as muscle spindles, which are found at the belly of muscles, as well as tendons and the fibrous capsules in joints. Incidentally, the very existence of these specialized organs can be regarded as one of the reasons for considering proprioception a *sense* (like vision and other traditional senses). Information is taken from those organs and integrated in the brain with information from what we call the vestibular system (a sensory system responsible for balance and spatial orientation) to form an overall sense of body position, movement, and acceleration. Later I come again to this and other non-traditional senses specific of the first-person – we can take other examples, such as thermoception (the sense of temperature), nociception (the sense of pain) and other internal senses generally classified under the tag of "interoception". Cases of impaired interoception (commonly associated with Autism Spectrum Disorder) may be indicative that the access we have to our own sensations and feelings may be imperfect; there seems to exist mechanisms that must be working properly in order for that access to be successful, and for us to be able to have certain kinds of *de se* thought. Looking at certain symptoms of Schizophrenia may also give us reason to believe that the supposedly necessary ability to keep track of objects through time in order to hold on to thoughts about them also apply to the first person. But I shall say more on all of this later on.

for example). If this happens, to use Recanati's terminology, one goes from an "accidental *de se* thought" – which is just a thought about oneself in the form of a *de re* thought –, to a genuine "explicit *de se* thought". In the case of an accidental *de se* thought, the thinker is *not* aware of being the object of her own thought. There is no self-counsciousness involved. In the case of an explicit *de se* thought, on the other hand, she is already aware that the thought she has is about herself. It is interesting to remark that in both cases, the thoughts have the same truth-conditions, since they are about the same individual. Still, as in Frege cases, we can say that the thoughts are different because we have different modes of presentation.

Now, it may be that some person happens to be thinking about herself and is aware of this fact (or so becomes), but for all that awareness, she still does not exploit a sensory mechanism typical of the first person, such as proprioception. If that is the case, her *de se* thought can indeed be considered a special case of a *de re* thought, if by that we mean that her thought has a conceptual representation of herself which is not based on any particular property (which could be spelled out descriptively). Having that in mind, yet another distinction proposed by Recanati may come in handy: the one between *explicit* and *implicit de se* thoughts.

Explicit *de se* thoughts are those which I have just described. If I see my own reflection in a mirror and take the image to be someone else's, I will entertain a thought about myself without realizing that I am actually the object of my thought, so I will arguably have a *de re* (maybe demonstrative) thought about myself. If I realize I am actually seeing myself in the mirror, and become aware that I am the object of my thought, I will have a properly *de se* thought about myself, since I know to be thinking about myself, but such a *de se* thought will still have a demonstrative quality to it, so to speak, being probably based on my sense of vision. This is the kind of thought Recanati calls explicit *de se*, because it involves an explicit self-identification. Now, in the case of implicit *de se* thoughts, "the subject serves as 'circunstance of evaluation' for the judgement, rather than being a constituent of content."

Following Recanati's line of reasoning,[35] in the case of implicit *de se* thoughts, no "complete" thought (or proposition) in the traditional sense is

---

[35] Recanati gives the credit of the idea he expresses here about implicit *de se* thoughts to other authors, such as CHISHOLM, R. 1981. *The First Person*. Minneapolis: University of Minnesota Press. But also LEWIS, D. 1979;

expressed, that is, what is expressed by that kind of thought is not the ascription of a property to an object, but rather the property alone. The content of the thought is just the property. In this kind of *de se* thought, he argues, there is no identification of the object to which some property is attributed, because the object is not even part of the thought. That is why this specific kind of first-person thought can be said to be *always* immude to error through misidentification, a characteristic we shall explore in a little more detail in the next section.

According to this idea, an implicit *de se* thought, which can be a very basic kind of thought – based solely on some form of interoception, is always immune to an error in identifying the referent, simply because there is no referent to be identified. There exists only the state perceived from the perspective of the first person. The thinker herself is not represented, but still, her being involved in the thought she has is guaranteed by "the mode of the grounding experience", as Recanati puts it. The content of an implicit *de se* thought "corresponds to that of a predicate, and the subject of which it is predicated remains implicit".[36] On the other hand, explicit *de se* thoughts may be both immune and vulnerable to error through misidentification. They are vulnerable to that kind of error because of the characteristics they share with plain *de re* thoughts. However, they too can be immune if they are grounded not *only* on what we may regard as being a demonstrative mode of presentation, but *also* on some form of proprioceptive or interoceptive experience. This is, I think, Recanati's basic idea. I think these distinctions he makes are worth taking into account, and parallels can be obviously drawn with what Evans has to say on the matter, as we shall see.

We have now began to see that 1. Thought comes in different forms, and these depend mainly on the way the object of the thought is presented to the thinker (the nature of the object itself and the relation between it and the thinker – more specifically the thinker's mental representation of it). And 2. Differences can be drawn to distinguish *de se* or first-person thoughts from other kinds of thoughts, such as the ones we usually call *de re* and *de dicto*. Arguably, there exists still other kinds of thought, but my objective here is to concentrate on thoughts held from the first-person perspective, *de se* thoughts, and it is common to constrast them to the *de re* variety (and by extension, to the *de dicto* too). It is possible to think about

---

[36] Which does not mean that she does not exist, of course, or even that there exists self-identification even in the case of so-called implicit *de se* thoughts. RECANATI, F. 2009, p. 259.

oneself in a *de re* or even in a *de dicto* way (which again makes it clear the great variety of possible kinds of first-person thoughts), but there seems to exist special ways in which I can be presented to myself that are unique to the first person. The great variety of possible ways of thinking about oneself is something we will explore again in the fourth chapter. Now I want to look more closely at the evidence in favor of the specialness of *de se* thoughts.

1.3.
De se thoughts in the literature

1.3.1.
Introduction

In this section I present and discuss some examples or cases from the literature meant to establish the existence of irreducibly *de se* thoughts. It is important, however, to keep it in mind that thoughts about oneself come in more than one variety — as already pointed out. My purpose in this main section is to review the cases and go through the theories behind them, with a view to pinpointing the distinguishing features of *de se* thoughts.

1.3.2.
Dr. Lauben

1.3.2.1.
Frege

Frege is partially responsible for initiating much of the debate in contemporary analytic philosophy around the first-person. Presumably the most famous passage from Frege's writings on the first person is the one below, taken from "The Thought". As a result of running a thought experiment involving Dr. Lauben, Frege makes two significant claims. First, he claims that each of us is presented to him-/herself in a "particular and primitive way". Second, he claims that I-thoughts are incommunicable to others – that is to say, I cannot communicate to other people the particular and primitive way in which I am presented to myself.

In fact, Frege seems to advocate in this essay a kind of "subjectivism" concerning both proper names and certain pronouns. I shall not dwell upon Frege's account of proper names here, but I think it is important to take a closer look at what

he has to say on the sense of "I" when contrasted with the sense of the name of the utterer of "I" (or some other term serving to identify such a person). Frege argues that different persons may associate different senses with a certain proper name, given that they have been presented to the bearer of the name in different ways.[37] Frege's remarks on this particular matter paves the way for what he has to say about the first-person:

> Consider the following case. Dr. Gustav Lauben says, "I have been wounded". Leo Peter hears these words and remarks some days later, "Dr. Gustav Lauben has been wounded". Does this sentence express the same thought as the one Dr. Lauben uttered himself? Suppose that Rudolph Lingens were present when Dr. Lauben spoke and now hears what is related by Leo Peter. If the same thought is uttered by Dr. Lauben and Leo Peter then Rudolph Lingens, who is fully master of the language and remembers what Dr. Lauben has said in his presence, must now know at once from Leo Peter's report that the same thing is under discussion. But knowledge of a language is a separate thing when it is a matter of proper names. It may well be the case that only a few people associate a particular thought with the sentence "Dr. Lauben has been wounded". In this case one needs for complete understanding a knowledge of the expression "Dr. Lauben". Now if both Leo Peter and Rudolph Lingens understand by "Dr. Lauben" the doctor who lives as the only doctor in a house known to both of them, then they both understand the sentence "Dr. Lauben has been wounded" in the same way, they associate the same thought with it. But it is also possible that Rudolph Lingens does not know Dr. Lauben personally and does not know that he is the person who recently said, "I have been wounded". In this case Rudolph Lingens cannot know that the same thing is in question. I say, therefore, in this case: the thought which Leo Peter expresses is not the same as that which Dr. Lauben uttered.[38]

Frege's conclusion is that the proper name "Dr. Lauben" has a sense which differs from the pronoun "I" when pronounced by Dr. Lauben himself, although both refer to the same individual. He presents us here with a case in which we can apply a famous criterion of difference for thoughts often ascribed to him. We shall analyze the criterion in more detail when we get to the discussion of the problems raised by cognitive dynamics in the next chapter. But in a nutshell, according to this particular criterion — Frege had many different criteria for establishing the identity of thoughts, two sentences express different thoughts if and only if a rational subject can hold divergent attitudes towards them.

In the passage just quoted, Frege tells us that it may be the case that Lingens does not know Lauben personally, although he was present when Lauben uttered: "I have been wounded". Because Lingens does not know who Lauben really is, it may be that he is unaware that the person he has heard uttering: "I have been

---

[37] PERRY, J. 1977, p. 488.
[38] FREGE, G. 1918, p. 297.

wounded" is indeed Dr. Lauben. So, he might perfectly well take the thought expressed by "I have been wounded", uttered by Lauben, to be true while taking the thought later expressed in his presence by the words: "Dr. Lauben has been wounded" to be false. "I have been wounded", uttered by Dr. Lauben, and "Dr. Lauben has been wounded", uttered by someone else, have different informative values. The thought expressed through the use of the first-person pronoun and the one expressed through the use of the proper name are clearly different, as borne out by the criterion of difference for thoughts just mentioned. But the deep reason for the difference in sense becomes conspicuous in a passage right after the one just quoted where Frege entertains the idea that each of us has a priviledged access to his/her own internal states – a particular and primitive way of presenting oneself to oneself. He seems to believe that the sense that each of us associates with the first-person pronoun is a unique and special mode of presentation that that he equates with this primitive way of presenting oneself to oneself:

> 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.[39]

Frege adds that one cannot communicate one's own I-thoughts the *same way* one grasps them. So, if Lauben is to communicate his first-person thought, "he must use the word 'I' in a sense which can be grasped by others, perhaps in the sense of 'he who is speaking to you at this moment'".[40]

The thesis Frege defends at this point is problematic when seen in the light of his philosophy of language because, to begin with, senses and thoughts in Frege's system are supposed to be objective and communicable to anyone. They are *not* subjective or private as ideas (*Vorstellungen*) and thinkings are, for instance. They are by nature public and accessible to everyone. What Frege says about the sense of "I" stands in an obvious tension with his view about the objectivity of senses. Arthur Sullivan labels Frege's view "individualism", meaning that Frege believes the following: competent speakers each have an individual grasp on the conditions for the correct application of the terms of their language.[41] This view is compatible

---

[39] *Ibid*., p. 298.
[40] *Ibid*.
[41] SULLIVAN, A. 2003, p. 204-205.

with the contention that those conditions, although objective, are grasped individually by speakers. For each individual person, they are "inside his/her mind".

It could also be argued that although my "I-thought" is objective like all other thoughts and could be held by someone else, it turns out that only I is in a position to have access to it. There are other examples of objective thoughts that we humans are not in a position to grasp: think of a thought involving a huge number or further mathematical thoughts that human minds are not in a position to hold. For all that, they do not cease to be objective.

Maybe we can accept an argument along these lines: one that justifies the existence of possible restrictions on the grasping of certain objective senses. The view according to which each of us has a particular and primitive way of presenting oneself to oneself is, nevertheless, perfectly challengeable, as Perry has it. "There are, I hope, ways in which I am presented to myself that I am presented to no one else, and aspects of me that I am aware of, that one one else is aware of. But this is not sufficient for Frege's purposes",[42] as he puts it, meaning that the fact that certain aspects of me of which only I can be aware of is not enough to advocate the thesis that there is a special and primitive mode of presentation of the first-person. Actually, Perry holds that these private ways of accessing oneself do not suffice to constitute a sense which would have a form similar to "*the M*", for example (i.e. the form of some definite description), which would uniquely determine *myself* as the referent of "I". Besides, for Perry, there is no reason to believe in the existence of any primitive aspect of *me* that only *I* has access to: the idea cannot really be proved.

1.3.2.2.
Recanati's take on Frege

A better way to understand Frege in the passages just quoted or, at least, to make sense of his idea of a "particular and primitive way" in which I am (supposedly) presented to myself is suggested by Recanati in the context of his theory of mental files. Even if one is not confortable with Recanati's talk of "mental files", I think he touches upon a crucial point when it comes to accounting for the specialness of first-person thoughts. As seen earlier, a first step towards understanding the special status of *de se* thoughts is to acknowledge the specialness

---

[42] PERRY, J. 1977, p. 490.

of the mechanisms of self-knowledge – some of which being unique to this kind of knowledge.

Recanati argues that the main ideas behind the mental file metaphor are the following: 1. During his cognitive life, a subject encounters a variety of objects and enters in various kinds of relation to them 2. Some of these relations are *acquaintance* relations, which means that they involve some kind mechanism that allows the subject to gain knowledge from them. They are, as Recanati puts it, "epistemically rewarding" relations 3. The role played by a mental file based on an acquaintance relation is to store information acquired in virtue of those relations 4. Mental files based on acquaintance relations are temporary: they subsist only as long as the relations themselves obtain.

So, here is how we can illustrate Recanati's mental files metaphor to explain how we get information from things and think about them. I restrict myself to the case of properly indexical or demonstrative thinking, which in any case is the best example to understand Recanati's ideas. When a subject perceives an object through his senses (when he sees or hears an object, for instance), he is capable of thinking about that object demonstratively. This particular way of thinking can only exist while the subject is perceiving the object about which he is thinking, because, as Recanati explains, demonstrative thinking involves the activation of a kind of mental file that is dependent on the perceptual relation: "When the relation is broken, the temporary file based on it disappears. (The information in the file is not lost, of course, but transferred into other files.)".[43] Files based on more "general-purpose tracking relations" are called "encyclopedic files".[44] Later we shall see how this kind of mental files story can be used to explain cognitive dynamics.

Now I present a possible scheme for the creation of demonstrative mental files that underlay demonstrative thinking and serve as temporary repositories of information on some object, later transferred to other kinds of files:

---

[43] RECANATI, F. 2010, p. 157.
[44] *Ibid.*, p. 158.

t_1 – context 1: the object is sensorily perceived	t_2 – context 2: the object is no longer sensorily perceived
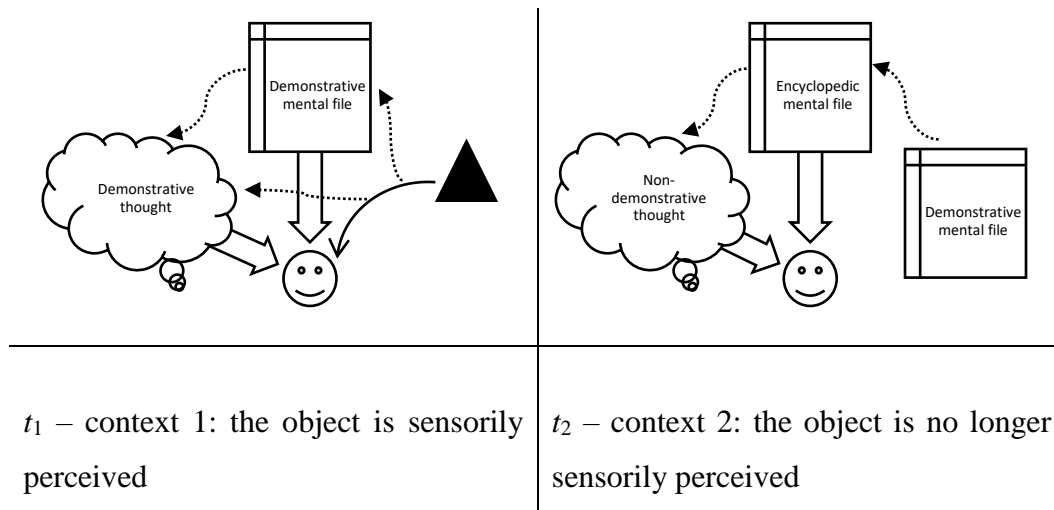
- The minimalist representation of a human face stands for a subject;
- The black triagle stands for an object perceived by the subject;
- The unbroken arrow stands for a (pre-conceptual) information link holding between the subject and the object she perceives (at time t_1 there is an information link; at time t_2 the information link has been severed – the object is no longer perceived);
- The large arrows symbolize whatever relations hold between 1. The subject and a thought and 2. The subject and a mental file (for example: grasping or having a thought and opening a mental file);
- The dotted arrows symbolize dependence relations (whereby information is transmitted, for example, from the information link and into a file).

It seems that in the case of a demonstrative thought, both the thought itself and the corresponding mental file (which supposedly feeds the thought with information) are dependent on the existence of an adequate perceptual relation (that is the reason for the dotted arrows in the picture). Besides, the thought depends on the mental file to exist, since it is supposedly through it that it gets information. So, the demonstrative mental file depends on the information link (made possible by the perceptual relation) and the thought, in its turn, is doubly dependent, depending both on the demonstrative file and on the information link. Now, once the subject changes from the original context in which she was perceiving the object to some other in which the object is no longer perceived, the demonstrative file ceases to exist, since the perceptual relation ceases to exist as well. The information is then transferred into an encyclopedic file, which stores informations from various kinds of sources, or "epistemically rewarding" relations, as Recanati puts it.

It seems to me that one could still find it strange and unnecessary to talk about a temporary mental file which stores information in such a transitory manner, and which disappears once the perceptual relation is interrupted. The idea of mental files certainly sounds appealing, but simply posing the existence of a mental file still seems to leave some questions open, leaving room for doubt concerning its explanatory power. How exactly is it that the transference of information happens? And what role does memory play in this picture? Memory is, after all, a cognitive faculty that seems to be closely connected to perception and, above all, to the re-identification of objects.

But we have digressed far enough from our discussion of Frege's ideas on the first person for now. I shall return to the details of Recanati's mental files theory later when I discuss cognitive dynamics more directly (for he proposes to account for cognitive dynamics with the help of his mental files metaphor). It was important to outline Recanati's theory because my purpose is to draw the reader's attention on the weight give in his theory to the notion of *acquaintance relations*, which are just channels through which we can get information from objects. They are what we might also call (following Evans) information links, though Recanati's notion is very broad. The importance of information links to demonstrative thinking in particular is explicit in Evans's work. As he puts it, "demonstrative thoughts take place in the context of a *continuing* informational link between subject and object".[45] In the specific case of pure demonstrative thoughts such information links seems to be pre-conceptual.[46]

Recanati argues that there exist a great variety of such *acquaintance relations*. I think that, even if we do not appeal to the notion of mental files in our theory, and even if we do not have a very broad notion of "acquaintance relation" or "information link", we can follow Recanati in this particular line of thought without raising too many problems.[47]

Now we get to the point that interests us most in the context of our current discussion: Recanati holds that among the various possible "acquaintance

---

[45] EVANS, G. 1982, p. 146. His italics.
[46] *Ibid*., p. 144.
[47] Carpintero talks about Recanati's view on acquaintance calling it a "liberalization of the acquaintance view", which Carpintero himself does not find fully convincing. *Cf.* GARCIA-CARPINTERO, M. 2013, p. 191. I agree with Carpintero on this, but like Carpintero in that paper, it is not yet the time to rebuke Recanati's opinion on that particular matter.

relations", some are special and, for that reason, are associated with very specific types of mental files. This is the case of the EGO or self-file.

The kind of explanation Recanati gives for the specialness of the self-file and the specialness of the thoughts that one can hold with the help of such a file can be seen as a way of explaining the Fregean thesis on the first person presented in "The Thought". Recanati holds (and I think he is right in so doing) that an individual obtains information about herself in ways in which she is not able to obtain information about others. The idea is that I have special acquaintance relations to myself that I have with nothing or no one else. He mentions, at least, three of such relations in the case of the first person: proprioception, kinaesthesis, and introspection.

Note that the distinction drawn by Recanati between *indexical* and *encyclopedic* files matters for understanding the nature of the self-file. Indexical files are based on specific contextual relations and have a short life span (the only exception being the self-file, but that is a very special case, because the self-file is not purely indexical). On the other hand, encyclopedic files are based on information originating from many different kinds of sources. They are the mental files that store the information gathered in demonstrative relations once these are severed. As files are meant to determine the referent, in the case of encyclopedic files what determines the referent is an "overarching tracking relation", as Recanati puts it, between the file and the object it is about (or which it tracks). The mental file responsible for storing information on the first person is a hybrid kind of file because it is based both on contextual relations (information links that are similar to the ones involved in indexical files), and also on information gained from other sources such as descriptive information (read somewhere or heard from someone, for instance). In the case of the first person, communication can also play a crucial role in finding out more about oneself: it is often by talking to one's own parents that one finds out one's own date of birth, for example. So, on analyzing the self-file we get to see that there can be files based on epistemically rewarding relations of more than one kind, involving mixed types of information:

> […] a file based on a certain [epistemically rewarding] relation contains two sorts of information: information gained in the special way that goes with that relation (first-

person information, in the case of the SELF file), and information not gained in this way but concerning the same individual as information gained in that way.[48]

On explaining Frege's example, Recanati suggests there is a linguistic mode of presentation for the word "I", which in communication happens to be the same for the producer and the listener. But this is not to be conflated with the psychological mode of presentation that occurs in the mind of the speaker – that is, with the concept the speaker has of herself made possible by the internal relations hold by each one to oneself.

Again, even if we are not confortable with the mental files metaphor, I think we can still agree with Recanati (and with Frege, if Recanati's construal of Frege's claims in on the right track) that the very concept one has of oneself is special mainly because it is based on certain information links one can only have with oneself. It is my opinion that therein lies the explanation for much of what we take to be the special characteristics of the first person. Maybe we can think of mental files in terms of concepts so that the self-file, for instance, is taken to be equivalent to the concept one has of oneself.[49]

1.3.3.
The messy shopper

Perry introduces what he calls "the problem of the essential indexical" with one of his most celebrated examples:

> 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.[50]

What is interesting in this case is the fact that Perry (in the example) seems to have *two distinct beliefs*: first, he believes that *the shopper with a torn sack* was making a mess, later he comes to believe that *he* was making a mess. This is puzzling, because both beliefs seem to have the exact same content: they attribute the property of making a mess to *John Perry*. So, in Perry's own terms, they express the same thought.

---

[48] RECANATI, F. 2010, p. 157.
[49] This has been suggested by Carpintero. *Cf*. GARCÍA-CARPINTERO, M. 2013, p. 197.
[50] PERRY, J. 1979, p. 3.

The kind of situation Perry describes is not uncommon, and can sometimes even be found in the literature (not of the philosophical kind, or at least not *primarily* philosophical). As it has been recently pointed out by Torre,[51] we find a similar example in "Pooh and Piglet Go Hunting and Nearly Catch a Woozle", one of the tales of *Winnie-the-Pooh*, written by A. A. Milne. According to the story, Pooh the stuffed bear follows tracks on the snow that he takes to belong to a dreaded creature he calls a "Woozle". After walking in circles for a while as the tracks multiply, however, he realizes he had been making the tracks himself and promptly gives up his hunt for the "Woozle".[52] Similar examples can be found in sources as old as Greek mythology, as Ovid's version of in the myth of Narcissus. Ovid tells us that young Narcissus kindles the desire of everyone he meets, men and women, but loves no one in return. He afflicts so much pain upon others that he is eventually cursed by the gods to fall madly in love with himself and suffer from the same agony. As it is well known, that happens when he chances upon his own image in a pool of reflecting water. But at first he is not aware that the image he is seeing is his own – despair and grief only attain its limits and drains him to his death when he comes to the realization that the object of his desire is himself and thus entertains a *de se* thought.

Note that in all of these examples there is a change in the subject's behavior after the switch from a *de re* to a *de se* thought. Besides, they show that we can think of ourselves without knowing it, without even *intending* it. It is interesting to note that sometimes it seems as if the content of my thought is determined independently of my intending to think it. There seems to be a shift in perspective when I go from thinking about myself in an impersonal or unconscious way to entertaining a self-conscious thought, even though the object of which I am thinking remains the same: myself. It can still be said that I am *thinking about myself* even if I am not aware of it – at least in the sense I can still entertain a thought which has myself as a constituent, in spite of my awareness of that fact. That kind of thought about oneself would be a *de re* thought, taking the individual as part of the

---

[51] In GARCÍA-CARPINTERO, M.; TORRE, S. 2016, p. 1.

[52] MILNE, A. A. 1926. "Pooh and Piglet Go Hunting and Nearly Catch a Woozle". In *Winnie-the-Pooh*. London: Methuen.

thought.[53] Only after having taken that into account, we may refine further our general notion of "thought about oneself" the way Perry does (and the way Recanati does as well) by distinguishing thoughts about the person *I happen to be* and thoughts about oneself *as oneself*. In Recanati's terms, "accidental *de se* thoughts" and "explicit" and "implicit *de se* thoughts" respectively.[54] The terminology that we can employ here is vast, of course, as are the treatments that can be presented. I just want to draw the reader's attention to the idea that it is not enough to entertain a *de re* thought about oneself to have or entertain a *de se* one. Something else is required: in Perry's theory, what makes the thought about oneself a full-fledged *de se* one is *a way of believing* (contemplating the singular content) that is not part of the (*de re*) content itself. Sometimes it is not up to the thinker to entertain a *de se* thought in that sense, because at least sometimes we do not *choose* to be in such (self-locating) belief state, as in the case of the messy shopper, Winnie-the-Pooh or Narcissus.

So, Perry explains his own example of the messy shopper in terms of change in *belief*. He explains the shift from a *de re* to a *de se* thought by saying that there is a change in the *belief state*, not in the *thought* held. Now, as I will try and explain, Perry arrives at that conclusion after coming to the realization that neither Frege's semantic theory nor his own reformulation of it provide us with enough elements to explain the problem of the essential indexical. The problem is this: in examples such as that of the messy shopper, the indexical "I" cannot be eliminated nor substituted by any other expression. Substitution of the indexical "I" in the examples above by any other designator (not only definite descriptions like "the shopper with a torn sack", but presumably also directly referential terms like "he" pointing to the subject's image in a mirror) results in a loss in the explanation of the subject's behavior and the role played in this explanation by (self-)locating beliefs. Another famous example put foward by Perry makes this point clear:

> An amnesiac, Rudolf Lingens, is lost in the Stanford library. He reads a number of things in the library, including a biography of himself, and a detailed account of the library in which he is lost. He believes any Fregean thought you think might help him. He still won't know who he is, and where he is, no matter how much knowledge he piles up, until that moment when he is ready to say,
> This place is aisle five, floor six, of Main Library, Stanford.

[53] Incidentally, that is not a kind of thought usually associated with the original Fregean doctrine, in which there existed only *de dicto* thoughts, but I think *de re* thinking can be accommodated into a neo-Fregean theory – Evans has already attempted this.
[54] *Cf.* section 1.2

I am Rudolf Lingens.[55]

Being lost and not knowing who he is, Rudolf could not possibly find his identity or his location even if he read all the books in the library, including a biography of himself. All that descriptive knowledge (or in other words, all those *de dicto* thoughts) he might entertain would not be of any help. We will deal again with the kind of problem posed by Perry's example of the lost amnesiac Rudolf Lingens in the next section, when we discuss Lewis's similar example of the two all-knowing gods and his own theory to account for this kind of case, which is quite different from Perry's. Now let us try to understand why exactly Perry takes the essential indexical to impose a problem for both Frege's semantic theory and Perry's own alternative to it.

Put simply, Frege's idea about indexicals and demonstratives seems to have been this: indexicals have incomplete senses that are only completed in the presence of elements of the contexts in which they are employed. Non-indexical sentences (also called "eternal sentences") express "eternal thoughts" which, in Frege's terms, are complete in every respect. A non-indexical sentence will always express the same complete eternal thought, independently of the specific contexts in which it may be used. In this case, the sense the non-indexical sentence *has* is always the same as the sense it *expresses* in every specific context. In the case of indexical sentences, only the sense expressed can really be complete, that is, once the indexical sentence is used in a context.

But within Frege's semantic theory, the senses of indexical terms such as "today" should also contribute to the senses of the sentences in which they appear (that is, the *Gedanke* expressed by those sentences). However, as Perry argues, while indexical terms can serve as sense completers, they do not contribute to the thought expressed by the sentences in which they appear with *completing senses*, as one would expect within a Fregean theoretical framework. *Pace* Frege, it is more appropriate to say that they contribute rather with *objects*: "when we understand a word like "today", what we seem to know is a rule taking us from an occasion of utterance to a certain object", Perry writes.[56] So, the meaning that we really understand when we understand indexicals and demonstratives seems to be those rules. And those rules Perry calls *roles*. Again, the rule is like a function that takes

---

[55] PERRY, J. 1977, p. 492.
[56] PERRY, J. 1977, p. 479.

us from *occasions of utterance* to the referents of the indexicals. Now, "the role of a demonstrative does not seem reducible to other notions available from Frege's philosophy. Senses do not carry us from context to references, but directly to references".[57] Frege's idea was that indexicals provided us not with objects directly, but with *completing senses*. But in Perry's perspective, the eternal role (the linguistic rule) associated with indexicals do not provide us with a sense, and the changing objects it denotes depending on the occasion of utterance does not give us a sense either – it provides a reference possibly corresponding to many senses. Since Frege insisted that only senses can be part of senses, Perry suggests that it is hard to think of a way in which we would have a thought perfectly composed of small entities of *Sinn* in such (Fregean) picture, given that we are working with indexicals.

Perry claims that if Frege's account of indexicals and demonstratives was right, we could at least in principle reproduce the same sense of an utterance containing an indexical/demonstrative. Can other people apprehend the thought David Hume might have entertained by expressing to himself that "I wrote the *Treatise*", for example? Perry asks us to suppose that a crazy man named Heimson believes himself to be David Hume. If Heimson expresses to himself the same sentence Hume might have expressed, *i.e.*, "I wrote the *Treatise*", will he be able to grasp the same thought Hume had? Even if the states of the minds of both men are very similar, the fact remains that only David Hume himself will have grasped a *true* thought.

By using that example, Perry's goal is to show that Frege's ideas on indexicals and demonstratives are wrong, and that the senses of indexicals and demonstratives cannot be Fregean senses. "There is no reason to believe we are on each occasion each equipped with some non-demonstrative equivalent of the demonstratives we use and understand. This goes for 'I' as well as 'today'".[58] There must be a different account for indexicals and demonstratives, and that pushes Perry towards his own account in terms of *roles*. He proposes we split the Fregean notion of *Sinn*. In Perry's scheme, the sense of a demonstrative sentence is a role, and the thought it expresses is individuated by an object (the referent of the demonstrative) together with an incomplete sense, as in "( ) is the author of the *Treatise*", for instance. This is how Perry explains the Heimson/Hume case. What we have in that case is

---

[57] *Ibid.*
[58] *Ibid.*, p. 488.

Heimson and Hume entertaining the very same sense – that incomplete sense just mentioned. But they think different thoughts, because the incomplete sense they entertain is completed with the rule (*role*) of the pronoun "I", which in each case gives a different individual, Hume in the case of Hume and Heimson in the case of Heimson. In each case, the contribution the first-person pronoun makes to the resulting thought is the individual directly.

Perry's idea is that the utterance of an indexical/demonstrative sentence in a context does not yield a complete sense appropriate to be a Fregean thought. it yields instead an incomplete sense and an object (the latter being given directly by the role of the indexical/demonstrative in question). Hence his need for splitting the Fregean notion of *Gedanke* in two: a thought proper (which has as a component part the object denoted and can be seen as a unit of information) and the role it plays, which in the case of indexical/demonstrative thought, varies with the indexical/demonstrative expression through which the thought is held.

Now, the part of meaning Perry calls "role" helps us explain behavior. Whenever the roles of the contents in which I believe change (when I go from thinking of myself in an impersonal way to having a self-conscious thought, for instance), my behavior also changes. However, appeal to the distinction between the role of the expression and the content believed in (or the thought held) is still not enough to account for the so-called problem of the essential indexical. Remember, the problem is this: in examples such as that of the messy shopper, the pronoun "I" seems to be irreplaceable. It cannot be substituted by any other designator (of Perry, the subject of the thought) without loss in the explanation of his belief and subsequent behavior. So, not only the role of "I" helps explaining Perry's behavior (making him stop following the trail around the counter and rearrange the torn sack in his cart), it is *the only thing* that does this job. Explaining why that happens is the problem.

Perry argues that the essential indexical poses a problem for all accounts of attitudes that see them as relations between subjects and thoughts of the traditional sort, such as Fregean thoughts. Perry calls the traditional way of dealing with belief

(Frege's theory included) "the doctrine of thoughts".[59] He sees that doctrine as having three tenets:[60]

1. Belief is a relation between a subject and an object which is denoted by a that-clause and is taken to be a thought;

2. Thoughts have a truth-value in an absolute sense (they are not simply true or false for particular persons at particular times).

3. They are individuated by their truth-value (two thoughts must have the same truth-value to be deemed identical) and by their truth conditions (two thoughts must attribute the same relation to the same objects if they are to be identical).

Now, the problem that the essential indexical poses to such traditional accounts of belief is this: such accounts cannot provide us with an explanation of what Perry came to believe when he grasped the thought "I am making a mess". First, indexically essential belief does not seem to be a relation between a subject and a traditional thought. Besides, indexical sentences cannot be thoughts as thoughts are characterized in that traditional kind of theory, because they do not seem to possess absolute truth-values – they only seem to be true or false when they are uttered by someone. Perry's idea is that, from the perspective of a traditional theory of thought, indexical thoughts miss a "*conceptual ingredient*: a sense for which I am the reference, or a complex of properties I alone have, or a singular term that refers to no one but me".[61] That lacking ingredient is indexical. And what Perry calls "locating beliefs" (which "refer to one's beliefs about where one is, when it is, and who one is"[62]) are yet different, they are "essentially indexical". As such, they could not be accounted for as thoughts in a traditional (Fregean) sense. This is basically the same problem Lewis identifies with what he also calls "self-locating beliefs", as we will see in the next section, in a paper Lewis also published in 1979. Lewis argues, like Perry, that such beliefs cannot be *propositional*.

We shall see in the next chapter that with the help of his distinction between character and content (which was just another way of splitting Frege's notion of *Sinn*), Kaplan is able to explain how we can use the same indexical in different

---

[59] Perry actually uses the term "proposition", because he takes it to be more neutral. I shall continue to use the terms "thought" and "proposition" interchangeably, though in general I prefer to follow Frege in using the word "thought". Whenever a clearer distinction between thoughts and propositions is needed, I shall provide it.

[60] PERRY, J. 1979, p. 6.

[61] *Ibid.*, p. 7.

[62] *Ibid.*, p. 5.

contexts to mean different things, and different indexicals in the same context to mean the same thing. So, depending on the context in which they are used, indexicals may present either different or the same content (referent). Perry has a very similar strategy. If we take Perry's example of the messy shopper, different indexicals ("he" and "I", which possess different *characters*, as Kaplan puts it) present the same referent – Perry himself. As we have already seen, in Perry's theory, an indexical character becomes its *role*. The role or character played by an indexical presents its fixed *linguistic meaning*, a kind of rule that gets us from contexts to contents.

Perry argues that the traditional doctrine of thoughts cannot accommodate *de se* belief into its picture, but even belief of the *de re* kind is kept out. If we wanted to include *de re* thoughts in Frege's theoretical framework, we would have to alter its third tenet (as described above) and allow thoughts to be individuated by objects and a part of thought as thoughts stand in the traditional account. This is exactly what Perry does. And by doing that, he feels able to allow for substitutivity of coreferential terms in belief contexts, thus making sense of quantification into those contexts. The sort of belief that would allow that would involve a relation to a proposition consisting of an object (or objects) and a conceptual ingredient is a *de re* belief, while "the sort of belief and the sort of proposition that fits the original doctrine [is called] '*de dicto*'".[63] That is all the traditional doctrine would be able to give us: *de dicto* thoughts.

Perry's new way of conceiving propositions allows him to accommodate in his theory *de re* thoughts or attitudes, but that is still not enough to explain the *de se*. *De se* belief states (such as the one presented in the example of the messy shopper) show us that there is a specific kind of thought – *de se* thought, which carries an extra layer of informative significance, so to speak, not explained by its "*de re* quality". *De se* belief is not reducible to *de re* belief any more than it is reducible to the *de dicto* variety.

"Even if we suppose – as I think we should – that when I said "I believe that I am making a mess" I was reporting a *de re* belief, our problem will remain", Perry argues.[64] It will not be Perry's having a *de re* belief about himself that will *explain his action*. If in the example of the messy shopper, there were mirrors at either end

---

[63] *Ibid.*, p. 10.
[64] *Ibid.*

of the counter alongside which Perry was pushing his cart, so that he could see a reflection of the messy shopper and really have a *de re* belief about himself, that would still not explain his change in behavior. The difference in the channels of acquaintance is clear in this second version of the example. First, he has a visual acquaintance with the referent – which is just himself. Later, he refers to himself through some other information link. That would help us explain the difference in Perry's beliefs. *De re* thoughts are still non-indexical, Perry claims. "Propositions individuated in part by objects remain as insensitive to what is essential in locating beliefs as those individuated wholly by concepts".[65]

Perry's solution to the problem of the essential indexical consists in making "a sharp distinction between *objects of belief* and *belief states*".[66] He argues that we must distinguish between the *content* of a certain belief state (which is to be taken as the kind of reformed *de re* proposition) on the one hand, and the *belief state* itself (in which the subject finds herself in) on the other. In the example of the messy shopper, Perry does believe in the same proposition, but he believes it in two distinct ways, that is, he grasps that proposition while being in two different belief states. The first could be said to correspond to a more traditional Fregean thought, and we could call it *de dicto*; the second is not reducible to either *de dicto* or *de re* thoughts (although it is *de re* in the sense of involving Perry himself as a constituent): it is *de se*.

The way Perry conceives of thoughts/propositions, these help us explain how a subject's attitudes are important in accounting for his rationality, but also for the adequacy of his beliefs to evidence, for example. Still, we need a more fine-grained notion of a thought if we want to fully account for human behavior. It will not do to have just those *de re* thoughts, we must also specify the belief states in which the thoughts are held. Thoughts taken as objects and properties, as Perry wants us to understand them, will not help us solve the problem of explaining behavior. We must take belief states to have a kind of meaning as well.[67] If we want to understand the essential indexical, traditional Fregean *de dicto* thoughts will not help us, and *de re* thoughts by themselves will not help us either. Instead of looking just at indexical thoughts, we should consider the belief states in which people find

---

[65] *Ibid.*, p. 12.
[66] *Ibid.*, p. 4. Emphasis mine.
[67] For a similar account, *cf*. GARCÍA-CARPINTERO, M. 2013, p. 194.

themselves when grasping indexical thoughts. Perry's proposal consists at analyzing the relationships between belief states and the thoughts in which we believe. We should not, as a traditional thought theory would recommend, take belief states to be individuated by the thoughts that are their contents. In the traditional scheme, "whenever we have believers in the same belief state, we must expect to find a proposition they all believe, and differences in belief state lead us to expect a difference in proposition believed". And now we return to Frege's theory of inaccessible and incommunicable senses. The kind of reasoning that consists in saying that belief states are individuated by their contents is what, according to Perry, led philosophers like Frege to posit thoughts with limited accessibility (such as I-thoughts).

That is not to say that a subject's belief states can (and sometimes should) be characterized by a kind of structure consisting of *de re* and *de dicto* thoughts/propositions (*de dicto* if the thought in question is not indexical, but not only then, of course, since one could also have a *de re* thought about something by using a proper name, for example). That is important, for example, to understand how a subject's belief state is to be adjusted as she moves from one context to the other, given that her belief is to be preserved.[68] From Perry's point of view, that shows that a subject may have different ways of believing the same content (the same *de re* or *de dicto* thought), "and to these different ways of believing the same thing, different actions are appropriate".[69] So, in the end, "anyone at any time can have access to any proposition. But not in any way", he argues.[70] Perry's conclusion is that, in the example of the messy shopper, only John Perry himself could access the *de re* thought that "John Perry is making a mess" *by way of the first person*, by being in the belief state corresponding to the sentence "*I* am making a mess".

---

[68] That something to which Kaplan has drawn our attention, calling it *the problem of cognitive dynamics*, as we will see in the next chapter.

[69] PERRY, J. 1979, p. 19.

[70] *Ibid.*

1.3.4.
The two gods

1.3.4.1.
Introduction

David Lewis famously defended that we should not take propositions (or thoughts) as the objects of our attitudes (such as belief). He proposed we take the contents of our attitudes to be properties instead of propositions, meaning by "property" a kind of set – "the set of exactly those possible beings, actual or not, that have the property in question".[71] For him, in some cases, propositions will not do well as candidates for being the objects of our attitudes, but not in *all* cases, while properties will *always* do. I will now try to sketch the basic line of reasoning that made Lewis feel capable of arguing for that thesis and also put forward another very interesting idea:

I think the most important part of Lewis's argument starts when he points out to us the following: First, that the subjects of propositional attitudes are spread out in physical space. For example, people living Brazil or in China or even orbiting the moon can hold beliefs. They are also spread out in time, as people have had beliefs in the past, have them now and will have them in the future. And they can be found all over in logical space too – which just means that subjects have possibly many different beliefs, they believe things in many possible worlds. But here is the second important point: *each particular* subject of attitudes inhabits only one *particular* world.[72] Now, as we have a great population of subjects of propositional attitude spread throughout both time and space (both physical and logical space), Lewis asks: what happens when *one particular member* of this vast population actually has an attitude (such as belief)?

His first answer is this: she locates herself in logical space. And here we find one of the most interesting ideas of Lewis's paper, the one that interests us the most here: when a particular subject holds one such attitude as belief, she also holds an attitude about herself, that is, she also has a belief about herself. That is the belief of inhabiting one specific possible world. Since Lewis wants to defend that the objects of our attitudes are properties, not propositions, that translates as self-

---

[71] LEWIS, D. 1979, p. 515.
[72] *Ibid.*, p. 517.

ascribing the property of inhabiting the possible world in question. And the self-ascribed property corresponds exactly to the proposition in which the subject believes (given a more standard view on the nature of propositional attitudes). So:

> To believe a proposition is to self-ascribe the corresponding property. The property that corresponds to a proposition is a locational property: it is the property that belongs to all and only the inhabitants of a certain region of logical space.[73]

But what is the point of substituting propositions (or thoughts) for (self-ascribed) properties when analyzing attitudes such as belief? The idea behind this move becomes particularly clear when we examine the kind of thought one can have about oneself.

1.3.4.2.
Rudolf Lingens (again)

Lewis takes up Perry's example of the amnesiac Rudolf Lingens lost in the Stanford library. We have seen that, being a lost amnesiac, Rudolf may as well read all the books in the library, including his own biography, and still be incapable of knowing *who he is* and *where he is* at that moment. According to Lewis's interpretation of Perry's thought experiment, Rudolf may acquire a massive quantity of knowledge, learning about the world he lives in and coming to believe more and more propositions about that world. As he does that, he gradually discards other possible worlds in which he could be living in. Still, as Lewis points out, such propositional knowledge is not enough for someone like Rudolf to *locate himself* in *ordinary space* – to do that he needs to self-ascribe a property, and that property does not correspond to a proposition, not any proposition he may learn from the books of the Stanford library.

1.3.4.3.
The two gods

At this point Lewis puts forward a thought experiment of his own, meaning to further illustrate his thesis and present us with an ever more dramatic situation:

> Consider the case of the two gods. They inhabit a certain possible world, and they know exactly which world it is. Therefore they know every proposition that is true at their world. Insofar as knowledge is a propositional attitude, they are omniscient. Still I can imagine them to suffer ignorance: neither one knows which of the two he is. They are not exactly alike. One lives on top of the tallest mountain and throws

---

[73] *Ibid.*, p. 518.

down manna; the other lives on top of the coldest mountain and throws down thunderbolts. Neither one knows whether he lives on the tallest mountain or on the coldest mountain; nor whether he throws manna or thunderbolts.[74]

As I have said, Lewis example of the two gods is really an exaggeration of Perry's example of Lingens, the lost amnesiac. The two gods know absolutely all the propositions about the world over which they reign, but the idea is that it is still possible to imagine that they lack some kind of knowledge: they do not know where they are, and who they are. "If the gods came to know which was which, they would know more than they do. But they wouldn't know more *propositions*",[75] Lewis explains.

With this, he believes to have shown that at least in some cases (such as the case of the two gods or of Lingens, the lost amnesiac), positing propositions as the objects of our attitudes will *not* be a good thing, but positing properties will – more exactly, self-ascription of properties. Lewis further suggests we call self-ascription of properties "*de se* attitude" (or "*de se* belief", or also "knowledge *de se*"). Again, the idea is that, when a subject thinks something, and thus has an attitude of belief which would normally be represented as a relation between the subject in question and a proposition, what actually happens is that she selects a set of *centered* possible worlds. There exists always this element of locating oneself in a certain place (logical and physical) and at a certain time. In Perry's example (which we discussed in the previous section), when he goes from something like "He is making a mess" to "*I* am making a mess", he locates himself in that particular world and time. With this, Lewis is able to distinguish between the content of that attitude, which can be the same for many subjects living in many possible worlds, from the actual act of believing it. When it is done by Perry, for example, upon spiling sugar all over the supermarket floor, when one is evaluating the truth or falsity of that particular thought, it is Perry's *attitude*, not its content, which comes into the picture. His attitude brings him as a subject, as well as the place and time relevant for such an evaluation. In Lewis's account, it is the act of believing that does that, not a supposedly believed proposition.[76]

Independently of accepting or not a Lewisian view on propositional attitudes, what Lewis remarks with his examples is important for us because he believes to

---

[74] *Ibid.*, p. 520-521.
[75] *Ibid.*, p. 521. Emphasis mine.
[76] *Cf.* GARCÍA-CARPINTERO, M. 2013, p. 193-194 for a similar analysis.

have found a kind of attitude or knowledge, i. e., *de se*, which is not reducible to any other kind of knowledge, such as knowledge *de dicto*. (In fact, as we will see, it is rather the contrary, knowledge *de dicto* is for Lewis reducible to knowledge *de se*.) A traditional Fregean thought would be an instance of knowledge *de dicto*. If we are working within a more liberal neo-Fregean theoretical framework, we could also say (with Recanati, for instance, as seen above) that certain thoughts involving a direct reference (or a direct epistemic link with the referent) can be called *de re*. None of these are what Lewis believes to have found (and what Perry found with the example of the messy shopper and of Rudolf, the lost amnesiac). He claims to have found irreducibly *de se* belief, which Perry calls "self-locating belief". But following his account to the end, he suggests we go further than Perry and analyze all knowledge as self-locating knowledge, or knowledge *de se*. In the end, he proposes a reduction of the other categories into the *de se*. So, it can be said that for Lewis *de dicto* and *de re* thoughts are a species of *de se* thought.[77]

Lewis believes to have shown that we can envisage the reduction of *de dicto* forms of knowledge into the *de se* category with the aid of examples like the one he gives involving the two omniscient gods. But he also argues for reduction of de re beliefs by an internalization of the acquaintance relations, as explained by Recanati.[78] We will get to that in a moment.

The point that both Perry and Lewis are trying to make with their examples (the two gods and the lost amnesiac) is clear enough: there exists a class of thoughts irreducible to others. I accept it, although I side more with Perry in what concerns the nature of the objects of our attitudes and the thoughts we have about things in the world. Lewis gets to the point of giving an advice to those who would prefer to work not with properties as the objects of attitudes, but with something that presents some kind of syntactic structure. "Be prepared to use predicates, open sentences, indexical sentences or meanings thereof – something that can be taken to express properties rather than propositions".[79] But he declares to prefer his own account over Perry's for reasons of simplicity: he thinks Perry's theory is too complicated.

---

[77] Following Recanati's interpretation. RECANATI, F. 2009, p. 261.
[78] *Ibid.*
[79] LEWIS, D. 1979, p. 536.

The way Lewis explains it,[80] in Perry's theory, we can regard our attitudes as having two kinds of objects: traditional (Russellian) propositions consisting of indivuals and properties, and also *functions* taking subjects (or objects) as arguments and delivering traditional propositions (thus, it delivers the first kind of object of attitude. In practice, we just end up with propositions.

But what is interesting when we see things under this light is the following. With the help of the first kind of object, Perry is able to explain why Heimson, the crazy fellow who believes himself to be David Hume, does *not* believe *the same thing* Hume himself believes when the latter thinks himself to be David Hume. We can explain the difference in their beliefs even if we imagine a scenario in which their minds' inner workings are the same or very much alike. Since what they believe in are *propositions* which have as constituents individuals and properties, Heimson's belief has Heimson as a component, while Hume's has Hume as a component. Thus, only in Hume's case the proposition is true. Now, the second kind of objects helps us explain people's behavior, because it is a kind of content which we can think of as being in people's heads, as Lewis puts it. Taking the objects of belief to be *functions*, we get the sense in which Hume and crazy Heimson have *the same belief*: they belief in *the same function* that assigns a given object (a person) to the property of being David Hume. Only when David Hume himself believes that function, the function will assign David Hume to the property of being David Hume, and thus deliver a true proposition.

1.3.4.4.
Lewis on the *de re*

Even though from Lewis's perspective, Perry's theory on the nature of the objects of our attitudes lack the simplicity of his own account, one of its advantages seems to be the possibility it opens to explain not *only* "self-ascription of properties", or belief *de se*, but also other-ascription, as Lewis puts it, or belief *de re*. In Lewis's own terms, belief *de re* can be simply taken to mean ascription of properties to individuals. And he thinks that, according to Perry's theory, belief *de se* is just a special case of belief *de re*. I ascribe a property to the individual I am. Of course this ascription has a special quality which deserves a different name, but

---

[80] *Ibid.*, p. 536-537.

it would still be a case of ascription of properties to individuals. This is of course the exact opposite of Lewis's own account, in which there is a reduction of other kinds of knowledge to *de se* knowledge.[81] Lewis suggests we treat *de re* ascriptions not as beliefs properly speaking, but as states of affairs:

> Beliefs are in the head; but I agree with Perry that beliefs *de re*, in general, are not. Beliefs *de re* are not really beliefs. They are states of affairs that obtain in virtue of the relations of the subject's beliefs to the *res* in question.[82]

He takes Perry's theory to represent more than just beliefs, which in Lewis's theory are, of course, reducible to self-ascription of properties. But how exactly does he proceed to reduce *de re* thoughts to the *de se* variety? At first, Lewis considers the possibility of finding a kind of description that could characterize belief *de re*. The idea here is this: when I entertain a *de re* thought about something, that thought should capture the description of some property of whatever I am thinking about. According to Lewis, we would not need to limit ourselves to "descriptions" that would be statable in some form of language, "we might take descriptions as properties", he argues, that is, "not as particular *expressions* of properties in thought and language".[83] In order to account for both belief *de se* and for belief *de re*, Lewis proposes we take belief about other things (*de re*) to involve those descriptions in a sort of reflexive manner. So, a subject *a* would ascribe some property *F* to a certain individual or object *o* under a certain mode of presentation *m* (which could be stated as a description) based on a relation R_m,[84] if and only if *a* has the relation R_m to *o*, and a *self-ascribes* the property of bearing such relation (via its description) uniquely to something having the property *F* in question.[85]

In the end, Lewis appeals to *acquaintance relations* to explain what goes on between a subject and a given object when the subject thinks of that object in a *de re* manner. After a brief consideration of "essential properties" (and after deciding that we only very seldom know the essences of individuals), Lewis comes to the conclusion that such acquaintance relations would be the best suitable candidates for the descriptions involved in belief *de re*. (Essences would be good candidates

---

[81] *Ibid.*, p. 537.
[82] *Ibid.*, p. 538.
[83] *Ibid.* Emphasis mine. Personally, I would not like to call that form of knowledge descriptive. It seems to me that it would be more like non-concpetual knowledge of the kind discussed by Evans, for example, in EVANS, G. 1982, chapter 5, section 5.2. But there is no need to discuss that issue here, it may be just a matter of terminology anyway.
[84] To borrow from Recanati's interpretation of Lewis's thesis: RECANATI, F. 2009, p. 261.
[85] LEWIS, D. 1979, p. 539.

too, if only they were easy to come by.) And as Recanati puts it (as I have already mentioned above), in Lewis's picture, an acquaintance relation of that kind that must be internalized – by internalizing them, we can also account for the *de se* quality of our thoughts. Thus, *de re* belief itself, which in principle is exterior, becomes internalized through a self-ascription of an exterior property (as in the scheme I reconstructed in the previous paragraph). Interestingly, in Lewis's picture too the relation of acquaintance creates a causal link between the subject and the object she is thinking of, and through that channel the subject can receive information from the object in question. The possibility exists that the acquaintance relations might be faulty, of course, feeding the subject with false information. But it is interesting to note that Recanati follows Lewis in considering that even what we might call "information chains" are also acquaintance relations.[86]

I would like to finish this section by pointing out that there is yet another aspect in which Recanati's conclusions are not unlike Lewis's, and in this I agree with them both, in spite of the details of their respective theories. The aspect is this: there is more than one possible (form of acquaintance) relation I bear with myself. In the example of a person who sees herself in the mirror, unaware that she reflection she is seeing is her own, she will have a *de re* thought about herself – she will think of herself as a *res*, an exterior object. That is one of the available ways one possesses of having an acquaintance relation to oneself that happens to be similar to the way people have acquaintance relations to exterior objects. In that specific example, the thought one is having only becomes *de se* when one makes use of some other acquaintance relation to receive information on one's own states.

1.3.5.
Evans on I-thoughts

The Varieties of Reference, chapter 7 is dedicated to *de se* thoughts, that is, in Evans's overall terminology, thoughts involving self-identification. He warns us at the beginning of the chapter that he will only be concerned with self-conscious *'I'-thoughts* (as he also calls them). There are many ways one may think of oneself, the idea is, not all of them involving self-consciousness.

---

[86] *Ibid.*, p. 542.

Remember our discussion in section 1.3 of Recanati's ideas on *de se* thoughts and his *accidentally*, *explicitly*, and *implicitly de se* taxonomy. Evans is only interested in what Recanati labels *explicitly* or *implicitly de se* thoughts, both of which involve self-consciousness, though in different ways. Accidentally *de se* thoughts are not self-conscious. So, Evans is not interested in them. These thoughts could be described, following Perry, as thoughts about the person one happens to be. In this case, I think of myself without realizing it, for example through a description (so, what I do have is a *de dicto* thought about myself), or in a demonstrative (*de re*) way, as when I see a reflection of myself without being aware that it is my own. Explicitly *de se* thoughts occur when self-consciousness kicks in, so to speak. When, for example, upon having an accidental *de se* thought, I become aware that I am actually thinking about myself – awareness may come through my deciding that the reflection is my own after all. Of course, coming to believe that the reflection is mine can be the outcome either of a fortuitous realization or of a deliberate choice on my part. Be it as it may, in this kind of case, I go from an accidentally to an explicitly *de se* thought. Note that both can still be reduced to either *de re* or *de dicto* thoughts. This does not seem to be the case when we consider implicitly *de se* thoughts. In this case, as I have already explained in section 1.3, I exploit some internal sense, such as proprioception or one of the other senses gathered under the general heading *interoception*. One example would be the thought expressed by an instance of "I am in pain". For Recanati, only *de se* thoughts of this kind (i.e. implicitly *de se*) are always immune to error through misidentification, because there is no identification whatsoever, and the subject would not enter as a component into the content of her own thought. *De se* thoughts of the other kinds are still vulnerable to that kind of error because the subject herself is somehow represented her own thought as a component part of it. In this kind of situation, there can be an error in the identification of the subject of the thought. In short, there is actually something to be identified, while in the case of implicitly *de se* thoughts there is nothing.

As I said, Evans is only interested in self-counscious I-thoughts, of the kind Recanati would call either explicitly or implicitly *de se*. However, for Evans, "the

essence of self-consciousness is self-reference"[87]. So, it is clear that his ideas on self-identification are not totally on a par with Recanati's.

Evans's investigation of thoughts about oneself is motivated by the role an account of such thoughts has to play in a more general theory of reference (and of thought), and by the parallels that can be drawn, as well as the discrepancies that can be found, between I-thoughts and thoughts of other kinds. Evans is particularly interested in drawing parallels between thoughts about oneself and what he calls "here-thoughts" and "this-thoughts".

I will present here a summary of some of the most striking ideas put forward in chapter 7 of *The Varieties*, particularly those relating to theses of other authors earlier discussed. This is not meant to be a comprehensive account of what Evans has to say on the issue of I-thoughts. Chapter 7 of *Varieties* is way too complex to be dealt with here in its entirety, and it has many connections with intricate ideas outlined by Evans in earlier parts of his book. I ignore some of them as for now and most of the time confine myself to a simpler vocabulary. Later on, we shall have the opportunity to explore some of his ideas more thouroughly (particularly when I tackle the issue of the dynamics of *de se* thoughts).

For Evans, I-thoughts resemble other kinds of thoughts in non-negligible ways. In particular, two other kinds of thought are closely related to thoughts about oneself: thoughts about *this* and thoughts about *here*. This is, of course, not very different from what Perry and Lewis had already pinpointed as "self locating beliefs".[88] But for Evans here-thoughts and I-thoughts are particularly similar. The obvious parallels involve: 1. Certain ways of receiving information in some modality, through information channels, or the kind of information grounding the thought 2. The subject's behavior, or the role played by both kinds of thought in action.

There are also important differences, of course, with respect to the quality and the complexity of the sensitivity we have to our own states. The sensitivity is richer in the case of I-thoughts than in the case of here-thoughts, for two reasons: first, I-thoughts also depend on the knowledge we have of our past states through memory; second, self-reference is always involved in I-thinking.

---

[87] EVANS, G. 1982, p. 213.

[88] Evans gives an example similar to the ones discussed by Perry and Lewis, in EVANS, G. 1982, p. 206.

Evans also argues against the view that thoughts about oneself are grounded primarily on some kind of inaccessible subjectivity or private mental states in an exclusive way. Against Thomas Nagel, for instance, he argues that in order to think about oneself one must be able to identify oneself in the objective structure of the world. This must be so if I-thoughts are to obey what Evans calls the Generality Constraint. We shall deal with this idea in more detail later on, particularly in connection to cognitive dynamics. Roughly speaking, the Generality Constraint states that if one is to entertain a thought such as "*a is F*", one must also be able to entertain other thoughts involving the exercise of the same set of conceptual abilities: the generality ability to think F, G, H of the same object in predications like *Fa, Ga*, *Ha* and the generality ability to think F of diferent objets in predications like *Fa*, *Fb*, *Fc*, etc.

To know that "*δ* = I", where *δ* is some objective description of oneself, is equivalent to entertaining a self-locating thought. This is in a way similar to the suggestion made by Perry and Lewis. In this, Evans seems to agree with Frege as well, and boldly defends the thesis that it is possible for there to be a thought that can be grasped *only* by the person who formulates it – a self-locating thought. He writes: "we cannot state in non-indexical terms what it is for the identity-proposition to be true; but why should we suppose that everything that is true can be represented in [a non-indexical way]?"[89]

In his discussion of the phenomenon of immunity to error through misidentification when it comes to I-thoughts, Evans makes it clearer why he takes thoughts about oneself to be more closely related to here-thoughts than to this-thoughts. In the case of I-thoughts and of here-thoughts, the information links that ground the thoughts are possibly always there (at least in standard cases), so that the subject can act whenever she receives the appropriate information. It is because of certain special information links one has to oneself that I-thoughts may present the curious property of being immune to error through misidentification. This can be explained by the fact that certain information links give rise to I-thoughts that are identification-free – and being free of identification is a condition for being immune to error through misidentification.

---

[89] *Ibid.*, p. 211.

Curiously, Evans claims that immunity to error through misidentification may not apply to certain cases of mental self-ascription of properties, while it may apply to certain cases of bodily self-ascription. This is of course contrary to what one might at first think: normally, we would think it is easier to be mistaken about the ascription of bodily properties, never about our own internal states. So, immunity to error through misidentification in the case of I-thoughts is generally taken to concern self-ascription of mental predicates. But for Evans it is actually tied to the ways one gains information about oneself. Some ways of gaining information about oneself are immune to error through misidentification, others are not, and that dichotomy happens in the case of information grounding both physical and mental self-ascriptions of properties. This opens up the possibility for certain self-ascriptions of mental predicates to be vulnerable to error through misidentification, and for certain bodily self-ascriptions to be immune to that kind of error.

We may contrast Evans's view with Wittgenstein's. Contemporary discussion on the phenomenon of immunity to error through misidentification began with a famous remark made by Wittgenstein in *The Blue Book*. He argues that when we express thoughts about ourselves, we make two distinct uses of the pronoun "I". One is what he calls "the use as object" and the other "the use as subject". Examples of the first kind can be found in cases of physical self-ascription of properties. Wittgenstein claims that these "involve the recognition of a particular person", so there exists the possibility of being mistaken in that recognition. The second kind of use of "I" (the use as subject) does not involve recognition of anyone. Examples of this second kind of use can be found in cases of self-ascription of mental predicates, such as seeing or hearing something, thinking that so and so, or feeling pain. In Wittgenstein's own words, in the case of someone who utters "I have a toothache", "to ask 'are you sure that it's *you* who have pains?' would be nonsensical". So, "it is impossible that in making the statement 'I have a toothache' I should have mistaken another person for myself, as it is to moan with pain by mistake, having mistaken someone else for me".[90]

Cappelen and Dever take Wittgenstein to be wrong on the point he makes in that passage. Wittgenstein argues that it is nonsensical to ask whether it is really I, rather then someone else, who is in pain when I say "I am in pain". Of course that

---

[90] WITTGENSTEIN, L. 1958, p. 66-67.

in normal situations the natural answer to that question is "yes". Even though it is "epistemically ill-advised" or "pointless" to wonder if I am in pain, that does not by itself render the question nonsensical. Cappelen and Dever's conclusion is that we need a more interesting claim to show that thoughts expressed by sentences such as "I am in pain" have a special epistemic status, being immune to error through misidentification. I think Evans does just that, providing us with a more sophisticated argument.

Evans also takes Wittgenstein to be wrong on the point he makes, though his reasons are very different from Cappelen and Dever's. There are two problems Evans sees in the ideas Wittgenstein expresses in *The Blue Book*.

The first problem is Wittgenstein's thesis that "I" does not refer to anything. According to Evans, one of the reasons for arriving at such an idea lies in a possible ambiguity in the use of the word "identify". Judging by what Wittgenstein writes in *The Blue Book*, he does not seem to be aware of that ambiguity. Evans argues that all thought about an object (including thoughts about oneself) requires the subject to identify the object of her thought (Evans adheres to Russell's Principle[91]). This is the first sense of the word "identify". But that does not mean that all thought (including all thought about oneself) involves an identification component and is thus "identification-dependent", as Evans puts it.[92] This is the second sense of the word "identify". Thoughts which involve identification in this second sense involve, among other things, criteria of recognition of its object. Now, if a thought does not involve an identification component in the second sense (and is thus "identification-free"), that does not mean that it does not involve an identification in the first sense. So, to sum it up: all thoughts which involve an identification of its object in the second sense (being thus "identification-dependent") also involve an identification of its object in the first sense. But not all thoughts which involve an identification of its object in the first sense must also involve an identification of its object in the second sense. Again, whenever they do not, they are called "identification-free". For Evans, identification-free thoughts are immune to error

---

[91] "The principle is that a subject cannot make a judgement about something unless he knows which object his judgement is about", writes Evans: EVANS, G. 1982, p. 89. He then refers to Russell himself in a footnote, to *The Problems of Philosophy*, p. 58. Evans interprets the kind of knowledge required by Russell's Principle to be a discriminating knowledge of the object. If we are to say that a subject is really thinking about a given object, he must be capable of distinguishing the object of her thought from all other things.

[92] EVANS, G. 1982, p. 218.

through misidentification. There cannot be misidentification if there is no identification in the first place.

The second problem Evans identifies is this: for Wittgenstein, in standard cases of bodily self-ascriptions (what Wittgenstein calls "uses of 'I' as object"), there is an identification component involved, so these ascriptions would be always vulnerable to error through misidentification. Now, against that view, Evans claims that immunity to that kind of error extends even to cases that Wittgenstein would call "use of 'I' as object", that is, to bodily self-ascriptions of properties. Evans argues that immunity to error through misidentification is a property of judgements made upon certain grounds, not simply of certain kinds of propositions. He gives stronger, epistemological reasons for a thought to be immune to error through misidentification or not. My I-thought's having that property or not will depend upon the way of gaining knowledge about myself that is at the basis of that thought. In other words, it will depend upon the kind of information link I have to myself.

Evans's idea is that in bodily self-ascriptions based on proprioception, for instance (a very special kind of information link one has to oneself, as seen earlier), it does not make sense to think that there is an extra identification component involved. There seems to be no sense in saying that the person in question will doubt the identity of whoever is feeling what she is feeling, that is, she cannot doubt that she is herself having a proprioceptive sensation. We have seen that other authors, such as Recanati, defend something similar (precisely under the influence of Evans, I suppose).

For Evans, we possess certain sensory channels and other links through which we receive information that is both immediate and identification-free – that information grounds some cases of our self-ascription of bodily properties. But at the same time, I do not think we can deny that we also have channels or links through which we receive information that is *not* identification-free and that also grounds some of our self-ascriptions (be them physical or mental). In short, it is a fact that there are cases in which we may be mistaken about the source of the information that grounds our self-ascription of properties.

In spite of that, Evans argues, the fact remains that in normal cases self-identification of physical properties does not involve an identity claim. The possibility of there being deviant causal chains of information, for example (possible science-fiction scenarios in which the subject would be misinformed

about her own body states), cannot be taken to ground the conclusion that one can error in identifying oneself in standard cases.[93] In fact, a similar line of reasoning holds not only for physical, but also for mental self-ascriptions. I shall return to this issue in connection to the dynamics of *de se* thoughts in the next chapter. Evans has also much to say on memory and on how memory relates to the first person and to thoughts one has about oneself. But those ideas too must be discussed in connection to cognitive dynamics.

1.4.
Conclusion

The two main objectives of this chapter were: 1. To flesh out the idea that the *de se* is a special kind of thought, irreducible to the *de re* and the *de dicto*; and 2. To find some of the (presumably) distinguishing features of *de se* thought.

I went through some of the main examples and arguments that can be found in the philosophical literature in defense of that idea that *de se* thought is a special kind of thought. I also discussed some of the main accounts put foward to explain it. Now I want to close the chapter by listing the features of *de se* thought encountered so far. I shall also raise some doubts about whether they are really distinctive of *de se* thoughts (at least for some of them).

The first and perhaps most important feature of *de se* thoughts we have discussed concerns the apparent impossibility of substitution of the first-person pronoun. Most of the time, *de se* thoughts are expressed through sentences containing the pronoun "I". Some of the examples discussed above seem to indicate that the first-person cannot be eliminated or replaced by other designator of whoever expresses the thought in question. From the point of view of grammar, it seems we cannot equate the meaning of "I" with that of a name, a description or any other singular term. These substitutions fail to preserve meaning, and alter the modal status of the sentence.

But that is not the only source of evidence for the specialness of *de se* thinking. Another feature of *de se* thoughts (including not only I-thoughts, as we might call them, but also "here-thoughts" and "this-thoughts", to use Evans's terminology) is the fact that they carry what we might call a "self-locating quality"

---

[93] *Ibid.*, p. 222.

which is not to be found in thoughts of other kinds, such as *de dicto* and *de re* ones. This level of content, it is argued, cannot be captured by a traditional *de dicto* (Fregean) thought. This is demonstrated by Perry's example of Rudolf Lingens and Lewis's thought experiment of the two gods. The irreducibility of the *de se* to the *de re* is shown in the example of the messy shopper.

The first-person also seems to be irreplaceable or "essential" in our explanation of behavior. That is also apparent in the example of the messy shopper. In that kind of example, the grasping of a thought through the pronoun "I" (or, more generally, the grasping of an actual *de se* thought, bearing that self-locating quality mentinoned above) seems to be the only way we can explain the subject's (change in) behavior.

Another interesting characteristic of I-thoughts is the fact that their object is also their subject. That is reflected in a distinction originally made by Wittgenstein in *The Blue Book*, as we have seen: that between thoughts about oneself as object and thoughts about oneself as subject. These two aspects are intertwined in normal situations of *de se* thinking. But they can be completely alienated, as in cases of my having a *de re* thought about myself (that is, thinking of myself in an impersonal way), and going from there to entertaining a full-fledged *de se* one (once again, a sort of situation exemplified by the messy shopper).

This takes us to another characteristic of the *de se*: it comes in many forms. It seems that we can identify various kinds of *de se* thought, and that at least some of them have a special epistemic status. Evans and Recanati, for example, defend that we all have unique epistemic mechanisms (such as introspection, proprioception and interoception more generally) that enable us to know our own internal states. *De se* thoughts grounded on those special forms of self-knowledge are supposedly immune to error through misidentification. According to Pascal Ludwig,

> It seems then that there exists a form of self-knowledge possessing the particular characteristic of conducting the *de se* thoughts proceding from it to be immune against error through misidentification. […] This form of self-knowledge is based on particular ways of acquiring knowledge about oneself. It is not possible, for instance, to acquire knowledge of oneself through proprioception if that knowledge does not imply a thought immune againt error through misidentification.[94]

---

[94] PASCAL, L. 2005, p. 103. "Il semble donc exister une forme de connaissance de soi possédant la caractéristique particulière de conduire les pensées de se qui en procèdent à être prémunies contre l'erreur d'identification. […] Cette forme de connaissance de soi repose sur des manières particulières d'acquérir une connaissance sur soi. Il n'est pas possible, par exemple, d'acquérir une

But what is immunity to error through misidentification, really? We can say that there are two ways in which one can be mistaken when making a judgement of the kind "*a* is *F*". The person in question may be mistaken in ascribing the property *F* to *a*; or she can be right about the property ascription, but mistaken in thinking that whoever has that property is actually the individual *a* (rather than some other individual *b*). When holding thoughts that are immune to error through misidentification, the thinker would be protected against making the second kind of mistake, that is, a mistake in the identification of the object of her thought. An example of thoughts with that immunity are *de se* thoughts formed through one type or other of perception which is taken to be particular to the first-person, such as proprioception. When, for example, based on proprioception, I hold the thought that I am sitting down, that would be a thought immune to error through misidentification, because I could not be mistaken about who is actually sitting down. As we have seen, it may be suggested that in this kind of thought I am not myself a constituent of the thought – the thought I hold is not of the kind "F*a*", it is rather composed solely of the property I feel myself as possessing (such as sitting down). In Evans's terms, that kind of thought would be "identification-free".

However, the question as to whether immunity to error through misidentification is a *defining* characteristic of *de se* thoughts is controversial in so far as 1. It seems only some types of *de se* thought present that immunity; 2. Not only first-person thoughts are immune to misidentification; and 3. Arguably, we can identify more than one kind of immunity to error through misidentification. Nonetheless, if our objective is showing the specificity of *de se* thoughts, perhaps it is already enough to prove that there is a specific kind of immunity to error through misidentification that originates due to specific mechanisms of knowledge that exists only in the case of the first person.

Carpintero suggests what he sees as a "subtler" connection between *de se* thoughts and immunity to error through misidentification. He argues that we should take I-thoughts that are so immune as being "fundamentally *de se*". On the other hand, I-thoughts that do not possess that immunity would be "derivatively *de se*". When someone entertains a derivatively *de se* thought by making use of the first-person concept, she would then identify herself as the object of another *de se*

connaissance sur soi par l'intermédiaire de la proprioception sans que cette connaissance implique une pensée prémunie contre l'erreur d'identification".

thought which is fundamentally *de se*.[95] A *de se* thought I might have based on my seeing my own reflection in a mirror is only derivatively *de se* according to Carpintero's terminology – it is not immune to error through misidentification, since I could be mistaken in identifying that image as mine (it could be someone else's). On the other hand, if I identify myself by looking at my own body or by (directly) perceiving through vision a situation in which I find myself in, I would have a *de se* thought immune to error through misidentification. However, even though such a thought would be *de facto* immune, it is possible to imagine weird situations in which it does not possess that immunity. A situation in which, for instance, a science-fictiony gadget for augmented reality creates certain environmental illusions.[96] In this picture, the only kinds of *de se* thoughts which truly deserve the classification of immune to error through misidentification are, again, such I-thoughts based on senses like proprioception. But would they really?

We could imagine a device for augmented reality with enough power to affect our senses, disturbing even proprioception and interoception more generally.[97] It is possible to imagine a mechanism through which all sensory stimuli (even interoceptive ones) of a given person would be redirected to me. Could I then doubt the immunity to error through misidentification of my thoughts based on such senses as proprioception? I think the answer is yes. The mere possibility of those science-fiction scenarios points to the fact that even my I-thoughts based on proprioception are not *de jure* immune to that kind of error, even if they are *de facto* immune in our world as it is. Though I return to this issue in the next chapter, it is not even necessary to turn to science fiction scenarios to see the fallibility to which even our internal senses are liable. In well-known cases of what we call "thought insertion", some schizophrenic patients self-ascribe thoughts which are not their own. Although there is no denying that some of our senses (proprioception, introspection and the like) they are unique to the first person and form the basis for certain kinds of thoughts we have about ourselves (and about nothing else), like all other senses, they are not completely and always reliable.

---

[95] GARCÍA-CARPINTERO, M. 2013, p. 199.

[96] This line of reasoning was originally put forward by Carpintero. *Ibid*., p. 200-201.

[97] Such a device has already appeared in the popular science fiction series *Black Mirror*, notably in the episode entitled "Men Against Fire", in which soldiers have a neural implant called "MASS" that enhances their senses, provides data via augmented reality and induces dreams and memories.

In spite of their failibility, we can still resort to them to classify different types of *de se* thoughts. I think the first great divide that can be made is that between those *de se* thoughts which are self-conscious on the one hand, and those which are not self-conscious on the other. We can use Recanati's terminology and call "accidental *de se* thoughts" those *de se* thoughts which are *not* self-conscious.[98] This is the kind of thought about oneself Perry grasps before he realizes that he himself was making a mess at the supermarket. These are always vulnerable to error through misidentification. Self-conscious *de se* thoughts come in two varieties. To employ Recanati's terminology again, the first variety can be called "explicit *de se* thought". The explicit *de se* can be either vulnerable or immune to error through misidentification. They are vulnerable when grounded solely on senses such as vision, which tipically give rise to thoughts that possess that kind of vulnerability. They are immune when grounded not only on senses such as vision, but also on senses such as proprioception, which tipically give rise to thoughts that possess that kind of immunity. *De se* thoughts that are not self-conscious and self-concious *de se* thoughts which we call explicit can both be regarded as special cases of *de re* thought. The second variety of self-conscious *de se* thoughts can be called "implicit *de se* thought". The implicit *de se* are always immune to error through misidentification, because they are grounded on senses such as proprioception or interoception, which tipically give rise to thoughts possessing that immunity. However, note that the idea that self-conscious *de se* thoughts can be immune to error through misidentification because of the mechanism of self-knowledge grounding them is not free of controversy. It can be argued that those mechanisms do not necessarily give rise to thoughts which possess that immunity. We shall discuss that in the next chapter. Be it as it may, I think it still makes sense to distinguish these three basic kinds of I-thought. And of course other variations that cross through these three types can also be identified, as when one talks about self-ascriptions of bodily predicates as opposed to self-ascriptions of mental predicates. We can also talk about *de se* thoughts based on past self-ascriptions, or even on pre-conceptual information gathered in the past and retained through memory. These

---

[98] In reality, these are not properly speaking *de se* thoughts, if we take true *de se* thoughts to be self-conscious. Here I am only reproducing Recanati's terminology. Accidental *de se* thoughts are only *de se* in the sense of being about oneself, though properly speaking they *de dicto*.

forms of I-thinking are particularly relevant for a general account of the dynamics of I-thoughts.

The following table sums up the most basic distinctions between the different kinds of *de se* thoughts:

Not self-conscious	Self-conscious	
Accidental *de se* thoughts: • Always vulnerable to error through misidentification.	**Explicit *de se* thoughts:** • Sometimes vulnerable to error through misidentification (when based on senses such as vision); • Sometimes immune to that kind of error (when based not only on senses such as vision, but also on senses such as proprioception).	**Implicit *de se* thoughts:** • Always immune to error through misidentification (being based solely on proprioception, interoception or introspection).
Can be taken to be special cases of *de re* thoughts		

2
Cognitive dynamics and the indexical "I"

It is tempting to say that cognitive dynamics is concerned not with retention and change in what is believed, but with retention and change in the characters under which our beliefs are held.

David Kaplan, "Demonstratives"

Nun ist jeder sich selbst in einer besonderen und ursprünglichen Weise gegeben, wie er keinem anderen gegeben ist.[99]

Gottlob Frege, "Der Gedanke"

2.1.
Introduction

My aim in this chapter is twofold: to gain a proper understanding of the phenomenon of cognitive dynamics and to discuss the extent to which the issue of cognitive dynamics applies to the case of thoughts held from the first-person perspective, expressed *via* a token of the first-person pronoun "I". In order to achieve that goal, I have divided the chapter into two parts, in which I tackle each issue in turn.

The next section, 2.2, is dedicated to cognitive dynamics. I briefly go through its history in the philosophical literature, outline the main problems it raises and assess some of the attempts to account for the phenomenon. In Sect. 2.2.1, I explore how cognitive dynamics was presented by Kaplan, namely as the issue of explaining what it means to say of a person who at a certain moment sincerely

---

[99] "Now everyone is presented to himself in a particular and primitive way, in which he is presented to no one else". English translation: FREGE, G. 1918, p. 298.

asserted an *indexical sentence* that at a later time she has (or has not) changed his mind with respect to her assertion.[100] In Sect. 2.2.2, I discuss a very specific problem posed by cognitive dynamics which has to do with what we call the "Intuitive Criterion" of difference for thoughts. In Sect. 2.2.3, I evaluate the merits of the most significant solutions given to the problems raised by cognitive dynamics in the philosophical literature. It is worth noting that there are diverging opinions on what cognitive dynamics really is and how we should approach it. Even though the approaches available to us are not *completely* satisfactory, I think each have the merit of shedding light on a different facet of the phenomenon, given a particular possible way of understanding it.

In Sect. 2.3, I sketch of some of the ideas that might help us account for the mechanism involved in the dynamics of thoughts typically expressed through the use of a token of "I". The problem of cognitive dynamics, as originally formulated by Kaplan, does not at first seem to arise for all kinds of indexical thoughts. It does not seem to arise for a subject's I-thoughts when she thinks of herself under the character of the first-person pronoun and under that character only. It would seem that in the case of that kind of I-thought, no dynamic mechanisms are at work. In the first place, there is no change of the indexical used, so there is no dynamics at the linguistic level. Besides, it can be argued that a subject who grasps an I-thought through time does not need to make an effort to keep track of herself, since she cannot be mistaken about who the referent of her own I-thouhgts is. Against that idea, I argue that dynamic mechanisms similar to that of the mental activity of tracking the referent can still be recognized. That, I think, points to a necessary broadening of the scope of the original problem.

2.2.
Cognitive dynamics

2.2.1.
What is to be found in this section

In this section my aim is to deal with cognitive dynamics as it has traditionally been understood in the philosophical literature. The phenomenon of propositional-attitude retention, change or substitution over time has been usually associated with

---

[100] KAPLAN, D. 1988, p. 538, footnote.

specific problems originally formulated by David Kaplan in his seminal essay entitled "Demonstratives". Kaplan has had a lasting influence on the study of a phenomenon to which he has himself drawn our attention because even though different authors have suggested different ways of approaching cognitive dynamics, most of them share a certain understanding that has as its source in Kaplan's ideas.

For most authors cognitive dynamics is, first and foremost, the study of the retention or change of specific types propositional attitudes, namely attitudes towards *indexical* thoughts (or propositions).[101] In this perspective, one of the goals of cognitive dynamics is to tell us what is necessary for someone to properly retain (or change) an attitude towards a given *indexical thought* through time (or, more generally, given contextual changes). But, as I shall argue, in its original formulation cognitive dynamics concerns a particular class of indexical thoughts, and that particular class only: indexical thoughts the re-expression of which requires an adequate adjustment of the indexical terms used.

2.2.2.
Perspective-tracking: the phenomenon as viewed by Kaplan

A natural starting point to deal with cognitive dynamics and get clear about what it really is is Kaplan's seminal essay in which he coined the very phrase "cognitive dynamics" for the first time in the literature. What Kaplan defends in "Demonstratives", both concerning indexicals and cognitive dynamics, can be seen in turn as a response to the theses put forward by Frege in "The Thought".

Kaplan's main objective in "Demonstratives" was to investigate the semantics of demonstratives and indexicals. To put it simply, it might be said that indexicals (also sometimes called "deictic terms"[102]) are terms that express a different *content* (have a different meaning or semantic value) depending on the context in which they are employed. That is the reason why they are said to be

---

[101] Propositional attitudes can be understood as mental relations between subjects and thoughts (or propositions). They are grammatically associated with verbs such as believing, hoping, wishing, etc., such verbs as can be connected to "that-clauses", those phrases preceded by the preposition "that". An example of such an expression is this: "Arthur believes that Marvin is a robot". What comes after "that" is the expression of a thought (or proposition). Russell noticed this kind of verb in his text "The Philosophy of Logical Atomism", in 1918, suggesting we call them "propositional verbs". *Cf.* RUSSELL, B. 1918.

[102] The word "deictic" is often used in linguistics like we use "indexical" in philosophy (as it is pointed out by Nunberg in his classic "Indexicality and Deixis" (p. 2).

context-sensitive: their semantic value of some expressions systematically depends upon the context of utterance. They are also called *token-reflexive*, for their context-sensitivity is encoded by a linguistic rule that tells us, for each particular *token* or *occurrence* of the expression-type, how its semantic value is determined with respect to features of the context of utterance. In this sense, indexicals are context-sensitive expressions and have a token-reflexive linguistic meaning. Common examples of indexicals are: pronouns ("I", "you", "my", etc.), adverbs of time ("now", "today", etc.), adverbs of place ("here", "there", etc.) and tensed verbs.[103] Given that we understand indexicals that way, the category of indexicals encompasses demonstratives. So-called "true demonstratives" are context-sensitive expressions that require some kind of supplementation from the part of the speaker to successfully refer. In Kaplan's terminology, they require an "associated demonstration".[104] Demonstrative pronouns are common examples of "true demonstratives": words such as "this" and "that", employed either alone or together with some conceptual expression as in "that tree". Either way they require some gesture or special intention of the speaker to successfully refer. As Kaplan himself makes clear, the third-person pronoun "he" can also play the role of a demonstrative if some kind of demonstration is *associated with it to refer to a* (male) individual. "Pure indexicals", by contrast, are indexicals for which "no associated demonstration is required, and any demonstration supplied is either for emphasis or is irrelevant".[105]

Kaplan calls *character* the aspect of the meaning of indexicals and demonstratives that can be represented as a rule that determines the expression's semantic value with respect to relevant aspects of the *context*. The expression's semantic value is what Kaplan terms its *content*.[106] Of course, his notions of character and content generalize over all meaningful parts of speech. The *content* of a sentence is to be understood as a structured proposition (which in turn can be understood as an ordered pair of, for instance, an object and a property). The content of a demonstrative such as "this", for example, is its referent, the object denoted,

---

[103] Other less orthodox examples could be cited, such as the adjective "local", but more important than the extent of the list here is to call attention to the common feature of that kind of expression: the context-sensitivity they share.
[104] KAPLAN, D. 1977, p. 490.
[105] KAPLAN, D. 1977, p. 491.
[106] *Ibid.*, p. 490.

given that indexicals refer directly. And the character of an indexical or demonstrative is to be understood as its conventional linguistic meaning – what is known by a competent speaker. Where Frege associates the *cognitive significance* of a term with its sense (and in the case of sentences, to the thought), Kaplan associates it with the expression's conventional meaning, in short its character.

The character-content distinction explains, among other things, two of the main features of indexicals. First, we can use the *same* indexical expression in *different* contexts to say *different* things (or to express different contents). Second, we can use *different* indexicals in *different* contexts to say the *same* thing. Of course, explaining those two phenomena amounts to explaining a lot about the semantics of indexicals. By employing his distinction, Kaplan is able to say that sometimes the same thought (not the referent, but the *thought*, the content of a sentence) can be presented in different ways, through different characters: "a given manner of presentation – a character [...] – will, in general, present different objects (of thought) to different persons (and even different Thoughts to the same person at different times)".[107] Whenever the same content is actually held under a different character, the holder of that content will find herself in a different cognitive state, which shows itself as a difference in behavior (be it linguistic or not). Kaplan's theory can explain how it is possible for two individuals to be in the same cognitive state (apprehending the same character) and grasp different thoughts: being different individuals, the context is different, so the thoughts yielded by the character are different. A similar thing can happen to the same person at different moments in time or in different circumstances (possible situations): if there is a difference in context and not in character (the very same character/cognitive state is at play), there will be a difference in the thought held.

At this point the connection of Kaplan's theory with cognitive dynamics becomes conspicuous. The way Kaplan construes cognitive dynamics depends on his account of the semantics of indexicals. Indication of this is actually given in one of the very first sections of his text (section VI), where he discusses some of the terminology he uses. A little further ahead (in section XVII), he directly addresses the connection between indexicality and cognitive dynamics. Kaplan's theory of the meaning of indexicals and demonstratives can be employed to explain the

---

[107] KAPLAN, D. 1977, p. 530. He writes "Thought" with a capital "T", alluding to Frege's *Gedanke*.

possibility of re-expressing the *same content* through *different characters* in *different contexts*, particularly given the passage of time. But, more importantly, thanks to his theory Kaplan can also account for the retention of cognitive states associated with the cognitive significance of characters. A famous excerpt from "The Thought", already quoted in the introduction of this work, is one of the motivations for Kaplan's discussion of that particular feature of indexicals:

> If someone wants to say the same today as he expressed yesterday using the word "today", he must replace this word with "yesterday". Although the thought is the same its verbal expression must be different so that the sense, which would otherwise be affected by the differing times of utterance, is readjusted. The case is the same with words like "here" and "there".[108]

Kaplan regards Frege's ideas expressed in this passage as basically correct, except for the fact that Frege did not see indexicals and demonstratives as devices of direct reference. Besides, Kaplan sees Frege as conflating what would be the equivalent of Kaplan's own notion of content with that of character. In Frege's theory, the notion of sense/thought plays both roles. But, for Kaplan, content corresponds to the thought, and is to be understood roughly as a *de re* or Russellian proposition. Cognitive significance should *not* be associated with content, but rather with character. Thus, cognitive dynamics as exemplified by Frege's remark should be understood as the study of the preservation of *the same content* (the same indexical thought) expressed by an *indexical sentence* through time, while *the characters change*.

For clarity's sake, let us take a look into the details of the way Kaplan introduced the cognitive dynamics-talk in the philosophical literature. Kaplan asks us to imagine the following situation. Assuming that, on a given day, a subject expresses the belief "It is a nice day today", Kaplan asks: What does it mean to say that she has retained her belief the day after? "It seems unsatisfactory to just believe the same content under any old character", he points out, because there seems to be no retention if the subject in question tries to re-express her belief through the same original character. In terms of Kaplan's theory, if the same character is employed in different contexts, it yields different contents (in the example, different thoughts) as semantic values. So maybe Frege is right in saying that there are standard adjustments to be made to the character for the same thought to be expressed given that the context has changed (e.g. when the time passes). And from that, it seems

---

[108] FREGE, G. 1918, p. 296.

right to conclude that preservation of the indexical thought through different (standardly adjusted) characters would be an indication of belief retention. Such adjustments would be, at least in principle, necessary for belief retention.[109]

This is the point where Kaplan introduces in a quick remark the case of Rip van Winkle and the problem it raises.[110] Those more or less familiar with U.S.-American literature (or with the problems of cognitive dynamics) will probably know that Van Winkle is a character in one of the most famous of Washington Irving's short stories. According to the story, one day Rip van Winkle goes hunting in the forest to run away from work at home. By the end of the afternoon, Rip falls asleep and sleeps for twenty years, during the whole of the American Revolution. But he does not realize he has slept for so long – he completely *loses track of time*. If we imagine that, on the day he went into the forest and fell asleep, Rip held the belief that "It is a nice day *today*", upon waking up he might try and re-express his belief by saying "It was a nice day *yesterday*", without realizing that in fact many years have passed. The problem is this: even though Rip lost track of time and made a mistake in the re-expression of his belief, our intuition is that he has *not* really *lost* his belief. Most of us would be inclined to say that Rip can still (somehow) remember the thought he had about that day when he fell asleep. But if, as Kaplan suggests, there are standard adjustments to be made in order to re-express indexical beliefs at different moments in time, we would be driven to the conclusion that someone like Rip, who loses track of time, is not in a position to retain any indexical beliefs. This is strange, as Kaplan himself puts it, in view of the intuitions we have concerning Rip's case. So, the question is: in a case such as Rip's, is there or not retention of attitude? This is how Kaplan unfolds what has since been taken to be the central problem of cognitive dynamics.

He goes on writing about the issue by introducing a further example. On this example, he sees his own reflection in a glass and is not able to tell if he is seeing his own image or someone else's:

> I first think, "His pants are on fire." I later realize, "I am he" and thus come to think "My pants are on fire." Still later, I decide that I was wrong in thinking "I am he" and conclude "His pants were on fire." If, in fact, I *am* he, have I *retained* my belief that my pants are on fire simply because I believe the same content, though under a different character? (I also deny that content under the former, but for change of

---

PUC-Rio - Certificação Digital Nº 1313515/CA

[109] Belief or whatever propositional attitude, of course. I will only talk of belief here and assume that whatever is said of belief generalize over other propositional attitudes.
[110] KAPLAN, D. 1977, p. 537-538.

tense, character.) When I first thought "My pants are on fire," a certain singular proposition, call it 'Eek', was the object of thought. At the later stage, both Eek and its negation are believed by me. In this sense, I still believe what I believed before, namely Eek. But this does not capture my sense of *retaining a belief*: a sense that I associate with saying that some people have a very rigid cognitive structure whereas others are very flexible. It is tempting to say that cognitive dynamics is concerned not with retention and change in what is believed, but with retention and change in the characters under which our beliefs are held. I think that this is basically correct. But it is not obvious to me what relation between a character under which a belief is held at one time and the set of characters under which beliefs are held at a later time would constitute retaining the original belief. Where indexicals are involved, for the reasons given below, we cannot simply require that the very same character still appear at the later time. Thus the problem of cognitive dynamics can be put like this: what does it mean to say of an individual who at one time sincerely asserted a sentence containing indexicals that at some later time he has (or has not) *changed his mind* with respect to his assertion? What sentence or sentences must he be willing to assert at the later time?[111]

There is much to consider when reading this passage, but I want to start with something very basic: the very meaning of the phrase "cognitive dynamics". In my opinion, Kaplan gives the name "cognitive dynamics" to the phenomenon in question because he understands it primarily as having to do with the dynamics of what he calls the "cognitive significance" of sentences containing indexicals. Cognitive significance is, on Kaplan's picture, connected to character. And the kind of activity performed on adjusting the characters in the expression of some indexical thought is closely related to the unfolding of the subject's cognitive state, considering that this state co-varies with the variations in character and cognitive significance of the linguistic expression of the thought in question. Because each character is associated with a certain cognitive significance (which in turn prompts a given cognitive state), cognitive dynamics should be seen as the problem of explaining how a given content and a given attitude towards it can (or cannot) be retained through different contexts under different characters. We know that Kaplan is particularly interested in indexicals, and his ultimate goal in "Demonstratives" is to account for the logic of indexicals.

We have seen that in *most* of the cases of indexical thoughts for the same thought to be re-expressed in different contexts, it must be re-expressed under different (appropriate) characters. Given that a certain cognitive significance is attached to each character, cognitive dynamics *in the case of indexical thoughts* must include a study of the *relations among characters*. It takes the form of a study

---

[111] *Ibid.*, p. 537-538, footnote 64.

of the kind of cognitive significance carried by those different characters and the adaptation of that cognitive significance in each new context. Explaining that should prove useful for understanding how a belief state, for instance, remains the same or unfolds or simply ceases to exist.

Thus, on Kaplan's way of understanding cognitive dynamics, it is closely related to *character* change, because Kaplan is concerned on the whole with the dynamics of *indexical* thoughts and attitudes. I think that judging by what Kaplan writes in the above quoted passage from "Demonstratives", he seems to believe that the main problem of cognitive dynamics has to do, above all, with the semantics of indexicals; more specifically, with cases of indexical thought *the linguistic expression of which must be adjusted if the same thought is to be re-expressed.*[112]

This can be seen as a straightforward (yet roughly construed) interpretation of what Kaplan says on cognitive dynamics in "Demonstratives". Kaplan's intentions when coining the phrase "cognitive dynamics" become conspicuous if one observes how, in the passage above, he explicitly claims that cognitive dynamics is concerned with retention and change in the characters under which some belief is held, not with belief change or retention as such. We know that there are linguistic rules that govern the way the same indexical thought must be presented (which characters must be used). According to Kaplan, we can identify relations between the way the a given indexical thought is expressed at one time and the ways it must be presented at some later time if the same thought is to be re-expressed. Given this, shall we say that those relations constitute *belief retention*? For Kaplan, it is the business of cognitive dynamics to answer that question.

In this perspective, the main problem of cognitive dynamics has to do with the unfolding over a period of time of the attitude of a given subject towards an indexical sentence to which she has previously assented and which also changes over time. So, the kind of indexical sentence at stake here seems very specific: sentences whose indexical components *must be* adjusted if we are to express the

---

[112] Clearly, and as I point out in what follows, Kaplan originally thought of cognitive dynamics in relation to indexicals, and indexicals only. He was particularly interested in indexical thoughts that can be subject to change over time. And that idea stuck. Later in the present chapter and subsequently in chapter 4, I will try to show that we can reasonably raise the question as to whether the phenomenon can be generalized beyond the scope of indexical thoughts. In the present chapter, suggestion will be made to the effect that dealing with the dynamics of first-person thoughts can already indicate a way of going beyond the seemingly narrow scope of how Kaplan has understood cognitive dynamics – as a problem arising only in the context of indexical thoughts and attitudes.

same thought. The examples entertained by Kaplan suggest that the type of indexical thought considered and shown to be problematic for a theory of propositional content is very specific. To explain its dynamics, it seems necessary to take into account the fact that the ways we have of expressing it change over time.

Take the example Kaplan gives in the long passage quoted. Seeing the image of a man with burning pants, Kaplan thinks, at a given time t_1, "His pants are on fire!". At a later time t_2, he realizes he is actually seeing his own reflection, and comes to think "My pants are on fire!". Still later, at t_3, he decides to go back to his original belief expressed by an utterance of "His pants are on fire!". If, unfortunately for Kaplan, he is indeed seeing his own reflection, has he retained his original belief just because he has retained the original thought-content – the thought that Kaplan is wearing burning pants? He actually expressed the same content (the same thought) under different characters, but he was not even aware of this fact. His re-expression of the same thought does not seem to capture the intuitive sense of retaining a belief towards it.

In the case of Rip van Winkle, he seems to have retained his original attitude – if nothing else, at least in the sense that he probably *remembers* it somehow. But Rip re-expresses it the wrong way: that is, he does not follow the rules that govern the way the indexical thought should be re-expressed. He does not *compensate for the change in perspective* that comes with context change. In Rip's case the rules would no longer be of any help – they would no longer even apply, because the time interval is too long and Rip completely loses track of it. Still, he seems to have *somehow* retained his belief (again, probably through memory).

I think Kaplan makes use of such examples to show that, when it comes to the *dynamics of indexical belief*, the mere preservation of the content of one's belief (a given thought or proposition) is not enough for us to say that there has been retention. At least not retention of *the same indexical belief*. It is not enough for it to be a case of belief retention that a subject to still remember, for example, his original belief, or even the thought in which she believed (understood as a Russellian or a *de re* proposition). Remembering such a content devoid of its indexical character (or under a different, inappropriate indexical character) won't do if we are trying to establish the conditions for retention of the same indexical belief. We saw in the previous chapter that it can be argued that indexicals are

essential (particularly, the first-person pronoun "I"). This means that they cannot be substituted by or reduced to any other form of thought. If this is right, maybe one of the consequences of Kaplan's line of thought is the following: if a subject who believes an indexical thought at one time is to retain that indexical belief and that indexical thought, she has to exercise a capacity to track not only the thought content understood as an eternal *de re* proposition, but also the kind of perspective provided by that very indexical, and that indexical only. She has to compensate for the possible loss in perspective that comes with contextual change.

When we see matters that way, we are in a better position to understand why indexical thinking is especially interesting to the cognitive dynamics theorist. In the case of indexical thoughts, the dynamics of our cognitive states is at the forefront, since they present an extra level of perspective. And if we are to say that there has been retention, the content of the thought must be preserved along with the (indexical) perspective from which it had originally been held. If we are to say that a given subject has really retained her indexical belief (towards an indexical thought), the very same indexical belief must be retained. In other words, the perspective must be tracked and adapted into the new contexts.

Perhaps it makes sense to talk about belief retention *simpliciter* – as the re-expression of a certain thought (content) under a different "guise", even a guise of a different nature (other than indexical, for instance). In what follows, I consider that possibility in connection with suggestions made by John Perry and others.[113] According to that kind of suggestion, in a case such as Rip's, there is obviously still belief retention in *some* sense: in the sense that a person like Rip can adapt an indexical character through memory into a non-indexical one, for instance. But again, by employing examples such as that of Rip van Winkle or of the man who sees his own image reflected in a glass, perhaps Kaplan's intention was to show that, in the case of indexical thought and indexical belief, it is not enough to preserve content. One must also update the indexical character under which that content was originally held. Following Frege, the thought (and, consequently, the belief) must remain indexical, and all the indexicals employed during the retention must keep on being those of a unique "family". Otherwise, either some *other* indexical content will be expressed, or some other *non-indexical* one.

---

[113] Especially in PERRY, J. 1997c. Also BRANQUINHO, J. 1999.

Maybe we can say that there are many ways to talk about thought/belief retention, and each of them can be regarded as shedding light on a particular feature of cognitive dynamics. I believe cognitive dynamics is a vast phenomenon and there are many aspects of it that we can explain – one of them is the way indexical sentences are adapted through changing contexts. In the next section I shall present and analyze other problems raised by cognitive dynamics, as well as some of the solutions given to them in the literature. Most of the authors who have attempted to give solutions to the problems raised by cognitive dynamics have also accepted Kaplan's perspective on the phenomenon and his formulation of the problem it raises. From that perspective, at least certain kinds of *de se* thoughts would not pose a problem in what concerns their dynamics, because in most cases the thoughts I express about myself are re-expressed over time with the same indexical "I".

2.2.3.
The Intuitive Criterion of Difference for thoughts

In the previous section we saw how Kaplan presented cognitive dynamics as the problem of accounting for retention of indexical thoughts and the attitude held towards those thoughts. At the same time, I tried to show that, according to a certain interpretation of the way Kaplan viewed the phenomenon, it is a matter of perspective-tracking, not only of linguistic adjustment, not even of content-retention. Another fundamental problem for the cognitive dynamics theorist is that of showing that the same thought can actually be grasped at different times. Of course, intuition tells us that indexical thoughts can be re-expressed through different sentences at different moments in time, which is not the case of non-indexical thought, since a non-indexical thought can be expressed through the same sentence regardless of context. So, once again, the problem of the dynamics of our thoughts seems particularly pressing in the case of indexical thoughts, due to the "flighty" character of their linguistic forms.

We saw in the previous section that Kaplan followed Frege in arguing that the re-expression of an indexical thought can happen through different appropriate characters (of the same "family" of indexicals, as Evans later put it). So, we know *how* a given indexical thought *should* be re-expressed in a different context, but the fundamental question remains as to whether that is even possible – if it is really the

same thought which is re-expressed or a different one. It seems we need a criterion of difference for thoughts to settle the issue.

Throughout his career, Frege himself established several criteria of identity for thoughts (not all of which are compatible with each other[114]). Among them, the (perhaps) most frequently used is the one referred to by Evans as "the Intuitive Criterion of Difference" for thoughts.[115] Frege employs it in "On Sense and Reference", for instance, but it appears in numerous other occasions throughout his work.[116] *Grosso modo*, the criterion tells us that we have different thoughts if it is possible for a rational individual, a person who has her cognitive faculties in perfect shape and is also fluent in a given language, to adopt incompatible attitudes in relation to the tokens of two given sentences of that language.[117] For example, if it is possible for such a person to believe in the first token and not in the second, or to believe in the first and refrain from assenting to the second. Here are Frege's own words in "On Sense and Reference":

> [...] the thought in the sentence "The morning star is a body illuminated by the Sun" differs from that in the sentence "The evening star is a body illuminated by the Sun". Anybody who did not know that the evening star is the morning star might hold the one thought to be true, the other false.[118]

So, according to the Intuitive Criterion, an individual entertains two different thoughts if it is possible for him/her to hold antagonistic attitudes to them. We have identical thoughts just in case, for example, it is impossible for a rational individual to ascribe falsity to one of its tokens and truth to the other.

Basically, as it has already been pointed out by Evans,[119] the problem with the Intuitive Criterion lies in the fact that it is limited in two ways. On the one hand, it cannot be applied to differentiate thoughts intersubjectively. On the other, the criterion does not work if applied to differentiate thoughts that an individual has *over a certain period of time* or, in other words, thoughts an individual grasps at

---

[114] To know more about this particular debate, the reader might want to start by checking this excellent paper by Carlo Penco: PENCO, C. 2003. "Frege: Two thesis, two senses". In *History and Philosophy of Logic*. Vol. 24, n. 2, p. 87-109.

[115] The criterion is discussed by Evans in the very first chapter of his *Varieties of Reference*. EVANS, G. 1982, particularly p. 18-22 and 32-33.

[116] This may have been the most recurrently used criterion of difference for thoughts in Frege's writings. *Ibid.*, p. 19.

[117] Mind that, in Frege's as well in Evans's examples: 1. The sentence-tokens are tokens of sentences of different types; 2. They contain *co-referential* tokens of singular terms (e.g., of distinct definite descriptions in Frege's example below); 3. They are (type-)identical in all other respects (that is, with respect to the predicate).

[118] FREGE, G. 1892, p. 162.

[119] EVANS, G. 1982, p. 21-22.

different moments within a certain period of time. The Fregean criterion needs to be either complemented or reformulated if it is to be used by a neo-Fregean cognitive dynamics theorist, since it has as a condition that the rational individual who holds the thoughts in question hold them *simultaneously*. The rationality requirement for the individual is *synchronic*, while the phenomenon of cognitive dynamics would need a *diachronic* criterion, since it has to do with thoughts held at different moments over certain periods of time.[120]

It may be helpful to give an example that clearly shows why the Intuitive Criterion in its original form cannot be an appropriate tool to tell whether (coreferential) thoughts held at different times are or aren't identical. The apprehension of thoughts is an epistemic fact, and Frege's Intuitive Criterion exploits that feature of senses, namely the epistemic facts involved in grasping them. What is used to differentiate them is the cognitive significance attached to them by whoever grasps them and is fluent in the language in which they are expressed.

If we try to use the Fregean intuitive criterion to individuate thoughts grasped at different times (that is, diachronically), we get an objectionable result. To use once again the commonest kind of example in the discussions on cognitive dynamics, imagine someone believes in what is expressed by a token of "It is a nice day today", produced a certain day, say on Saturday. This means she believes in the thought expressed by that sentence on that particular Saturday, maybe because the weather was fine, and she spent a great day with her girlfriend. But then, on the evening of that day, that person has a terrible fight with her girlfriend that ruins the good impression she had of her day. By the following day, Sunday, she has changed her mind about the thought she held the day before, and expresses her new, modified belief in uttering: "It was not a nice day yesterday". We know from our earlier discussion of Kaplan's understanding of the phenomenon that, for both Kaplan and Frege, the same indexical thought can be re-expressed in different linguistic forms. This is, of course, also a presupposition for the cognitive dynamics theorist, as stated above. So, a token of "It is a nice day today" on a given day, and a token of "It was a nice day yesterday" on the following day express one and the same thought (in accordance to semantic rules concerning indexicals). The criterion of difference

PUC-Rio - Certificação Digital Nº 1313515/CA

---

[120] See DOKIC, J. 1997, p. 4.

for thoughts should help us confirm that fact. But if it was applied to the example just given, in which the girl holds conflicting attitudes towards what we would normally take to be the same thought, it would yield an incorrect result. On Saturday she believed that which is expressed by an utterance of "It is a nice day today", then one day later, on Sunday, she changes her mind and does not believe that which is expressed by an utterance of "It was a nice day yesterday" – she comes to believe in the negation of that sentence. So she holds antagonistic attitudes towards the two expressed thoughts. If applied here, the criterion would tell us either that the thoughts are different or that there is a problem with the subject's rationality, because she has changed her mind.

Change of mind concerning the instances of two given sentences does not (necessarily) mean that the sentences express different thoughts, nor does it entail that the person in question is irrational. Still, if we tried to use Frege's criterion, that is exactly what it would tell us. The criterion only works when applied to thoughts held at the very same moment in time (synchronically), and fails to account for any sort of change *in the epistemic attitude* from the part of the thinker, if that attitude is held towards the same thought. If one accepts the criterion, change in attitude just means change in the *content* of the attitude (*i.e*, in thought-content). So, the way it is formulated, the Intuitive Criterion cannot be trusted to evaluate thoughts held over time. It seems impossible to reconcile it with the simple fact that people may change their minds in what concerns the same thought without becoming irrational because of that.

João Branquinho has suggested a diachronic version of Frege's criterion, or as he puts it, a diachronic extension of it which is supposed to help us account for the dynamics involved in the preservation and change of propositional attitudes.[121]

First of all, he argues that the diachronic version of the Intuitive Criterion he proposes is, as he puts it, "likely to rest on" notions applying to *tokens* of attitudes (such as, for example, real instances of belief retention) instead of *types* of

---

[121] BRANQUINHO, J. 2000. As far as I know, Branquinho's proposal is the only truly comprehensive attempt to provide a diachronic version of Frege's criterion, though some authors had already pointed toward the need of doing so in the context of an investigation of cognitive dynamics (or even outside it, given the natural limitations of the synchronic criterion). Evans was the first to emphasize the limitations of the original criterion, and we can also mention the already quoted DOKIC, J. 1997 (in his introduction); and in the same volume LUNTLEY, M. 1997 and HOERL, C. 1997.

attitudes.[122] Why? Because the objects of retention are *mental particulars*, not mental *types*. "Token attitudes will be concrete mental states in which subjects may be for certain periods of time, states that have particular propositions as their contents and belong to certain types",[123] Branquinho rightly points out. So, he makes use of the type-token distinction both in his reconstruction of Frege's original version of the Intuitive Criterion and in his proposed diachronic extension of it. Here is the criterion in its synchronic (that is, in its original) version, as reconstructed by Branquinho:[124] "a sufficient condition for propositions to be distinct is that it is possible for a rational subject to take token attitudes of antagonistic types towards them at the same time".[125]

Note that on that reconstruction, the synchronic character of the criterion is explicitly stated by the phrase "at the same time". Branquinho argues that the restriction of the criterion may be lifted if we are given an appropriate notion of retention of attitudes, such as the attitude of belief. So, in a way, developing a diachronic extension of the intuitive criterion is also part of Branquinho's own account of cognitive dynamics. Indeed, he also argues that special attention must be given to the notion of retaining a propositional attitude if we are to be successful in explaining the dynamics of our cognitive states.

He believes that attitude retention must satisfy at least two principles: 1. An absence of change of mind and 2. Preservation of information. Although these are relatively obvious observations, I think it is important to discuss each of the two principles in more detail.

According to Branquinho, when we think of cognitive dynamics and the cases in which a direct epistemic link with the object of the thought cannot be guaranteed (for example, in the case of thoughts about moments of time), memory should take

---

[122] BRANQUINHO, J. 2000, p. 4.

[123] *Ibid*., p. 1.

[124] In his paper, Branquinho also presents versions of the criteria in logical language. I will not present these formal versions here.

[125] *Ibid*., p. 5. The following issue is not essential at this point, but it is probably worthwhile to draw attention to what Branquinho writes about attitudes that, from his point of view, are to be taken as examples of antagonistic belief: 1. Failing to believe a certain thought; 2. Disbelieving a certain thought, which is usually construed as *believing the negation* of that thought; and 3. Suspending belief in a certain thought, which can be exemplified by a situation in which someone neither believes nor disbelieves the thought in question. *Ibid*., p. 3. I think the notion of failure of belief is a little obscure, and maybe this short list is not exhaustive, but I do not think we need to worry about this particular issue just now. It suffices to keep it in mind that the criterion establishes that the "subject take token attitudes of antagonistic types" towards the thoughts in question, what counts as antagonistic attitudes can be established at an immediatly later stage.

care of the second principle. More specifically, the relevant kind of memory here is the one that can be called propositional memory, but the sense of propositional memory which Branquinho wants to convey does not seem to be simply that of memory of propositions (such as remembering that *p*, where *p* can be substituted by a sentence expressing a thought). Branquinho seems to be concerned with propositional memory in the stricter sense of preservation of previously acquired information (supposedly in the form of propositions), in contrast to acquisition of new information.[126] The kind of proposition Branquinho has in mind is, I think, a Russellian one, possibly defined as an ordered pair of an individual and a property. In any case, this seems to cover the preservation of certain kinds of information, but not information of any kind. It should also be noted that the preservation through memory that happens in the case of attitude retention is, above all, preservation of the content of the attitude. So, the sole content needs to be preserved. In order to retain a given belief in some thought *p*, the subject does not need to remember that she believes *p*,[127] she just has to believe *p* itself.

In connection with the first principle, which states the obvious requirement that there must be no change of mind in attitude *retention*, Branquinho argues that the change of mind in question should not be conceived as a higher order state – for example, a belief state about a belief. If we were to endorse that view, we would be committed to the idea that a change of mind is always epistemically transparent to the subject. This is not to say that it *cannot* be transparent; just that we have to allow for the possibility of an individual's changing his mind without being conscious of it.[128]

Here is how Branquinho formulates his diachronic extension of the Intuitive Criterion: "a sufficient condition for propositions to be distinct is that it is possible for a rational subject to take, at (possibly) different times, conflicting attitudes towards them, provided that she retains at later times the attitudes previously held"[129].

---

[126] BRANQUINHO, J. 2000, p. 7-8. I think this point remains obscure in the paper.
[127] *Ibid.*, p. 8.
[128] *Ibid.*, p. 7.
[129] *Ibid.* p. 6.

Note that the fact that the attitude is retained from time *t* to time *t'* is essential if the criterion is to give us a reasonable result about the identity of thoughts held over time. According to Branquinho, we need to assume that:

> the attitude had by the subject at a certain time, say *t*, towards proposition *p has been retained* by her at the later time, say *t'*, at which she holds her attitude towards proposition *q*. In other words, one must suppose that her original token attitude persists throughout the period which goes from the instant *t* at which that attitude was taken to the instant *t'*.[130]

This looks awkward, because it seems, then, that even the diachronic criterion cannot account for change of mind *towards the same thought* – if change of mind happens, it just means we have different thoughts. Branquinho claims that "the persistence of an attitude involves a continued relation to a content, and any change of mind concerning that content would have the effect of interrupting such a relation".[131] But I think it is not necessary that change of mind happen that way. There need not be a radical interruption of the link between the subject and the content of his attitudes (the thought), even if these attitudes evolve over time. If a certain attitude is understood (like all propositional attitudes) as a relation between a subject and a thought, when the subject changes his mind, the original relation is, of course, interrupted in that it *ends*. But the overall relation between the subject and the thought can persist if we consider that the original attitude did not end in the sense of being lost and substituted by a different one, but rather in the sense that it evolved or *changed into* the new one.

Following Branquinho, if we are to evaluate the identity of two given propositions (*via* the sentences that express them), say *p* and *q*, "one needs to assume that the attitude had by the subject at a certain time, say *t*, towards proposition *p* has been retained by her at the later time, say *t'*, at which she holds her [possibly conflicting] attitude towards proposition *q*".[132] If the attitude held towards *q* at *t'* conflicts with the one held towards *p* at *t*, then *p* and *q* are different. We have seen how Branquinho accounts for the conditions for attitude retention, namely with the help of the two principles presented just a couple of paragraphs above: absence of change of mind (that is, retention of the same attitude from *t* until *t'*) and information preservation (that is, preservation of the *thought* in question: *p*). Branquinho himself seems to acknowledge the problem on which I am drawing

---

[130] *Ibid*. p. 5.
[131] *Ibid.*, p. 7.
[132] *Ibid*. p. 5.

attention here: the diachronic version of the Intuitive Criterion put forward here does not do a better job than its synchronic counterpart. Indeed, if certain suppositions are added to the synchronic criterion, the diachronic one could be derived from it. This is why the diachronic criterion is taken by Branquinho to be just an extension of the original one.[133]

So, it seems that Branquinho's diachronic extension of the Intuitive Criterion does not satisfyingly account for the identity of thoughts (continuously) grasped at different times by the same subject in the case of a subject who *changes her mind* about the thought in question. What Branquinho can say is that such cases are cases of irrationality, because a condition on the criterion is that we are dealing with a rational agent. Another possibility is that we have two thoughts. But again, is it really the case that change of mind would necessarily imply the existence of two thoughts? If there is a problem left here, it seems that we need either an independent (diachronic) or another form of extension of the original synchronic one. But there may be another way out.

Remember the question with which we started: can the thought expressed by a token of "It is a nice day today" be re-expressed the day after by a token of "It was a nice day yesterday"? Frege, Kaplan, Evans and other important authors share the opinion that the sentences do express the same thought. Branquinho's diachronic extension of the Intuitive Criterion would corroborate the opposing view. That is, in a case of change of mind towards a thought expressed *via* a token of "It is a beautiful day today" said in *d* and re-expressed in *d*+1 *via* a token of "It was a beautiful day yesterday", Branquinho's criterion would just tell us that the thoughts are different.[134]

Branquinho's diachronic extension cannot accommodate the possibility of a subject's changing her mind toward *the same thought* over a certain period of time – in such cases, he can say that either the subject is irrational or we have different thoughts, which seems to me to be a rather counterintuitive conclusion. I think we need to find a way of appropriately modifying Frege's criterion, or perhaps come up with an independent criterion of difference for thoughts held over time. But there may be another way of seeing the issue at hand.

---

[133] *Ibid.*, p. 9.
[134] *Ibid.*, p. 9-10.

The *fact* that thoughts *can* be re-expressed and grasped at different times is actually a *presupposition* made by anyone working with cognitive dynamics. If not, no problem is raised by the phenomenon. Actually, most authors dealing with cognitive dynamics seem to have followed Frege and Kaplan in assuming that the same thought can be re-expressed through different linguistic forms. Jérôme Dokic tells us that cognitive dynamics starts with two basic questions: First, is it possible to grasp the same sense at different times? Granted it is, the next question is *How* can someone grasp the same sense in a diachronic way or, in other words, in a continuous way?[135] Now, by addressing the second question, we will have an answer to the first. But, if give the former a negative answer, cognitive dynamics does not even get off the ground. In other words, we can simply focus on the problems cognitive dynamics raises, for as soon as we have a solution to them, we will also have the elements necessary to account for the fact that thoughts can be grasped at different times.[136] As I understand it, the main problem of cognitive dynamics is actually that of accounting for attitude retention and, more generally, for our capacity to continuously grasp a given thought through time. The possibility of multiple expressions of the same thought at different times is a necessary presupposition for that kind of investigation.

Even though it seems clear that there are some epistemic constraints on the grasping of an indexical thought, nothing forces us to claim that it is impossible for that thought, once it has been grasped, to be retained and re-expressed at a later time, even though the epistemic conditions for its grasping change over time. Perhaps it is not necessary to be in exactly the same epistemic position for the thought to be retained. Following an idea taken from Evans (to be discussed in more detail in the next section), it might be that we have to compensate the differences in perspective that come with context change through the employment of a specific kind of mental activity. That is in tune with my understanding of Kaplan's view of the issue of cognitive dynamics as being a matter of perspective-tracking. Perhaps that would allow us to say that in practice the same unfolded thought is preserved. The whole business of cognitive dynamics hinges upon our capacity to *continuously grasp thoughts* over certain periods of time (or as we move through space), and to

---

[135] DOKIC, J. 2001, p. 72.
[136] *Ibid.*

hold attitudes towards certain thoughts during certain periods of time (or as we move from one place to another).

2.2.4.
The main accounts of cognitive dynamics

Ever since Kaplan's "Demonstratives", the general consensus is that the kind of thought which is problematic for an account of cognitive dynamics is indexical. If we understand cognitive dynamics that way, the so-called "eternal" or simply "non-indexical thoughts" do not raise a serious problem for the cognitive dynamics theorist, if they raise a problem at all. To take an example of a non-indexical thought from Kaplan himself, consider the one expressed by the sentence: "All persons alive in 1977 will have died by 2077". As Kaplan puts it, the thought expressed by that sentence can be true or false depending on what he calls "the circumstances of evaluation" (the possible situations with respect to which the thought is evaluated as true or false), but its *meaning* (its *linguistic character*) does not vary according to the context in which it is used. That meaning is not sensitive to context, though its truth-value is sensitive to the circumstances of evaluation. In that sense, and to use Kaplan's terminology, non-indexical sentences have a *fixed character*, in contrast with the *context-sensitive character* of indexicals. This means that non-indexical sentences express the same content (the same thought) in all contexts, or in other words, no matter in what context they are used.[137]

To see how this relates to cognitive dynamics and how, approaching the phenomenon from Kaplan's perspective, it is uninteresting to investigate it in the case of non-indexical thoughts, we just have to imagine the following situation. If someone believes, at a certain time t_1, in the thought expressed by an utterance of "All persons alive in 1977 will have died by 2077", and then continues to believe it until some later time t_2, that person will not have a problem re-expressing her retained belief with another token of the very same sentence. She will, at t_2, just re-express the same thought by uttering again: "All persons alive in 1977 will have died by 2077". With this, we see how the problem of cognitive dynamics is, for Kaplan, tied to the expression of indexical thoughts, and indexical thoughts only. From Kaplan's perspective, if there is no problem with the expression and re-

---

[137] The example is taken from: KAPLAN, D. 1977, p. 506.

expression of the thought (for there is no need to adjust the character), there seems to be no problem in what concerns cognitive dynamics. So, in dealing with cognitive dynamics, our attention should be focused on indexical thoughts. But again, not any kind of indexical thought: only indexical thoughts the expression of which must be adjusted through changing contexts.

This section is dedicated to assessing some of the most important accounts of cognitive dynamics to be found in the literature after Kaplan's "Demonstratives". For now, I shall focus on ideas put forward by Perry and Evans who hold what I take to be the two most important opposing views on the problems formulated by Kaplan. My aim is to show that the restriction imposed by Kaplan on the range of thoughts worth being investigated by the cognitive dynamics theorist (namely, indexical thoughts the retention of which requires some adjustment over time) can be reasonably lifted, at least up to a point.

Perry argues that changing one's mind must mean coming to take a contradicting attitude towards a thought to which one had been previously related in taking a certain attitude to it. It is in relation to the thought itself that we change our minds – as Perry understands them, thoughts are composed of the objects denoted in a direct way.[138] In this, he agrees with Kaplan, and he also draws inspiration from Kaplan's notions of content, character, and context and of their relations.

What is Perry's answer to the issue raised by Kaplan? First, he re-interprets the Kaplanian notion of character using his own notion of *linguistic* or *utterance-relative role*. The expression's linguistic role is often tied to other (epistemic, pragmatic) roles an indexically referred object can play in our lives, although sometimes all it is available to us to think about some object is the linguistic role such object can play. If we want to understand Perry's notion of linguistic role, the key notion is that of an utterance. The expression's character is thought of by Perry in terms of the role played by the *denotatum* relative to the utterance (e.g., the role of speaker in the case of the indexical "I"). As Korta and Perry put it, "the role provides an identifying condition of the referent, but one that is utterance-bound".[139] Kaplan himself did not put forward a theory of utterances, but rather a theory of "sentences in context", because the latter was better suited for his main

---

[138] PERRY, J. 1977, p. 496.
[139] See KORTA, K.; PERRY, J. 2011, particularly p. 22-23.

goal in "Demonstratives", i.e., the development of a logic of indexicals. "Utterances take time, so the premises of a spoken argument won't all occur at the same time, but for the purposes of logic we want them to occur all in the same context", as Korta and Perry put.[140] But thinking of (Kaplanian) characters in the terms of (utterance-relative) roles is desirable if our goal is to develop a theory of meaning from the point of view of pragmatics. Then the notion of utterance is central. Note that the split between the notions of role (or sense, or Kaplanian character) and thought (or content, in Kaplan's terminology) help us make sense of the fact that if someone accepts a given linguistic role at one time, and its negation at another, we must not be inclined to say that such a person has necessarily changed her mind.[141]

Perry argues that if we want to solve the problem of cognitive dynamics such as it has been formulated by Kaplan, we have to distinguish between linguistic characters or roles from doxastic ones.

Indeed, Kaplan speaks of the need of adapting the linguistic expression of a sentence for there to be retention of the cognitive state in the case of an indexical thought. But it seems that adjusting the linguistic character is not enough. When belief retention occurs, dynamic mechanisms other than the adjustment of the linguistic expression of belief must be at work. Kaplan does not draw a clear distinction between linguistic characters, on the one hand, and mental or psychological characters, on the other hand. Both can be understood as ways of presenting that undergo ajustments as time passes by, so that the thought held and the belief had may remain the same.[142] Perry takes a step in that direction when he urges us to distinguish between doxastic and linguistic characters.

Perry argues that instead of taking belief to be a simple attitude towards a sentence, we should understand them as *mental structures* which draw (information) from various kinds of epistemic relations (such as perception) and have propositions as contents. On his view, beliefs have 1. *Contents* – thought of as propositions because a belief is a *belief that* such and such is the case. And they also have 2. *Causal roles*: they are *caused* by certain kinds of perception and they *cause* certain actions.[143] These two elements should mesh. As Perry's argument

---

[140] *Ibid.*, p. 23.
[141] PERRY, J. 1977, p. 495.
[142] *Cf.* SOUTIF, L. 2014, p. 16-17.
[143] PERRY, J. 1997c, p. 19.

goes, each instance of any kind of epistemic state (not only belief) has its own specific causes (pertaining to that instance), but there are patterns – general combinations of factors that bring about the state (as well as causing other things as a consequence of the state). These patterns, Perry argues, are the causal roles of the states.[144]

Now, the examples Kaplan presents us with in "Demonstratives" show that the meshing of those two elements is *not* as simple and obvious as one might think it is. Perry explains this by pointing out that the typical causes of the belief state held by Kaplan when he uttered the sentence "His pants are on fire" are quite different from the typical causes of the belief state expressed by "My pants are on fire", though both express the same proposition. So, the two different belief states do not line up directly with their content – the proposition they are about. This suggests that a person's being in a certain belief state does not necessarily entail she believes some proposition we might expect her to believe. As Perry points out, many people can be in the same belief state while the contents of those states differ.

Perry argues that belief states have characters that are *independent* of language. The meaning of beliefs is analogous, but not *derived from* the meaning of sentences. Besides, the linguistic characters or roles of beliefs must be different from the characters or roles of sentences/linguistic terms. While linguistic characters are utterance-relative, belief characters are *thought-relative*, so Perry suggests we call them *cognitive* roles. "These roles are based on relations that an object can have to a given episode of thought or a particular belief, such as being the owner of the thought (the self role)", he writes.[145]

From Perry's point of view, Kaplan finds himself in two different belief states: the first when he thinks that some person other than himself has her pants on fire, and then when he realizes that his own pants are on fire. He acts differently in each of those cases, and the sentences he may utter are different too – so both his actions and the sentences he utters *reflect a difference in beliefs*.

> The concept of believing under a character is intended to capture this difference. The difference in belief would be there, even if Kaplan didn't say anything. And *the difference between the two cases is not the proposition that is believed*, but the character under which it is believed.[146]

---

[144] *Ibid.*, p. 20.
[145] *Ibid.*, p. 21.
[146] *Ibid.*, p. 23. Emphasis mine.

From this passage, we can understand that Perry argues for a clear split between the proposition and the belief state, so that the two levels are completely independent. Does that mean that to be in a certain belief state does not seem to involve being disposed to assert a certain proposition? I believe it does involve being disposed to produce certain linguistic constructions, but perhaps not necessarily following the rules of the proper re-expression of a given thought, for example.

To solve the problem of cognitive dynamics, Perry appeals to his notion of information games. He writes:

> An information game involves the acquisition and later application of a belief about an individual. That is, at some time one comes to believe something about some person or object. Then, later, that belief guides one's behavior towards that object or at least an object that one takes to be the same as it.[147]

Information games thus help us understand how beliefs guide our actions. Perry describes eight information games in total. The first and simplest one he calls the "straight-through" game. In this information game, recognizing the source of the belief is not difficult, since there is no interruption of the epistemic link one has with it. The source of one's belief continues connected to the subject. As the name of the game suggests, the identification is made in a direct way. If the source of the belief is a person, she can answer to questions, for instance.

Things get a little more complicated with second kind of information game, which is called "tracking". In this case, recognizing the source of the belief is more complicated because one needs to exercise an ability of keeping track of it as it moves through space. The task is still fairly easy because it does not involve re-identification. However, the subject of the belief does have to exercise her ability of making sure that a single object or individual has (possibly) played various different roles for her (as it has moved in space).

The third kind of information game is called "detach-and-recognize". It involves re-identifying the source of one's belief, and there are at least three steps in the process. First, the subject gets acquainted with the object or individual in question; then a certain belief about it/her is formed; after that, the subject spends a certain amount of time without any contact with the source of the belief (during this time, the epistemic link is obviously broken, which does not happen in the two

---

[147] *Ibid.*, p. 24.

previous games); and finally, at a later time, the subject re-encounters the source of her belief and recognizes it. This game involves an important notion: memory. While the subject does not have an epistemic access to the source of the belief, the source can still play a role to the subject through memory. Through memory the subject has a way of thinking about the object – and that way of thinking has a very loyal character, Perry argues.[148] For him, the characters we access through memory are loyal ones. Curiously, Perry speaks at this point of "(mental) files". He argues that for it to be useful, a "detached belief" must be part of a larger file about the object of one's belief. For instance, in case the object of the belief is a person, additional elements about that person to be found in the file provide facts about her that can enable the subject of the belief to recognize her later on.[149]

Now, if a person acquired a belief while having the source of her belief directly available to her, under what kind of character will she retain her belief once the epistemic link is broken? Perry's answer is this: under what we might call a "memory demonstrative". As an example, suppose I have been introduced to a man and learned various things about him. Perry's suggestion is that later on I can think of him through a demonstrative, as in "That man (the source of my belief) was F (that is, any property I ascribe to him) then (at the time the belief was acquired)". "A belief like this one, totally useless at one time [the time during which I have no way of re-accessing the source of the belief], may become useful later", writes Perry[150]. This gives us a hint about Perry's answer to the riddle posed by the case of Rip van Winkle. Detached beliefs are important for all the five remaining information games, though I shall not go into the details of each of them, because they are less urgent when trying to re-construct Perry's answer to the riddle posed by the case of Rip van Winkle. Perry calls the other games "recollection", "inference", "updating", "communication" and "planning".

Perry compares how we usually keep track of our own experiences relative to time and space with how we specifically keep track of particular points in space and time. It seems easier and more natural to keep "track of whether our experiences lie in the past, the present or the future. This is fairly easy because there is an exact correspondence between the mode of thinking about the experience (remembering

---

[148] *Ibid.*, p. 27.
[149] *Ibid.*, p. 28.
[150] *Ibid.*

it or planning to have it or having it) and its position in time relative to the present moment".[151] So, if I had an experience in the past, the mode of thinking about it will be through memory. If I have the prospect of an experience in the future, I usually think about it by planning it. However, argues Perry, the keeping track of those events is usually not very precise. If I have an appointment tomorrow, I do not usually retain my belief about that appointment by constantly updating the hours our minutes remaining.

As Perry points out, when thinking about some past time, we are no longer pragmatically or epistemically attached to it. We are no longer in a position to explore that time, though we can explore our own memories, maybe forming new beliefs as a result.

Something similar holds for the future: we cannot change anything in a future time; all we can do is change things in the present hoping they will have an effect on that future time (I can, for example, now plan a future event). Neither can we now perceive events in the future. Perry writes: "Can next July 4th be a source of my thoughts? No, for it lies in the future [...] But I can have a *sourceless* detached belief about next July 4th".[152]

At this point, Perry makes a very important remark that sets him apart from Evans. While Evans has an account of keeping track of time drawing on an analogy with keeping track of an object or individual in space, Perry contends that an important distinction has to be drawn between the two cases.

> To keep track of time is usually not to pick out a day and track it, but to be aware of which day it is. That is, to be aware of the important properties of the day that plays the "today" role – what the date is, what events are planned, etc. To lose track of what day it is, means not knowing that the day that plays that role has some other important attributes, like being one's anniversary or the day a philosophy paper was promised to an editor.[153]

Perry concludes that with days (and, more generally, moments in time) we can play the straight-through and the tracking information games, but we cannot play the detach-and-recognize game, because we cannot experience a day again. However, we can have sourceless beliefs about them due to their special metaphysical status, which is reflected in our system of dates.[154]

---

[151] *Ibid.*, p. 33.
[152] *Ibid.*, p. 34.
[153] *Ibid.*, p. 35.
[154] *Ibid.*

Now what does Perry have to say about Rip van Winkle? The day Rip fell asleep, the held a thought about that day under the character "Today [the day of his thought] is nice". Then, he sleeps for many years without realizing it. Upon waking up, he thinks he has slept for just one night. Here is how Perry explains the situation: when Rip falls asleep, he loses the epistemic contact he had with that day, but he has ready-made *back-up characters* in his mind to keep on thinking about that day – so, he can just use one of those back-up characters and hold his belief under it. In that case, the following character is available to sustain his belief: "That day [the day I remember] is or was a nice day".

Perry accounts for the case of Rip van Winkle by saying that although Rip tries to express his belief the day he woke up and fails because he expresses it with the wrong character: "Yesterday [the day before the day of this thought]", he still retains his original belief under *various other back-up characters*.

Perry describes what he thinks is the way Kaplan understands the problem in question to be: "retaining belief consists in moving from flighty character to flighty character in ways that reflect change in context". With his idea on back-up characters and his information games he thinks he has refuted the Kaplanian conception of belief retention.[155]

There is a clear sense in which Rip seems to remember *something* of his original belief, as Perry suggests. João Branquinho defends another version of that idea, in these terms: what Rip retains, by means of memory and not of his particular position in space-time, is the same *Russellian* proposition, a proposition which corresponds to an ordered pair consisting of at least an object (in that case, the day he fell asleep) and a property (that is, the property being beautiful).[156] Rip does not seem to be able to retain the indexical thought. So, the suggested account is basically this: there is a transformation of the original indexical thought into that kind of more "neutral" (in the sense of "de-indexicalized"), so to speak, "Russellian proposition". The transformation could also be from the original indexical thought to a non-indexical thought, to another kind of indexical thought or to any such "back-up character", of course. Interestingly, this is roughly what we find in the following example, also taken from a work of fiction, this time a short story by the famous science-fiction writer Philip K. Dick. In the story, a man called Jennings

---

[155] *Ibid.*, p. 37.
[156] *Cf.* BRANQUINHO, J. 1999.

has the memory of the last two years of his life wiped clean by the corporation he had been working for. Here is how Dick introduces Jennings's situation:

> "Where are we?" Jennings shook his head, trying to clear the dull ache. "Or maybe I should ask that a different way." Already, he could see that it was not late fall. It was spring. Below the cruiser the fields were green. The last thing he remembered was stepping into an elevator with Rethrick. And it was late fall. And in New York. "Yes," Rethrick said. "It's almost two years later" [...]
> "Where are we going?"
> "Back to the New York Office. Where you first met me. Remember? You probably remember it better than I. After all, it was just a day or so ago for you."[157]

Note that, although set in a science fiction scenario, the kind of thing Jennings's companion Rethrick says sounds perfectly reasonable. Jennings does remember the thoughts he had two years ago, and being aware of his memory loss (unlike Rip van Winkle), he is even able to locate those thoughts in time, though the *continuous* mnemonic link to them has been severed (like in Rip's case). Indeed, he is probably even able to remember his thoughts more clearly than his companion Rethrick, who has consciously kept track of time for the two years of Jennings's induced amnesia. So, he does remember his thought *somehow*, in spite of the memory gap, and maybe in a way that is similar to Branquinho's or Perry's account. But if he had a "today-thought" moments before the period of his memory erasure, it is fair to say he would not be able to remember *that indexical* thought. That particular kind of informative value which an indexical thought has (a kind of information that has to do with perspective) is lost. And we can probably argue that the dynamics of his thought is indeed radically changed, because the continuity that links his present thoughts to those he had two years before has been broken. If his attitude of belief also had a perspectival quality (because it was directed towards an indexical thought), then arguably that *indexical* belief is also lost. As we shall see in what follows, Evans would be ready to argue that neither Rip from Irving's short story, nor Jennings from the example retained their original beliefs.

Still, I think we could rightly say that there is belief retention in a case such as Rip's, through retention of the content of that thought. At the same time, we could say, also rightly, that there is *no* retention in that case, since the complete indexical thought *cannot be retained* by someone who has lost the original perspective she had on the object. (Such perspective should be understood as being

---

[157] DICK, P. K. *Paycheck and other classic stories by Philip K. Dick*. Introduction by Roger Zelazny, foreword by Steven Owen Godersky. Citadel, 2003.

extended in time, of course.) I think it is possible to talk of different *levels* of attitude/thought retention. Suppose all a given subject can remember is her original thought *content* (a *de re* proposition, one might call it) – in such a case, there is already some level of retention. Perhaps the next step or next level of thought retention would correspond to remembering the content of one's thought as well as the perspective from which that thought was held (its original indexical character). Yet another level of retention could be exemplified in one's capacity to retrace, if asked, the steps in the evolution of one's thought (from one's present state back to the original perspective held). So, maybe a truly complete case of retention of an indexical thought would be characterized by such a capacity of retracing the evolution of one's thought given that no mental abstraction has been used to perform such recollection and that the disposition to appropriately express that evolution in language exists. But then again, maybe in everyday life we can already (correctly) talk of thought retention if all one remembers is a *de re* thought detached from its original *de se* or indexical character.

And in some situations (probably in most situations of everyday life), it could be that if someone like Rip van Winkle were to regain her lost perspective and were able to *recognize* the object of which she originally thought in an indexical way, that would be enough to say that there has been thought retention – even if in a weaker, more flexible sense. But we can be more demanding in our examination of those cases and say that, since there has been some discontinuity in the "connection" the subject had to that (indexical) thought, the existence of that break by itself suffices to say there has not been retention. This is more or less the path Evans choses to take.

Evans sides with Kaplan in taking Frege's intuition to be right when he wrote in "The Thought" that if a subject wants to re-express an indexical thought in different contexts, she must substitute the indexicals appropriately. But for there to be attitude retention, the appropriate substitution of the indexical terms, or at least the inclination to substitute them, when re-expressing the thought is not enough. Evans argues that it is also necessary to keep track of the very object of one's thought, in analogy with keeping track of a moving object in our field of vision. If a person happens to be lost in time and does not realize she has moved from one context to the other, she will automatically fail to live up to that cognitive constraint, an error in appropriately substituting the indexicals is likely to follow, and,

following Evans, we would be led to say that such a person is not capable of grasping indexical thoughts (or at least that indexical thought) through different contexts (through time, for example). Evans's way of dealing with the case of Rip van Winkle is opposed to Perry's: Rip is not capable of retaining his indexical thought/belief because he has lost track of time. The temporal continuity (be it conscious or not) of his thought has been broken since his awareness of the passage of time is blurred, and that awareness is necessary if we are to say that a temporal indexical thought has being retained.

We immediately see that, like Perry, Evans tries to find conditions for belief retention other than the proper substitution of the linguistic characters, for the latter is not enough to explain cases such as Rip's. But when he talks about the need to exercize a mental ability of tracking the object, he gets to conclusions very different from those of Perry. We saw that the latter speaks of doxastic and back-up characters that can be employed in belief retention. He rejects Evans's conclusion that Rip did not retain his belief in favor of saying that Rip's belief could have been retained under a different "guise" – other than the ones we would normally expect. (In fact, I alluded to that possibility in the end of section 2.2.1.)[158] Evans, for his part, defends an idea that is not diametrically opposed to that of Kaplan – I think we can see Evans's proposal as a necessary expansion of Kaplan's. If we follow Evans, the ability exercised by the subject when having an indexical thought should not be understood in a *static* way – for example, as the simple act of thinking about an object at a given time t, and at t only, abstracting from the passage of time. Such an act should be literally seen as an abstraction of a more general capacity of keeping track of the object from time t_1 to t_2. Thinking about the object at t_2 also depends on the thought held earlier, and the same is true for many tokens of a given

---

[158] Now, I talk about Kaplan and Evans's conclusion as being "natural", but it could be easily argued that their conclusion is not *that* natural after all. On the contrary, to claim like Evans that Rip has not retained his belief because he has lost track of time could be taken as rather counterintuitive. Why should retention of belief, (which is a cognitive phenomenon) depend on external and epistemically opaque facts such as *the fact* that the token of "today" (on the day Rip falls asleep) and the token of "yersterday" (on the day he wakes up) *do not* corefer? Can't he be mistaken about the contexts of utterance while retaining his belief from one occasion to another? Well, yes, I think he can (and in that respect I tend to side with an idea we can find in Perry). But the point is not so much that belief retention depends on external, epistemically opaque linguistic facts, for example. Those external factors notwithstanding, I think Evans's conclusion is "natural" in the sense that proper retention of an *indexical thought* involves the retention of the indexical perspective the thinker originally had. And for that, a very specific kind of cognitive activity must be exercised. When Rip loses track of time, his original indexical perspective is also lost. I discuss Evans's position in more detail below.

indexical thought. When a subject has a thought in this manner, Evans argues we can call it a "dynamic Fregean thought".

Evans's idea is that, when we consider thought/belief retention, at least at a certain level of our explanation we can understand each token of a given indexical thought (expressed, for instance, through sentences containing such words as "today", "yesterday", etc.) as a single continuous indexical thought. We can take it to be the same overall way of thinking about the object: a dynamic way of thinking about it. Frege uses the image of "modes of presentation" of the reference to explain sense. Evans reinterprets modes of presentation as "ways of thinking" about the reference, and so goes on to argue that:

> […] the *way of thinking about an object* to which the general Fregean conception of sense directs us is, in the case of a dynamic Fregean thought, a *way of keeping track of an object*. This permits us to say, after all, that a subject on d_2 is thinking of d_1 in the same way as he did on d_1, despite lower-level differences, because the thought-episodes on the two days both depend upon the same exercise of the capacity to keep track of a time.[159]

The idea is literally expressed when he writes that in the case of someone who has an indexical temporal thought about a day, "there is some level of description at which he is thinking of the same day in different ways".[160] The difference could be exploited to show how different attitudes could be held by a given subject towards tokens of the sentences through which the thought is expressed at different times. And that would "preclude their being the same thought", as Evans puts it. Indeed, this seems to be a natural consequence in a Fregean theoretical framework, and we can reasonably argue for it.[161] But following Evans's own line of thought, upon closer examination, the consequence is unwarranted. For him, behind the argument lies an unjustified use of Frege's Intuitive Criterion of difference for thoughts: unjustified because the criterion cannot be applied to thoughts held diachronically (as we saw in the previous section). Evans does not propose a diachronic version of Frege's criterion, but rather seeks to avoid the problem through his own theory.

Evans argues in favor of a theory of thought that we could label "holist", and which he himself associates with Frege,[162] in opposition to what he calls the

---

[159] EVANS, G. 1982, p. 196. Further justification for his opinion on the case of Rip van Winkle is given in a footnote on the same page.
[160] *Ibid.*, p. 193.
[161] *Cf.* PROSSER, S. 2005, particularly his arguments in section 4.
[162] *Cf.* SOUTIF, L. 2014, p. 20.

"atomist conception" of (indexical) thought.[163] Evans argues that, from the perspective of the atomist, indexical thoughts can only be apprehended once – in the initial context of apprehension. They cannot be retained through time and are not susceptible of being grasped again at any time following that of the original apprehension (say, under some other character). Of course, there are relations between certain indexicals. But according to the atomist conception of thought, when a subject expresses a belief on *d* through an indexical sentence such as "It is a nice day today", even though there are relations between the indexical "today" and the indexical "yesterday" (among others), the original thought can only be grasped in that form on that day – on *d*. On *d* + 1, I have at my disposal a mode of presentation of the day which is different from that which I had on *d*, even if I express my belief by uttering "It *was* a nice day *yesterday*".

Against such an idea, Evans defends that "one's thought *at* a time is dependent upon an ability which is necessarily manifested only *over* time".[164] Instead of conceiving of each token of a given indexical thought (as it is expressed with terms such as "today", "yesterday" and so on) as independent of one another (even if possibly related), Evans suggests that the only way an individual can even understand each instance is through her capacity of thinking all of them (both in the past and in the future). In Evans's own words:

> No one can be ascribed at *t* a belief with the content 'It is now *Φ*', for example, who does not have the propensity, as time goes on, to form beliefs with the content 'It was *Φ* just a moment ago', 'It was *Φ* earlier this morning', 'It was *Φ* this morning', 'It was *Φ* yesterday morning', etc.[165]

Evans suggests that the properly Fregean view on cognitive dynamics would be to defend that beliefs persist over time despite local differences that the changing circumstances impose on the agent of the propositional attitude. "Our ability to think of a place as 'here' is dependent upon our general ability to keep track of places as we move about",[166] he writes. And for that we also need to have the ability to "know when we are moving". This necessary psychological ability is central to Evans, who believes that for us to have thoughts like "It is ξ here", we should be able to think "It is ξ there", and so on.

---

[163] For the whole argument *cf.* EVANS, G. 1982, p. 194-196.
[164] *Ibid.*, p. 195.
[165] *Ibid.*, p. 194.
[166] *Idem*, 1981, p. 308.

Evans seems to think that Kaplan chose to refuse such a view for taking it to be too problematic, though the view could be defended, given good arguments. For Evans, "a capacity to keep track of the passage of time is not an optional addtion to, but a precondition of, temporal thought".[167] The existence of such a capacity, if it is working properly, would be a replacement for the synchronic Fregean Criterion.

Now, we can denounce Evans for failing to explain all the facts of the problem formulated by Kaplan,[168] since his answer to the case of Rip van Winkle betrays our intuition that Rip is still capable of retaining his original belief *somehow*. Evans's thesis is not able to accommodate that fact – it may seem counterintuitive to say that belief retention depends on external, epistemically opaque facts such as the fact that the word "today" uttered on the day Rip falls asleep and the word "yersterday" employed on the day he wakes up *do not* corefer. (See footnote 157 above.) In spite of that, I think it is reasonable to follow Evans in saying that, in a way at least, Rip does not retain his indexical thought. Evans's account makes sense: there must be continuity of the indexical aspect of the thought for the same indexical thought to be retained. That continuity does not seem to allow for a situation in which the subject loses track of the object of his thought and its whereabouts (in the case of a day or a moment in time, its location in the course of time). At the same time, perhaps we could somehow incorporate in an Evansian theory something similar to Perry's idea of "back-up characters".

The problem of cognitive dynamics formulated by Kaplan does not seem to be that of accounting for the retention of the *contents* of the thoughts in which we believe, at least not in the sense of *de re* or Russellian propositions. So understood, thought contents can be expressed through different characters (or roles, or "guises"). Indeed, the very fact that the same thought content can be re-accessed at different times (and thus in different ways) is a presupposition of cognitive dynamics. We learned with Kaplan's examples that preservation of the same thought content is not enough to characterize attitude retention.

Now, if the problem of cognitive dynamics is really that of accounting for the retention of our *belief states* (or any other kind of epistemic state we can find ourselves in), it seems we would have to presuppose an explanation of what a belief state is, similar to the one Perry articulates. Perry seems to think that the main way

---

[167] *Idem*, 1982, p. 194.
[168] As did SOUTIF, L. 2014, p. 21.

to characterize a belief state is through the actions it prompts, that is, through its impact on our behavior. What changes through time is the *way* the proposition in question is believed (that is, the character), not the proposition itself. Changes in the ways we believe certain propositions have the power of affecting our behavior. If this is what cognitive dynamics is, then Perry's account of it seems to be on the right track, at least in what concerns the intuitions behind it.

But the task of the cognitive dynamics theorist is not only that of accounting for the possibility of re-expression of the same proposition, but of retention of a certain attitude towards a certain thought. That thought or mental state, I take it, should be seen as the proposition along with the character or role, both *linguistic* and *mental*. And I think Perry is right in saying that all of those elements do not always *mesh*. Such as it has been understood by Evans (and as it should have been understood by Kaplan), only the perfect harmony of all of those elements as time passes would constitute the "perfect" retention.

But when it comes to retaining temporal indexical thoughts, for example, since we are no longer epistemically linked to the source of our thought, the way we think of it necessarily changes, even if slightly. Both Kaplan and Evans want to allow for the possibility of thinking about something under a "family of characters" – that amounts to thinking about that thing the same way, under a single character that evolves through time. Based on Frege's intuition, Kaplan seems to be willing to allow for that possibility at first, but then, in face of Rip van Winkle's case, he starts eyeing it with suspicion. Evans, on the other hand, fully embraces it. This seems to me to be a good idea, because I think it correctly reflects what happens on all levels: the world, our minds and language. Things change and evolve in the world; our relations to them and the ways we think about them also change; and all of that is reflected in language. In particular, families of indexicals such as "the day before yesterday, yesterday, today, etc." reflect the way we can sometimes preserve our perspective towards something in an evolving reality. In the spirit of Evans, things change, and even our perspectives change, but somehow it remains the same, because we can *track* those changes and make our own perspectives keep up with them.

Is the phenomenon of cognitive dynamics multilevelled? I think it is. And for that reason, we can look for an account of a full-blown dynamics, scenarios in which everything is properly preserved (the proposition, the way we think about it and our

attitudes) and scenarios in which there is preservation at one level, but not in the other(s). Can we talk of weaker and stronger forms of belief or thought retention when investigating cognitive dynamics? If we can, anyone working on what we call cognitive dynamics would have as his task to account for all of its forms. What I want to say is this: Perry has a certain way of understanding the problem and Evans has another, but perhaps in the end both are right – their accounts can be used to explain different *facets* of the more general phenomenon of the dynamics of our mental states and our thinking.

I tend to agree with an intuition to be found in Perry, for example:[169] there is a strict sense in which one speaks of "belief retention" and that sense is rarely useful in our everyday cognitive experiences. Our capacity to recognize objects and previously held concepts or ideas is essential to us, and we lose track of things all the time, but in many situations. I also agree with Kaplan's intuition, whose ultimate consequences Evans tried to explore. I think each of these alternatives can be used to explain some cases in contrast to others, depending on our purposes.

Perhaps the constraint imposed on thought retention depends on *the kind of retention* one is trying to pinpoint, as well as the role that such a retention is meant to play. Perhaps in certain situations our account of thought retention will require us to say that several constraints must be met if we are to say that there has actually been retention. But we can also be more flexible, and I think most of the time, when explaining thought retention in *everyday life,* we are. This is probably what Kaplan meant on claiming that the sense of belief retention he was interested in is a sense associated "with saying that some people have a very rigid cognitive structure whereas others are very flexible".[170]

2.3.
The dynamics of *de se* thoughts

The clearest way in which the phenomenon of cognitive dynamics manifests itself is in the form of indexical thinking. I think there are two main reasons for that. First, in the case of indexical thought, the dynamics is conspicuous in the way the linguistic expressions of the thought are adjusted over time. And second, if we

---

[169] Again, in PERRY, J. 1997c.
[170] KAPLAN, D. 1977, p. 537, footnote 64.

follow Evans, one must have certain cognitive skills that need be working properly if we are to say that a thought can be continuously entertained through a given stretch of time. Some of those skills are purely mental in nature, others may exploit our senses and perception, some are linguistic.

In the case of indexical thinking, a very particular kind of mental skill must be exercised: that of keeping track, as Evans puts it, of the object of one's thought. One must (either consciously or with the aid of some unconscious mechanism, or maybe both) compensate for the changes in the environment surrounding oneself and the object of one's thought. In other words, one must mentally (and sometimes also physically) compensate for the changing contexts. In the case of a demonstrative thought, for example, one must presumably keep sight of the object as one moves through space, and thus adjust to the changing perspective as one's position relative to the object evolves (in certain situations, the adjustment may be physical).

But it seems clear that, in the case of the dynamics of indexical thoughts, that kind of mental skill is not enough: the subject must also be capable of updating the thought *expression* as context changes. It might be said that the linguistic skill is in some way dependent on the mental one – at least in the sense that a competent speaker of a given language will naturally update the expression of her thought if she does not lose track of its object. It might also be said, following Evans, that even the capacity a speaker has of *understanding* an indexical term depends on her capacity of tracking objects denoted by those terms, among other dispositions.[171]

I have tried to show that one of the consequences of Kaplan's and Evans's ideas is this: if we want to establish the conditions under which preservation of an indexical thought takes place, preservation of the thought-content under any other "guise" (as a memory demonstrative, for instance) will not be enough. If we are to say that the same indexical thought has really been preserved, the indexical element of the thought has to be retained as well – we have to somehow preserve that special perspective we had in thinking indexically.[172] So, indexical thoughts present the most interesting cases of attitude/thought retention because their dynamics is

---

[171] EVANS, G. 1981, p. 305.
[172] It is not the case that the same character needs to be preserved (actually, it cannot be, otherwise we would have completely distinct thoughts), but the characters must be connected in a way that they reflect the retention of that perspective.

evident through language. Besides, the mental effort needed to preserve them conspicuous in this case and that effort seems to be the hallmark of indexical thoughts. Now, how do thoughts about ourselves get into the picture?

Two sorts of cases are invoked by Kaplan in "Demonstratives" when it comes to the dynamics of I-thoughts, The first is related to the issue of communication. This is the problem of accounting for the expression of the same indexical thought by two persons who naturally have different perspectives, like when *e,g.* I express a thought about myself using the pronoun "I" while my interlocutor uses "you" to re-express the same thought from her own perspective. Note that in this case, some adjustment of the linguistic characters occurred to express the original thought, and this is why it matters to Kaplan.

The second kind of case is illustrated by the example of a subject who sees her own image reflected in a looking-glass and is not sure as to whether she is seeing her own image or someone else's. I cannot be said to have retained a belief about myself held under the character of the first-person pronoun "I" if, for example, I go on from thinking of myself under such a character (if I take the reflection I see to be my own) to thinking of myself under some other character (if I take the image to be someone else's). There is a fairly intuitive sense in which there is no belief retention if, thinking of myself as "I", I replace the character of "I" either with that of the third-person pronoun "he" or of a demonstrative such as "that man" before getting back to using the first-person pronoun to express an explicitly *de se* thought. Still, there is arguably preservation of a *de re* thought because throughout the steps I have been thinking of the same referent first self-consciously, then from a third-person perspective, and again self-consciously. And, again, preserving the *de re* thought is not enough to account for belief-retention, given that our purpose here is to tell whether the indexical *de se* thought has been retained.

Clearly, the issue of the communication of I-thoughts and that of the internal retention of thoughts about oneself should be kept separate, although they are each other related in Kaplan's account – for his aim was to provide a comprehensive account of the functioning of indexicals. Here I shall only be concerned with the latter since it is nothing but the issue of cognitive dynamics properly speaking. As I see it, Kaplan's choice to deal with the issue of the communication of I-thoughts as part of the problem of cognitive dynamics is another indication that, for him, cognitive dynamics concerns the re-expression of indexical thoughts *when a*

replacement of indexical expressions occurs – in short, when characters need be updated.

So, in Kaplan's original strict sense of cognitive dynamics, *no* dynamics occurs when a subject continuously grasps a *de se* thought through time, assuming of course that she thinks about herself under the character of "I". No dynamics occurs because no character change is needed: the subject of the thought will just keep on using the first-person pronoun "I" to refer to herself. Kaplan was not interested in instances of belief retention without character change. He does not seem to see a problem in explaining how the subject's internal continuity is secured when the subject's attitudes are not expressed by indexical sentences requiring some updating of the indexical expression(s). However, I believe that, leaving aside Kaplan's idiosyncrasies, the problem is a real one for it generalizes over all instances of indexical thought.

All indexical thoughts have flighty characters and exhibit a dynamics. As the characters under which some thought is held change, the cognitive significance of that thought also changes. Indeed, we have seen that Kaplan associated the problem of cognitive dynamics with at least *some* cases of I-thought. But I suggest that a possible general re-statement of the central problem of cognitive dynamics could be as follows: how to account for our ability to retain or change our epistemic attitudes towards *indexical thoughts* as we move from one context to another? One might wonder whether there is any difference between this formulation and Kaplan's. Now remember Kaplan's own words in "Demonstratives". He suggests that the task of cognitive dynamics is to answer the following questions: "What relation between a character under which a belief is held at one time and the set of characters under which beliefs are held at a later time would constitute retaining the original belief?"[173] It is worth noting that in the above re-statement of the problem no restriction (apparent or otherwise) is set on indexical thoughts according to whether their expression must updated or not. One of the lessons to draw from Rip's case is that accounting for the *linguistic* dynamics involved is not enough to explain *cognitive* dynamics.

Be it as it may, we know that in the case of the dynamics of *de se* thoughts the subject need not update the indexical "I" when entertaining I-thoughts over

---

[173] Both are taken from the long quotation above, in: KAPLAN, D. 1977, p. 537-538, footnote 64.

time. Besides, as seen earlier, she does not even seem to need to keep track of herself through time as the referent of the pronoun "I". The very meaning of the pronoun guarantees that every time a given subject says "I", she is referring to herself. And it does not seem plausible to think that I can be replaced by someone else the way an external object can be without being aware of it. So, can we identify an analogue of the ability of keeping track of the object of one's thought in the case of I-thoughts? Do we ever even need one in the case of "I"? Can we be mistaken in identifying ourselves, so that the issue of whether Rip was able to retain his belief in the case of the use of "I" can be raised? I think we can and we do, and I shall explain why in what follows. But note that if I can be mistaken in identifying myself, then a sort of psychological mechanism of *retention of the self* must function properly so that I can be credited with retention of thoughts about myself, such mechanism is undoubtedly different from the one at work in the case of demonstrative and other indexical thoughts. If we suppose it exists, our next move would be to acknowledge that the problem of cognitive dynamics must be re-stated beyond its original scope and be presented in a different way if our objective is to account for the dynamics of *de se* thoughts.

A thought I have about myself is based on a few ways I have available to me to get information about the person I happen to be. I have access to those special information channels because I find myself in an epistemically enriching relation with myself that may be similar to the kind of relation I have with external objects (I can discover new things about myself by looking at my own image in a mirror, for example), but which may also be unique to the first-person. As pointed out earlier, I can find out things about myself by looking at my reflection in a mirror, for example – that is, I use vision, a sense which is also employed to find out things about external objects. But we also have senses that have developed solely for the purpose of receiving data on our own states, such as proprioception, introspection and other senses gathered under the heading *interoception*. Through the information I receive from those senses I can, for instance, make self-attributions of both physical and mental predicates. Moreover, as Evans argues, *de se* thoughts can also be based on both the relations I have with the external world and on memory. Examples of the first case would be self-locating thoughts relating to objects of the world, such as "I was looking at the burning house"; In the second case, I would explore autobiographical memories, as in "I had a terrible headache (that day)".

The role played by memory for the dynamics of first-person thoughts would be twofold: it can make the retention of cognitive states (beliefs) possible, but can also retain the information on which such states could be later based.[174] For Evans, the dynamics of *de se* thoughts is special because the retention of this kind of thought happens through memory and without effort. For him, that type of thinking is based on information and other mental states retained through memory without danger of misidentification of the referent. If I have a thought of the form "I am *F* (now)" at a given moment t_1, my mnemonic capacity will ensure that at a later time t_2, I am inclined to think "I was *F* (at that time)", without the need of a supplementary judgement for identifying the reference (myself). For Evans, the nature of the thought I had about myself in the past does not matter, nor does the nature of the information retained. Even thoughts which in the present are open to error through misidentification of the referent (as in the case of someone who sees her own image reflected in a glass and is not sure if she is seeing her own image) will produce, through memory, thoughts that do *not* exhibit such feature. The kind of information gathered in the past which is at the basis of a current thought does not matter either. It may be that the information comes from a hallucination and is thus better characterized as *misinformation*. In any of these cases, memory will play its role in retaining either the thought (open to error through misidentification or not) or the information (be it real information or misinformation), later producing a thought which is free of the need for identification of the referent, which for Evans means immune to error through misidentification.

However, it is perfectly conceivable that an individual have memories she associates with herself that have someone else as their causal source.[175] Memories of this kind could happen in at least two cases (though there are others, as I shall argue). In the first one, we imagine that a *mental* subject is split into two – the two resulting subjects would be different from the original one, and would therefore have autobiographical memories having another person as their causal source. In the second case, memories of a given person are transplanted into someone else's mind or brain – also in this case the subject receiving the transplant has autobiographical memories having a source other than herself. Both cases belong

---

[174] EVANS, G. 1982, p. 235.
[175] The argument is Shoemaker's, and Evans contemplates it: *cf. Ibid.*, p. 241-242.

in a science fiction scenario, but they serve to show that the error through misidentification in the case of the first person is possible.

On the other hand, we cannot ignore the fact that we live in a (possible) world in which splitting people's minds in two and performing memory transfers are not a reality (at least not as yet). So, restricting our investigation to this world (and other similar worlds), we can say that from the logical possibility of error of identification, we should not conclude that *de se* thoughts based on autobiographical memories are *de facto* open to such error, for in general we do not think they are – at least not in our world.

Nevertheless, two things are worth being noted. First, the fact that in normal cases the dynamics of *de se* thoughts is unconscious and "effortless" (and that such thoughts are not open to error through misidentification) does not force on us the conclusion that there is *no* mechanism responsible for keeping track of ourselves that needs to function well so that we can apprehend and properly retain *de se* thoughts. Second, I do not think we need to appeal to science fiction scenarios to glimpse the possibility of that conclusion's not being forced on us, and to show that the possibility of being mistaken in the identification of oneself exists. This second point will be clear in what follows, when we analyse cases of people who suffer from mental disorders such as schizophrenia.

It seems natural to think that my connection with the referent of "I" cannot be broken, while in the case of other indexicals, the connection I have with the referent of the term can – if my perception of it is lost, for example. It also seems natural to think that everytime I use the first-person pronoun I refer to myself, while I can refer to many different persons and things by employing other indexicals. But things are not always such as I have just described them.

There are situations in which the connection I have with myself does not guarantee knowledge which is reliable. In certain cases of schizophrenia, for example, patients attribute to themselves actions performed by others.[176] Studies have suggested a connection of Austism Spectrum Disorder (ASD) and Asperger Syndrome with both impaired proprioception and interoception.[177] But impairments

---

[176] These are cases of a phenomenon we may call "(schizophrenic) thought insertion". In the philosophical literature, *cf.* CAPPELEN, H.; DEVER, J. 2013, p. 131. Also VIGNEMONT, F. 2005, p. 12-13 and CAMPBELL, J. 1999.

[177] Take, for instance, WEIMER, AMY K. M. D. et al. 2001, a study that provides us with some evidence of proprioceptive deficit in Asperger Syndrome patients.

in those senses are not always *caused by* some kind of mental disorder – they can actually be among the causes of such pathologies. That is suggested by an even more a recent study which indicates that impaired interoception could be connected with a disorder known as Alexithymia, often found in people presenting symptoms of ASD, but not necessarily connected to autism.[178] These studies give support to the idea that whatever grounds I have for my supposedly immune I-thought are in reality susceptible to failure. And that renders the thoughts based on them susceptible as well. That includes our interoceptive senses. That suggests the existence of cases in which my use of "I" will not be guaranteed to refer to myself. But we do not need to suffer from any (serious) psychological disorder for our memories to somehow get muddled and for us to make a mistake in the identification of ourselves through time. In short, I believe that situations similar to schizophrenic thought insertion might generate cases in which there is either a failure in the apprehension of a given *de se* thought or confusion in its retention (in our self-ascription of certain predicates through time).

If, being faithful to Evans's ideas, one is to preserve the same kind of indexical belief about oneself through time, one has to rely on memory and on our senses (that includes proprioception and other internal senses generally gathered under the heading of interoception). But neither memory nor our senses are always reliable (even if we are talking about proprioception or interoception). What is special in the dynamics of I-thoughts is the fact that the retention mechanism seems to be effortless.[179] In normal cases where there is no memory loss, nor any other problem of self-identification, memory by itself ensures the retention of our I-thoughts. I think Evans is right in defending such a view. But it does not show that there is no mechanism of retention that must be at work when someone thinks about herself over a certain period of time. Perhaps the existence of cases in which such a mechanism fails serve as evidence for its existence.

Ordinarily we exercise that dynamics pretty well – we are probably trained to do so as we grow up, or perhaps that capacity is hardwired into natural intelligent systems, rather than being the product of a training. (In this case, perhaps it is the correct *use* of the indexical "I" which is a matter of training.) Be that as it may, the dynamics involved in *de se* thinking – that is, the updating of the information

---

[178] SHAH, P. et. al. 2016.
[179] EVANS, G. 1982, p. 237.

associated with the use of "I" – seems to be a kind of dynamics that we naturally perform in normal situations (when there is no memory loss nor anything of the sort). We are usually capable of keeping track of ourselves effortlessly and unconsciously, precisely because the kind of dynamics at stake here is not to be equated with the conscious activity of retaining the information gained through the usual information channels, but with the updating of information that need not be neither conscious, nor require effort. It is in this sense that the kind of dynamics we perform when entertaining temporal indexical thoughts is usually different, because some "effort" and adjustments are required. If I acquire information about some object about which I think indexically, updating that information will usually require from me a conscious activity of updating the indexical perspective I originally had. Usually the same does not happen in the case of I-thoughts.

But even though it may be rare, sometimes an analogue of a tracking activity in the case of the self has to be conscious. This may happen because it is possible that we make mistakes in identifying ourselves – because the information channels I rely on are faulty, for example. It is also possible that we have difficulty retaining information about ourselves through time. I think we are always vulnerable to those situations, in the sense that one or more of them may happen to anyone at any time. But of course they are not so frequent, so most of the time the dynamics of I-thoughts will indeed seem effortless. On the other hand, perhaps cases in which the activity is unconscious might exist in the case of temporal indexicals too. We do not put too much thought into the updating of an indexical such as "today" to "yesterday" when we want to re-express a certain thought about day d on $d+1$. Rip would not have thought twice before trying to re-express his thought. But we can make mistakes if we lose track of time. What I suggest is this: I may lose track of myself too, if I grasp a *de se* thought at some time t_1 by ascribing myself some property F and, come the time to update that thought at t_2, I wonder if I was right in ascribing F to myself rather than to someone else.

There is possibly a variety of situations in which that natural dynamics can be disturbed and one comes to be mistaken even when thinking about oneself as the source of a given state. I think we can draw a distinction between at least two kinds of cases in which the dynamics of our *de se* thoughts can fail and in which a conscious effort of keeping track of the referent (myself) may need to be performed.

The first is illustrated by canonical cases of I-thoughts which are vulnerable to error through misidentification. Think of the case of a person who relies on visual stimuli to have an I-thought and cannot be sure if the body she sees is really her own. That person will explicitly be in doubt as to whether the possessor of a given property is really herself or someone else. In this kind of case, only a short period of time elapses.

The second possibility is this: suppose I entertain a *de se* thought in the present. I either make a mistake in the identification of the referent (myself) without realizing it (maybe because it was a case of an explicit *de se* thought in which the relevant information channel I relied upon was vision), or I correctly identify myself as the source of whatever information prompted my I-thought. And then I retain that I-thought over a longer period of time. In a way, and to agree with Evans, it no longer matters if that original thought was immune to error through misidentification or not, because I somehow come to "own" that thought. But that does not mean that the retention mechanism itself will work properly over a longer period of time. It can be that memory somehow prompts me to doubt if it was really I who was in the state I remember.

In the sort of case I am describing, I think we can say that I will have retained at t_2 my *de se* thought grasped at t_1 if at all times in between t_1 and t_2 I still remember it (even if not consciously, but I should be able to re-express it if asked). In a way, it is as simple as that. Can we imagine a situation in which I doubt if the I-thought I am grasping at some time t_2 is the same thought I grasped at some ealier time t_1? I think we can, if at t_2 I come to doubt that the thought grasped at t_1 was really "immune to error through misidentification" – that is, if the thought grasped then really had myself as a source.

This kind of situation seems to be opposite to the one described in the example of the messy shopper. In the messy shopper case, Perry goes from a *de re* thought to a *de se* one. In the case I am trying to describe, if I come to think that the *de se* thought I have is fruit of some kind of identification mistake (or thought insertion, or something else), I will move from a *de se* thought to a *de re* one (or at least wonder if I should make that move). This would be a case similar to the one described by Kaplan in "Demonstratives", in which Kaplan wonders if he is really seeing his own image in the glass. The difference is that the time elapsed is longer, and that gives rise to the possibility that in between t_1 and t_2 there can be some kind

of memory insertion or other problem that makes me wonder if the thought I am having is really a thought about myself. I think that can happen even if the thought is supposedly immune to error through misidentification, because based on interoception, for example. I would then wonder if I should change the indexical from "I" to "he" or some other kind of expression designating someone else. Perhaps this is the key to understanding the dynamics of *de se* thoughts: if I wonder if I should change the indexical through which I am expressing my *de se* thought, that would mean that I have not retained it. The first condition that must be met for there to be retention of a *de se* thought is this: I must be right in identifying myself as the source of some property which prompts my grasping of a *de se* thought describing my having that property and I must not come to doubt the truth of that ascription over time. Suppose I am somehow wired to someone else and can feel everything this other person does at any moment (as if I were this person). She can also feel everything I do, exactly the same way. Perhaps it is not so hard to imagine that in such a situation retaining *de se* thoughts over time (particularly over olong periods of time) would at least sometimes require from both of us a kind of conscious tracking activity of our own selves so that we are able to differentiate one another.[180]

So, I think that it is not as if there were no dynamic mechanisms responsible for the retention of *de se* thoughts and of tracking the object of my *de se* thought through time. There is a sense in which cognitive dynamics must involve an activity from the part of the individual retaining the thought/concept in question. We cannot lose that from sight. But I think that the fact that sometimes that activity seems automatic or "effortless", as in the case of *de se* thoughts, does not mean that no actitvity is performed. It just means it may be performed without our being aware of it. The activity can be "automatic", but that does not mean it does not exist. It is not always automatic, and mistakes can happen. Perhaps in the case of *de se* thoughts the well-functioning of the mechanisms of self-knowledge and of memory are the correlates of the tracking activity that must exist in the case of indexical thoughts if we are to say that there has been retention.

---

[180] We could also imagine this kind of situation in a world like the one portrayed in the science fiction series Sense8. Perhaps there is a sense in saying that if I retain a thought ascribing to me a property which I did not really have, I will neither have grasped a thought about myself nor about someone else, I will have retained a confused thought.

The dynamics of I-thoughts may consist in our capacity of continuously associating the same individual to my use of "I". The pronoun I employ to think of myself will in general always be the first-person pronoun "I". However, the kind of information on the basis of which I formulate my I-thoughts will not always have myself as a source. And there may be cases in which memory does not work very well.

2.4.
Conclusion

We saw that Kaplan was the first to draw our attention to the issue of cognitive dynamics, taking as starting point the aforementioned Frege´s notorious remark in "The Thought". Kaplan presented cognitive dynamics as being concerned, first and foremost, with the evolution of our attitudes towards indexical thoughts whose modes of expression (linguistic characters) must be adjusted over time. This restriction precludes us, at least in principle, to straightforwardly apply the Kaplanian model to the special case of indexical I-thoughts.

Some authors have elaborated on Kaplan's ideas on cognitive dynamics and tried to settle Kaplan's issue. Two of the most important attempts of accounting for the phenomenon are Perry's and Evans's. Evans embraces the (counterintuitive) consequences of a thesis Kaplan seemed hesitant to defend: someone who loses track of time, for example, cannot grasp temporal indexical thoughts. In spite of the counterintuitiveness of some of his claims, the positive and valuable aspect of Evans'a proposal is to meet the demand to retain the original perspective via his theory of keeping track where the perspective *itself* is tracked. This fits with my interpretation of Kaplan (cognitive dynamics as a matter of perspective-tracking), but it is radically questioned by *de se* skeptics like Dever and Cappelen, whose ideas I shall discuss in the next chapter.

Now, Perry criticizes Evans and holds that there are many ways in which one's belief can be retained out of the original context. For Perry, even if someone loses track of time, such a person could still retain her belief under a "back-up character" from memory, even if such a back-up character is of a completely different nature when compared to the one under which the thought was originally held. I think both authors have a limited perspective on the phenomenon and only

explain the problem pointed out by Kaplan in a partial way. But although their accounts do not fully explain cognitive dynamics, I think they shed light on particular facets of the problem, and there may be a way of accommodating both in a more general theory of the dynamics of our thinking.

I have also tried to very tentatively indicate that there is a sense in studying the dynamics of *de se* thoughts. That kind of thought has a very special dynamics, because its expression does not present the same behavior seen in the case of other kinds of indexical thought. In particular, I do not need to substitute the indexical "I" in re-expressing thoughts about myself over time. What about the psychological skill of keeping track of the object of one's thought? In the case of *de se* thoughts, either that skill is not necessary or its exercise is effortless. In my opinion, it is obvious that the same tracking ability we exercise in entertaining a demonstrative thought, for instance, will not be at play in the case of I-thinking. But cases of memory loss, of I-thoughts which are vulnerable to error through misidentification and psychological conditions such as schizophrenia seem to indicate that a certain kind of tracking mechanism must be working properly if someone is to be credited with having *de se* thoughts through time.

**3
Is "I" an essential indexical and do *de se* thoughts have a
special status *qua* indexical thoughts?**

> *An amnesiac, Rudolg Lingens, is lost in the Stanford
> library. He reads a number of things in the library,
> including a biography of himself, and a detailed account of
> the library in which he is lost. […] He still won't know who
> he is, and where he is, no matter how much knowledge he
> piles up, until that moment when he is ready to say "This
> place is aisle five, floor six, of Main Library, Stanford. I am
> Rudolf Lingens".*

John Perry, "Frege on Demonstratives"

> *[…] there is no such thing as essential indexicality,
> irreducibly* de se *attitudes, or self-locating attitudes.*

Herman Cappelen and Josh Dever, *The Inessential
Indexical*

3.1.
Introduction

Chapter 1 was meant to show that *de se* thoughts have a special status and are
irreducible to other kinds of thoughts, notably *de dicto* and *de re*. In this chapter I
turn to objections that have been raised against this widespread assumption. My
objective is to discuss and argue against a certain form of skepticism that found
expression in the literature concerning the specialness of *de se* thoughts *qua*
indexical thoughts and the relevance of a study of the phenomenon of indexicality
on the whole.

In chapter 2 we saw that the issue of cognitive dynamics was originally raised (by Kaplan) for indexical thoughts. This restriction of the scope of cognitive dynamics is not arbitrary. The natural dynamic character of human thought becomes particularly conspicuous in the case of indexical thoughts. As Dokic puts it, "it is no doubt in relation to indexical thoughts that we can better appreciate, at an initial stage, the presence of a cognitive dynamics"[181]. The expression of an indexical thought usually undergoes changes over time. Besides, it can be argued that any change in the expression of an indexical thought must be accompanied by a mental effort of adjustment on the part of the thinker. Adjustment to new contexts at the linguistic level is not enough to explain thought (attitude) retention. Following Evans, it might be said that an effort of adjustment is also required at the psychological level, which perhaps amounts to the exercise of a capacity to keep track of the referent of the indexical. At least in principle, it seems that indexical thought raises the most interesting and difficult challenge for the cognitive dynamics theorist because the phenomenon to be explained is multilevelled.

In the case of first-person thoughts, their re-expression does not require any change in the indexical employed – namely, in most of the cases[182] the first-person pronoun. Besides, at least in principle, it seems that no mental tracking effort is required, although this point is not uncontroversial, as seen in the previous chapter. I believe there is an analogue in the case of I-thoughts of the capacity to keep track of the referent involved in the subject's retention of the indexical thought. Explaining that mechanism in the case of I-thoughts is equivalent to explaining their dynamics. In any case, it seems that all this presupposes that indexicality *is* a relevant philosophical phenomenon and that a study of the dynamics of *de se* thoughts and of cognitive dynamics must take it for granted. The indexical element in our indexical thoughts must be thought of as a deep phenomenon worthy of explanation if there is any point in explaining their dynamics.

Not everyone endorses the assumption, although it seems to follow naturally from Perry's and Lewis's thought experiments. Accordingly, not everyone grants the existence of irreducibly *de se* (or self-locating) thoughts. The most influential

---

[181] DOKIC, J. 2001, p. 67.

[182] Remember, the quasi-indicator appearing in oblique contexts such as de se belief reports behaves like the first-person pronoun outside those contexts: it cannot be eliminated or substituted by any coreferential singular term without that affecting the sentence's content and truth-value. This has been pointed out by KAPITAN, T. 1999, p. 3.

authors who refuse to grant their existence are Cappelen and Dever. In their 2013 book entitled *The Inessential Indexical*, they defend a variety of *de se* skepticism. They provide a battery of arguments in support of the view that the whole debate about indexicality is an illusion and that, the phenomenon of *de se* thoughts being philosophically shallow and uninteresting, it cannot play any explanatory role.

Perhaps the first thing to be pondered when assessing Cappelen and Dever's proposal is their initial claim that *de se* and indexicality are *philosophically shallow*. They make that claim in the opening pages of the book, using related terms at various points.[183] In face of that, one of the first questions to ask seems to be this: what does it mean to say of something that it is philosophically *deep* or *shallow*?

Cappelen and Dever do briefly address the question, writing that the "weasel words" they employ to argue for their thesis are only employed because "there's clearly some role played by the indexical, *i.e.,* non-constant, characters",[184] immediately adding a footnote to say that they do not deny that changes in indexicals engender changes in behavior. But right after that, once again they employ one of those weasel words to argue that they deny that the notion of character as such is an *interesting* one, claiming that only the broader notion of Fregean sense has any real interest.

It is not at first easy to see exactly why the Kaplanian notion of character is "philosophically uninteresting". After all, Kaplan came up with the notion as a way of accounting for the meaning of context-sensitive expressions in language. I shall analyze here some of the arguments put forward by Cappelen and Dever in support of *de se* skepticism and the idea that Kaplan's notion of character and, more generally, indexicality are uninteresting from a philosophical point of view.

They take it that the whole talk of indexicality and *de se* can actually be reduced to a myriad of different phenomena[185], and we do not need to appeal to irreducible indexicality to explain them. So, each chapter of the book is dedicated to one of these phenomena. I shall not discuss all of them here, but only those that are of particular interest to our own investigation – the dynamics of *de se* thoughts.

---

[183] "This book is an extended exploration and defense of the view that perspectivality is philosophically shallow" (CAPPELEN, H.; DEVER, J. 2013, p. 2); "The considerations coming out of the Perry-Lewis tradition [...] provide no evidence that there are philosophically interesting or important roles played by non-constant characters" (*Ibid.*, p. 16).

[184] *Ibid.*, p. 17.

[185] *Ibid.*, p. 3.

Cappelen and Dever's first strong complaint against the so-called *essential indexical thesis* is the overall lack of substantial arguments for it. I think that in this they are right, and their calling our attention to that state of affairs is already a significant contribution to the debate on the subject. I mean to provide at least tentative arguments in support of essential indexicality, thus supplying Cappelen and Dever with what they think is lacking in the debate on the issue.

In section 3.2 below I try to figure out what the problem of the essential really is and I discuss the skeptical view on which there would actually be *no* problem raised by so-called *essential indexicality*. In section 3.3. I discuss Cappelen and Dever's idea that *de se* thoughts do not have a relevant role to play in explaining agency. In section 3.3 I address the issue of immunity to error through misidentification. Cappelen and Dever argue that such immunity is not a distinguishing feature of *de se* thoughts and cannot be meaningfully connected to indexicality on the whole.

3.2.
What is the problem of the Essential Indexical?

As Dilip Ninan rightly points out,[186] while the belief that *de se* thoughts have a special status is widespread, the exact nature of the problem supposedly raised by them is not at all clear. As he puts it: "Perry and Lewis, for example, attempted to motivate novel theories of propositional attitudes as a response to that (putative) problem".[187] But what, exactly, *is* the problem?

In chapter 1 we saw that the problem posed by the essential indexical, as originally formulated by Lewis and Perry, is that traditional theories of propositional attitude cannot account for attitudes expressed via that indexical: *de se* attitudes. Traditional theories regard propositional attitudes as relations between agents and abstract objects: usually either Fregean thoughts or Russellian propositions. As Stojanovic puts it, Perry and Lewis "were convinced that propositions could not provide an apparatus powerful enough to model [essentially indexical] belief, granted that belief helps in accounting for behavior".[188] But even if "*de se* attitudes have certain distinctive features that differentiate them from non-

---

[186] NINAN, D. 2016.
[187] *Ibid.*, p. 87, footnote 3.
[188] STOJANOVIC, I. 2001, p. 304.

de se" ones,[189] that by itself does not (automatically) entail that they raise a problem for propositional attitude theories.

Ninan argues that we should not introduce the special status of *de se* thoughts by saying things like "A *de se* attitude is a thought about oneself when one thinks of oneself via a special, first-person mode of presentation that no one else has access to."[190] The suggestion then could be that we characterize the notion of *de se* attitude ostensively, by showing examples of *de se* thoughts contrasted with non-*de se* ones. Then we provide a sort of preliminary method to recognize *de se* thoughts, such as this: if there is a belief one would normally express through a sentence containing "I", "me" or "my", for instance, that is probably a *de se* belief. Starting with that, we can later work on some way of showing that those beliefs have a distinctive character or a philosophically or psychologically interesting feature.

Recall that one of the examples used by Perry to introduce the thesis of the essential indexical is that of Rudolf Lingens, the amnesiac lost in the Stanford Main Library who reads all the books in the Library, including his own biography, and is, in spite of this, unable to comprehend who and where he is. Is there a difference between the sentence "Rudolf Lingens has been to Paris" on the one hand, and "I have been to Paris" when uttered by Rudolf Lingens, on the other? Do they convey different pieces of information? That would partially explain why Lingens believes the first, but not the second – he does not believe *de se* that he has been to Paris.

Cappelen and Dever argue that we do not need to appeal to essential indexicality or the *de se* to explain the difference. They argue that in cases such as that of Rudolf Lingens, we are actually facing an instance of Frege's famous puzzle concerning the difference in cognitive value of sentences involving coreferential terms. The problem is that of explaining why a sentence of the form "*a = a*" (such as "Holland is Holland") is relatively uninformative, while a sentence of the form "*a = b*" (such as "Holland is The Netherlands") is informative for someone who ignores that the latter terms are coreferential. To put it differently, Frege's puzzle is the problem of explaining why someone who accepts a sentence like "The Morning Star is the brightest object in the night sky" as true may take "The Evening Star is the brightest object in the night sky" to be false or be agnostic about it. This, of course, occurs provided the subject is not aware that the Evening Star = the

---

[189] NINAN, D. 2016, p. 87.
[190] *Ibid.*, p. 89.

Morning Star. Frege employs his notion of sense to explain this phenomenon. In the sentence of the form "$a = b$", each of the terms, a and b, expresses a different sense, although both designate the same thing in reality. This is the case with Holland and The Netherlands and with the Evening Star and the Morning Star. So, the sentence expressing their sameness becomes "epistemically enriching" and can, thus, be a source of new information for an individual who is not aware of the truth of the sameness expressed by the proposition in question. Lingens reads in his own biography that "Rudolf Lingens has been to Paris" and accepts it as true, but is not in a position to believe the thought expressed by "I have been to Paris", even though "I" and "Rudolf Lingens" happen to be coreferential. Cappelen and Dever argue against the specialness of the *de se* by saying that the same puzzle arises in the case of coreferential terms involving "I" and the proper name "Rudolf Lingens" as in cases involving non-indexical singular terms (e.g. definite descriptions like "the evening star" and "the morning star"). So, the substitution failure observed in Lingens's case would have nothing to do with indexicality or the kind of thought expressed by the use of "I".

Even though cases such as that of Rudolf Lingens or of the messy shopper can be correctly regarded as instances of Frege's puzzle, that does not change the fact that they raise a problem for traditional theories of attitude. So, the so-called problem of the essential indexical remains. These see the objects of attitudes as having an absolute truth-value. And when such objects are taken to be structured propositions, these are supposed to be either singular or descriptive. The *de se* seems to elude these two categories.[191] Still, as Cappelen and Dever point out,[192] the problem raised may not be specific to the *de se*.[193]

But there is arguably still a problem, and it is specific to the *de se*. To show this, Ninan puts foward an extended version of Perry's account of the "traditional doctrine of thoughts" and argues that *de se* thoughts such as those exemplified in cases such as that of Lingens cannot be explained by that kind of theory. We discussed Perry's version of the doctrine of throughts in section 1.3.2. For him, the doctrine has three tenets:

---

[191] *Ibid.*, p. 95.
[192] CAPPELEN, H.; DEVER, J. 2013, section 3.1.
[193] NINAN, D. 2016, p. 95.

1. Belief is a relation between a subject and an object denoted by a that-clause and taken to be a thought (or proposition);

2. Thoughts/propositions have absolute truth-values.

3. They are individuated by their truth-value, by their truth conditions, and by their being composed of the same concepts.

Ninan puts forward a version of that doctrine in a form sufficiently general to encompass doctrines with either Fregean, Russellian or Lewisian roots. To achieve that goal, he adds two more tenets to the doctrine:

4. Agreement is relation between a group of individuals and a thought/proposition, and it happens when all individuals take that thought to be true.[194]

5. "If two agents have all the same (relevant) beliefs and desires, then, other things being equal, they will behave in the same way".[195]

Now, as Ninan argues, the *de se* would pose a problem to this expanded doctrine, because the two added tenets are inconsistent in the face of a *de se* attitude.[196] According to tenet 4, two subjects find themselves in agreement when they believe and have other attitudes towards the same set of thoughts. But tenet 5 tells us that if two agents have the same beliefs and other attitudes towards the same set of thoughts (that is, if they find themselves in total agreement), other things being equal, they will act the same way. To see the problem, we may appeal to one more of Perry's famous examples. In this one, Perry is taking a walk in the woods with a friend and they find a wild bear. Perry happens to be nearest to the bear and is thus in immediate danger. His friend is a few meters away. So, Perry explains:

> When you and I entertain the sense of "A bear is about to attack me", we behave similarly. We both roll up in a ball and try to be as still as possible. Different thoughts apprehended, same sense entertained, same behavior. When you and I both apprehend the thought that I am about to be attacked by a bear, we behave differently. I roll up in a ball, you run to get help. Same thought apprehended, different sense entertained, different behavior.[197]

In this sort of scenario, the two friends agree that Perry will be mauled by the bear if he does not curl up into a ball and play dead. It seems that they both believe the same contents – taking this notion in a sort of generic way. So from tenet 5 we

---

[194] *Ibid.*, p. 100.
[195] *Ibid.*, p. 102.
[196] His whole argument is to be found in *ibid.*, section 3.1.
[197] PERRY, J. , p. 494.

would predict that both friends would behave the same way, all things being equal. However, they behave differently. In this sense, the *de se* do pose a problem for this sort of traditional theory of attitude/thought construed in a generic way. The role played by the *de se* in action explanation seems to be one of its most relevant distinguishing features.

In the next section I will discuss Cappelen and Dever's argument against that very point: the idea that the *de se* has an important role to play in explaining our behavior. Their strategy could be interpreted as follows. Perry and his friend have a different behavior in the bear attack example, even though they believe the same contents. If their beliefs and desires all have the same contents, all other things being equal, they should act the same way. But according to Cappelen and Dever, the problem is in the intermediary clause of the previous sentence: "all other things being equal", for they are not. Although their beliefs and desires are the same, different actions are available to Perry and his friend.[198]

Ninan points out that there is an important distinction to be made between two types of action: an agent-specific one and an agent-neutral one. An example of an agent-specific kind of action is "the action that John Perry curls up", while its agent-neutral counterpart would simply be the action of curling up. This distinction is very important for Cappelen and Dever. In their assessment of the example, of course Perry's friend cannot (at least not in principle) perform the the agent-specific action that John Perry curls up. But if we work with agent-neutral types of action, it is not hard to see that both friends could perform the action of curling up. They do not, though. So the example remains the same and still gets its appeal.

The problem then remains: *de se* thoughts cannot be explained by traditional theories of propositional attitude, as stated above. This is the problem of *de se* thoughts and attitudes. And there are reasons to think that it is a problem raised by *de se* thoughts and *de se* thoughts alone. Appealing to the notion of functional role played by beliefs, Ninan argues that non-*de se* attitudes do not pose the same problem as do *de se* ones. If a pair of non-*de se* thoughts in which two individuals agree are similar – if, in other words, they are expressed via the same sentence (an eternal sentence), then probably they will also have a similar functional role. That means that they will probably induce the same sort of behavior in both subjects. As

---

[198] NINAN, D. 2016, p. 105-106.

Ninan explains, "If, for example, I want to visit Mark Twain's grave, my belief [expressed through the sentence "Mark Twain is buried in Elmira"] might cause me to seek routes to Elmira. But if you too want to visit Mark Twain's grave, your belief [expressed via the same sentence] would likely have similar sorts of effects on your behavior. Arguably, this does not happen with *de se* beliefs.

3.3.
The role of indexicality and the *de se* in action

We have seen how Perry argued for the irreducibility of *de se* attitudes and for the need to account for the role they play in the explanation of behavior.

There are many ways in which we can think of ourselves at a given time and through time, and there can be situations in which we are thinking of ourselves without realizing it – in that case it might be said that we are grasping accidental *de se* thoughts. In general, coming to realize this fact seems to shed a new light on our thought, so to speak. Perry argues that when the messy shopper goes from entertaining a *de re* thought (or an accidental *de se* one) about himself to entertaining an explicitly *de se* one, his *belief state* changes while the "core" content of his belief remains the same. He also argues that, as a consequence of this *belief* change, his *behavior* changes as well.

We have seen that Perry illustrates his idea about the roles our belief contents can play in explaining our behavior with various examples, such as that of the messy shopper. Cappelen and Dever also revisit Perry's example of the two friends who are are taking a walk in the woods and find a wild bear. In the thought experiment both Perry's and his friend's seem to believe the same thing: that Perry is about to be attacked by a bear. But they are not in the same "belief state", as Perry puts it, because they believe that content in different ways – to use Kaplan's terminology, they believe it under different characters. Their behavior will depend on the role played by the different ways in which they grasp that content (or on the character under which that content is believed). According to Perry, the difference in their behavior can only be explained by the fact that each of them holds a different *de se* thought. Neither of the two wants Perry to be hurt, but that shared desire does not suffice, Perry argues, to explain the courses of action they choose. What explains Perry's friend's behavior is *his* own belief that *he* can run away to get help, while

Perry rolls into a ball because *he* thinks that by rolling up into a ball *he* will be safe. Both thoughts are *de se* because in both cases the indexical "I" cannot be eliminated or substituted by any other designator of either Perry or his friend.

Cappelen and Dever claim that the difference in their behavior is not enough to motivate us to introduce "new" *de se* thoughts or beliefs in our explanations of their actions. If we were to connect indexicality to action explanation that strongly, we should at least provide a better account of the connection. They think Perry does not meet the requirement. They take Perry to be defending that indexical terms are opaque in action explanation contexts, meaning that they cannot be replaced *salva veritate* in such contexts. More specifically, their reconstruction of the essential indexical thesis construes it as endorsing that which they label the "impersonal incompleteness claim". According to this claim, "Impersonal action rationalizations (IAR) are necessarily incomplete because of a missing indexical component".[199]

However, Cappelen and Dever hold that the claim about indexical opacity is a particular case of a more general phenomenon of opacity in action explanation contexts. According to this "generic opacity claim", co-referential expressions, be them indexical or not, cannot be substituted *salva veritate* in those contexts.[200]

Now, if we assume, following Cappelen and Dever's suggestion, that indexicals are not particularly opaque in contexts of action explanation, how are we to explain the difference between Perry's and his friend's behavior? Cappelen and Dever think they can explain it by appealing to a simple fact: there's a difference in the actions available to Perry and his friend. To make sense of our actions without appealing to first-person judgements, they endorse what they themselves call the "action inventory model". According to this model, we all have an inventory of impersonal beliefs and desires which generate impersonal intentions which can, in turn, motivate our actions. But these will only move us into action if we are actually capable of performing the desired actions in the first place. On this picture, there is no need to resort to first-person beliefs.

They apply a similar line of reasoning to Perry's example. On their view, both friends want John Perry to be safe. But Perry's friend cannot make Perry roll up

---

[199] CAPPELEN, H.; DEVER, J. 2013, p. 37. They explain in a footnote that although the examples considered involve first-person indexical thoughts, the point generalizes over other indexicals (or "locating" terms), such as "here", "there", "now", "we", etc.
[200] *Ibid.*, p. 33.

into a ball and stay still. Given the situation at hand, the only option available to Perry's friend is running away to get help, while Perry has the choice of rolling up into a ball. That should be enough to explain why Perry's friend behaves the way he does (why he runs away to get help). Now, Perry wants John Perry to be safe. Perry will be safe if Perry rolls up into a ball. Perry can do that, and so does. Supposedly, no *de se* belief needs to be involved.

As Stephan Torre rightly points out, however, the view that the availability of actions alone would explain the behavior of the two friends seems implausible because in any case they also have different beliefs about the actions available to them. Those beliefs must include self-beliefs like Perry's belief that *he himself* is being attacked by a bear and that *he* has the chance of rolling up into a ball to be safe. In an interesting variation on Perry's thought experiment, Torre asks us to imagine that Perry's friend has magical powers without being aware of it. He could cast a spell to make Perry roll up into a ball and be very quiet, so that he can perform the intended action, namely make Perry safe. But if he is not aware that he can do it, and thus cannot think he can do it, he will *not* do it. The point is that we do not act solely on the basis of the availability of some actions to us. We also act on the basis of the thoughts and beliefs we have about the actions that are available to us.[201] I think this idea complements some of the ideas defended by Ninan, discussed in the previous section. Even if we do not agree with Ninan's argument, it is still likely that in order to act we all must have at least the predisposition to have certain beliefs concerning what actions are available to us, even if those beliefs may remain implicit or unconscious.

3.4.
Immunity to error through misidentification

Evans deals with immunity to error through misidentification by the end of the sixth and the beginning of the seventh chapters of the *Varieties of Reference*. He claims that immunity to error through misidentification is not only a feature of first-person thoughts. He suggests that it is above all a characteristic of some types of demonstrative thought.

---

[201] Cappelen and Dever consider a similar argument on p. 54 and 55 of their book.

Chapter 6 of *Varieties* is dedicated to demonstratives. One of the conclusions drawn by Evans is this: in absolutely normal cases, demonstrative thoughts are what he calls "identification-free". Evans's idea is that being identification-free is essential for being immune to error through misidentification. Suppose I see in the distance a guy kissing someone else and take him to be my boyfriend. In this case, I will probably entertain a belief like "My boyfriend is cheating on me". But in such a case, before coming to entertain that particular belief ("My boyfriend is cheating on me"), I must identify the guy I see (of whom I think via some mode of presentation like "that guy") with my boyfriend. I thus make an identity claim (even if unconsciously) which could be expressed by a sentence like "That guy is my boyfriend". Only after that I come to believe that which can be expressed by "My boyfriend is cheating on me". The resulting thought is not immune to error through misidentification because it is not devoid of an identification component – before I come to the belief expressed by "My boyfriend is cheating on me", I go through an identity claim (equating "that guy" and my boyfriend). For Evans, normal cases of demonstrative thought do not involve that. If I grasp a thought such as "This pen is blue", I cannot, in principle, risk being mistaken in the identification of the referent of the demonstrative.

In order to understand the way Evans deals with this special characteristic of some kinds of judgments or thoughts (the characteristic of being "identification-free" and thus immune to error through misidentification), it is important to understand the Evasian notion of an Idea.

Evans sets out to investigate identification-free thoughts by analyzing what he calls "here-Ideas" and "this-Ideas", which according to him give rise to identification-free thoughts. *Grosso modo*, Evans's notion of an *Idea* stands for a *general* capacity to think of particular objects (e.g. places or material objects) exercized in a series of interconnected thoughts like the thought "that a is F", "that a is G", "that a is H", etc. Naturally, an individual can have different *kinds* of Ideas: he can have, for example, descriptive Ideas (when he is thinking about a given object via a description), demonstrative Ideas or indexical Ideas.

Evans distinguishes two kinds of knowledge, or two kinds of thought: one is free of identification and the other is dependent on it. Roughly speaking, a thought that would take the form "b is F" would be dependent on the identification of its referent if the judgement made by the thinker is involves two steps. Just like in the

boyfriend example, the thinker would first judge that "*a* is *F*" and then would entertain a thought of the form "*a* is *b*". The resulting identification-dependent judgement that "*b* is *F*" can be false because the subject who makes it might be mistaken in what concerns the identity represented by "*a* is *b*".

In the case of the so-called "here-thoughts" and "this-thoughts", the Ideas that sustain them are connected to certain ways of gathering information about the referents of the indexical or demonstrative in question. The relevant notion Evans uses here is that of an information link, a notion also used by other authors like François Recanati (who, by the way, embraces Evans's ideas on identification freedom, at least partially). For example, an Idea about "here" (a "here-Idea") is grasped in virtue of the subject's keeping to herself the information she gathers about the place she thinks about as the place she is at that moment. When a here-Idea is behind a subject's thought, it is connected to the place the thinker is occupying via an information link in such a way that any information concerning the possibility of a predicate's being instantiated or not is directly connected to the conception the subject has of that Idea. The result is the following: the subject in question will dispose of information about the predicate's being instantiated and her thoughts will not involve an identification component. A possible example of such a situation is the thought expressed by an utterance of "It is cold here". Such a thought usually does not depend on two judgements being formed, like "It is cold at *p*" and "*p* is here". Evans's idea seems to be that, in a case like this, when the information links that constitute the subject's here-Idea provide her with the information that the predicate "( ) is cold" is instantiated, such information is directly connected to the Idea. So, the subject is automatically disposed to think that *it is cold here*. This would not happen if the thinker were seeing an apparently cold place through a monitor screen and had to judge, based on the information received through the monitor, that the place she was seeing was also the place where she was. In this case the information link with the individual through the monitor would not be one of the links that constitute her here-Idea.[202]

Evans compares his own notion of identification freedom with Sydney Shoemaker's famous notion of immunity to error through misidentification.[203] Both Evans and Shoemaker think that it is essential for the notions they develop that the

---

[202] *Cf.* EVANS. G. 1982, p. 179-181.
[203] *Ibid*,., p. 188.

two following kinds of thought be distinguished: thoughts whose formation involve a judgement of identity about the referent and thoughts that do not involve such a judgement. The difference is the criterion that the authors use to differentiate those thoughts. *Grosso modo*, according to Shoemaker, the criterion is the following: the judgement "*a* is F" is immune to error through misidentification whenever it makes no sense to raise the following question: something is F, but "Is it *a* that is F?" Such as it has been formulated by Shoemaker, the criterion is not sufficient for Evans, since it eliminates some cases which, in Evans's view, should not be eliminated (should be taken as immune to error through misidentification of the referent). Shoemaker's criterion eliminates all cases in which it is possible to raise the question as to the correctness of the applicability of the predicate to the object (or place) the judgement is about. Evans seems to want to eliminate only cases in which the correct applicability of the predicate can be questioned because there is an identity clause which intervenes and which can be put into question (such as the clause "That guy is my boyfriend" in the example given above).

Evans suggests the following reading of Shoemaker's criterion:

> What we should say is that a judgment is identification-free if it is based upon a way of knowing about objects such that it does not make sense for the subject to utter 'Something is F, but is it a that is F?', when the first component expresses knowledge which the subject does not think he has, or may have, gained in any other way. A way of capturing the point of this revised criterion is this: the utterance 'Something is F, but is it a that is F?' needs a special background, in the view, of course, of the person who utters it; he has to suppose that the knowledge expressed in the first component was not gained, or may not have been gained, in the way with which the Idea involved in the second component is associated. If the situation is perfectly normal, and the subject does not take it not to be normal, the utterance does not not make sense.[204]

Evans seems to ground his argument on this: the subject's informational system must be working perfectly – that is one of the pressupositions for the application of his criterion. Following Evans, it does not matter for our account of normal cases that the object of my (demonstrative) thought can be replaced without my being aware of it, perhaps because my perception is somehow deceived. This constant possibility of my perception's being deceived is not sufficient for coming to the conclusion that in normal cases demonstrative thoughts are immune to error through misidentification. In normal cases demonstrative judgements do not normally come about as a result of an identity claim involving the referent of the

---

[204] *Ibid.*, p. 189-190.

demonstrative. Evans thinks that something similar happens in the case of I-thoughts. And I think this can be problematic.

As I tried to explain in the previous chapter, while it is true that in general *de se* thoughts are immune to error through misidentification because of the way our world is, they do not seem to be immune in a *de jure* way. This idea is not so far from what Cappelen and Dever argue for in *The Inessential Indexical*. They argue against the idea that if immunity to error through misidentification is manifested in certain types *de se* thoughts, that does not give us a sufficient reason to believe that those thoughts have a special epistemic status. Since the immunity is not a feature of all kinds of I-thoughts, it cannot be used to establish the unique status of *de se* thoughts *on the whole*.

Examples of immunity to error through misidentification seem to have led most authors to think there is a connection between that kind of immunity and indexicality. But the phenomenon of immunity, Cappelen and Dever argue, is not in any way philosophically interesting.

They set off to prove their point by claiming that the most interesting way of characterizing immunity to error through misidentification is by appealing to epistemology. If we do that, a possible way of explaining the phenomenon would be this: an utterance which is immune to error through misidentification has a certain epistemic privilege by virtue of which that kind of error is "impossible, in the sense of *a priori* ruled out".[205] However, Cappelen and Dever claim that no judgement whatsoever can be epistemically privileged that way. We do not need to consider science-fiction cases in which a subject has been wired to someone else to receive stimuli from that person since these are not actualized in our world. There are other real possible cases in which a subject comes to think *de se* thoughts that do not have herself as the source. We have already mentioned them: an example is the phenomenon of "thought insertion", common to some schizophrenic patients.

A possible explanation of problematic cases for the supporter of immunity to error through misidentification is that the subject comes to own the states that do not have herself as the source – she does that by taking them to be her own. But Evans has a different approach. He suggests that in cases such as those of "deviant wiring", any thought the subject could have would be a "shot in the dark", since she

---

[205] CAPPELEN, H.; DEVER, J. 2013, p. 130.

would not be in a position to know who is the person whose stimuli she is receiving, if those are real or even if there really is anyone who is experiencing something in the first place. The problem is, as Cappelen and Dever rightly point out, that if we try to justify the existence of immunity to error through misidentification on epistemological grounds, then we face the challenge posed by the fact that the information channels are never completely reliable. As they put it, "the guarantee of immunity is only as strong as the guarantee that proprioception (for example) only ever represents the proprioceiving agent".[206] So, the conclusion seems to be that *de se* thoughts are not *a priori* immune to error through misidentification. They are at best immune *de facto*, since in normal cases in our world as it is they present that immunity. Now, if immunity is not a distinctive feature of I-thoughts (as can be reasonably assumed), so what is? I think we can take the information channels or senses which feed our I-thoughts (and our I-thoughts only) as initial evidence to consider our I-thoughts unique.

3.5.
Conclusion

I think *de se* thoughts do have a special status *qua* indexical thoughts, and their status seems to be at least up to a point similar to that of what we may call "here" and "this-thoughts". *De se* thoughts seem to play an essential role in action explanation, and in that respect they pose a problem to traditional theories of thought and attitudes. *De se* skeptics are right in claiming that the "arguments" put forward by Perry and Lewis in the form of thought experiments are not strong enough for proving the truth of the essential indexical thesis. But I think they give us a good place to start our investigation and look for a more complete account of the characteristics of *de se* thoughts. Besides, I think *de se* skepticism, even in its most radical form (such as it has been defended by Cappelen and Dever), fail in its turn to provide us with knockout arguments against the philosophical significance of the *de se*.

In my opinion, *de se* thoughts are grounded on a distinctive epistemology, because the way we think of ourselves is grounded on several very distinct information channels. Some of them are supposed to give rise to thoughts which

---

[206] *Ibid.*, p. 132.

are immune to error through misidentification of the referent (in the case of *de se* thoughts, oneself). This is an idea defended by Evans, for example. But I think we can agree with both Evans and *de se* skeptics like Cappelen and Dever in that immunity to error through misidentification is not a feature belonging solely to first-person thoughts. It is clear that not all kinds of *de se* thought possess it, and even the idea that some of them do is controversial. At the same time, I think in normal cases *de se* thoughts can be considered immune to that error *de facto*, and we can still think that a cognitive mechanism takes care of their actually being so immune. If that mechanism fails, as it can, our I-thoughts will be vulnerable to mistakes in the identification of their referents (ourselves).

That does not change the fact, I think, that I-thoughts constitute a particular set of thoughts, with a functioning and characteristics which differentiate them from related mental acts (other kinds of self-locating beliefs). The way we gather information about ourselves and articulate first-person judgements is unique in making use of various kinds of information links, some involved in other forms of thought, some unique to the first-person. This is explained by Evans and, after him, Recanati with his theory of mental files. I shall return to that theory in the next chapter, in which I explore its details and also claim that some features of *de se* thinking are not fully captured by theories of indexicality, even ones, like Recanati's, that account for a purely mental or conceptual level of meaning.

4
Beyond Indexicality

> *Former un jugement égologique, c'est donc certes penser à quelque chose et y penser d'une certaine manière, mais cette pensée est aussi, pour le penseur, un mode d'être ou une certaine manière d'exister que l'on peut appeler "subjective".*[207]
>
> Stéphane Chauvier, *Dire "Je"*

> *Le problème de la dynamique cognitive qui nous intéresse ici concerne la définition appropriée de la continuité interne, c'est-à-dire la capacité de saisir continuellement le même sens ou concept.*[208]
>
> Jérôme Dokic, *L'esprit en mouvement*

4.1.
Introduction

The reader may not be convinced by what I said in the previous chapter against *de se* skepticism. So, in this chapter my main goal is to argue that, although *de se* thoughts can and should be seen as a special kind of indexical thought – in this, I hold fast to my position against Cappelen and Dever, some of their peculiarities cannot be wholly captured by a theory of indexicals. So, even if you do not agree with me and favor some form of *de se* skepticism, I suggest *de se* thoughts are not only indexical and self-locating, and in the context of an account

---

[207] So, to form an egological thought is certainly *to think* of something and to think of it in a certain way, but that thought is also, for the thinker, a mode of *being* or a certain way of *existing* which we can call "subjective".

[208] "The problem of cognitive dynamics which interests us here concerns the proper definition of our internal continuity, that is, our capacity of continually grasping the same sense or concept".

of how they are apprehended through time, it is essential that we explain their non-indexical charater.

As claimed in chapter 2, in order for the dynamics of *de se* thoughts to be fully explained, the problem has to be reformulated beyond its original scope. If that is the case, we can at least ask if cognitive dynamics as a whole does not extend beyond its original scope – that is, beyond the study of indexical thoughts – to a dynamist investigation of all kinds of thought.

So, these are my two chief aims in this chapter. First, I want to show that the thoughts we have about ourselves are not only indexical and that this fact should be taken that into account when explaining their dynamics. Second, this granted, I suggest we understand cognitive dynamics as the broader study of our internal (mental) continuity more generally, that is, involving indexical and non-indexical thoughts alike.

Section 4.2 is dedicated to justifying the idea that a theory of indexicality is not enough to explain all the characteristics of *de se* thoughts. First, in Sec. 4.2.1, I briefly discuss the indexical model for thought proposed by Recanati (already introduced in chapter 1), and ask if there is any characteristic of *de se* thoughts that is not captured by such a theory of mental indexicality. Then, in section 4.2.2, I try to identify the many different ways in which a subject can think of herself, that is, the different possible types of *de se* thought one can have. This will go beyond the preliminary list we have established in chapter one. In the process, I list and discuss some of the possible ways of grasping *de se* thoughts that do not involve indexicality and, thus, escape a theory of indexicality, whether old or new.

In section 4.3, I return to the dynamics of *de se* thoughts, keeping in view all that has been said about the non-indexical aspects of that kind of thought. I suggest we approach the retention of *de se* thoughts over time taking into account their non-indexical features.

Section 4.4 is dedicated to a discussion of "cognitive dynamics" understood as having a broader scope when compared to the way the phenomenon is usually addressed in the literature. I suggest we approach the phenomenon as the internal continuity of human intentional states. Drawing on a short passage from Frege's "The Thought", Kaplan pinpointed one of the central problems of cognitive dynamics as being related first and foremost to indexicals. I shall claim that nothing forbids us to see the phenomenon as having larger scope than the one given to it by

Kaplan. The alternative perspective on the dynamics of thinking presented in this section is mainly inspired by the work of Dokic who elaborated himself on the work of Evans.

4.2.
The non-indexical aspects of *de se* thoughts

4.2.1.
Further kinds of *de se* thoughts and levels of explanation

There are several kinds of *de se* thoughts. Wittgenstein famously distinguished between thinking about oneself "as subject" and "as object",[209] which gave rise to the investigation of what Shoemaker called "immunity to error through misidentification". Perry distinguishes I-thoughts about "the person I happen to be" (as in "The shopper with a torn sack is making a mess" when I happen to be that shopper) from I-thoughts that really characterize "genuine" *de se* attitudes (as in "*I* am making a mess", when I realize that *I* am the shopper with a torn sack). Recanati grants these distinctions, and refine them even further, speaking of *de se* thoughts of the non self-conscious kind, coined "accidentally *de se*", and self-conscious ones, coined either "explicitly *de se*" or "implicitly *de se*".

I take it that these distinctions are significant and that further ones can still be drawn. But there may be a sense in which we talk of "*de se* thinking", not captured by those distinctions, that hinges upon the kinds of epistemic links we have to ourselves. With respect to this issue, I shall draw on ideas put forward by Chauvier.[210] To begin with, take the idea behind Recanati´s notion of so-called "implicitly *de se* thoughts". Evans would articulate the notion saying they are identification-free because when I grasp them, I am not represented in thought as a component that can be identified. Or as Recanati puts it, when I entertain an implicit *de se* thought, I am not represented in my thought content as a constituent.

Such an idea (or related ones) has led some authors like Wittgenstein to claim that a more appropriate way to express these thoughts would involve the elimination of the first-person pronoun. On this view, the pronoun "I" would be kind of a

---

[209] WITTGENSTEIN, L. 1958, p. 66-67.
[210] CHAUVIER, S. 2001.

"linguistic fiction". So, instead of saying "I am exhausted", for instance, I should say something like "There is exhaustion".[211]

Now, we may believe that if there is something which is not the object of *apperception*, it cannot be part of an egological thought (a *de se* thought).[212] If that is so, we are led to completely discard apperception in those cases. But that is not necessary.[213] Chauvier suggests we grant (at least, provisionally) the so-called "thesis of the elusiveness of the self". According to this thesis, the "I" is always part of all our experiences, even when it has no correlate in the thought expressing those experiences (such as in cases of implicitly *de se* thoughts). Inspired by Kant, many consider that certain *a priori* concepts guarantee, or serve as conditions for the *objectivity* of experience. By analogy, Chauvier's account of the cases in which "I" is elusive (which may be understood as instances of implicitly *de se* thinking) suggests that the concept we have of ourselves serves as a condition for the *subjectivity* of experience.

Two things are noteworthy in that proposal. First, saying that the concept I have of myself serves as a condition for the subjectivity of my experience should not be equated with saying that the word "I" is necessarily connected to that special form of thinking. Sometimes words bring about certain ways of thinking, but what is important is the way of thinking itself, not the word contingently connected to it. Besides, there's no need to believe that to entertain a *de se* thought we must (always) have an epistemic or apperceptive relation to ourselves. According to Chauvier,

> […] an egological judgment, prior to being a certain theoretical act through which we become aware of a certain fact, is a certain event in the Being, an event which comes about in the mind of a subject and through which the subject becomes aware of what he is or of what he does.[214]

It is customary to draw a distinction between a thought endowed of a certain content, such as the thought to the effect that a certain event is occurring (*e.g.,* the thought *that* I am having a headache), and the event itself (my having the headache). But the very thought is, so it seems, a further event. Most of the time the distinction

---

[211] We can identify the influence of Lichtenberg and Hume on Wittgenstein's 'egoless' view on I-thoughts.

[212] "Apperception" is a term used by classic philosophers such as Descartes, Leibniz and Kant. The latter used it in more than one sense, but the sense which matters to us here involves self-awareness.

[213] *Ibid.*, p. 13.

[214] *Ibid.*, p. 15. The original in French reads as follows: "[…] un jugement égologique, avant d'être un certain acte théorétique par lequel nous prenons conscience d'un certain fait, est d'abord un certain événement dans l'Être, un événement qui se produit dans l'esprit d'un sujet et par lequel celui-ci prend conscience de ce qu'il est ou de ce qu'il fait".

above is suitable. But it should not be uncritically applied to *de se* thoughts. Chauvier argues (I take it, rightly) that *de se* thoughts are special because the thought is also the event. In the latter sense, to entertain a thought of the form "I am *F*" is part of what it takes for me to be *p*. Grasping that thought would be part of the overall subjective event that very thought consists in.[215] This is, I take it, a very interesting insight, although it might be said more prudently that the equation between my thought and the event in which it consists is not, in the case of *de se* thoughts, the whole story. The equation only partially explains *de se* thinking. At another level of explanation, I-thoughts should see them as *part* of the experiences they describe or events they consist in.

Chauvier argues that the source of our so-called "egology" does not lie in the informational structure our I-thoughts are based on – be it mental (accessed with the help of introspection) or corporeal (accessed through our senses). "The 'I' is not derived from the informational component of our experience", he claims, although we do possess information links to ourselves, and these feed the content of our *de se* thoughts. I take Chauvier's insight to be that the information accessed in a variety of ways is not the whole story about *de se* thoughts. Even though *de se* thoughts have information-based, their "egological articulation", as Chauvier puts it, cannot be explained by their informational base. Again, I believe we need not be as radical as Chauvier, but there surely is an interesting insight here. Of course, *de se* thoughts are an instance of (what Evans calls) "*information-based* thoughts", but Chauvier is right when he claims that the informational basis does not wholly explain what it takes for them to be *I*-thoughts.

In order to be brought about each different type of *de se* thought must involve a distinct type of cognitive operation. In Chauvier's opinion, the "*summa divisio*" of "egological" thoughts is that between egological thoughts about current (or occurring) properties, on the one hand, and egological thoughts about permanent properties, on the other hand. The former kind seems to involve self-awareness or apperception, for in order to think those properties, I need to perceive them directly. The latter kind take the form of utterances with a conceptually identifiable nature, such as "I am a nice guy". These constitute a sort of "quasi-concept" of ourselves.[216]

---

[215] *Ibid.*, p. 15. In reality, Chauvier's account of *de se* thoughts is part of an *ontology* of events.

[216] *Ibid.*, p. 26. Chauvier's own example of this kind of thought about oneself is this: "je suis un garcon courageux".

It is noteworthy that so far, among the authors discussed, Chauvier is the first to speak of this kind of thought. Would thoughts of this kind be instances of implicit *de se* thoughts, according to the taxonomy put forward by Recanati? It is far from clear to me, but in any case, Chauvier deals primarily with the first kind of egological thoughts, the one involving some kind of perception. They can be, in turn, divided into three sub-kinds.
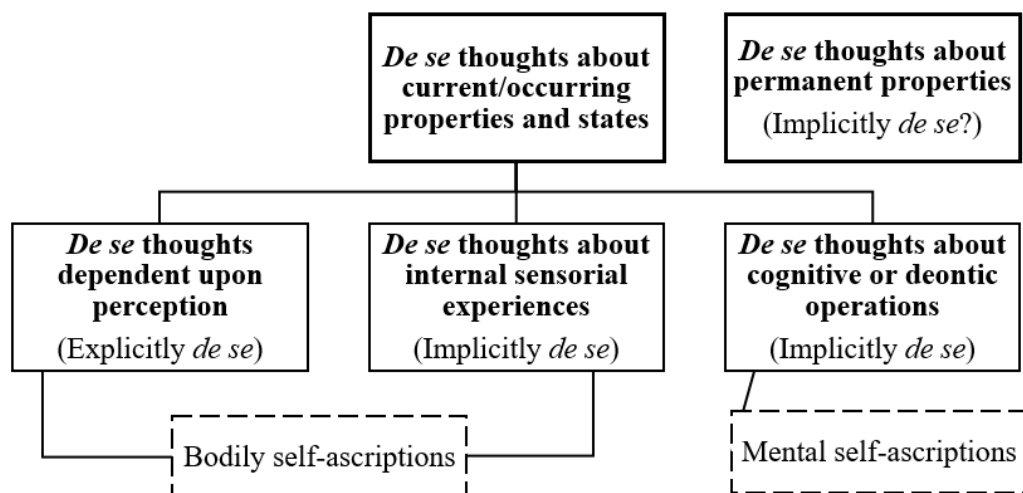
First, thoughts that are very dependent on perception. We could adopt an Evansian vocabulary and call these "self-ascription of location". They involve "external" senses, such as vision and touch, but they also involve "internal senses", such as proprioception. One example is the thought expressed by an utterance of "I am sitting opposite to him".

Second, thoughts about inner sense experiences. Call them "self-ascriptions of bodily properties", if you want. But it should be kept in mind that they are *special cases* of bodily self-ascriptions, for they involve senses such as proprioception and interoception more generally. Possible examples are thoughts expressed by utterances of "I'm in pain" or "I raise my left arm".

Third, thoughts about cognitive or deontic operations. They might also be coined "self-ascriptions of beliefs, desires and the like", *e.g.* the thought expressed by an utterance of "I want to finish this work". These thoughts do not seem to require of us that we perceive that we are thinking or wanting something. In this sense, they seem to be perfect illustrations of Chauvier's idea that grasping a *de se* thought is just part of the overall subjective event in which the thought consists.[217]

It might be said that of these three subkinds of *de se* thoughts about current or occurring properties, the first two ones are instance of what Evans calls "bodily self-ascriptions of properties". The third one would be an instance of mental self-ascription. The first could be taken to be explicitly *de se* (according to Recanati's terminology), for they involve "external" senses such as vision. The second and third ones would be implicitly *de se*. Overall, the new taxonomy is roughly as follows:

---

[217] For Chauvier's own presentation and discussion of these different kinds of *de se* thought (which is slightly different from mine), see *ibid*., especially p. 25-28.

```
┌────────────────────────┐   ┌────────────────────────┐
│ De se thoughts about   │   │ De se thoughts about   │
│ current/occurring      │   │ permanent properties   │
│ properties and states  │   │ (Implicitly de se?)    │
└────────────────────────┘   └────────────────────────┘
```

┌──────────────────┐ ┌──────────────────┐ ┌──────────────────┐
│ De se thoughts   │ │ De se thoughts about │ │ De se thoughts about │
│ dependent upon   │ │ internal sensorial   │ │ cognitive or deontic │
│ perception   │ │ experiences   │ │ operations   │
│ (Explicitly de se)│ │ (Implicitly de se)   │ │ (Implicitly de se)   │
└──────────────────┘ └──────────────────┘ └──────────────────┘

┌ ─ ─ ─ ─ ─ ─ ─ ─ ─ ┐   ┌ ─ ─ ─ ─ ─ ─ ─ ─ ─ ┐
   Bodily self-ascriptions   Mental self-ascriptions
└ ─ ─ ─ ─ ─ ─ ─ ─ ─ ┘   └ ─ ─ ─ ─ ─ ─ ─ ─ ─ ┘

Even though the three subkinds of *de se* thoughts about occurrent properties are based on information having the I as its source, this fact does not (fully) explain their "egological" character. We can and are justified in distinguishing between two kinds of judgement: inner-sense judgement – issued about the things we "perceive" in ourselves through interoception (or by introspection); and outer-sense judgement –typically issued about the things we perceive "outside of" us with the help of outer senses such as vision or audition. However, Chauvier sustains that inner-sense judgements should not be regarded as based on a kind of perception, for, as he puts it: "it is not in a perceptual way that we acquire information about ourselves".[218]

Our proprioceptive sense, while it plays a central role in our lives or maybe because of that, often goes unnoticed. Somehow similarly to the sense of smell that may disappear over time, proprioception gradually becomes unconscious. That is, as a matter of fact, a necessary evolutionary feature as the proprioceptive impressions must keep on working in the background while we focus our attention on other tasks.

The idea is to distinguish between perceiving (*percevoir*) and feeling (*ressentir*). Perception proper is characterized by our being in the presence of a certain property. Feeling is what characterizes *de se* states: it means being in a certain sensorial state – being the property, not perceiving it. Take pain, for instance. We have the impression that pain is located in specific parts of our bodies, Hence, our tendency to objectify them and treat them as perceptions. But, as

[218] *Ibid.*, p. 29. "Ce n'est pas de manière *perceptive* que nous acquérons des informations sur nous-mêmes."

Chauvier rightly points out, in reality nothing forces us to take the fact that pains seem to be located in our bodies to constitute a form of objectification of them: "We only have that impression because we add to the bare experience of a localized pain, not just the mental map of our body, but also our perception of that body as an object located in space".[219]

From that perspective, no object, apart from myself, is perceived by me when I find myself in a certain state like when I am feeling pain (either physically or psychologically). However, perceiving some state of the world surrounding me is always an intentional act. The "conceptual articulation of intentional information", Chauvier has it, does not occur in the case of sensory states like being in pain. If it did, we would be forced to say that we first feel the pain and then perceive the feeling of pain.

The idea that we can "perceive our own feeling pain", for instance, is not alien to the most widely accepted kind of theory of self-consciousness. That theory predicts that judgements and thoughts are always intentional and that intentionality must precede or ground the information which allows us to issue the judgements themselves. Chauvier suggests that first-person judgements could be explained in a more "economical" way. In the case of *de se* thoughts, it makes sense to speak of both aspects. In any case, this economical account of the nature of I-thoughts goes in the way of explaining Wittgenstein's comparison of an utterance of "I am in pain" with a cry of pain. In the case of such judgements, there is a special relation between the judgement and the sensory state that it supposedly describes: the judgement is part of the state, prolonging it. In Chauvier's opinion, that feature of *de se* thoughts makes it possible for us to take them to be "immune to the risk of hallucination" (*immunisé contre le risque d'hallucination*). "If we needed to perceive our suffering to judge it, it would be at least logically possible for us to be mistaken in judging that we suffer", Chauvier argues. Given that "the judgement is internal to the state it expresses, its occurrence suffices to bear out the reality of that state".[220]

---

[219] *Ibid*., p. 37. "Nous n'avons cette impression que parce que nous ajoutons à la nue expérience d'une douleur localisée, non seulement la carte mentale de notre corps, mais notre perception de ce corps comme un objet dans l'espace."

[220] *Ibid*., p. 41. "S'il fallait apercevoir sa souffrance pour en juger, il serait au moins logiquement possible qu'on se trompe en jugeant qu'on souffre. […] Le jugement est interne à l'état qu'il exprime, son occurrence suffit à attester de la réalité de cet état".

Chauvier is right, I take it, in claiming that the feeling of pain I may experience is real – it is there, even if I am hallucinating. Analogously, we can feel all sorts of things while dreaming. Again, at one level of explanation, the feeling seems real in those cases too. But from this it does not follow that we cannot be mistaken about the source of those states. What I mean is this: at some level of explanation, of course, it might be said that a subject's pains are the outcome of hallucinations. But I agree with Chauvier when he claims that, in spite of that, the feeling is real. Chauvier's conclusion is noteworthy: because the the occurrence of an egological state is a way of being, *de se* thoughts have special epistemic status; they are *part* of the states that justify them. They make up those very states and are, accordingly, a form of our consciousness.[221]

Now, if this is how we understand *de se* thoughts, it seems to me that no available theory of cognitive dynamics, not even the latest and most elaborated ones like Recanati's or Dokic's, would be of any use to account for their dynamics. The level of explanation of *de se* thoughts at which Chauvier's account operates is presumably beyond the scope of the theories of (linguistic or mental) indexicality and the available theories of cognitive dynamics. Perhaps Chauvier is right in identifying this extra level of explanation of *de se* thoughts (identifying them with events, with being in certain states). And if he is right, that level of explanation seems to be outside the scope of all available theories dealing with first-person thoughts (and indexical thoughts more generally), both in language and in thought. In the next section I shall discuss Recanati's theory of mental files, and hopefully the point I am trying to make here will be clearer. To be fair, as we have seen above, Recanati does speak of "implicitly *de se* thoughts", drawing inspiration from Evans. But I think Chauvier sheds a different light on that kind of *de se* thought. And he the aspect of *de se* thoughts he identifies cannot be ignored in a more complete account of *de se* thinking.

Perhaps the identification of *de se* thoughts with events (with being in certain states) can help us explain the essential character of those thoughts without appealing to the notion of indexicality. I think this is a natural consequence of Chauvier's line of reasoning. Indeed, he explicitly argues for the non-indexical character of *de se* thoughts: "the form of thinking which is connected to the use of

---

[221] *Ibid.*, p. 42. This presumably holds , at least, for so-called "occurrent" *de se* thoughts.

'I' cannot be completely assimilated with indexical thinking",[222] since the latter will always be a thought about a given object, in the sense of that object being represented as an identifiable component when we represent the thought itself. In this sense, Chauvier rightly characterizes indexical thoughts by calling them "intentionally referential thoughts".[223]

4.2.2.
Indexicality in thought?

Competing theories are available on the market to account for the phenomenon of indexicality in language. But one of the questions to ask concerns the scope of the phenomenon; is it purely linguistic or is it more emcompassing? On some understanding of the phenomenon, when contemporary philosophers of language talk about indexicality, they touch upon an issue that is in a way as old as philosophy itself, namely the issue of perspectivality.[224] If this is so, assuming indexicality is a philosophically interesting phenomenon, it might be conceded that we are talking about a phenomenon that goes way beyond language since it has implications for epistemology, psychology and other disciplines. Following that line of thought, it seems natural to ask whether there is a mental or psychological kind of indexicality, that is, whether the phenomenon encompasses *e.g.* how we perceive and think the world, I take it the answer is yes. And I suspect that the fact that indexicality is so widespread in human cognition and in our ways of expressing ourselves is by itself a sign of the essential dynamicity of human thinking. As Evans once argued, indexical modes of thinking of objects are necessarily dynamic.

In view of what has already been discussed, we can say that the study of cognitive dynamics is situated at the interface of the philosophy of language and the philosophy of mind (though, in my opinion, those areas of philosophical investigation are hardly separable). The problems raised by cognitive dynamics will not be settled (at least not wholly) if the solutions proposed concern exclusively or primarily language, or the mechanisms of expression of our thoughts. This is one of the lessons to draw from Kaplan's "Demonstrative" and the problem raised by

---

[222] *Ibid.*, p. 123. "La forme de pensée qui s'attache à l'emploi de 'je' ne peut être complètement assimilée à la pensée indexicale".
[223] *Ibid.*
[224] CAPPELEN, H.; DEVER, J. 2013, p 2.

Rip van Winkle. Thus, the transition from the linguistic to the mental realm seems to be absolutely necessary (unless you think like me they are not separate realms). We can hardly escape the need to take into account human psychology when dealing with the very process of thinking of anything whatsoever.

In this section I explore the possibility of making that move from language to thought in dealing with indexicality. I will briefly present Recanati's suggestion of accounting for mental indexicality through his theory of mental files. In spite of the fact that the mental files theory is very sophisticated and presents many positive points, I think it is not capable of accounting for *de se* thoughts in a completely satisfactory way. There is still something about *de se* thoughts which escapes the theory: the character of *de se* thoughts highlighted by Chauvier and discussed in the previous section.

Recanati claims that his mental files theory functions like an "indexical model for thought", which is a way of accounting for the phenomenon of indexicality in the psychological realm. In a way, the theory can be seen as an attempt to answer the fundamental issue of explaining what goes on in our minds when we use indexicals.

Remember the discussion in chapter 1. According to Recanati, mental files are mental devices for the storage of information. The type of the file varies according to the type of relation in which the subject stands to the object of his thought. That relation, whatever specificity it may present, is what Recanati calls an "epistemically rewarding" relation, because it typically involves gaining information from the object. Mental files are supposed to play, at the mental level, the roles played by Frege's notion of sense. First, they are meant to determine the reference of expressions through the initial act of acquaintance. They are also meant to explain the cognitive value associated with certain expressions. When a subject opens different mental files to store information about one and the same object, it is because she does not realize that the files are about the same object. Recanati calls this the "FC role" played by mental files, that is, the role they play in obeying Frege's Constraint (which is just the name Recanati gives to what we have already called Frege's Intuitive Criterion of Difference for thoughts). Mental files are also meant to explain the possibility of trading upon identity (abbreviated: TI), that is, the possibility for a subject's to go from premises like "*a* is *F*" and "*b* is *G*" to the

conclusion: "There exists a *x* which is *F* and *G*", without the need of any extra intermediary premise of identity (*a* is *b*).[225]

To see how the mental files theory can be characterized as an indexical model for thought, we just have to consider the fact that the mental files change depending on the changes in the (epistemically rewarding) relations between the subject and the object. Recanati argues that something very similar to the adjustment of indexical expressions in language happens at the level of thought – that is, at the level of mental files.

One of the first difficulties that a theory of mental indexicality like Recanati's mental files theory faces is this: the notion of conventional linguistic meaning (which in Kaplan's theory is represented by the notion of character) cannot be employed at the mental level. Mental files are not classified into types according to a conventional linguistic meaning. On Recanati's model, the mental file is typed by the type of epistemically rewarding relation it exploits. For example, each of us is capable of entertaining certain thoughts about him/herself because he/she stands in certain epistemically rewarding relations to him/herself that are very peculiar and are only available to him/her. By means of those relations, each of us can store information about him/herself in a specific file type: the *self* file.

It seems that one of the most important motivations for the development of an indexical model for thought based on, or similar to the existing standard model for language lies in the possibility of applying the distinction between type and token (of a type) not only to language, but also to thought. Besides, it seems that contextual relations play an important role in the two realms (language and thought). The idea is roughly this: when they are tokened, both mental files (types) and indexical terms exploit specific contextual relations to refer to some entity in reality. Now, if we are to apply in the realm of thought the indexical model originally designed for language, we need to find some kind of surrogate for the functional role of linguistic meaning (that is, for the *character* in Kaplan's terminology).

The mental surrogate for linguistic meaning is the funcional role of files. So, the central role played by context is preserved at the mental level, and so is the type-token distinction. To see this, let me sum up the features of Recanati's theory as

---

[225] RECANATI, F. 2016, p. 71-72.

follows: the primary function of mental files *qua* types is to store information obtained through a contextual relation holding between a mental-file token (of the mental-file type) and the reference of the term.[226] In Recanati's own words:

> opening a file of type α in a context *c* presupposes that there is, in *c*, a unique object *x* such that the subject stands in relation $R_α$ to *x* and, in virtue of standing in that relation to *x*, is able to gain information from it which the role of the file is to store.

So, the type of the file will hinge upon the type of relation in which the subject stands to the object she is thinking of, and the token of the file will typically exist as long as the contextual relation holds. A demonstrative-relation exploiting mental file, for instance, will only endure as long as the relation holds. Afterwards, the information is transferred to another file, of another type. We saw that idea in chapter 1. The requirement that a subject stand in a suitable epistemically rewarding relation to the object for a given file type to be tokened is normative. As Recanati puts it,

> there may be no object to which one is *R*-related, or the information channel which the file token exploits may put us in relation to several objects instead of exactly one. In such cases the tokening will typically be infelicitous; but the file will be tokened nonetheless.[227]

Now, if a subject tokens two mental files of different types, "Frege cases" are easily entertained in which the subject takes antagonistic attitudes towards the same object. For mental files to meet Frege's Constraint, they must obey a fine-grained individuation criterion. Examples of fine-grained files are demonstrative and memory files, because they are based on specific epistemically rewarding relations. "The subject can always wonder whether, e.g., the object he sees is the object he encountered in the past and remembers. In such a case there has to be two distinct modes of presentation".[228]

There is an intuition behind Recanati's theory that I take to be right. The split between the mental and the linguistic levels is only arbitrarily stipulated. It may be justifiable for theoretical reasons. However, it should not be taken too literally. Recanati's theory has the merit of trying to account for two important facts of our cognitive lives. First, language and the mind are intrinsically connected. Second, information is constantly being gathered, and concepts are created, stored, modified and retained through time. I think the most basic role of cognitive dynamics is

---

[226] *Idem*. 2012, p. 60.
[227] *Ibid*., p. 63.
[228] *Idem*. 2016, p. 72.

precisely to explain the unfolding of concepts, ideas, and thoughts *in our minds*. This occurs, to some extent, regardless of the way we express them in language.

4.3.
The dynamics of *de se* thoughts (again)

As suggested earlier, if cognitive dynamics is understood from the perspective of the standard Kaplanian-inspired view, the relevance of an investigation of the dynamics of *de se* thoughts is questionable. For one thing, in the case of first-person thoughts, there is no need to change the pronoun (the original token of "I") in the re-expression of the same thought after some period of time. Moreover, it can be argued that at least certain kinds of I-thoughts present the peculiarity of being immune to error through misidentification of the referent. This means that upon having such thoughts one can in no way be mistaken about their object (*i.e.* the reference of 'I' in context). So, given that it seems impossible for us to lose track of ourselves through time *and* that there is no need to adjust the main relevant indexical (the first-person pronoun) – these being the main issues raised by the standard kaplanian view of cognitive dynamics for indexicals, what is left for us to explain?

Remember that I argued in chapter 2 that there still *is* something left to explain since the mechanisms of apprehension and retention of *de se* thoughts over time are not free of failure and must be working properly if we are to say that there has been apprehension and subsequent retention. From the fact that their dynamics is *prima facie* automatic and effortless, it does not follow that there is *no* dynamics at play.

In chapters 2 and 3 we saw that the view that certain I-thoughts are immune to error through misidentification is controversial. I-thoughts grounded on proprioception, for example, can be at best *de facto* immune. To see that, there is no need to appeal to science fiction scenarios like the ones in which a subject is wired to someone else and receives proprioceptive stimuli from him/her. Our own world already comprises cases of so-called "thought insertion", a symptom presented by certain schizophrenic patients. Besides, there are related cases of impaired proprioception and interoception reported in the literature, which are not caused by mental disorders, but are themselves among the causes of certain

symptoms of psychological pathologies.[229] In face of such evidence, one might simply want to drop the "immunity to error through misidentification" or even "*de facto* immune" vocabulary, having in view that even I-thoughts based on senses such as proprioception or interoception are not really immune to that kind of error. I shall keep on using the term, nevertheless, to stay in tune with the literature on the topic, as I believe that all this points towards the view that the apprehension of *de se* thoughts is actually neither effortless, nor as simple as one might take it to be.

Something similar happens with the retention of I-thoughts. If we want to settle the issue as to whether a given first-person thought grasped (by me) at time *t* has been retained (by me) at a later time *t'*, one may naturally agree with Branquinho (1999) and say that memory took care of the retention. Remembering my I-thought at time *t'* seems to be enough to characterize retention. In that respect, the retention of an I-thought would be similar to the retention of an eternal thought: if at all times between *t* and *t'* I remember my I-thought (or if I am at least disposed to bring it into consciousness, possibly re-stating it), this simply means that I have retained it. So, in a way what happens in the case of I-thoughts is almost the opposite of what happens in the case of other indexical thoughts: if no need is felt to change the first-person pronoun in the re-expression of the thought, that is a sign of retention. If I do not wonder now whether I was mistaken in ascribing a certain state to myself (in the past) and do not wonder either whether I should replace now the token of "I" with a token of, say, "he" or some other expression referring to someone else, that is surely a sign that retention occurred. In a way, it is indeed as simple as that. At the same time, this should not be taken to mean, apparences notwithstanding, that no mechanism is responsible here for the retention of *de se* thoughts or, else, for the tracking of the *de se* thought's object over time – namely myself. It makes sense to say, following Evans, that the mechanism at play is straightforward and "effortless". On the other hand, the retention mechanisms involved in other kinds of thoughts (including temporal indexical thoughts) perhaps can be considered effortless as well in standard cases. This is so as we normally do *not* consciously exercise our capacity to keep track of the referent. Sometimes we

---

[229] WEIMER, AMY K. M. D. et al. 2001 provides us with evidence of proprioceptive deficit in Asperger Syndrome patients. SHAH, P. et. al. 2016 indicates that impaired interoception could be connected with a disorder known as Alexithymia, often found in people presenting symptoms of ASD, but not necessarily connected to autism.

do, in situations in which we run the risk of losing track of the object of one's thought. But this is unusual, for most of us have mastered the cognitive hability well enough to perform it without giving it a "second thought", so to speak. If I am right, this holds true for the competent-speaker's capacity to suitably update temporal indexicals as time passes.

This said, we make mistakes. With respect to *de se* thoughts, the problem of cognitive dynamics seems to be that of explaining how it is possible for an individual to retain or change his/her self-beliefs (or, for that matter, any kind of *de se* attitude) through contextual changes, particularly as he/she moves through time and space. This way of putting the problem is, of course, faithful to the tradition initiated by Kaplan. In that case, the problem is, as expected, connected to the question of understanding now "I"-sentences uttered in the past. Their understanding, particularly of the pronoun "I", seems to require of the subject that he/she use the (new) token of "I" along with the information he has at his/her disposal through some sort of information-storing mechanism like memory. One of the relevant issues here concerns the possibility of being mistaken on self-ascribing (mental or physical) predicates based on information available through memory. There exists some evidence concerning the possibility of such mistaken self-ascriptions.[230] Our *de se* thoughts are, accordingly, as open to error as the mechanisms of self-knowledge employed and the information-storing mechanisms we have at our disposal. All this points, I suspect, towards the existence of a mechanism analogue to that of reference tracking, mechanism that must be working properly for retention to occur in the case of *de se* thoughts.

These are, in a nutshell, the ideas entertained in chapter 2. Now I want to take a further step and claim that in order to account for the dynamics of *de se* thoughts, we must go beyond the Kaplanian formulation of the problem. The problem must be reformulated beyond its original scope primarily because *de se* thoughts do not require us to explain the same things. In that minimal sense, their dynamics is special. Still, there is a dynamics at play here too. So, perhaps cognitive dynamics itself should not be restricted to indexical thoughts. I shall explore this idea throughout the next section, taking as my guiding thread the following insight: in

---

[230] I am thinking about cases of thought inserstion (associated with schizophrenia) discussed above, as well as malfunctioning of our internal senses (usually associated with other mental disorders). I think we can also imagine cases of memory insertion, or of memory loss or confusion.

order to account for the dynamics of *de se* thoughts, it is necessary to deal with self-knowledge within the framework of an encompassing dynamist theory of our mental lives. As we shall see, this means (following Dokic's proposal in *L'esprit en mouvement*) that we need to take into account the subject's disposition to manifest certain patterns of behavior when she grasps thoughts about herself through time.

On the other hand, thinking about oneself is not merely a cognitive act (unfolded over time), but an ontological event that takes place *in* time – as seen earlier, following Chauvier. To self-ascribe a state is to *be* in that state, no matter who is the source of the state (myself or someone else). If, as earlier claimed, the "subjective quality" of the state is not captured by the accounts of indexicality available on the market, the question to ask is: is there some kind of cognitive dynamics (also) involved here? If so, how are we to account for it?

I do not have a detailed account of the dynamics of *de se* thoughts to offer here. But I think that a complete theory of the dynamics of that kind of thought pertains ultimately to a theory of personal identity. I suspect that one of the main functions of our internal dynamics is to give unity to the constantly evolving entities that we are.

4.4.
Cognitive dynamics beyond its original scope

4.4.1.
Introduction

I have argued that *de se* thoughts are more than instances of indexical thoughts, for some of their features outstrip a theory of indexicality. However, their dynamics is explainable. This suggests that the problem of cognitive dynamics should be reformulated beyond its original, Kaplanian scope. Working with a broader notion of cognitive dynamics, the dynamics of *de se* thoughts should naturally fit into the broader picture. Not only *de se* thoughts, but further kinds of thoughts too, even non-indexical ones. Let us see, in the sections below, if we can make sense of that idea.

My aim is to sketch an argument in support of the relevance of an investigation of mental states in general (and thoughts more specifically) from an

essentially *dynamist* perspective. One of the sources of inspiration for that idea is the work of Jérôme Dokic on cognitive dynamics,[231] which is, as far as I know, the only extensive work in the philosophical literature clearly and exclusively dedicated to the phenomenon.

4.4.2.
A dynamist take on intentional states and thoughts

Cognitive dynamics is generally understood as the study of how subjects relate to indexical thoughts over time and how they re-express them (and their relation to them – the epistemic states they find themselves in) at different times. In chapter 2 I traced the view back to Kaplan. To take a step further, we need to realize that *any* intentional act has its own dynamics. Let me flesh this out.

Cognitive dynamics is the study of how attitudes towards thoughts persist over time. We commonly call them "propositional attitudes", because they are directed towards propositions – just another name for Fregean *thoughts*. Those attitudes are also called *cognitive* or *epistemic*, and Kaplan himself speaks of *cognitive states*. They can, indeed, be regarded as states of the mind. And cognitive dynamics should be understood as the study of how they persist or change over time. I believe the idea should be taken even more seriously.

We commonly take belief to be a paradigmatic case of a propositional attitude. And belief can be view as a special kind of representation, just like other sorts of attitude directed to (the content of our) thoughts like desire, hope, doubt, etc.[232] Every belief is belief *in* something, it is a belief *that* something is such and such. Linguistic signs have the power of indicating/representing/pointing to things which are external to them. So do beliefs and other intentional states. The talk of intentionality is derived mostly from Brentano[233] (though it originally dates back to medieval philosophers[234]), and Brentano defended that the mental is characterized by intentionality: all that is mental is intentional, and all that is intentional is mental.

Following Brentano, the distinguishing feature of mental phenomena is the property of being directed onto something else. In his famous *Explaining Behavior*,

---

[231] Already mentioned: DOKIC, J. 2001.
[232] DRETSKE, F. 1988, p. 52.
[233] See, for instance, his *Psychology from an empirical standpoint*.
[234] See Peter Geach, *Reference and generality: an examination of some medieval and modern theories*, p. 181.

Dretske goes as far as defending that those intentional states in the mind have a more fundamental representational power than that of linguistic signs (the mind is the source of all intentionality). Be it as it may, the power of *intending to*, *pointing to* something else is the reason why we call those states "intentional". We can thus say that intentionality is not only a feature of language, but also of thought, even without entering the debate as to what kind of intentionality is more primitive (that is, the debate concerning its origin). In short, and refraining from giving any opinion on the origins of the phenomenon, we can say that when talking about intentionality, we are talking about the property of pointing to exterior objects. The phenomenon arguably pervades both the linguistic and the mental realms.

It is also clear to me that our intentional relations to objects (in the broadest possible sense) always take place in time and take *some* time. I agree, then, with Dokic when he claims that a dynamist treatment of intentionality has priority over what we may call a "static" account of intentional acts, which in turn abstracts away from the passage of time and more generally context changes. As he puts it, "the instantaneous apprehension of sense must be conceived through abstraction from the apprehension extended in time".[235]

A dynamist study of human intentional states in general would naturally include not only indexical thinking, as seen in chapter 2, but *thought* in general and even perception (though I shall not directly deal with perception here): in short, any exercize of the capacity to "point" towards external things. Thoughts are but a particular type of intentional act, although it is the main focus of our research here..

Now, there is just one more question I would like to raise. If cognitive dynamics understood more broadly is the study of intentionality, can we really talk of cognitive dynamics in the case of thoughts about oneself that are immune to the risk of hallucination, for example? Do they still involve an intentional relation if one defends something along the lines of what Chauvier view according to which that kind of thought is more like a form of being or existing than of knowing something?

Raising that question gives us a chance of exploring another feature of *de se* thoughts which, I think, distinguishes them from other kinds of thought: they seem to be both instances of intentional states (states directed towards an object, in this

---

[235] DOKIC, J. 2001, p. 88. See also p. 2.

case ourselves) and *also* as forms of being, as Chauvier puts it. Again, *de se* thoughts are unique in this respect: they unite characteristics which we do not usually see coexisting in other forms of thought. So, to answer the question posed above, at one level of explanation, I think we can regard I-thoughts as involving an intentional relation (against Chauvier). But there is another level of explanation in which it makes sense to say that entertaining certain *de se* thoughts just means being in certain states. I tend to agree with Evans when he defends that whenever we think of something, we always have to somehow identify the object of our thought (in other words, all thoughts obey Russell's Principle in some way[236]), but I also think Chauvier makes a good point, as we discussed in section 4.2. So, whenever I have a thought about myself, no matter the nature of that I-thought, there is an identification of the object of my thought, which happens to be myself. At the same time, at another level of explanation of our I-thoughts, we could claim that there is no intentional identification, nothing we identify.

Most importantly for us, there is no single unchanging object that persists over time.[237] Of course, that is not saying that nothing exists, or that we cannot identify anything constituting the whole set of evolving states and properties. That set must be a person. In any case, I think the dynamics of *de se* thoughts may have an important role to play in a theory of personal identity. Evans's idea that there must be an object which we identify when we think of ourselves is right even if that object is not fixed. And I think those ideas may not be incompatible with the existence of an aspect of *de se* thoughts which escapes all identification, as Chauvier claims.

---

[236] Evans calls "Russell's criterion" a test for indentifying truly referring expressions. The criterion establishes that an expression can only be a properly referring term if the object to which it supposedly refers actually exists. Besides, as far as Evans's reading of Russell goes, Russell believed that it is not possible to think of something unless one knows what particular thing one is thinking of. There are two ways of meeting that requirement: either by being directly acquainted with the object of one's thought, or by thinking of it as the unique satisfier of some description.

[237] One is reminded here of a thesis developed by David Hume (particularly in *A Treatise of human nature*, book 1, part 4, section 6.). Hume believes we have no impression of what he calls the "self". According to him, for us to have a clear and inteligible idea of something, we must have the impression of that thing. But when we turn our mind's eyes inwards (using introspection), all we find are sensations, thoughts and perceptions, never anything separated from them, like an entity without these impressions. A similar view is present (though not explicitly asserted) in Montaigne, in how he sees (what we may now consider to be) the subject of the first person in his *Essays*.

4.4.3.
The dynamist model for thought

Here I want to argue for the idea that *any* intentional act can be the object of a study from a dynamist perspective, provided we are interested in studying ist unfolding over time. This holds not only of indexical thinking, but of *thought* on the whole and also perception (though I shall not directly deal with perception here). For reasons stated above, I believe a dynamist account of mental acts has a priority over what we may call an "static" account of intentionality.

In this section I draw a contrast between the Kaplanian-inspired view of cognitive dynamics and what I take to be the sweeping issue of explaining how each particular thinking episode (or, more generally, intentional act) depends upon its evolution through time. What is at stake here is not necessarily some sort of diachronic criterion of identity for mental states. Perhaps the identity of those states would be determined by an account of their continuity, without the need of an independent identity criterion.

The dynamics of thinking is conspicuous in the case of indexical thoughts, for the expression of an indexical thought over time makes it clear that some kind of effort is needed to preserve one's relation to it . In this case, a proper adjustment of the indexical term is *part* of both the required effort and what goes on in the subject's mind. I think Evans was right in signaling that when we entertain an indexical thought over time, more occurs than the mere substitution of the indexical because the substitution presupposes a mental activity that has bearings on the expression of the thought in language. And I believe that, even when *no* proper linguistic adjustment is required (as it happens when the thought expressed is non-indexical or eternal), an explanation of the dynamics taking place *at a purely cognitive level must be provided.* This may even be a matter of internal (intentional) continuity, excluding e.g. external facts like (successful) reference.

As Kaplan (influenced by Frege's remark in "The Thought") understands it, the problem of cognitive dynamics concerns first and foremost the (re-)expression of indexical thoughts. Restriction to that kind of investigation can be reasonably justified. Explaining the dynamics of indexical thoughts is certainly one of the most important tasks of the cognitive dynamics theorist. But at the same time, someone who understands cognitive dynamics the way Kaplan did or deliberately chooses to

impose restrictions on her investigation may fail to appreciate the following fact: although it is multifaceted, the phenomenon concerns *human thinking* and our *intentional capacities* on the whole. The specific issue as to whether and how we are able re-express indexical thoughts over time is important and, perhaps, critical because the challenge of explaining the *dynamics of our cognition* is particularly dramatic in the case of indexicals. But even if a wholly satisfactory explanation were available with respect to the specific issue, this alone would not settle the overall issue of cognitive dynamics. I'd venture to say that no proper solution to the issues raised by cases like Rip's will be available unless one starts tackling the issue with the broader picture of the unfolding of mental states over time in mind.

Although Kaplan coined the term "cognitive dynamics", meaning both the dynamics of the cognitive significance attached to linguistic terms and the corresponding dynamics of the cognitive states brought about by the linguistic term's significance, I believe we should understand the adjective *cognitive* in *cognitive dynamics* slightly differently. I think we must understand cognitive dynamics a study of an important part of our *cognitive lives*. While the phenomenon was (partially) brought to light when Frege was investigating language and the peculiar behavior of indexical expressions, there is no reason to suppose that the study of cognitive dynamics does not have deeper roots. I think that on trying to account for it we should put more weight on an investigation of the mental realm.

As already pointed out, Evans was one of the first to undertake the kind of investigation needed more seriously as a way of accounting for cognitive dynamics. He tried to explain how the retention of indexical thoughts occurs over time, of course. But he laid more emphasis on the special kind of cognitive activity we need to exercise in order to retain that thought throughout the changing contexts. Evans argues that the kind of thought needed to even *understand* a sentence containing an indexical must be Fregean in the sense of involving a specific *way of thinking* about the object denoted by the indexical. And in the case of attitude retention or change (such as belief towards an indexical thought) over time, he claimed that it is necessary to exercise the cognitive ability of *keeping track* of the object in space and time.

The broader picture of the phenomenon of cognitive dynamics I shall argue for in this chapter is largely inspired by Jérôme Dokic's 2001 work, *L'esprit en mouvement. Essai sur la dynamique cognitive*, which in turn is largely inspired by

Evans. Dokic's main working hypothesis is that the best way to approach a study of mental acts is from a *dynamist perspective* whose premise is that one can only make reference to a real object (at any given time) in virtue of diachronic process of referring to that object throughout a certain period of time. Roughly speaking, the meaning of a given thought at a given time is dependent on its "dynamic meaning", *viz.*, its meaning taking into account the way it is grasped and the way it unfolds through changing contexts. I think this is not very distant from Evans's talk of "dynamic Fregean thoughts". I want to explore that influence here.

Dokic's dynamist model is clearly inspired by Evans: he elaborates on Evans's idea that the exercise of tracking abilities over time is a necessary condition for the retention of the indexical thought. While the application of the dynamist model to the case of demonstrative thinking is straightforward (for it was originally designed to model that instance of thinking), it should generalize over all kinds of thinking, even if not exactly in the same terms. According to Dokic, we can think of a more encompassing definition of the phenomenon of cognitive dynamics, and to that end "the notion of change of mind provides us with a possible departing point".[238]

Dokic sets general conditions for a change of mind. The first condition is this: if we are to say that an individual has changed his mind with respect to a certain thought, she must have *contrary* attitudes in relation to the thought at different times. In Dokic's words, "When a rational person changes her mind, first she has a certain epistemic attitude towards a given thought […] and later has a *contrary* attitude towards the same thought".[239] There may be a difficulty with the notion of rationality employed to account for change of mind in this case. According to the standard definition of Fregean sense, that notion should be transparent from a synchronic point of view. The idea is not unproblematic, but, problems aside, nothing forces us to take Fregean sense to be also *diachronically* transparent: "It is obviously possible to grasp the same sense at different moments without being aware of it".[240]

---

[238] *Ibid.*, p. 69. "La notion de changement d'avis nous fournit un point de départ possible".

[239] *Ibid.*, p. 69. "Lorsqu'une personne rationnelle change d'avis, elle a d'abord une attitude épistémique déterminée à l'égard d'une pensée […] et a ultérieurement une attitude *contraire* à l'égard de la même pensée".

[240] *Ibid.*, p. 71. "Il est manifestement possible de saisir le même sens à des moments différents sans le savoir."

The second condition for a change of mind is as follows: the subject must apprehend the same sense/thought continuously during the relevant period of time. According to Dokic, "if that condition is not satisfied, we cannot say that the subject has intentionally modified her epistemic attitude towards a certain thought".[241]

I think the most important idea to be extracted from the discussion is the idea that the problem of cognitive dynamics is not so much about mind changing as about internal continuity. And accounting for internal continuity is fundamental when we formulate a theory of meaning, given that we accept the idea that a theory of meaning should include a theory of understanding. We do not grasp senses or thoughts at once and in an eternal form. At a more fundamental level, a theory of cognitive dynamics should help us explain not the conditions for a change of mind, but the conditions for an enduring apprehension of thoughts.

In order to explain what it means "to continuously grasp a sense", one must distinguish two types of apprehension of a sense:[242]

A) Episodic apprehension: this kind of apprehension involves a certain mental *activity*, such as the act of judging a thought, for example, as true or false and expressing it in a sentence.

B) Dispositional grasp: it would be more appropriate to say that this kind of apprehension involves not a mental activity, but rather a capacity of apprehending a sense / thought.

If someone apprehends a thought episodically (judging it as true, for example), he/she must be able to apprehend it dispositionally, that is, she must be able to grasp it continuously in time. And, according to Dokic, for a subject to continuously grasp a thought over time, the dispositional capacity to apprehend the thought must be maintained without interruption. But the disposition can be "hidden", so to speak, in the sense that it may not be explicitly shown: "depending on the circumstances, the subject can actualize [the disposition] or not (for example, by letting others know about her thought)".[243]

Now, Dokic makes use of an idea of Evans's to argue that our ability to retain thoughts/concepts through time is a condition of possibility for our general capacity

[241] *Ibid.* "Si cette condition n'est pas satisfaite, nous ne pouvons pas affirmer que le sujet a modifié *intentionnellement* son attitude épistémique à l'égard d'une pensée déterminée".

[242] *Ibid.*, p. 71-72.

[243] *Ibid.*, p. 72. "[…] le sujet peut actualiser [la disposition] ou non (par exemple, en faisant part de sa pensée à autrui) selon les circonstances."

of reasoning.[244] Among other sources, Dokic uses Evans's observations on dynamic Fregean senses to establish what he takes to be the five principles of cognitive dynamics. Based on these principles, he sketches a "dynamist model for thought".[245]

The principles of cognitive dynamics outlined by Dokic are meant to at least partially define the notion of internal continuity.

The first principle is that which we may call the "principle of continuity". According to this principle, in order to explain how a subject grasps the same sense/thought/concept over time, it is not enough to say that she has grasped the same sense at different times, even if we stipulate that she has done so at every particular moment of the time stretch. According to Dokic, "the grasp of the same thought at two different times does not guarantee the grasp of that thought from one moment to the other".[246]

However small the lapse between the two times may be, a question about the continuity of the subject's grasp of the thought or concept in question can always be raised. Of course, we may ask ourselves to what extent it is reasonable to defend this. When I keep sight of a bird flying across the sky, is the eye blinking enough to wonder if I am still looking at the same bird? If we follow Dokic, it seems that the answer is affirmative, though the example of the use of a sense such as vision is not exactly equivalent to that of the retention of a concept. In any case, the most important point is this: a theory of cognitive dynamics is meant to give an explanation of the *diachronic relations* holding between the (successive) apprehensions of a thought at different times, not only with instances of apprehension of the thought at different times.

The second principle of cognitive dynamics is coined "the principle of ideal diachronic positioning".[247] According to this principle, for a subject to properly grasp the meaning of a term, she must be in an ideal epistemic position with respect to the referent. In this case, the meaning will be related to and dependent upon that position the subject occupies. And for the same meaning to be apprehended over a

---

[244] *Cf.* DOKIC, J. 2001, for example, p. 74, though the thesis can be found in many places throughout his work.

[245] For Dokic's complete discussion of the principles, *cf. Ibid.*, p. 77-83.

[246] *Ibid.*, p. 77. "La saisie de la même pensée à deux moments différents ne garantit pas la saisie de cette pensée d'un moment à l'autre".

[247] The names given to the principles here are not of Dokic's own coinage, though I draw from what he writes about them to give them these labels.

certain time stretch, the subject must remain in an epistemic position (or a series of epistemic positions) which is (are) at least equally ideal, in the sense of being properly related to the original one.

The second principle is a direct response to the first, which already stipulates that in order to explain the continuous apprehension of a sense through time, we need more than the simple apprehension by the subject of the same sense at different times. One suggestion to account for the mechanism described by the second principle is this: the subject must exercise a cognitive capacity to stay in the same ideal epistemic position in relation to the referent of the term whose sense she is grasping. As Dokic explains, the apprehension of a sense depends on a certain disposition on the part of the subject. And such disposition has to be maintained so that the subject can continuously grasp the sense in question. For it to be maintained, Dokic argues, the subject must continually adjust her relation to the semantic value of the sign – this allows her to retain the necessary ideal epistemic position. Besides, "we do not master a concept once and for all, we must rather keep it in our cognitive repertoire".[248] Dokic holds that the continuity that the cognitive dynamics must explain is an internal continuity in the sense that the thinker must somehow exercise a capacity to maintain the sense she has grasped in her mind, the idea being that in order to maintain an internal continuity the subject must be active. Continuity is not undergone passively by the subject.

According to Dokic, cognitive dynamics presupposes the existence of two levels of thought – the timeless synchronic level and the diachronic level of the apprehension of thought. This is what the third principle establishes. It is indeed an important step if our aim is to broaden the scope of cognitive dynamics. The idea here is that "the notion of Fregean sense is not enough to account for cognitive dynamics, [...] we must introduce a substantial notion of the *apprehension* of a sense".[249] In a way, this is obvious: we need some sort of explanation of a certain mechanism of thought retention which must be at work when a subject retains a thought through time.

---

[248] *Ibid*., p. 79. "On ne maîtrise pas un concept une fois pour toutes, mais on doit le mantenir dans son répertoire cognitif".

[249] *Ibid*., p. 80. "La notion de sens frégéen n'est pas suffisante pour rendre compte de la dynamique cognitive, [...] il faut introduire une notion substantielle de *saisie* du sens".

The fourth principle is coined the "principle of inferential determination". According to this principle, establishing the conditions of the grasp of the same sense through time help establishing the normative conditions of inference of a certain conclusion from a set of premises that employ the same sense. This functions as a sort of meta-principle about the other principles.

The fourth and fifth principles have the function of responding to a possible objection to the problem raised by the phenomenon of cognitive dynamics such as formulated by Dokic. The objection is this: in Dokic's formulation, the problem of cognitive dynamics would be a purely psychological problem, it would have nothing to do with semantics, since the scope of the phenomenon is that of the diachronic apprehension of thoughts.

In his justification of the fourth principle, Dokic alludes to Lewis Carroll's famous paper "What the Tortoise Said to Achilles".[250] Carroll's example shows that inferential reasoning is not simply something that imposes itself on us: it is also an activity. In Carroll's short story, the tortoise refuses to draw the conclusion of a given argument, even though she has good reasons for doing so (she understands and accepts the premises). If, unlike the tortoise, an individual had a good reason to draw a conclusion from a set of premises and did so, that reason could not be explained in the form of a proposition, which in turn could occupy the place of a new premise in the argument in question. For Dokic, reason lies in the way the subject grasps the premises of the argument. That process is necessarily extended in time. So, in a case such as that of the tortoise, at least two conditions should be met if we are to say that a subject has successfully drawn a conclusion from a certain reasoning (which necessarily happened over a certain period of time):

1. The conclusion must follow logically from the premises;

2. The subject must continually grasp the sense of the elements that appear more than once in the premises throughout the reasoning.

"It is in this sense that cognitive dynamics contributes to the determination of the conditions of justification of our reasonings", Dokic argues.[251]

---

[250] CARROLL, L. 1895. "What the Tortoise Said to Achilles". In *Mind, New Series*. Vol. 104, 416. Oxford University Press, 1995, p. 691-693. URL = <http://www.jstor.org/stable/2254477>.
[251] DOKIC, J. 2001, p. 82. "C'est en ce sens que la dynamique cognitive contribue à déterminer les conditions de justification de nos raisonnements".

The fifth principle of cognitive dynamics is coined "the principle of determination [of truth conditions]". According to this principle, the conditions of the apprehension of the same sense through time determine the contribution of that sense to the truth conditions of the thoughts of which they are a constituent part. The idea is roughly this: "upon explaining the conditions under which we grasp a thought through time, we already give a semantics of the sentence in question".[252]

Presumably the most important among those principles is the one that states that the apprehension of the meaning of a term over time occurs only if, throughout the time stretch, the subject maintains an ideal epistemic position in relation to the referent of the term.[253] For that to be the case, the subject must continually adjust its relation to the referent. The thinker must be sensitive to a group of possible transformations of the reality around the object, which must trigger in her specific "responses" that result in the appropriate cognitive adjustments. Dokic calls such responses "compensatory transitions": they allow the subject to compensate for the changes in reality so that an ideal epistemic position is maintained.[254] They allow, for instance, for the persistence through time of the information link the subject has with the object of his thought. We can imagine that the maintenance of an ideal epistemic position must happen also in the case of *de se* thoughts. We have already discussed how it can fail in cases of malfunctioning of our "internal" senses or troubles a subject may have with memory.

To summarize the dynamist model:

1. In order to grasp the same sense over time, one must grasp it at every instant of the time interval in question;[255]

2. To grasp the same sense at every moment is not enough: it is also necessary to appeal to the notion of transformation. In the case of senses of singular terms, one must be sensitive to the transformations of reality surrounding the referent of the term, and make appropriate compensations if any transformation happens.

The above-listed principles give the guidelines to any account of the subject's preservation of her ideal epistemic position in relation to the referent of a sign

---

[252] *Ibid*., p. 83. "En dégageant les conditions dans lesquelles on peut saisir une pensée à travers le temps, on donne déjà une sémantique de la phrase concernée".

[253] *Ibid*., p. 78.

[254] *Ibid*., p. 84-85.

[255] At least in the sense of being disposed to appropriately re-express it at any moment during the time interval in question.

whose sense she has apprehended. She has to maintain that position if she is to retain the sense grasped. If we want to stay within a Fregean theoretical framework, we need a dynamist model of the notion of grasping a Fregean sense. That should allow for the principles discussed above to be compatible with Frege's ideas. The first step towards a characterization of such a model is the establishment of what Dokic calls "an ideal epistemic position", as we have already discussed. Dokic defends that a subject will only stay in such a position if she achieves a a sort of epistemic stability. There are two conditions for achieving that stability:

1. The subject must be sensitive to a group of possible transformations of the reality surrounding the object of her thought (the referent of the term whose sense she is grasping). If a given transformation happens, it must trigger a response in the subject that prompts her to make a cognitive adjustment.

2. Depending on the transformations that actually take place, the subject "refocuses her intentional network", that is, she makes a transition that compensates for the transformations and allows her to remain in the ideal epistemic position described in 1.[256]

4.5.
Conclusion

If cognitive dynamics is to be understood as (roughly speaking) the study of the unfolding of intentional acts over time, and if the intentional object in the case of I-thoughts is the *self*, or the person I take myself to be, an account of my first-person thoughts presupposes an account of the relation with that intentional object that I happen to be over time. But when we discussed Chauvier's ideas, we saw that there is an aspect of *de se* thoughts that oustrips an explanation in terms of an intentional relation I should have to myself. Grasping a *de se* thought which consists in a certain state I find (or found) myself in. An account of the dynamics of *de se* thoughts must take that aspect of *de se* thoughts into account. I think that a theory of their dynamics in those terms would amount to a theory of personal identity.

Now, the kind of dynamics involved in *de se* thinking is very specific and goes beyond the scope of the problem of cognitive dynamics such as it has been originally formulated by Kaplan. Given that it is still possible to study the dynamics

---

[256] *Ibid.*, p. 84.

of those thoughts (as I have showed in chapter 2), in order for their dynamics to be fully explained, the problem has to be reformulated beyond its original scope. That being the case, we can at least ask if cognitive dynamics as a whole does not extend beyond its original scope – beyond the study of indexical thoughts – to a dynamist investigation of all kinds of thought. A suggestion of such an extension of the phenomenon is represented by Dokic's dynamist model for thought. I think that such a model could be used at least in its most general lines to initiate a study of the dynamics of several different kinds of thought.

If cognitive dynamics can have a large scope such as the one Dokic's model suggests, I think we have a reason to believe that a comprehensive theory of cognitive dynamics can only be elaborated by taking into account and analyzing more deeply all the particular dynamic processes of human thinking. That would include not just "*de se* thinking", but other sorts of thinking mechanisms as well, even those that are normally left out in a theory of cognitive dynamics in Kaplan's terms. I believe all processes of human thought share one feature: that of being dynamic processes. But I think each of them may present different characteristics that would render them unique.

5
Conclusion

Dokic sketches an account of the dynamics of I-thoughts in the specific case of the communication of those thoughts:[257] cases in which my *de se* thought expressed using the first-person pronoun "I" will be expressed by my interlocutor using the second-person pronoun "you". Dokic defends that the same sense can be expressed by "I" and by "you in both cases, and so my original *de se* thought can be expressed using different expressions to designate myself. He defends a hypothesis which he calls "the hypothesis of the deictic ego" (*l'hypothèse de l'ego déictique*), according to which

> The sense expressed by the pronoun "I" is the same as that which can be expressed by the deictic term "this person" in particular circonstances, which include the fact that "I" and "this person" designate the same person.[258]

Dokic's hypothesis is not free of controversy, as attested by the current research on the communication of *de se* thoughts.[259] But as I have tried to argue above (particularly in chapter 2), I think that accounting for the dynamics of *de se* thoughts in contexts of their communication is one problem. A different and relatively independent issue is that of accounting for the experience of my own (internal) continuity.

I tried to show that *de se* thoughts, usually expressed using the pronoun "I", are a species of indexical thought, but they raise problems of their own. And one of the reasons why they raise peculiar problems is this: although they indexical, they present characteristics which are not shared by other indexical thoughts. I-thoughts are different from other indexical thoughts in important ways.

We saw that Dokic draws inspiration from Evans to argue that the grasp of an indexical thought is based on the subject's capacity to keep track of the object about which she thinks in an indexical way. The exercise of this capacity is usually connected to perception: we keep track of the object in spite of the fact that our

---

[257] DOKIC, J. 2001, chapter 7, sections 10 and 11.
[258] *Ibid.*, p. 343.
[259] Take for example, the second half of GARCÍA-CARPINTERO, M.; TORRE, S. (Ed.) 2016.

(perceptive) relation to it changes as we move through time and space. The same does not seem to be necessary in the case of *de se* thoughts: "the subject who grasps an egological thought does not need to perceptively keep track of himself".[260]

We can try to draw parallels between the way we perceive and gain information about exterior objects and the way we "perceive" and gain information about ourselves. We have ways of gaining information about ourselves which are unique to the first-person perspective (proprioception, introspection, etc.). Arguably, when we grasp first-person thoughts, these special information links do not (always) play the same role as the information links we have to exterior objects when we grasp thoughts about these objects. The main difference between the two cases is the following: the capacity we have of keeping track on an object we think of indexically is fallible, while in the case of the first-person, it is not. In the case of "I", a substitution of the object of my thought without my being aware of it is not intelligible. As Dokic explains, "the capacity of keeping track of an object imposes itself precisely because this object is distinct from us",[261] so there is a permanent risk of that object escaping our grasp. Arguably, the person I am cannot escape my own grasp the same way. But we have seen that we can make mistakes when thinking about ourselves (there may be cases of thought insertion, malfunctioning of our internal senses, etc.). Perhaps the exercise of the capacity of "keeping track of the referent" just . Of course, in general possible mistakes we can make in ascribing properties to ourselves does not prevent us from grasping a sense which determine us,[262] or from building a concept which we identify with the persons we are. Be it as it may, the dynamics of *de se* thoughts still share an important characteristic with the dynamics of other kinds of indexical thoughts: both involve an experience of continuity in relation to their objects.

But we have seen with Chauvier that grasping a *de se* thought is also in a sense part of the event of being in the state the thought is supposed to describe. (At least for certain cases of *de se* thought, such as "implicitly *de se* thoughts", to use Recanati's terminology. But arguably that feature is present in all cases of *de se* thought, be them explicit or not.) Dokic arrives at the same conclusion. He writes:

[260] DOKIC, J. 2001, p. 345.
[261] *Ibid*. "La capacite de garder la trace d'un objet s'impose précisément parce que cet objet nous est distinct."
[262] *Ibid*.

The apprehension of an egological thought *qua* a determined mental act is itself a *component* of that experience. The subject who apprehends an egological thought becomes *ipso facto* aware of himself *by the fact that he apprehends that thought*.[263]

This is a difference at the level of apprehension of the thought. Dokic further argues that such a difference explains (at least in part) why the reference to ourselves when we grasp *de se* thoughts is *quasi infallible*.[264]

Following Dokic's line of thought, the experience of my own continuity when thinking about myself must be dynamic, because it is extended in time. It must include my physical and mental activities. Revisiting Perry's example of someone who is about to be attacked by a bear in the forest, Dokic argues that:

The transition from my judgement 'I should roll up into a ball' to the action of rolling up into a ball is reasonable because it is grounded on a dynamic experience of my continuitny which *includes* precisely that judgement and that action.[265]

The account of this kind of dynamics must be different from an account of the kind of dynamics involved in the communication of *de se* thoughts. The exact structure of that dynamics would probably take the form of a theory of personal identity.[266] Of course a theory of personal identity is a metaphysical theory. So, why should one of his tasks be that of explaining a cognitive phenomenon such as internal continuity? That will of course depend of the kind of theory of personal identity relevant here. I think it might take the form of a neo-Lockean theory of personal identity, in terms of psychological continuity.

---

[263] *Ibid*., p. 346. "La saisie elle-même de la pensée égologique, en tant qu'acte mental déterminé, est une *composante* de cette expérience. Le sujet qui saisit une pensée égologique prend *ipso facto* conscience de lui-même *par le fait qu'il saisit cette pensée*."

[264] *Ibid*.

[265] *Ibid*., p. 348. "La transition de mon jugement 'Je devrais me mettre en boule' à l'action de me mettre en boule est raisonnable parce qu'elle se fonde sur une experience dynamique de ma continuité qui *inclut* précisément ce jugement et cette action."

[266] *Ibid*.

Bibliography

Following the name of each author is the year of the first publication of his or her work. That is the year used for reference in footnotes throughout this work. When the year of the first publication diverges from the year of publication of the edition I used, the new year of publication is also informed by the end of each reference.

ASCOMBE, G. E. M. 1974. "The first person". In Guttenplan, S. (Ed.) *Mind and Language: Wolfson College Lectures*. Clarendon Press, p. 45-65.

AZZOUNI, J.; CRANE, T. 2011. "Singular Thought". In *Proceedings of the Aristotelian Society*. Supplementary volume LXXXV. The Aristotelian Society.

BACH, K. 1987. *Thought and Reference*. Oxford: Clarendon Press.

_____. 2010. "Getting a Thing into a Thought". In: Jeshion, R. (Ed.) 2010, p. 39-63.

BERMÚDEZ, J. L. 2003. "'I'-Thoughts and Explanation: Reply to Garrett". In *The Philosophical Quarterly*. Vol. 53, 212, p. 432-436.

_____. 2005. **"**Evans on the Sense of 'I'". In Bermúdez, J. L. (Ed.) *Thought, Reference and Experience. Themes from the Philosophy of Gareth Evans*. Oxford: Claredon Press, p. 164-194.

_____. Forthcoming. "Memory and self-consciousness". In Bernecker, S.; Michaelian, K. (Ed.) *Routledge Handbook to the Philosophy of Memory*.

BERNECKER, S. 2004. **"**Memory and Externalism". In *Philosophy and Phenomenological Research*. Vol. 69, 3, p. 605-632.

BOËR, S. E. 1989. "Neo-fregean thoughts". In *Philosophical Perspectives*. Vol. 3, *Philosophy of Mind and Action Theory*. Ridgeview Publishing Company, p. 187-224.

BRANQUINHO, J. 1999. "The Problem of Cognitive Dynamics". In *Grazer Philosophische Studien*, Vol. 56, p. 22-56.

_____. 2000. "On the Individuation of Fregean Propositions". In Kanamori, A. (Ed.) *Analytic Philosophy and Logic. The Proceedings of the 20th World Congress*

of Philosophy. Vol. VI. Ohio: Bowling Green, Philosophy Documentation Center, p. 17-28.

BROOK, A.; DEVIDI, R. C. (Ed.) 2001. *Self-Reference and Self-Awareness*. Advances in Consciousness Research. Amsterdam/Philadelphia: John Benjamins Publishing Company.

BURGE, T. 1977. "Belief *De Re*". In The Journal of Philosophy. Vol. 74, N. 6, p. 338-362.

BURGESS, N.; MAGUIRE, E. A.; O'KEEFE, J. 2002. "The human hippocampus and spatial and episodic memory". In *Neuron*. Vol. 35, p. 625-641.

CAMPBELL, J. 1997. "The Structure of Time in Autobiographical Memory". In *European Journal of Philosophy*. Vol. 5, 2, p. 105-118.

_____. 1999. "Immunity to Error through Misidentification and the Meaning of a Referring Term". In *Philosophical Topics*, Vol. 26, p. 89-104.

_____. 2001. "Memory Demonstratives". In Hoerl, C.; McCormack, T. (Ed.) *Time and Memory: Issues in Philosophy and Psychology*. Oxford: Oxford University Press, p. 169-186.

CAPPELEN, H.; DEVER, J. 2013. *The Inessential Indexical: On the Philosophical Insignificance of Perspective and the First Person*. Oxford: Oxford University Press.

CASSAM, Q. (Ed.) 1994. *Self-Knowledge*. Oxford Readings in Philosophy. Oxford: Oxford University Press.

CASTAÑEDA, H.-N. 1966. 'He': A Study in the Logic of Self-Consciousness. Ratio 8, p. 130-157.

CHATEAUBRIAND, O. 2001. *Logical Forms. Part 1. Truth and Description*. Campinas: UNICAMP, Centro de Lógica, Epistemologia e História da Ciência. *Coleção CLE*, Vol. 34.

CHAUVIER, S. 2001. *Dire "Je", essai sur la subjectivité*. Paris: Librairie Philosophique J. Vrin.

_____. 2003. *Qu'est-ce qu'une personne?* Paris: Librairie Philosophique J. Vrin.

CHRISTOFIDOU, A. 1995. "First Person: The Demand for Identification-Free Self-Reference". In *The Journal of Philosophy*. Vol. 92, 4, p. 223-234.

COLIVA, A. 2003. "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, 8, p. 416-431.

CORAZZA, E. 1995. *Référence, Contexte et Attitudes*. Montréal/Paris: Vrin/Bellarmin.

CORAZZA, E.; DOKIC, J. 1992. "On the cognitive significance of indexicals". In *Philosophical Studies*, 66, p. 183-196.

_____. 1993. *Penser en contexte. Le phénomène de l'indexicalité*. Combas: Editions de l'Eclat.

DOKIC, J. 1996. "The Dynamics of Deictic Thoughts". In *Philosophical Studies*, 82, p. 179-204.

_____. (Ed.) 1997. *European Review of Analytical Philosophy*. Vol. 2. *Cognitive Dynamics*. Stanford: CSLI Publications.

_____. 1997b. "La signification des expressions egocentriques". In Joëlle Proust. *Perception et intermodalité. Approches actuelles de la question de Molyneux*. Paris: PUF, p. 103-124. URL = <http://jeannicod.ccsd.cnrs.fr/ijn 00000270>.

_____. 2001. *L'esprit en mouvement: Essai sur la dynamique cognitive*. Collection Langage et Esprit. Stanford: Éditions CSLI.

_____. 2001b. "Is Memory Purely Preservative?" In Hoerl, C.; McCormack, T. (Ed.) *Time and Memory: Issues in Philosophy and Psychology*. Oxford: Oxford University Press, p. 213-232.

_____. 2012. "Indexicality". In Newen, A.; van Riel, R. (Ed.) *Identity, Language & Mind. An Introduction to the Philosophy of John Perry*. Stanford: CSLI Publications, p. 13-31.

DRETSKE, F. 1981. *Knowledge and the Flow of Information*. Cambridge/Massachusetts: The MIT Press.

_____. 1988. *Explaining Behavior*. Cambridge/Massachusetts: The MIT Press.

DUMMETT, M. 1973. *Frege. Philosophy of Language*. London: Duckworth.

_____. 1978. *Truth and Other Enigmas*. London: Duckworth.

_____. 1989. "More about thoughts". In Beaney, M.; Reck, E. H. (Ed.) *Gottlob Frege: Critical Assessments of Leading Philosophers*. Vol. IV. New York: Routledge, 2005, p. 330-350.

_____. 1991. "The Relative Priority of Thought and Language". In *Frege and Other Philosophers*. New York: Oxford University Press, p. 315-324.

EVANS, G. 1982. *The Varieties of Reference*. McDowell, J. (Ed.) New York: Oxford University Press.

_____. 1985. *Collected Papers*. Oxford: Clarendon Press.

EZCURDIA, M. 1997. "Dynamic and Coherent Thoughts". In Dokic, J. (Ed.) 1997, p. 105-139.

FODOR, J. 1975. *The Language of Thought*. New York: Thomas Y. Crowell.

FREGE, G. 1879. *Begriffsschrift, a formula language, modeled upon that of Arithmetic, for pure thought*. In van Heijenoort, J. (Ed.) *From Frege to Gödel. A Source Book in Mathematical Logic 1879-1931*. Cambridge/Massachusetts: Harvard University Press, 1971, p. 1-82.

_____. 1892. "On sense and Meaning". In McGuinness, B. (Ed.) *Collected Papers on Mathematics, Logic, and Philosophy*. Glasgow: Basil Blackwell, 1984, p. 157-177.

_____. 1893. *The Basic Laws of Arithmetic. Exposition of the System*. T.; Ed.: M. Furth. Berkeley/Los Angeles: University of California Press, 1964.

_____. 1906. "Letters to Husserl". In *The Frege Reader*. T.: Michael Beaney. Blackwell Publishing, 1997, p. 301-307.

_____. 1918. "The Thought. A Logical Inquiry". In *Mind, New Series*. Vol. 65, nº 259. T.: P. T. Geach. 1956, p. 289-311.

_____. 1918b. "Der Gedanke. Eine logische Untersuchung". In *Beiträge zur Philosophie des deutschen Idealismus*, 2, p. 58-77.

_____. 1979. *Posthumous Writings*. Ed.: H. Hermes; F. Kambartel; F. Kaulbach. T.: P. Long; R. White. Oxford: Basil Blackwell.

_____. 1980. *Philosophical and Mathematical Correspondence*. Blackwell Publishing, Oxford.

GARCÍA-CARPINTERO, M. 2013. "The Self File and Immunity to Error Through Misidentification". In *Disputatio*. Vol. V, 36, p. 191-206.

GARCÍA-CARPINTERO, M.; TORRE, S. (Ed.) 2016. *About Oneself:* De Se *Thought and Communication*. Oxford: Oxford University Press.

GARRETT, B. 1998. *Personal Identity and Self-Consciousness*. London/New York: Routledge.

_____. 2003. "Bermúdez on Self-Consciousness". In *The Philosophical Quarterly*. Vol. 53, 210, p. 96-101.

GEIRSSON, H. 2002. "Frege and Object Dependent Propositions". In *Dialectica*, 56 (4), p. 299-314.

GRIMALTOS, T.; MOYA, C. 1997. "Belief, Content and Cause". In Dokic, J. (Ed.) 1997, p. 159-171.

GRUSH, R. 2000. *Guide to Gareth Evans'* The Varieties of Reference. URL = <https://mind.ucsd.edu/misc/resources/evans/evansindex.html>.

_____. 2007. "Evans on Identification-freedom". In *Canadian Journal of Philosophy*. Vol. 37, 4, p. 605-618.

HARCOURT, E. 1999. **"**Frege on 'I', 'now', 'today' and some other linguistic devices". In Beaney, M.; Reck, E. H. (Ed.) *Gottlob Frege: Critical Assessments of Leading Philosophers*. Vol. IV. New York: Routledge, 2005, p. 262-287.

HOERL, C. 1997. "Cognitive Dynamics: An Attempt at Changing Your Mind". In Dokic, J. (Ed.) 1997, p. 141-158.

HOFFMANN, A. 2012. "Are propositions sets of possible worlds?" In *Analysis*. Vol. 72, 3, Oxford: Oxford University Press, p. 449-455.

HUME, D. 1739. *A Treatise of Human Nature*. Ed.: Ernest C. Mossner. England: Penguin, 1969.

_____. 1748. *An Enquiry concerning Human Understanding*. Ed.: Tom L. Beauchamp. New York: Oxford University Press, 1999.

JESHION, R. (Ed.) 2010. *New Essays on Singular Thought*. Oxford: Oxford University Press.

KAPITAN, T. 1999. "First-person Reference". In Castaneda, H-N. (Ed.) *The Phenomeno-logic of the I*. Indiana University Press, p. 3-16.

KAPLAN, D. 1977. "Demonstratives: An essay on the semantics, logic, metaphysics, and epistemology of demonstratives and other indexicals". In Almog, J.; Perry, J.; Wettstein, H. (Ed.) *Themes from Kaplan*. Oxford: Oxford University Press, 1989, p. 481-563.

_____. 1970. Dthat. In Cole, P. (Ed.) *Syntax and Semantics*. Academic Press, 1978, p. 221-243.

_____. 1989. "Afterthoughts". In Almog, J.; Perry, J.; Wettstein, H. (Ed.) *Themes from Kaplan*. Oxford: Oxford University Press, p. 565-614.

KORTA, K.; PERRY, J. 2011. *Critical Pragmatics, An Inquiry into Reference and Communication*. Cambridge: Cambridge University Press.

KRIPKE, S. 1977. "Speaker's Reference and Semantic Reference". In *Philosophical Troubles. Collected Papers, Volume 1*. Oxford/New York: Oxford University Press, 2011, p. 99-124.

_____. 2006. "The First Person". In *Philosophical Troubles. Collected Papers, Volume 1*. Oxford/New York: Oxford University Press, 2011, p. 292-321.

_____. 2008. "Frege's Theory of Sense and Reference: Some Exegetical Notes". In *Philosophical Troubles. Collected Papers, Volume 1*. Oxford/New York: Oxford University Press, 2011, p. 254-291.

LEWIS, D. 1979. "Attitudes *De dicto* and *De Se*". In *The Philosophical Review*. Vol. 88, 4, p. 513-543.

LOWE, E. J. 1989. "What is a criterion of identity?" In *The Philosophical Quarterly*. Vol. 39, 154, p. 1-21.

LUDWIG, P. 2005. "Variétés de la connaissance de soi". In Bourgeois-Gironde, S. (Ed.) *Les formes de l'indexicalité, langage et pensée en contexte*. Éditions Rue d'Ulm, p. 95-117.

LUNTLEY, M. 1997. "Dynamic Thoughts and Empty Minds". In Dokic, J. (Ed.) 1997, p. 77-103.

McDOWELL, J. 1984. "*De re* senses". In *The Philosophical Quarterly*. Vol. 34, 136, Special Issue: Frege, p. 283-294.

_____. 2005. "Evans' Frege". In Bermúdez, J. L. (Ed.) *Thought, Reference and Experience. Themes from the Philosophy of Gareth Evans*. Oxford: Clarendon Press, p. 42-65.

MEACHAM, C. J. G. 2008. "Sleeping Beauty and the Dynamics of *De Se* Belief". In *Philosophical Studies*. Vol. 138, p. 245-269.

MILLIKAN, R. G. 2000. *On Clear and Confused Ideas: An Essay About Substance Concepts*. Cambridge: Cambridge University Press, 2004.

MONTAIGNE, M. 1580. *Essais*. Ed.: Pierre Villey. Presses Universitaires de France, 1965.

MORGAN, D. 2009. "Can you think my 'I'-thoughts?" In *The Philosophical Quarterly*. Vol. 59, 234, p. 68-85.

NAGEL, T. 1974. "What is it like to be a bat?" In *The Philosophical Review*. Vol. LXXXIII, 4, p. 435-50

NINAN, D. 2016. "What is the Problem of *De Se* Attitudes?" In García-Carpintero, M.; Torre, S. (Ed.) 2016, p. 86-120.

NOONAN, H. 1984. "Fregean Thoughts". In *The Philosophical Quarterly*. Vol. 34, 136. Blackwell Publishing, p. 205-224.

O'BRIEN, L. F. 1995. "Evans on Self-Identification". In *Noûs*. Vol. 29, 2, p. 232-247.

PERRY, J. 1977. "Frege on Demonstratives". In *The Philosophical Review*. Vol. 86, 4. Cornell University, p. 474-497.

_____. 1979. "The Problem of the Essential Indexical". In *Noûs*. Vol. 13, 1, Oxford: Wiley-Blackwell, p. 3-21.

_____. 1980. "A Problem About Continued Belief". In *Pacific Philosophical Quarterly*, 61, p. 317-332.

_____. 1995. "The Self". In *Supplement to the Macmillan Encyclopedia of Philosophy*, 1998.

_____. 1997. "Indexicals and Demonstratives". In Hale, R.; Wright, C. (Ed.) *Companion to the Philosophy of Language*. Oxford: Blackwell Publishers Inc.

_____. 1997b. "Myself and I". In Stamm, M. (Ed.) *Philosophie in Synthetisher Absicht (A festschrift for Dieter Heinrich)*. Stuttgart: Klett-Cotta, 1998, p. 83-103.

_____. 1997c. "Rip Van Winkle and other Characters". In Dokic, J. (Ed.) 1997, p. 13-39.

_____. 2006. "Using Indexicals". In Devitt, M.; Hanley, R. (Ed.) *The Blackwell Guide to the Philosophy of Language*. Oxford: Blackwell Publishing.

_____. 2007. "'Borges and I' and 'I'". In *The Amherst Lecture in Philosophy*, 2, p. 1-16. URL = <http://www.amherstlecture.org/perry2007/>.

PROSSER, S. 2005. "Cognitive Dynamics and Indexicals". In *Mind and Language*, 20, p. 369-391.

PROSSER, S.; RECANATI, F. (Ed.) 2012. *Immunity to Error through Misidentification: New Essays*. New York: Cambridge University Press.

PROUST, J. 2000. "L'animal intentionnel". In *Terrains*, 34, p. 23-36.

QUINE, W. V. 1956. "Quantifiers and propositional attitudes". In: *The Journal of Philosophy*, vol. 53, n. 5, p. 177-187.

RECANATI, F. 1979. *La transparence et l'énonciation: pour introduire à la pragmatique*. Paris: Editions du Seuil.

_____. 1993. *Direct Reference, From Language to Thought*. Oxford: Blackwell Publishers, 1997.

_____. 1997. "The Dynamics of Situations". In Dokic, J. (Ed.) 1997, p. 41-75.

_____. 2001. "Loana dans le métro. Remarques sur l'indexicalité mentale". In Bourgeois-Gironde, S. (Ed.) *Les formes de l'indexicalité, langage et pensée en contexte*. Éditions Rue d'Ulm, 2005, p. 19-34.

_____. 2006. "Content, mode and self-reference". In S. Tsohatzidis. (Ed.) *John Searle's Philosophy of Language: Force, Meaning, and Thought*. Cambridge: Cambridge University Press.

_____. 2007. *Perspectival Thought A Plea for Moderate Relativism*. Oxford: Oxford University Pres.

_____. 2007b. "Imagining de se". URL = <http://jeannicod.ccsd.cnrs.fr/ijn 00160757>.

_____. 2008. *Philosophie du langage (et de l'esprit)*. Collection Folio Essais. Éditions Gallimard.

_____. 2009. "Le soi implicite". In: Longuenesse, B. (Ed.) *Revue de Métaphysique et de Morale: 'Le Moi/the Self/le Soi'*.

_____. 2010. "Singular Thought: In Defence of Acquaintance". In Jeshion, R. (Ed.) 2010, p. 141-189.

_____. 2012. *Mental Files*. Oxford: Oxford University Press.

_____. 2015. "Cognitive Dynamics: a New Look at an Old Problem". To appear in Korta, K.; Ponte, M. (Ed.) *Reference and Representation in Language and Thought*. Stanford: CSLI Publications.

_____. 2016. *Mental Files in Flux*. Oxford University Press.

ROVANE, C. 1993. "Self-reference, the radicalization of Locke". In *The Journal of Philosophy*. Vol. 90, 2, p. 73-97.

RUFFINO, M. 2007. "Fregean propositions, belief preservation and cognitive value". In *Grazer Philosophische Studien* 75, p. 217-236.

RUSSELL, B. 1903. *The Principles of Mathematics*. New York: W. W. Norton & Company, INC Publishers, 1903.

_____. 1905. "On Denoting". In *Mind* 114 (456). Oxford, 1905, p. 479-493.

_____. 1910-1911. "Knowledge by Acquaintance and Knowledge by Description". In *Proceedings of the Aristotelian Society*, 11, p. 108-128.

_____. 1912. *The Problems of Philosophy*. Oxford: Oxford University Press, 1980.

_____. 1918. "The Philosophy of Logical Atomism". In *Logic and Knowledge*. New York: Routledge, 2004.

SAINSBURY, R. 2002. *Departing from Frege. Essays in the philosophy of language*. New York: Routledge.

SALMON, N. U. 2005. *Reference and Essence*. Second edition. (Studies in analytic Philosophy. Series Editor: Quentin Smith.) New York: Prometheus Books.

SCHELLENBERG, S. 2012. "Sameness of Fregean sense". In *Synthese*. Vol. 189, p. 163-175.

SHAH, P. et. al. 2016. "Alexithymia, not autism, is associated with impaired interoception". In *Cortex*. Vol. 81, p. 215-220. URL = <http://creativecommons.org/licenses/by/4.0/>.

SHOEMAKER, S. 1968. "Self-reference and Self-awareness". In Cassam, Q. (Ed.) 1994, 80-93.

_____. 1970. "Persons and Their Pasts". In *American Philosophical Quarterly*.

_____. 1994. "Phenomenal Character". In *Noûs* 28: 1, Oxford: Basil Blackwell, p. 21-38.

_____. 1996. *The First Person Perspective and Other Essays*. Cambridge: Cambridge University Press.

SOUTIF, L. 2014. "Dinâmica cognitiva". In Branquinho, J.; Santos, R. *Compêndio em Linha de Problemas de Filosofia Analítica*. URL = <http://compendioemlinha.letras.ulisboa.pt/wp-content/uploads/2015/07/soutif_2014_dinamica_cognitiva.pdf>.

STAJONOVIC, I. 2001. **"**For whom is the problem of the essential indexical a problem?" In Akman, V.; Bouquet, P.; Thomason, R. H.; Young, R. A. (Ed.) *Modeling and Using Context Lecture Notes in Artificial Intelligence*. Springer Verlag, p. 304-315. URL = <http://jeannicod.ccsd.cnrs.fr/ijn 00460896>.

STRAWSON, P. F.; GOCHET, P. 1970. "Phrase et acte de parole". In *Langages*. 5ᵉ année, n° 17. *L'énonciation*, p. 19-33. URL = <http://www.persee.fr/doc/lgge_0458-726x_1970_num_5_17_2573>.

SULLIVAN, A. 2003. "'Paging Dr. Lauben! Dr. Gustav Lauben!': Some Questions About Individualism and Competence". In *Philosophical Studies*. Vol. 115, p. 201-224.

TEXTOR, M. 2007. "Frege's theory of hybrid proper names developed and defended". In *Mind*. Vol. 116, 464, p. 947-981.

VIGNEMONT, F. de 2003. "Retour au modèle observationnel de la connaissance de soi". In: *Philosophi@ Scienti@e*. URL = <https://sites.google.com/site/fvignemont/connaissancedesoi.pdf?attredirects=0>.

WAHL, R. 1993. "Russell's Theory of Meaning and Denotation and 'On Denoting'". In *Journal of the History of Philosophy*. Vol. 31, 1. The Johns Hopkins University Press, p. 71-94.

WEIMER, AMY K. M. D. et al. 2001. "'Motor' Impairment in Asperger Syndrome: Evidence for a Deficit in Proprioception". In *Journal of Developmental & Behavioral Pediatrics*. Vol. 22, Issue 2, p. 92-101.

WETTSTEIN, H. 1990. "Frege-Russell Semantics?" In *Dialectica*, 44, p. 113-135.

WILLIAMS, C. J. F. 1990. "Thoughts Which Only I Can Think". In *The Philosophical Quarterly*. Vol. 40, 161, p. 489-495.

WITTGENSTEIN, L. 1953. *Philosophical Investigations*. Oxford: Basil Blackwell.

_____. 1958. *The Blue and Brown Books*. New York: Harper & Row Publishers, 1965.

WOLFF, F. 2005. "Je et maintenant". In Bourgeois-Gironde, S. (Ed.) *Les formes de l'indexicalité, langage et pensée en contexte*. Éditions Rue d'Ulm, p. 1-17.

ZALTA, E. N. "Fregean Senses, Modes of Presentation, and Concepts". In *Philosophical Perspectives* (Noûs Supplement), 15, 2001, p. 335-359.