Projected self: the de se across dimensions and beyond pronouns

Rodanthi Christofaki

Corpus Christi College

University of Cambridge

This dissertation is submitted for the degree of

Doctor of Philosophy

September 2018

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text.

It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text

It does not exceed the prescribed word limit for the relevant Degree Committee.

Abstract

This dissertation is about how attitudes de se are encoded in content outside the assertive domain, as well as by expressions which are not really pronouns, such as nouns. First of all, it provides evidence that expressives are necessarily de se, which means that non-pronominal expressions, such as nouns, verbs etc. can be de se in a non-assertive dimension. Next, it examines in full detail the landscape of expressive (or as it is more recently known, use-conditional) meaning and the extant dedicated frameworks, arguing that they are misguided, due to conceptual issues and not being restrictive enough to be theoretically meaningful. I then move to the debate of whether expressives are presuppositions or a distinct kind of meaning, arguing for a conciliatory solution which proposes that expressives are ordinary presuppositions endowed with a compositionally irrelevant, indexical kind of meaning which I call 'associative'. This solution explains the ambivalent behaviour of expressives, i.e. that they may display filterability and non-displaceability simultaneously. The next topic is that of self-reference of Japanese, a language which lacks genuine personal pronouns and where speakers may use a multitude of expressions instead for this function: a long list of so-called personal pronouns which are actually nouns, then also common nouns and proper names. After a thorough description of the data, I offer a formal account of self-reference in Japanese, accommodating all the different kinds of expressions used. The main theoretical implication of this account is that the absolute distinction between indexicals and names is denied, as names in Japanese can encode the de se. Moreover, it is shown that Japanese so-called 1st person pronouns are de se in both dimensions, as they are used as self-referring expressions but they are also endowed with associative meaning.

Acknowledgments

First and above anyone else, I would like to thank from the bottom of my heart my unofficial, 'behind-the-scenes' (as he himself once said), but essentially real supervisor Dr. Yasutada Sudo from UCL. The importance of Yasu's support and help towards me as a student, a linguist, and as a result a person, cannot possibly be exaggerated, and only those who knew the truly dire situation I was in before he came into the picture can really understand it. I got to know Yasu when I was starting my 3rd year of PhD, when I hesitatingly contacted him to ask if I could attend his introductory course in formal semantics at UCL after yet another failed attempt to teach myself formal semantics, which truly fascinated me, and while experiencing great confusion about the direction of my work. Yasu not only allowed me to attend his course, but also treated me no differently than his other students, correcting my homework every single time, and giving me a real opportunity to learn and understand. A few weeks into his course everything fell apart in my PhD as my then supervisor withdrew, and I found myself in a state of despair and confusion in order to see how (or whether) to proceed. It is no exaggeration to say that during those dark few weeks I spent without knowing what would happen to my PhD, Yasu's course once a week was the sole reason I would get out of bed, get dressed and leave the house, my main motivation being to complete his course and understand what formal semantics was about even if I ended up having to quit my PhD. Eventually after a period of extreme anxiety due to negotiations with the department about a new supervisor, they accepted to appoint Yasu as my unofficial supervisor. It was January 2017, and I was already well in my 3rd year; I was very confused about my topic and the work I had done before, I had little time and even less funding left and I had just discovered formal semantics, yet I was so excited, so truly thrilled to have the opportunity to be supervised by someone like Yasutada Sudo, that I felt that I had won the lottery.

And I really had won the lottery as far as working with Yasu was concerned - everything else in my PhD continued extremely stressful and difficult. I still can't believe that he accepted to supervise someone so late in her PhD, so confused about what she had done so far and what could do from then on, so junior in her knowledge of formal semantics, so traumatised from her PhD experience so far, and he did this without anything to gain for himself except for extra work. Unlike most people at my local environment, instead of seeing me as a problem student and discouraging me from getting too deep into something as new and demanding as formal semantics so late in my PhD, Yasu saw the enthusiasm and hunger to learn in me, and kept feeding them with his guidance. He would always find ample time to meet me, and always answered my very often naive questions in full detail; he always challenged me to think deeper and further, pushed me towards the greater questions, and never treated me as a lost cause or a special case, but with the same respect and expectations as his other students. Thanks to his constant encouragement, support and availability, I managed to produce work in a field and of a standard I could never even dream of, to submit my PhD on time, and to be examined by the people I always admired and dreamed to academically interact with. Thinking back at how much I learned and was able to do when I was supervised by Yasu, I so wish I had had more time to enjoy the pleasure of doing linguistics under such a great supervisor, but given the circumstances I can only be grateful that this was even possible. I cannot express my gratitude enough, as the impact of Yasu's help extends well beyond the completion of my PhD thesis: it enabled me to realise my lifelong dream of becoming a linguist, and to build confidence in my abilities in the process. His support empowered me to pursue things I thought unattainable, such as my current job as a linguist in the industry. Thank you so much, Yasu, for being such an amazing supervisor!

Importantly, I would like to sincerely thank the Leverhulme Trust for funding the first 3 years of my PhD. I would have never been able to embark on a PhD without this

funding, and I am extremely grateful for this opportunity. I would also like to thank Corpus Christi College and the department of theoretical and applied linguistics in Cambridge for helping me with funding during the first half of my 4th year of PhD.

There are many people I would like to sincerely thank for their support and role during my PhD. First, many thanks are due to Dr. Minyao Huang, who acted as a co-supervisor during the first two plus years of my PhD. Minyao taught me so much, especially about philosophy of language, and she always warmly encouraged as well as challenged me. I would also like to specially thank people I was fortunate enough to meet and talk to at conferences, and who had a huge influence in my work: Stefano Predelli and Philippe Schlenker who eventually turned out to be my thesis examiners, and Elin McCready. The occasion of meeting Philippe in ESSLLI 2017 was extremely significant for me, as I had been strongly discouraged from spending time in summer schools so late in my PhD and I was not supposed to be there at all. That brief talk I was able to have there with Philippe about expressives and presuppositions proved to be a real turning point in my PhD. I would also like to sincerely thank Dan Lassiter, who I met at the same conference, for expressing a true interest in my situation at the time. That conference was a very difficult time for me given my academic isolation and general lack of support with the exception of Yasu, and it made a huge difference that I could talk to someone about it while I was there.

Moreover, I would like to sincerely thank everyone I interacted with at UCL, even if only briefly at seminars: Nathan Klinedinst, Daniel Rothschild and Richard Breheny. The UCL linguistics department is an extremely welcoming and open community, and as I started visiting it not only for my supervisions with Yasu but also for seminars and talks, I always felt at home and welcomed by everyone, which was extremely significant for me given how isolated I was feeling where I was supposed to belong academically. In a way I feel that UCL is my true alma mater, and I am proud to have been there even as an unofficial visiting student.

Lastly, I would like to express my sincere thanks to Sheila Watts and Brechtje Post at Cambridge, who found themselves responsible for sorting out my PhD situation, as well as appointing a new supervisor in order to help me continue my PhD. I know that it was an extremely difficult task for them given the complexity of the situation, and I am grateful not only because they helped me reach a good outcome out of a terrible situation, but also for doing so with true empathy and genuine interest in my well-being. Finally, I would like to thank Ianthi Tsimpli, who undertook the role of the official supervisor at Cambridge - this allowed me to continue being a PhD student there while working with Yasu.

The last part of the acknowledgments is dedicated to thanking my friends and the special people in my life who have been constant beacons not only during the dark times of my PhD, but continuously.

First of all, my most sincere thanks go to Alexis Dimitriadis - not the linguist, although he could easily have been a great one given his tremendous intellectual depth and curiosity. Alexis and I left Greece for Cambridge together with our two cats, Junior and Little Chief, when I got the PhD studentship; he grew from a long-time partner to forever family, and I am so grateful to have him in my life. He always believed in me, truly cared for me, actively supported me during the bad times of my PhD, and has been continuously the one person I could always count on no matter what life brought. Alexis, habemus dissertationem - we did it!!! I know how much effort you put into this PhD by keeping me sane and safe while at it.

Next, I want to thank my close and extremely beloved friends Eleni Savva, Luca Sbordone (who are also known as M.P.) and Federica Della Grotta. Elenaki, you have always been such a wonderful friend and a truly beautiful person that I feel blessed to have close to me. Our conversations, from the most trivial to the deepest, are still among the greatest pleasures in my life. Despite the fact that our differences are often a topic of

conversation and jokes between us, I feel a very deep connection to you and I can always be my true self with you. Luchino, I don't even know where to start talking about how special you are - your wit, your humour, your intellect, your craziness (which is so normal to me), and above all your genuinely good heart that makes you such a true friend are only a few of the things I love about you. Federica, you are a true force of good, kindness and strength that keeps the rest of us in balance and sanity. Guys, I love you so much. There is absolutely no way I would have survived the PhD without you. Thank you for everything you've done for me; I am truly lucky to have found you in my life.

Moreover, I want to thank the amazing Dimitris Michelioudakis - not only a great friend, but also the reason I wanted to become a linguist in the first place. Dimitris and I met as undergraduate students in different departments of the University of Athens, went to Cambridge together for our MPhils again at different departments, and supported each other throughout. My admiration for him drew me towards linguistics in the first place and eventually towards its more theoretical side for my PhD. Dimitris, thank you for always believing in me and for not letting me give up my dream of becoming a linguist. We've been through so much together all these years, both rough and fun, we have grown together, and I'm so grateful to have you in my life!

Next, I would like to thank my beloved friends from UCL, Ruoying Zhao and Gregor Williamson. Guys, I am so grateful to you for embracing me from the first moment, for making me feel like I belonged when I felt lost and isolated, and for always supporting and believing in me when I felt that I couldn't really make it. Your presence as my peers was very significant during the last 1,5 year of my PhD, and I feel so lucky to have you as my true friends even after its completion.

I would also like to express my warmest thanks to Napoleon Katsos, Nafsika Smith and their wonderful children, George and Michalis. Napoleon was my supervisor during my MPhil in 2006-7, and I pride myself to have been the first postgraduate student he supervised. Apart from being an amazing supervisor at the time, Napoleon's help and guidance were vital also for my PhD, both in the beginning and when everything was falling apart. I am so grateful the invaluable help and support Napoleon and Nafsika showed me and Alexis when we came to Cambridge and had to find accommodation that would accept two cats (a near-impossible task), but also throughout my time in Cambridge. I could never thank you enough for all your kindness!

Many sincere thanks go to good friends I made during my PhD, and who made everything more special and fun: András Bárány, Vaishali Mahalingam, Yves Moussallam, Mark Dsouza, Aretousa Giannakou (I never thought we'd end up doing our PhDs together!), Olénka Dmytryk, Rodrigo Retamal and Vladimir Rumleanschi. Also many thanks to the fantastic people I met at the PFILNU summer school in Brazil in 2016, one of the absolutely most fun times of life: Sandy Ritchie, Moreno Mitrović, Gerardo Manuel García Chinchay, Viswanath Naidu and Ernesto Cortés Teherán.

I also want to thank my beloved friends back home who are truly family: Sofia Giannopoulou and Anastasia Pavlakou, thank you for always being the wonderful friends that you are, for so many years that I can't remember how life was before I met you. You really make me feel the true joy of love and friendship, and even though we don't see each other as often as we used to, you really are an integral part of my life. Also, many thanks to Kostas Seretis, who I absolutely adore, and the wonderful Elena Sereti, Hara Sereti, Dimitris Seretis and Louisa Kouvatsou - I love you kids so much!

Before finishing, I want to thank my parents Stelios and Kalliopi, for always doing the best they possibly could for me, and for always supporting my efforts to pursue knowledge. I love you very much and I am very grateful for everything. I want to thank my brother Labis for always showing me true unconditional love, also his wife Konstantina, and, of course, my wonderful little niece Myrto who has brought such joy to all of us and has been a true blessing for our family.

Lastly, but most importantly, I would like to thank Ourania Giakoumaki. Ourania mou, my gratitude to you is beyond words. Thank you for enabling me to accept, forgive,

love and believe in myself; for helping me become myself. None of the many wonderful things I have in my life now that make me grateful to be alive would have been possible without your help.

This thesis is dedicated to my beloved cats: Little Chief, who I prematurely lost during the first year of my PhD and will miss forever, and Junior, the beautiful ginger lion king who still proudly reigns in Cambridge. I love you both so much!

Contents

List of tables and figures

Chapter 1. Introduction

1.1. Introduction

Understanding attitudes as mental states towards propositions, among different kinds of attitudes an individual may hold, there is a specific class of attitudes about the self qua self, also known as attitudes 'de se' (Latin for 'of oneself'). Since Kaplan's (1989) seminal work on indexicals, and Lewis' (1979) as well as Perry's (1979) work on attitudes de se and so-called essential indexicality, there has been growing interest in the topic especially in the last decades with authors exploring it in the domain of natural language expressions other than the classic indexicals, such as PRO (Chierchia 1989), shifted indexicals (Schlenker 1999, 2003, Anand 2006, Anand & Nevins 2004, Sudo 2012) and logophoric pronouns (Pearson 2012). Despite the diversity of these expressions, one common characteristic is that they are all pronominal. Moreover, as pronouns they all make semantic contributions to the assertive (also known as at-issue) content of the sentence.

This dissertation is about how attitudes de se are encoded in content outside the assertive domain, as well as by expressions which are not really pronouns, such as nouns. With regards to lying outside the assertion, it is known that there are expressions whose content seems to project across operators which usually operate on assertive content. A classical example of such an expression is the verb 'stop', which asserts that some activity is no longer taking place but also conveys the inference that it used to take place in the past. Interestingly, while the assertive part can be targeted by operators such as negation, question, modality, and the antecedent of a conditional, the inference about past activity projects through in all cases (Family of Sentences test, Chierchia & McConnell-Ginet 1990):

(1) a. Little Chief stopped chasing birds.

Assertion: Little Chief does not chase birds.

Projective content: Little Chief used to chase birds in the past.

b. Little Chief hasn't stopped chasing birds.

c. Has Little Chief stopped chasing birds?

d. Possibly Little Chief has stopped chasing birds.

e. If Little Chief has stopped chasing birds, we don't have to worry about rescuing them from his fury.

As it is widely known, this kind of inference constitutes a presupposition of the expression 'stop'. In the recent years, it has been found that there is a class of expressions which express emotive attitudes, and which also seem to manifest projection in a way similar to the one seen for presuppositions above (Potts 2005, 2007):

(2) a. That bastard Kresge is late.

Assertion: Kresge is late.

Projective content: Kresge is a bastard.

b. That bastard Kresge is not late.

c. Is that bastard Kresge late?

d. Possibly that bastard Kresge is late.

e. If that bastard Kresge is late, he should be fired.

This kind of expressions has been called 'expressives' (Kaplan 1999, Potts 2005, 2007), as they seem to express an emotive attitude rather than describe a fact. As can be seen from the above, expressives seem to pattern very closely with presuppositions in terms of their projective behaviour. However, there has been intense discussion about whether the projective content of expressives constitutes a presupposition (Schlenker 2007) or some distinct kind of inference in merit of its very own representational logic in the compositional semantics (Potts 2005, 2007, McCready 2010, Gutzmann 2015). This is the main topic of the present thesis, and it will be explored in elaborate detail below, with a critical examination of existing proposals and frameworks and a novel answer to be put forward. The main points are the following: first, the need for the Pottsian compositional semantics for expressives (especially the McCready 2010/Gutzmann 2015 type) is rebutted, and second, it is shown that there are several different types of projective content that have been lumped together, with some of them not being compositionally relevant and others amenable to a presuppositional analysis. Moreover, despite the intense debate on the precise ontology of the projective inferences of expressives, one much less discussed characteristic of these expressions is their great similarity to indexicals, which invites consideration of the possibility of de se. Of course, it is not the case that this similarity has not been noticed at all; Schlenker (2003, 2007) has elaborately drawn a direct link, while McCready & Wechsler (2012) briefly pondered on the question whether expressives are necessarily de se. This dissertation firmly advocates that the answer to this question is positive, and links the two topics of de se and expressives as inextricable. As a result, it complements the aforementioned work on the de se in expressions which are part of assertive content with evidence that it can also be encoded beyond the assertive dimension.

The realisation that the de se can also be expressed beyond assertive content also opens up the possibility that it can be found in expressions that are not pronouns, as previous work on the de se in natural language focused on pronominal expressions which form part of the assertive content of a sentence. Given the advocated link between the de se and expressives, this means that all kinds of expressions may be de se, provided that they are expressives, such as verbs, adjectives, adverbs, and of course nouns. But this thesis goes beyond a mere complementation of the de se in a non-assertive dimension, by also proposing that there can be de se instances of assertive content in expressions that are not pronominal, such as common and proper nouns. This claim is supported by data from the Japanese language.

In order to show this in practice, I take on the topic of Japanese so-called personal pronouns, and more specifically those of the 1st person (although the analysis can extend to 2nd person pronouns to show the 'de te' equivalent). Unlike languages such as English and Greek, but also non-Indoeuropean ones such as Finnish and Ilocano, Japanese is endowed with a multitude of expressions which are used as 1st person pronouns. Although the function of these words is very similar to personal pronouns, they resemble nouns much more than pronouns in terms of their combinatorial abilities (they can be modified by adjectives and determiners), their structural characteristics (they cannot easily be bound) and the fact that there are less constraints in entering this class (hence the fact that there are considerably more than one). Building on studies that have advocated that such words are essentially nouns which are used as pronouns as the language lacks genuine pronouns[1] (Hoji 1990, Noguchi 1997, Longobardi 2008), I examine them and show how they encode de se content in the assertive dimension. Interestingly, Japanese so-called pronouns also carry expressive content, which means that they are de se in both dimensions.

But the most significant implication of the proposed analysis of the Japanese data is that the absolute distinction between names (Kripke 1980) and indexicals (Kaplan 1989) is not preserved; specifically, given their ability to appear in de se positions, names in Japanese cannot conform to an account which assumes that names are entirely devoid of linguistic meaning, as such an analysis could not accommodate the de se. At the same time, I resist the claim that names are bona fide indexicals (Pelczar & Rainsbury 1998), which could potentially accommodate the de se by assuming a non-constant character. Instead, I adopt an analysis of proper names in Japanese as underlyingly nouns endowed with an indexical element (Matushansky 2008), which can account for their ability to appear in de se positions, but can also be extended to nouns as well as so-called Japanese personal pronouns, which are treated as 'names for the speaker'.

1.2. Structure of this thesis

In particular, chapters 2 and 3 are extensive reviews of previous studies, providing the reader with necessary background by critically reviewing the literature on indexicality, the de se, expressives and specific frameworks for expressive meaning (Potts 2005, 2007, McCready 2010, Gutzmann 2015), so they can potentially be skipped by readers already familiar with these topics. However, one novel claim in chapter 3 is that expressives are necessarily de se, which is supported by examples. Chapter 4 is a continuation of the literature review but undertakes a very thorough critical examination of the most recent frameworks of expressive meaning, presenting arguments for their demise. Next, in chapter 5 I offer my own novel analysis of expressives as presuppositions also endowed by compositionally irrelevant meaning, which I call 'associative'. Chapter 6 acts as a description of self-reference in Japanese, presenting so-called personal pronouns as well as proper names and nouns used as such, and it can be skipped by readers familiar with these data. This chapter provides the necessary data background for chapter 7, which presents a novel analysis of expressions

[1] This claim is about articulated pronouns, as there is indeed a null pronoun (Kuroda 1965).

used for self-reference in Japanese that can capture their de se uses, and which is also complemented by the analysis of expressive meaning of chapter 5 to give a full account of the data in question.

Chapter 2. The de se and its semantic relevance

This chapter will introduce the concept of attitudes de se, i.e. attitudes about oneself as oneself, and its relevance to the semantics of natural language expressions such as indexicals.

2.1. The beginnings: Frege

Gottlob Frege's seminal work 'On Sense and Reference' (1892) famously demonstrated that equating the meaning of a term with its reference gives rise to puzzling phenomena. First, although two different linguistic terms may stand for exactly the same object, it is possible for there to be a difference in 'cognitive significance' between them for different speakers.[2] To exemplify this, let's see the case of the famous graffiti artist known as 'Banksy', whose real identity is not officially known, despite great interest in the public and media to discover it. While speculating about the artist's identity, the statement 'Banksy is Banksy' happens to be not just undeniably true but also completely uninformative; it is a tautology, true by virtue of its structure and thus not providing any information about the world. Nevertheless, although the statement 'Banksy is Robert Del Naja' (assuming that it is true[3]) strictly seems to say the same thing as the previous one, namely that a certain individual is identical to himself, is informative and interesting enough to be circulating all over the media. In a similar vein, the substitution of terms with the same referent does not yield valid inferences when reporting propositional attitudes; for example, we cannot validly report 'Lois finds Superman attractive' as 'Lois finds Clark Kent attractive', as it would mean attributing to Lois an attitude she does not herself have.

In order to solve these puzzles, Frege introduced the notion of 'sense' as a mode of presentation of the referent, proposing that meaning involves both sense and reference. The idea is that although a certain expression can have a certain object as its referent (reference), its meaning is not exhausted by it but it also involves a certain manner in which the referent is presented (sense) to users of the expression, which could be understood as a description that speakers associate with the referent.[4] To use Frege's own example, while the names 'Phosphorus' and 'Hesperus' share the same referent (i.e. the planet Venus), they have different senses, which can be understood as the descriptions 'the morning star' and 'the evening star' respectively. As above, here too the difference in senses can be shown by contrasting the referentially equivalent sentences 'Phosphorus is Phosphorus' and 'Phosphorus is Hesperus', seeing how the first one is entirely uninformative yet the second one amounts to a major astronomical discovery. Importantly, senses are publicly available entities that are shared among all speakers,

[2] The notion of 'cognitive significance' (translation of German 'Kognitive Signifikanz') was introduced by Frege (1892) to explain 'the question of how identity sentences in which proper names flank the identity sign can both state truths and be informative' (Wettstein 1986: 185).

[3] http://www.independent.co.uk/arts-entertainment/music/news/banksy-identity-revealed-goldie-interview-robert-del-naja-massive-attack-art-a7804091.html

[4] Although 'mode of presentation' has been widely understood as a kind of description, it has been suggested that what Frege meant by this was some kind of access/acquaintance relation to the referent, and not necessarily a description (Evans 1981). This could potentially ease some of the problems Fregeanism is said to have, as those seem to presuppose a notion of pure descriptivism.

not private impressions that vary per person as is the case of ideas; the public character of senses is crucial to Frege's account as senses are employed in logical inferencing, which is subject to the laws of logic and cannot possibly vary according to individual psychology (Frege's anti-psychologism). The notion of sense as an intermediate entity between the mind and the object referred to allowed Frege to account for the puzzles as follows. The difference in cognitive significance between terms which share the same reference can be attributed to a difference in senses; in the above example, even though 'Banksy is Banksy' and 'Banksy is Robert del Naja' say the same thing, namely that a certain individual is identical to himself (assuming that these two different names refer to the same individual), the former is trivial but the latter is informative because of their difference in sense. When we say 'Banksy is Banksy', we are asserting a tautology because we are using the same expression twice, such that the same sense is involved in both cases. However, when we say 'Banksy is Robert Del Naja', the assertion reveals that two expressions that differ in sense actually share the same reference in the real world. As this statement says something about what the world is like, it is informative and interesting. Analogously for the second problem, the fact that the substitution of co-referring expressions fails to preserve truth value in the case of reporting someone's attitude can be attributed to the difference between senses; as two different expressions can share the same reference but differ in sense, the fact that they cannot be freely exchanged in certain sentences without changing the meaning suggests that what is involved in these constructions is an expression's sense, not reference. If Fregean senses are understood as modes of presentation of their referents as mentioned above, attitudes towards objects can be said to be entertained under specific modes of presentation and not directly towards the objects themselves.

2.2. Challenges for the Fregean account: proper names and indexicals

Despite the success of the Fregean account in accounting for these puzzles, it has been shown to face problems with certain kinds of expressions, such as proper names and eventually, demonstratives and indexicals. Specifically, the assumption that names and indexicals have Fregean senses seems to provide the wrong predictions of how they are actually used and understood.

2.2.1. Proper names: Kripke

Proper names were the first to be suggested as resisting a Fregean treatment, by Marcus (1961) and more famously Kripke (1980). Specifically, Kripke argued that descriptivist theories make wrong predictions about how proper names are interpreted and used by speakers, but also that there is a clear difference between how descriptions and names behave in modal contexts.

First of all, he pointed out that contrary to the idea that the meaning of a proper name corresponded to a description (or even to a cluster of descriptions), speakers are perfectly capable of using as well as understanding proper names even when they were unaware or even blatantly wrong about the supposed identifying description(s); these are Kripke's so-called 'argument from ignorance' and 'argument from error'. For example, the majority of people know Gutenberg as the inventor of typography, i.e. him

inventing typography turns out to be the most distinctive if not only identifying characteristic of Gutenberg for most people, at least as a historical figure. In this sense, assuming a descriptivist theory of names we might say that the meaning of 'Gutenberg' corresponds to the definite description 'the inventor of typography'. However, even though Gutenberg was the first person to employ movable type printing in Europe, it has been found out that this technique had been previously invented several centuries ago in China by Bi Sheng (990–1051 AD). So if the meaning of the name 'Gutenberg' really amounts to the description 'the inventor of typography', when people employ it they should in fact be referring to Bi Sheng, rather than the German individual known as Johannes Gutenberg. Moreover, it has been found that movable type was independently invented in Korea as well around the late 14th century, presumably by a monk called Baegun[5]. This seems to complicate things even more, because even though movable type had been invented in China much earlier, if Baegun indeed came up with the idea without prior knowledge of the Chinese mechanism, in a way we can still say that he invented the movable type. Considering these historical facts, we end up in a situation where the description 'the inventor of typography' either picks out Bi Sheng instead of Gutenberg, or it doesn't pick out anyone since there is no single person who invented typography, but several. However, we know very well that while most people are totally unaware about such historical details, they are perfectly capable of referring to Gutenberg himself by employing his name. Importantly, even if this particular example might seem exceptional because it involves a historical figure people may not know much more than a single thing about, this argument naturally extends to all names simply because different people may have different identifying descriptions in mind about a certain individual, yet they are somehow able to communicate information about the same individual. So if people can use and understand proper names effectively even though they might be wrong or ignorant about corresponding identifying descriptions, then the meanings of names cannot amount to descriptions, whatever those may be.

But apart from the aforementioned epistemological arguments, Kripke also offered another even more powerful argument, which concerns the very different behaviour of proper names and descriptions in modal sentences (known as the 'modal argument'). Specifically, definite descriptions which express contingent properties of the individuals in question may end up denoting different entities in different possible worlds, while the reference of proper names stays fixed even in modal environments. Let's use the example of Bucephalus, the horse of Alexander the Great; in this case it feels relatively safe to assume that the only reason this particular horse became and remained famous is because he was the horse of Alexander, so the description 'the horse of Alexander the Great' feels very suitable as the identifying one.

(3) Bucephalus might not have been the horse of Alexander the Great.

(4) The horse of Alexander the Great might not have been the horse of Alexander the Great.

(5) Bucephalus might not have been Bucephalus.

[5] https://www.livescience.com/43639-who-invented-the-printing-press.html

As the examples above show, there is a clear difference between the description 'the horse of Alexander the Great' and the proper name 'Bucephalus', even though Bucephalus was indeed the horse of Alexander the Great. This is because Bucephalus was the horse of Alexander the Great in the actual world, but we can easily imagine an alternative course of things in which Alexander ended up choosing another horse; for example, he may have preferred white or female horses, and as Bucephalus was known to be black and male, he may not have fitted Alexander's criteria. Considering such a possibility shows why sentence (3) is true, as it simply states that this specific horse may not have ended up as Alexander's horse. However, both sentences (4) and (5) are analytically false; this is simply because while being Alexander's horse was a contingent property of Bucephalus, i.e. one which he happened to have in the actual world (as far as we know) but which could perfectly lack in other possible worlds, being identical to himself is a necessary property of Bucephalus (which makes 5 false), as it is of any other entity (which makes 4 false).

Based on the contrast between proper names and descriptions in modal sentences such as the above, Kripke pointed out a clear contrast between descriptions and names: the former allow their reference to vary across different circumstances, while the reference of the latter seems strictly fixed, completely impervious to any counterfactual considerations. To point out this contrast with definite descriptions, he introduced the term 'rigid designator' for names. Obviously, such a stark contrast seems to suggest that we cannot give the same semantic treatment to these different classes of expressions.

However, there is still a question to be answered: if names do not have any other meaning than their reference, by virtue of what do they refer? At least with the descriptive theory the answer to this question was straightforward, as it was the speakers' knowledge of the identifying description(s) which enabled reference. The idea here is that lacking any descriptive properties, names fulfil their referential function by virtue of a so-called 'causal chain'. That is to say, after some kind of dubbing or 'initial baptism' (Kripke 1980) takes place, which involves some sort of direct acquaintance with the entity in question, a name is assigned. This act of dubbing not only inaugurates but also legitimises the adoption of the given name, as every use of it can be causally traced back to it: practically speaking, people repeat the name they hear from other people, and tracing back we are led to the initial baptism event. This idea explains both the semantic properties of names (i.e. how names achieve their reference), but is also epistemologically apt as it explains the kind of knowledge people have when they know a certain name: they know it corresponds to a certain person, and this correspondence has been made either by a dubbing of their own or by appealing to one made by someone else previously.

2.2.2. Indexicals

2.2.2.1. Kaplan (1989)

Turning now to indexicals, although they are obviously different from names, they seem to be equally problematic for the Fregean account, if not more. Specifically, Kaplan

(1989[6]) proposed a new conceptual framework together with a new logic for indexicals and demonstratives, which has been extremely influential. Kaplan noted the difference between 'pure indexicals' and 'true demonstratives' in that the former get their reference from parameters of the context while the latter are accompanied by demonstrations, but the general principles of his framework can be applied to both. Due to the focus of this work, here I will concentrate only on indexicals.

According to Kaplan, indexicals refer directly, without the mediation of a Fregean sense. That is to say, an indexical's contribution to the proposition is the referent itself, rather than any concept via which the speaker represents the referent, and it crucially depends on the context of utterance rather than any fixed linguistic knowledge. For example, let's compare the time expressions 'Monday 3rd of July 2017' and 'today', which can be both used to refer to the same day (assuming that the latter is uttered on Monday 3rd of July 2017). It seems that the difference between these two expressions cannot be accurately captured by saying that they have different senses; specifically, 'Monday 3rd of July 2017' and 'Al-Ithnayn Shawwal 9, 1438' have different senses (the former is calculated in the Gregorian calendar while the latter in the Hijri one) but the same reference. Nevertheless, although 'today' can have the same reference as the expressions above, this crucially depends on whether the utterance is made on that specific day. That is to say, contrarily to the expressions above, the reference of 'today' is not fixed in any way, but it is crucially determined by the context of utterance.

Therefore we see that 'today' is obligatorily interpreted as referring to the day of the current utterance, while 'the day of utterance' can easily refer to another day. In fact, there is no way 'today' can be shifted away to context different than the one of utterance, even embedded under an attitude verb:

(6) Ruoying believes that today there will be a Slavic pagan metal concert.

(7) Ruoying believes that on the day of utterance there will be a Slavic pagan metal concert.

We see that even though (6) reports another person's attitude, not the speaker's, we understand 'today' to unambiguously refer to the day the specific utterance is produced, and not to the day that Ruoying has in mind, whatever that is. Similarly to what we saw above, (7) can easily be understood as a different day, specifically in this example it is most likely to refer to the day that Ruoying herself would describe as 'the day of utterance', which could be any day that happens to be salient in the discourse (e.g. a day in which something very significant was uttered, such as a baby's first words or someone's revealing of a very important fact).

This is obviously similar to the Kripkean proposal of proper names as rigid designators to the same individuals, immune to counterfactual considerations that could affect sense and shift the reference. Nevertheless, there is at least one crucial difference between proper names and indexicals. As it was convincingly argued by Kripke (1980), proper

[6] A clarifying note on the dates: although Kaplan's 'Demonstratives' was officially published in 1989, it had started becoming known among philosophical circles much earlier (at least by 1977), either by talks given by Kaplan or by circulating manuscripts.

names seem to lack any meaning and just stand as linguistic props for individuals. However, indexicals do not stand for specific individuals; on the contrary their reference constantly changes from one entity to the other. Moreover, they do seem to have a kind of linguistic meaning. As noted above, eternal paraphrases such as 'the time of utterance' for 'now' in fact match speaker intuitions, despite their differences in how they secure their reference. Therefore indexicals are similar to proper names in that they refer directly without the mediation of a Fregean sense, but they differ from them in that they seem to have some kind of linguistic meaning.

In order to account for this, Kaplan proposes that indexicals have two kinds of meaning: character, which acts like a kind of instruction on how to extract reference from the context, and content, which given the circumstances yields the entity referred to in the specific context of utterance. Character is therefore the linguistic meaning that speakers understand indexicals to have, but as it is proper for devices of direct reference, this meaning does not enter the proposition. Instead, it is the content meaning that does so.

More formally, if we understand contexts as tuples of parameters such as speaker, addressee, time, location and world, then the character corresponds to a function from the context of utterance to the relevant parameter in that context (which can then yield the content and ultimately the referent of the indexical expression). This step is necessary as far as indexicals are concerned, since their reference depends on the context of utterance, but it is also sufficient, since we do not need to consider any circumstances given the directly referential nature of indexicals. But the contrary holds for non-indexical expressions: their reference does not vary by context, but by the circumstances under consideration, which Kaplan called 'circumstances of evaluation', which we can understand as a possible world. We saw this above with an expression such as 'the inventor of typography', which was shown to refer to different individuals when considering different possible worlds. This means that we should consider which possible world we are evaluating in, which in formal terms is captured by relativising the interpretation function to a world (as well as a time, which we can ignore for convenience).

(8) $[\![\text{the inventor of typography}]\!]^{w}$ = the person who invented typography in w = Gutenberg

(9) $[\![\text{the inventor of typography}]\!]^{w'}$ = the person who invented typography in w' = Gregorius

The introduction of character meaning by Kaplan means that we should also relativise to a context, which affects the reference only of indexical expressions:

(10) $[\![\text{today}]\!]^{c,w}$ = the day it is in c = (e.g.) 6th of July 2017

(11) $[\![\text{the day of utterance}]\!]^{c,w}$ = the day of utterance in w = (e.g.) 29th July 2017

Specifically, the reference of indexical expressions depends only on the context, not on the possible world. Contrarily, the reference of non-indexical expressions depends the possible world, and not on the context. As natural languages have both indexical and non-indexical expressions, we need to relativise the interpretation function both to a

context and a possible world in the interpretation function, which is why Kaplan's (1989) framework is characterised as 'two-dimensional'.

To summarise, character can be understood as a function from context to content, and content itself as a function from the circumstances of evaluation to an extension. Given their directly referential nature, the reference of indexicals is not affected by anything further than the context, such as the circumstances holding in the world of evaluation, so in this case content is a constant function. As for non-indexicals, given that they are indifferent to the context, their character is a constant function, and their content is a function from circumstances of evaluation to an extension (in those circumstances).

2.2.2.2. The so-called 'essential indexical' (Perry 1979)

As noted by Perry (1977), such an analysis of indexicals reveals major problems for the Fregean account, as thoughts expressed in indexicals can no longer be accommodated as Fregean thoughts due to the lack of corresponding senses for them. Although Frege's notion of sense is a perfect tool to explain cognitive significance between co-referring terms such as 'Banksy' and 'Robert del Naja', it seems that we cannot analogously propose the same for co-referring terms such as someone's name and the indexical 'I' when used by them, due to the context-dependent nature of 'I'; the public character of Fregean senses clashes with the fact that 'I' receives a different meaning each time it is uttered by a different person. Frege's solution out of this conundrum was to posit the existence of incommunicable thoughts as he proposed for thoughts about oneself:

> The same utterance containing the word "I" will express different thoughts in the mouths of different men, of which some may be true, others false. […] Now everyone is presented to himself in a particular and primitive way, in which he is presented to no-one else. So, when Dr. Lauben thinks that he has been wounded, he will probably take as a basis this primitive way in which he is presented to himself. And only Dr. Lauben himself can grasp thoughts determined in this way. But now he may want to communicate with others. He cannot communicate a thought which he alone can grasp. Therefore, if he now says 'I have been wounded', he must use the "I" in a sense which can be grasped by others, perhaps in the sense of "he who is speaking to you at this moment", by doing which he makes the associated conditions of his utterance serve for the expression of his thought. (Frege 1918: 296, 298)

As Perry informs us, it was the realisation that indexicals behave differently from other expressions (a fact that is accounted properly Perry's (1977) roles and Kaplan's (1989) character) that compelled Frege to adopt the above solution (Perry 1977: 474, 489, 497):

> He is led to say that, when one thinks about oneself, one grasps thoughts that others cannot grasp, that cannot be communicated. Nothing could be more out of the spirit of Frege's account of sense and thought than an incommunicable, private thought. Demonstratives seem to have posed a severe difficulty for Frege's philosophy of language, to which his doctrine of incommunicable senses

> was a reaction. [...] Frege was led to this view by his own philosophical work, in particular by some realization of the problems I have discussed for his general account, as they apply particularly to "I". All three problems turned on the failure to find a suitable description for the value of the demonstrative, whose sense would complete the sense of the sentence in just the right way. If the sense we are looking for is private and incommunicable, it is no wonder the search was in vain. [...] Frege's appeal to incommunicable senses in the case of "I," is probably an implausible attempt to deal with these problems.

Perry calls such thoughts and beliefs 'self-locating' (1977: 492) and proposes that they require not just the grasping of certain thoughts but their grasping them by means of sentences containing demonstratives. In subsequent work, Perry (1979) instantiates this line of thought even further to claim that indexicals are 'essential' for the apprehension of self-locating beliefs, using his famous 'messy shopper' example:

> 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 a torn sack to tell him he was making a mess. With each trip around the counter, the trail became thicker. But I seemed unable to catch up. Finally it dawned on me. I was the shopper I was trying to catch. […]I believed at the outset that the shopper with a torn sack was making a mess. And I was right. But I didn't believe that I was making a mess. That seems to be something I came to believe. And when I came to believe that, I stopped following the trail around the counter, and rearranged the torn sack in my cart. My change in beliefs seems to explain my change in behavior. […]The reason [things are not so simple] is the importance of the word "I" in my expression of what I came to believe. When we replace it with other designations of me, we no longer have an explanation of my behavior and so, it seems, no longer an attribution of the same belief. It seems to be an essential indexical. […] These indexicals are essential, in that replacement of them by other terms destroys the force of the explanation, or at least requires certain assumptions to be made to preserve it. (1979: 366)

As this illuminating example shows, Perry's critical realisation could not be captured in other than indexical terms; any other co-referring non-indexical term, even if it were uniquely picking Perry, such as his name or a description exclusively picking him, could not express what was essentially an indexical belief without presupposing knowledge of certain facts about oneself. In order for them to work, Perry would need to already possess the knowledge "I am John Perry" or "I am the bearded philosopher in the supermarket", which are also indexical beliefs. This story clearly demonstrates the crucial connection between indexical belief and action, as in this story it was Perry's coming to believe that he himself was making a mess that allowed him to put an end to it, which was what he was after all along.

This observation about the ineliminability of indexicals reveals a problem for the thesis of belief as relation to a proposition, if a proposition is understood as a timeless entity with specific conditions under which it refers to the value either true or false, absolutely. It is immediately understood that propositions thus defined cannot contain indexicals, given their context-dependent nature, but we also saw that some beliefs can solely be

expressed with indexicals. Therefore classical propositions are shown inadequate to capture indexical beliefs. As an alternative, Perry examines the idea of relativised propositions, namely propositions that are not timelessly true (or false) but true for a person or at a time. Initially, it seems that this idea could accommodate indexical beliefs given their change of truth value depending on the thinker/speaker, e.g. a proposition such as "I am making a mess" is true for John Perry in the above scenario, but false for the well-behaved reader. However, Perry rejects it on the basis that any relativised proposition can be believed by anyone, without it being a first-person belief, as the indexical seems to dictate. In our example, anyone can believe that the proposition "I am making a mess" is true for John Perry without believing it true for themselves; in fact, even John Perry himself can believe that it is true for John Perry, if he has heard that a certain John Perry is notorious for messing up supermarkets, having forgotten that this is his own name. So it seems that relativised propositions cannot represent indexical beliefs adequately.

Another alternative examined by Perry as a candidate for representing indexical beliefs is what he calls 'propositions of limited accessibility' (1979: 15), or else entities that are not publicly available but are essentially the private perspectives of believers. As Perry previously noted (1977), this is the solution adopted by Frege having preemptively diagnosed the problem of indexicals for his account, but Perry himself considers it undesirable and problematic for retaining a plausible metaphysics of mind. Eventually Perry proposes a solution based on the distinction between *what* one believes and *how* one believes it, which correspond to believing a certain proposition versus being in a belief state respectively. This allows him to retain the traditional notion of a proposition as the object of belief, but at the cost of putting the explanatory burden on the idea of a belief state. A belief state can be understood as the state one is in when they are able to say (or accept) a sentence containing an indexical, so in the original example the state one is in when they are able to say "I am making a mess". Note that this is a sentence, and not a proposition; it is being in a certain belief state, and not believing a proposition, that is linked to action and behaviour. For Perry, this dissociation between propositions and belief states suffices to solve the problem of the essential indexical, although he does acknowledge a certain degree of limited accessibility, which he considers innocuous (1979: 19):

> We have here a metaphysically benign form of limited accessibility. Anyone at any time can have access to any proposition. But not in any way. Anyone can believe of John Perry that he is making a mess. And anyone can be in the belief state classified by the sentence "I am making a mess". But only I can have that belief by being in that state.

Again, the analogy with Kaplan's (1989) character and content distinction, which we saw already in Perry (1977), is clear. The belief state in which John Perry is when he utters the sentence "I am making a mess" is captured by the indexical's character, while the proposition "John Perry is making a mess" is what we get when we extract the content of the indexical in the context where John Perry utters the aforementioned sentence. As character was defined by Kaplan (1989) as the lexical meaning of indexicals in the minds of speakers, it makes sense why it will be responsible for the

cognitive significance it seems to have, as it is shown to be adequate for expressing the being in a certain belief state.

Perry undeniably reveals a problem for the doctrine that propositions can serve as objects of belief. Moreover, his proposed solution ties in very well with Kaplan's framework for indexicals. Nevertheless, as noted by Torre (2016), it seems that the notion of a belief state leaves much to be desired; it is defined solely by whether one is able to say or at least accept a certain sentence. But a sentence happens to be a linguistic entity, and it feels a bit awkward to require it in order to characterise belief and action. Given that all the explanatory burden has been pushed to this notion of a belief state, it feels that it should be founded on more solid grounds. Moreover, although Perry considers the limited accessibility proposed for when someone believes a certain proposition in a certain belief state (namely, the first-person one) 'benign', it is not quite sure whether it is so. It seems that the kind of justification Frege (1918) needs to provide when he says that 'everyone is presented to himself in a special and primitive way, in which he is presented to no-one else' (1918: 298) is also needed by Perry when he says "only I can have that belief by being in that state" (1979: 19).

2.2.2.3. De se and centered possible worlds: Lewis (1979)

Lewis (1979) makes the same point as Perry (1979) about the inadequacy of propositions, but his solution involves relativised propositions, an idea that as we saw was rejected by Perry. As Lewis explains, there are attitudes that cannot be captured by traditional propositions understood as sets of worlds: these are attitudes that have to do with oneself, which he dubs 'attitudes de se'. Apart from the messy shopper we saw above, other examples illuminating the inability of classical propositions as sets of possible worlds to capture these attitudes is the story of Rudolf Lingens (from Perry 1977[7]).

Rudolf Lingens is an amnesiac lost in Stanford library containing all possible information about the world he is inhabiting – even a biography of a person called Rudolf Lingens lost in the Stanford Library (we assume it was written by an oracle of the past who had predicted the future at least up to the specific moment when Lingens finds himself in this specific place in Stanford). Even if Rudolf reads the contents of the entire library, he will not be able to find his way out of there; there is no propositional knowledge that can set him free. He will only be able to find his way out when he comes to the realisation that he himself is Rudolf Lingens, the person whose biography he just read, recognises his exact location from the biography and matches it to his surroundings in reality so as to find his way out. But as in the example of the messy shopper, these beliefs involve indexicals, thus cannot be rendered as propositions without distorting them.

[7] Although Lewis had an example of his own, that of the two gods, I will not use it here due to its conceptual complexity and possible incongruence, cf. Stalnaker (2008: 56): 'It is not obvious that the coherence of this story will survive close examination (can different agents perform different actions, without realizing, as they act, which one of them is the agent of which action?)'.

This example serves to contrast two different kinds of knowledge: the knowledge acquired from books versus a 'You are here'-kind of knowledge. The former provides Rudolf's self-location in logical space, i.e. what the world he inhabits is like, which can be captured in propositional terms, or else as belief that certain worlds are possible candidates for the actual. However, the latter kind of knowledge can reveal his self-location in physical space, i.e. where exactly he is in the world he inhabits. It turns out that this cannot be captured in terms of possible worlds, because they are too coarse-grained to distinguish between different individuals inhabiting them; in our example, Rudolf cannot possibly distinguish whether he is Rudolf Lingens, the person lost in Stanford library, or any other of other personages lost in libraries he read about. It thus seems that what we need is a way to mark location within a world, which Lewis proposes to do using the idea of centered possible worlds, which he borrows from Quine (1969). Specifically, a centered world is to be understood as a pair of a world and a designated inhabitant thereof, its centre. A class of centered worlds corresponds to a property, namely the property of constituting the centre of the specific centered world.

It can be shown that centered possible worlds are fine-grained enough to capture attitudes de se; Rudolf Lingens' realisation that he is Rudolf Lingens corresponds to identifying oneself as a member of the subpopulation of the actual world who has the property of being Rudolf Lingens, the person the biography is about. In other words, by self-ascribing the property of being Rudolf Lingens, the guy in aisle five, floor six of the Stanford Library.

So Lewis suggests that we can account for belief de se by considering also a centre where we were merely considering possible worlds. In this way we get centered possible worlds, which can be shown able to capture everything possible worlds could, but also the finer-grained beliefs involving self-location that were previously unaccountable for. Of course, we could keep possible worlds for non-de se attitudes, but for purposes of uniformity and elegance, we can adopt centered possible worlds across the board by recognising that in these cases the centre is simply irrelevant. For example, when it comes to realising his identity, Rudolf's knowledge can be represented as self-ascribing the property of being Rudolf Lingens. This demonstrates the superiority of properties to propositions: within the same world, different individuals may be distinguished by having a property or not, but as this is taking place within the same world this distinction cannot be captured by means of a proposition. Now, when it comes to Rudolf realising that in the world he inhabits elephants have two trunks, a piece of information he acquired by reading books, his knowledge can also be represented in property terms, namely as self-ascribing the property of inhabiting a world where elephants have two trunks. If we decide to use properties across the board, we can distinguish cases where self-ascription concerns the centre (being Rudolf Lingens) from those that it does not (inhabiting a world where elephants have two trunks) by saying that the former are interesting centered propositions whereas the former are boring centered propositions, following Egan (2006). This allows us to retain uniformity in characterising the objects of belief, but also to distinguish between the sub-cases.

Since believing a proposition was explained as believing that a certain set of worlds is possible, the belief that a set of certain centered worlds is possible is understood as the

subject's self-ascription of a certain property, namely being the centre of these worlds entertained as possible. In this way, the subject is in-built in the belief relation, which manages to render relativised propositions immune to Perry's objection that anyone could entertain any proposition relativised to anyone: while entertaining a Lewisian relativised proposition, the subject necessarily does so in a first-person way by the self-ascription of being in the centre. Therefore Lewis succeeds in offering an account of belief de se, but as it was the case for Perry, by means of a revision of the thesis that belief is a relation to a proposition: it is now a relation to a centered proposition, or else to a property.

2.3. The problem of de se communication

As we have seen so far, indexicals seem to require different treatment from non-indexical terms, and we have seen three different accounts for them. For Frege, indexicals differ from other expressions in that their senses are private and incommunicable. For Perry, indexicals do not have senses but are better analysed in terms of roles (1977) or else character (Kaplan 1989). They are essential and ineliminable for expressing beliefs concerning oneself, and since the classical doctrine of belief as a relation to a proposition cannot capture this, Perry complements it with the idea of a belief state, which is the way one believes a proposition. Lewis accommodates the essential character of indexicals by the idea of relativised propositions, which corresponds to the speaker's self-ascription of a property, namely constituting the centre of a certain possible world. Each of the above solutions has its merits and drawbacks, but they all seem to face one common problem: the problem of de se communication.

First, we need to establish what we mean by communication, and then see how de se information may fit into it. One of the most influential proposals is Stalnaker's (1978) model of communication. According to this model, conversation takes place against the background of a body of information that is taken for granted by all participants to be commonly shared, and it is commonly known that it is the case. As an assertion is understood as an act of proposing to add a proposition to the common ground, if it is unanimously accepted (not objected to), then it will be added to this common ground. As the conversation progresses, the more propositions are added to the common ground via the effects of assertions, the further conversationalists manage to narrow down the 'context set', which is the set of possible worlds compatible with what is presupposed.

To see an example of how this model works, consider the case in which Eleana and Georg are discussing cat breeds. During the conversation, Eleana utters the sentence "Burmillas originate from the United Kingdom". According to Stalnaker's model, the effect of Eleana's assertion was that the proposition which corresponds to the set of possible worlds in which Burmillas originate from the UK has been added to the common ground. As a result, all possible worlds in which Burmillas originate from anywhere else in the world have been eliminated from the context set. Now it is commonly presupposed by both Georg and Eleana that Burmillas are from the UK. Later in the conversation, Georg contributes by uttering the sentence "Cymrics as well,

specifically from Wales". Nevertheless, this is contested by Eleana who utters "This is not true; Wikipedia says they don't come from Wales, although the name sounds Welsh". In this case, although Georg's utterance attempted to narrow down the context set by eliminating worlds in which Cymrics come from anywhere else but Wales has been flagged down by Eleana's dissent. Due to the rejection of Georg's update, the context set has not been changed[8].

As the above shows, this model of communication can successfully account for a conversational exchange. But both the assumptions shared as well as the contributions made by conversationalists are represented as classical propositions, which as we have seen are inadequate for capturing de se attitudes. And it is without doubt that people constantly convey such attitudes in their communicative exchanges, which makes it imperative that any would-be successful model of communication should account for them. Let's see at which cost the different proposals we have examined can fit into this model of communication.

For Frege and Perry, who each proposed a kind of limited accessibility, albeit to a different extent, in theory this model of communication based on propositions can be retained, but at the cost of ceasing to tell the entire story of what communication in fact is: as Frege admittedly says for Dr. Lauben, 'He cannot communicate a thought which he alone can grasp' (1918: 298). In Perry's case, propositions were maintained to play the roles of the objects of belief, but were complemented by belief states as far as cognitive significance was concerned, which does not seem to be accounted for by this communication model. We know very well from everyday communication that when people use the indexical 'I' we understand them to be talking about themselves qua themselves, in the first person, and our actions confirm this: when someone says "My pants are on fire", we immediately understand why he is rolling on the ground, while we run to get the fire extinguisher.[9] Analogously, when people use the indexical 'you' we understand them to be talking to us directly, in the second person: when someone tells us "Your pants are on fire", we immediately fall down and start rolling on the ground, while we call for someone to grab the fire extinguisher. These different actions would remain mysterious if what was communicated was solely a classical proposition such as "X's pants are on fire"; while Frege and Perry do indeed account for the cognitive significance that is linked to action by private senses and belief states respectively, their additions render this model of communication incomplete at best.

Following our discussion about Lewis and his proposal for substitution of propositions by properties across the board, an idea is to replicate this model of communication employing this replacement of possible worlds by centered possible worlds. Let's examine an example in which Eleni and Luca are camping in the woods and they are trying to light a fire. While Luca is caught up in the effort, he suddenly utters "My pants

[8] Although to be more accurate, we should say that the context set has been updated minimally to exclude worlds in which Cymrics come from Wales, as this was blatantly rejected by Eleana and not protested by Georg. In order to have had no effect whatsoever in the context set, Georg would have to say something along the lines of "It is not clear at all where Cymrics come from, there is no information whatsoever", in which case every possibility is open as far as the origin of Cymrics is concerned, which guarantees that the context set has not been narrowed down.
[9] This well-worn but effective example is from Kaplan (1989).

are on fire!" Following our adapted model, the effect of Luca's assertion is that the centered proposition which corresponds to the set of centered possible worlds in which the centre's pants are on fire has been added to the common ground. As a result, all centered possible worlds in which the centre's pants are not on fire are eliminated from the context set. Now it should be commonly presupposed by both Eleni and Luca that the centre's pants are on fire. Unfortunately, this prediction turns out to be wrong. If Eleni actually does presuppose this, she will have to accept that her own pants are on fire, not Luca's. This happens because who the centre represents is a function of who the attitude holder is. That is to say, if Eleni accepts Luca's utterance, she will stand in the belief relation with the centered proposition that the centre's pants are on fire, which corresponds to her own pants being on fire. The fact that this does not happen is confirmed by the fact that as soon as Eleni hears Luca say that, she immediately rushes to get the fire extinguisher and spray Luca's pants with foam; it is thus clear to Eleni that it's Luca's pants and not her own that are on fire, contrary to the predictions of our adapted model.

From all the three de se accounts that we have investigated, the Lewisian one has been the most popular one as far as its adaptation by linguists is concerned. Unsurprisingly, it has also been offered modifications so as to render it compatible with a communication model. Let's see such a proposal by Ninan (2010).

First, Ninan starts by extending centered propositions to pair-centered propositions, i.e. propositions that are triples of a possible world, a centre, and the centre's addressee, e.g. <w, x, y>. This enables us to represent both the speaker and the addressee but does not yet help; similarly to before, the fact that each conversationalist him/herself represents the first centre of the centered proposition is the source of the problem. Specifically, when Luca uttered "My pants are on fire" Eleni was mistakenly supposed to presuppose the centered proposition that the centre's pants are on fire, namely her own. But as Ninan shows, the addition of a second entity in the sequence, the centre's addressee, could save us if we stabilise the sequence. Such a move would allow a pair-centered proposition to be transferred from Luca to Eleni via the communication process, but would keep references fixed. In order to accomplish that, instead of taking the first centre to be a function from the attitude holder, Ninan posits that the first centre represents whoever comes first in the conversational sequence. The order can be totally arbitrary, all that matters is that we make a selection and stick to it.

In our example, let's take it that Luca comes first in the conversational sequence, i.e. that the sequence is <Luca, Eleni>. Then if Eleni accepts and presupposes the pair-centered proposition that the first centre's pants are on fire, she will presuppose that her interlocutor's pants are on fire, who happens to be Luca (not hers, since she is the second centre). So she will presuppose this pair-centered proposition in a de te way, not a de se one, but not merely a de re either (which would be the case for a classical proposition)[10]. This matches well with her reaction of running to get the fire extinguisher and spraying the person she was talking to. This account works equally

[10] A note on terminology here: 'de te' means 'of you', so a de te way to understand an expression is to understand it as referring to the addressee in a given context (it is essentially the second-person equivalent of 'de se', which is first-person). 'De re' means 'about the thing', and amounts to the specific entity being referred.

well for de te utterances (e.g. “Your pants are on fire!”), which are to be accepted and presupposed as de se by the person they are addressed to, based on the fixed conversational sequence. And of course, it can distinguish between de se from non-de se (and de te from non-de te); in such cases, the entities referred to are simply not represented in the conversational sequence, e.g.

(12) Luca to Eleni: *I love you!* = {<w, x, y>: x loves y in w} (de se and de te)

(13) Luca to Eleni: *I love Eleni!* = {<w, x, y>: x loves Eleni in w} (de se, non-de te)

(14) Luca to Eleni: *Luca loves you!* = {<w, x, y>: Luca loves y in w} (non-de se, de te)

(15) Luca to Eleni: *Luca loves Eleni!* = {<w, x, y>: Luca loves Eleni in w} (non-de se, non-de te)

Following this solution, belief is understood as relation to a pair-centered proposition relative to a conversational sequence which is crucially fixed. According to Ninan (2010: 562):

> Note that the notion of believing a pair-centered proposition relative to a conversational sequence can be explained in terms of believing a centered proposition. When we say that John believes the pair-centered proposition {<w, x, y>: P(w, x, y)} relative to the conversational sequence <John, Mary>, we could understand this as saying that the following two things are true: (i) Mary is John’s addressee, and (ii) John believes the following centered proposition:
>
> {<w; x> : there is a y such that y is x’s addressee in w; and P(w, x, y)}

However, it may sound a bit cryptic that ‘someone believes something relative to a conversational sequence’. The idea of a fixed sequence as we defined it is a theoretical stipulation to help get this account off the ground, and not something of which we predicate any cognitive reality. It is thus important to emphasise here the difference between the theorist’s goal of providing an objective account and the actual beliefs in the conversationalist’s mind, which can never be stripped from perspective. Establishing a fixed order in the conversational sequence and sticking to it is a necessity for the theorist striving to account for the communication of attitudes such as belief – for the conversationalist, it is expected that the order is fixed in such a way that they are always first. Therefore, when this account presents belief as an attitude to a pair-centered proposition relative to a conversational sequence, it is important to clarify that this definition aims to be a theoretical, therefore objective one, and not a description of the attitude from the subject’s point of view.

With regards to the semantic representation of Ninan’s proposal, the content of a de se (or de te) attitude involving utterance is represented as a pair-centered proposition, which corresponds to a set of pair-centered worlds (the set of all pair-centered worlds in which the pair-centered proposition is true).

Now according to the Kaplanian model, indexicals get their reference from the context, while the extensions of non-indexical expressions, including propositions, are extracted against certain circumstances of evaluation. Technically, this is done by doubly

relativising the interpretation function both to a context and to a possible world. As we saw before, Kaplan defined content as a function from circumstances of evaluation to extensions, which in the case of propositions (even pair-centered ones) correspond to truth values. Let's follow Kindermann (2012) and call this standard kind of content the 'Kaplan horizontal' (2012: 92). Summarising, we can say that the 'Kaplan horizontal' of a sentence S corresponds to the set of worlds at which S is true.

Kaplan horizontal of S: {c, i: $[\![S]\!]^{c,i} = 1$}

But in the case of centered propositions, evaluation should take place against entities which are fine-grained enough to include individuals, or in the case of pair-centered propositions, against entities which include a pair of individuals. As Schlenker (2011) notes, such entities are in fact homologous to contexts. This means that in order to extract the extension of propositions involving de se attitudes, the interpretation function should be doubly relativised to a context (rather than to a context and circumstances of evaluation, as in the case of the Kaplan horizontal above). Or in other words, the content of a pair-centered proposition corresponds to the 'Kaplan diagonal' (Ninan 2010: 563) of the sentence uttered. Essentially, the Kaplan diagonal of a sentence S is the set of contexts at which S is true.

Kaplan diagonal of S: {c: $[\![S]\!]^{c,c} = 1$}

Let's see now how all this fares with our original scenario.

(16) Luca: "My pants are on fire!" = {<w, x, y>: x's pants are on fire in w}

The Kaplan diagonal of this sentence is the pair-centered proposition in which the first centre's (x's) pants are on fire. Since we previously agreed that Luca comes first in the conversational sequence, what is communicated is that Luca's pants are on fire. Thanks to the idea of a fixed conversational sequence consisting of only two individuals, in order to represent the second centre's de se attitudes, all we need to do is take the inverse of the Kaplan diagonal[11].

(17) Eleni: "My pants are on fire!" = {<w, x, y>: y's pants are on fire in w}

Since Eleni comes second in the conversational sequence, instead of the Kaplan diagonal which will give us Luca, we take the inverse of the Kaplan diagonal, which in this case will correctly give us Eleni.

Summarising, given a fixed conversational sequence <x, y>, if x utters a sentence S, the content of her utterance is the diagonal of S; if y utters a sentence S, the content of her utterance is the inverse of the diagonal of S. As a result, the Kaplan diagonal of a sentence is always the same irrespective of who the speaker (or the addressee) is; it always corresponds to the set of contexts at which the sentence is true. Faithfully to the original model, the content of an utterance is what the participants all come to presuppose if the utterance is accepted, namely what updates the context set.

[11] It should be clarified here that the inversion applies only to the sequence of individuals (i.e. <x, y> becomes <y, x>), and not to the triple which includes the world (<w, x, y>).

2.4. Stalnaker's pragmatic alternative

Besides the reformatory proposals by Perry (1979) and Lewis (1979), there is also a more conservative account offered by Stalnaker (1978, 1981), which consists in putting the explanatory burden of cognitive significance in the pragmatics. In order to understand it, we should first go into Stalnaker's model of communication with a bit more detail.

As we saw above, facts about the world (e.g. that Burmillas do indeed come from the UK) determined the truth value of Eleana's utterance "Burmillas come from the UK" and eventually its addition to the common ground. But according to Stalnaker, there is another way that facts about the world affect assertions, specifically in determining what was said by an assertion. This is more easily demonstrated by utterances containing context-dependent expressions such as indexicals, whose content cannot be determined without considering the relevant context. This double role of possible world facts in determining the content of an assertion, i.e. both what was said and whether or not it was true, can be captured by constructing a table for the utterance "This cat is ginger":

		(possible worlds as facts that determine truth-value)		
		i: actual world (Junior = ginger, Little Chief = Aegean cat, Soya = Russian blue)	j: world in which Junior = Russian blue, Little Chief = ginger, Soya = Aegean cat	k: world in which Junior = ginger, Little Chief = Russian blue, Soya = Aegean cat
(possible worlds as context that determines what is said)	i: *Junior is ginger.* (the cat pointed at is Junior)	**T**	F	T
	j: *Little Chief is ginger.* (the cat pointed at is Little Chief)	F	**T**	F
	k: *Soya is ginger.* (the cat pointed at is Soya)	F	F	**F**

Table 2.1: Stalnaker's (1978) propositional concept

What a table such as the above represents is called a 'propositional concept', as it maps a world to the proposition a certain utterance would have at this world. Assuming that every utterance token corresponds to such a propositional concept, Stalnaker claims that some operations that determine the interpretation and evaluation of assertions act on the propositional concept of an utterance, rather than merely on the proposition. One such operation is the 'diagonalisation' operation, in which interlocutors associate an utterance with the proposition that corresponds to the diagonal of the propositional concept. The motivation for doing so is to prevent a communication failure. Let's see how this works in some detail.

Successful assertion of an utterance such as "This cat is ginger" means that interlocutors were able to determine the relevant proposition and assign a truth value to it, so as to narrow down the context set accordingly. Since by definition the purpose of communication is to reduce the number of possible worlds compatible with what is presupposed, interlocutors do not know which world they are in, which means that they may not be able to determine what is said by the utterance. As this uncertainty would result in communication breakdown, Stalnaker suggests that in such cases, assuming that the speaker is being truthful as per Gricean principles, interlocutors understand the proposition as that one which is true at a world i for any world i, if and only if what is expressed by the utterance in i is true at i, which in fact corresponds to the diagonal proposition of the relevant propositional concept.

In our example, by uttering "This cat is ginger" the speaker could be interpreted as telling his interlocutors the following: if we are in world i (where the cat pointed at is Junior), eliminate world j from the context set; if we are in world j (where the cat pointed at is Little Chief), eliminate worlds i and k; and if we are in world k (where the cat pointed at is Soya), eliminate them all. But of course, the third option would be self-defeating, as the purpose is to narrow down the context set as much as possible, but eliminating it altogether would amount to massive communicative failure. That is to say, by taking the diagonal of the propositional concept corresponding to the speaker's utterance, interlocutors succeeded in narrowing down the context set from three to two worlds.

How does this all relate to our discussion about attitudes de se? Essentially, Stalnaker's proposal is to maintain classical propositions as the objects of attitudes, and explain the critical de se element in terms of the context in which the relevant attitude is ascribed, making use of the diagonal of the propositional concept. In this way, the self-motivating effect of an indexical belief such as "I am making a mess" would not be analysed as a relativised proposition (à la Lewis) or attributed to an obscure 'belief state' (à la Perry), but would be understood in terms of the diagonal proposition of the propositional concept in question. As the propositional concept of an utterance depends on the set of worlds that are possible in each situation, this means that the explanatory burden is moved towards the pragmatics, which seems like an inevitable move if we want to preserve the classical doctrine of propositions. Obviously, as this account is based on Stalnaker's model of communication, it is perfectly compatible with it. However, it may feel that such an account of attitudes de se does not do justice to them, given their prime importance in explaining agency and action. Despite the shortcomings of the proposals by Perry and Lewis, at least those succeeded in capturing the intuition that there is something special about de se attitudes[12], while Stalnaker's account shifts the attention to the circumstances in which the attitude is expressed rather than putting any emphasis on the attitude itself.

[12] Of course, we should here mention that there are theorists that deny the thesis that there is anything special about attitudes de se, the so-called 'de se skeptics' Millikan (1990), Cappelen & Dever (2013) and Magidor (2015). As I find that the skeptic argument is convincingly dispelled in Torre (2016), I will not concern myself with this line of thought in this dissertation.

2.5. De se in the LF

Although as we saw there is persuasive evidence from the philosophical literature that attitudes involving a de se element cannot be adequately captured as classical propositions, the relevance of this philosophical thesis to linguistic theory is not straightforward but needs to be independently motivated. Specifically, this issue extends to the semantics of natural language only if we can show that the distinction between reporting a de se attitude versus a non-de se one is encoded. In such a case, we would need our linguistic theory to provide different representations for sentences differing in the respect of encoding the de se element or not. Interestingly, it turns out that this is indeed the case.

To illustrate this, let's see an example of a belief report.

(18) Alexis believes that he is a creative guitarist.

Thinking about possible situations in which this could be judged true, we can certainly think of two different ones. In the first one, Alexis is listening to the recordings of his band and he is so impressed by how creative his playing is that he is certain that he will win the Psychedelic Riff competition organised by the Psychedelic Funk society. In the second one, he is listening to some recordings that his friend sent him, and without realising that these are actually the recordings of his own band, he is so impressed by the creative playing of the guitarist that he is certain that the guitarist of this band will win the competition.

The fact that the above sentence is marked true by two distinct situations does not by itself constitute a motivation of linguistic interest. It could well be the case that natural language does not explicitly encode the difference between de se and non-de se attitudes, which would mean that despite the topic being greatly interesting for philosophers, it has no linguistic importance. Nevertheless, it turns out that natural language does indeed encode such a difference; we can see that both in English as well as in other languages. Specifically about English, let's contrast the following sentences:

(19) Alexis expects that he will win the Psychedelic Funk society competition.

(20) Alexis expects to win the Psychedelic Funk society competition.

It turns out that (19), just like (20), is equally true in the situation where Alexis recognises the guitar playing as his own as well as in the situation where he doesn't know who the guitarist is but is perceptually acquainted with his playing in a way that he can judge that that person is sure to win the competition. Nevertheless, (20) seems to apply only to the first situation, in which Alexis would express his thought using the 1st person pronoun 'I', thinking to himself "I will win the Psychedelic Funk society competition". In other words, as (19) can apply to both scenarios it entails (20), but the opposite is not true. This contrast suggests that natural language, in our specific example English, has a way of unequivocally expressing the de se element, which calls for an appropriate semantic representation.

According to Chierchia (1989), the fact that control constructions such as the above have an unambiguously de se interpretation means that there is a designated LF. To

show this, Chierchia reconstructs (19) and (20) attributing an attitude to Alexis on the basis of Lewis' centered worlds account:

(21) expect (Alexis, λx [x is in a place where Alexis wins the competition])

(22) expect (Alexis, λx [x is in a place where x wins the competition])

In this formulation, since the subject is in-built in the relation expressed by 'expect', this relation is understood as self-ascriptive, in the sense that the subject self-ascribes the property of being located in a certain world. But there is a crucial difference between the two: (21) can be captured in possible world terms (Alexis expects that the set of worlds in which a certain individual wins the competition is possible), which following Chierchia we can understand as a 'constant property', or following Egan (2006) a 'boring' one. However, (22) slips a possible worlds treatment, since in this case the subject's expectation necessarily involves himself *qua* himself, which requires positing a centre in the set of worlds he considers as possible so as to make the property 'interesting' (in Egan's terms), namely co-vary with the subject of the attitude.

More formally speaking, the representation for the de se interpretation requires that the variable corresponding to the centre (in the interesting property version) is accordingly bound. Chierchia proposes to do this by introducing an extra lambda binder in the attitude relation, which syntactically speaking corresponds to a property abstractor operator in the left periphery of the clause:

(23) Alexis_i expects that [he_i will win the Psychedelic Funk society competition].[13]

(24) Alexis_i expects [O_i PRO_i to win the Psychedelic Funk society competition].

Apart from control predicates in English that we saw above, it has been discovered that there are other constructions which seem to express unambiguously de se readings, such as shifted indexicals (Schlenker 1999, 2003), long-distance anaphors (Huang and Liu 2001), but also dream reports in English (Lakoff 1972, Percus & Sauerland 2003) etc[14]. We would like our semantic representation to account for this fact, but it is an open question whether the same representation should underlie all these different phenomena.

According to Anand (2006), the answer is negative. He proposes that there are three different underlying representations that end up in a de se interpretation: a de re one (i.e. a non-de se LF), and two cases of de se LF, a syntactic and a semantic one.

First, we saw that in attitude reports pronouns can be read either de re or de se, depending on the context, as in sentence (18), 'Alexis believes that he is a creative

[13] It should be noted that (23) is also compatible with a de se interpretation, in which case the corresponding representation is as follows:
(25) Alexis_i expects [O_i that [he_i will win the Psychedelic Funk society competition]].
However, it is not clear whether a de se interpretation needs to correspond to a special de se LF as the above, or it should be treated in terms of underspecified meaning. One such proposal of deriving de se from de re is by Anand (2006), which will be presented below.

[14] Another relevant work is Pearson's (2012) study of logophors in Ewe, though she notes that they are not obligatorily de se for some speakers.

guitarist'. We saw that this could verify two different scenarios, one in which Alexis is admiring himself and another in which he admires someone else who unbeknownst to him, is himself. Given the optionality between the de re and the de se readings, it feels natural to say that in such constructions there is no de se-specific LF. But according to Anand, this line of thought could even be extended to cases with obligatorily de se readings, such as Chierchia's control structures that we saw above. Following Kaplan's (1968) proposal that de re is a kind of de dicto under a suitable description expressing a relevant acquaintance relation, as long as the description yields the object in the actual world,[15] Anand's argument is that de se can be analysed as a species of de re in which the relation of acquaintance is reduced to identity. In the case of pronouns, which allow for both de se and non-de se interpretations, this proposal appealing to a derived index will apply only in the first case; in the case of control structures, it needs to be complemented with the assumption that PRO is a lexical item that amounts to such a description analysed in a derived index. This enables Anand to account for the obligatorily de se interpretation of control predicates without making use of a dedicated LF à la Chierchia, which applies to pronouns as well when needed.

Nevertheless, Anand contends that other de se cases do seem to require special de se LFs, although even in such a case we have two different instances, one syntactic and one semantic. In the first one, the special de se LF is the product of a designated syntactic representation, in which the de se element is said to be bound by an operator in the scope of the attitude verb. An example of this syntactically induced de se are English dream reports, such as the well-known example by Lakoff (1972):

(26) Last night I dreamed I was Brigitte Bardot and I kissed me.

a. Last night I dreamed I was Brigitte Bardot and $I_{B.B.}$ kissed $me_{G.L.}$

b. * Last night I dreamed I was Brigitte Bardot and $I_{G.L.}$ kissed $me_{B.B.}$

As shown above, the sentence is strictly interpreted as that the dream-self Bardot (de se) performs the kissing on the actual-self Lakoff (de re), while the opposite interpretation in which the actual-self (de re) does the kissing on the dream-self (de se) is impossible. Anand suggests that the unambiguous interpretation of examples such as the above is evidence of a syntactic intervention effect called the 'De re blocking effect', according to which no syntactic de se anaphor (an anaphor bound by an operator) can be c-commanded by a de re one. Moreover, Anand also observes this syntactic effect in Yoruba logophors.

The second instance of a dedicated LF is purported in the semantics, and it accounts for the unambiguously de se readings of shifted indexicals observed in some languages such as Amharic (Schlenker 2003), Zazaki (Anand & Nevins 2004), Uyghur (Sudo 2012) and others. According to Anand, this is not a result of syntactic binding, but involves shifting the context parameter (which as we saw above is responsible for the reference of indexicals) from that of utterance to that of the reported speech act. Assuming Kaplan's

[15] The motivation for this proposal by Kaplan (1968) was to provide an account for Quine's (1956) so-called 'double-vision' problem (termed such by Klein 1979), according to which if de re is analysed in terms of scope, we may end up attributing contradictory beliefs to a rational subject in cases where she is acquainted by an object under different guises.

system explained above, according to which the interpretation function is doubly relativised to a context as well as an index parameter, the proposal here is that certain attitude predicates introduce an operator that forces overwriting the context parameter and copying the index one in its place. In this way, the indexical expressions under the scope of this operator fail to secure their reference from the current context of utterance, as bona fide indexicals do, and get it instead from the index parameter that has replaced it[16]. As Anand notes, this process is precisely what Stalnaker's (1978) diagonalisation operator does:

(27) $[\![OP_{\delta}\ [\alpha]]\!]^{c,i} = [\![\alpha]\!]^{i,i}$

Below is an example in pseudo-Zazaki (from Anand & Nevins 2004: 27):

(28) ZAZAKI: John said to Bill that I am mad at you.
$[\![OP_{\forall}$ [I am mad at you]$]\!]^{c,j}$ =
= $[\![$[I am mad at you]$]\!]^{j,j}$ =
1 iff AUTH(j) is mad at ADDR(j) in j =
= 1 iff John is mad at Bill in j

Of course, as Anand points out and as we saw above, Stalnaker proposes diagonalisation as a strategy interlocutors resort to in order to preserve the pragmatic principles and salvage communication from failing. Therefore, while for Stalnaker diagonalisation is a non-propositional operation that acts upon utterances and not sentences, Anand's implementation of it is clearly a semantic one, as it successfully accounts for the shifted reference of indexicals.

[16] As it can be seen by the rich cross-linguistic data analysed by Anand, indexical shifting does not work identically in all the languages that allow it; some allow all indexicals to shift (Zazaki), others only person indexicals (Amharic, Slavé), and there are even intra-language variations (some predicates in Slavé allow all person indexicals to shift, others only the 1st person). Since the shifty operators are posited as introduced by the relevant attitude predicates in this analysis, Anand notes that both the cross- as well as the intra-linguistic variation can be reduced to lexical idiosyncrasies of these predicates (Anand 2006: 110).

Chapter 3. De se across dimensions

In the previous chapter we reviewed the philosophical background on the special character of attitudes de se, as well as its relevance to the semantics of natural language in the case of de se attitude ascriptions. Crucially, all the linguistic accounts for the de se, whether it was analysed as a special kind of de re under identity, or attributed to syntactic binding or a context-shifting operator, all concerned the at-issue dimension of meaning. Nevertheless, it has long been shown that natural language meaning spans across different dimensions (e.g. Karttunen and Peters 1979, Büring 1999, Dekker 2002, Potts 2005, 2007, Beaver et al. 2017, among many others). If we are convinced by the philosophical arguments about the special character and importance of the de se, then we could anticipate that it could appear in other dimensions of meaning besides the at-issue one. Specifically, this section will examine its relevance to the so-called expressive dimension (Potts 2005, 2007).

3.1. The ontology of meaning: more than one dimension

3.1.1. Presupposition

As noted by Karttunen & Peters (1979), there are certain expressions or syntactic constructions conventionally prone to convey propositions separate from that corresponding to the sentence they are contained in; in other words, these expressions contribute 'an aspect of meaning distinct from the kind of semantic content that is the subject matter of ordinary truth-conditional semantics' (1979: 1). Traditionally, these additional propositions have been identified as presuppositions, i.e. preconditions that need to hold for a proposition to have a truth value; this idea has its origins in an intuition by Frege (1892) and got its explicit formulation by Strawson (1950).

The most identifying characteristic of presuppositions is projection, namely their ability to elude any operators that usually target the meaning of assertions. This can be readily diagnosed by the 'Family of Sentences' (FoS) test (Chierchia and McConnell-Ginet 1990), which shows that a presuppositional inference projects across the following operators: negation, question, modality, and the antecedent of a conditional:

(29) a. Junior stopped chasing squirrels.

Assertion: Junior does not chase squirrels[17].

Presupposition: Junior used to chase squirrels in the past.

b. Junior hasn't stopped chasing squirrels.

c. Has Junior stopped chasing squirrels?

d. Possibly Junior has stopped chasing squirrels.

[17] Note that another possible analysis of the assertive content here is 'Junior used to chase squirrels but does not anymore'. Specifically, such an analysis allows the distinction of presuppositions in entailed and non-entailed (Sudo 2012).

e. If Junior has stopped chasing squirrels, we don't have to worry about rescuing them from his fury.

In the above example, although the content asserted in 29a (that Junior does not chase squirrels) is negated, questioned, marked as possibly rather than definitely true and hypothesised as true, we can see that the inference that Junior used to chase squirrels in the past persists across all sentences. This persistence across semantic operators is called projection, and it has been considered as the hallmark of presuppositions[18].

As noted above, traditionally presuppositions were thought of as conditions imposed by certain expressions that needed to be satisfied for there to be a truth value in the sentences in which these expressions appeared. Due to this dependence between presuppositions and the sentence which included them, the idea that every proposition refers to only one of the truth values of True or False had to be complemented with the notion of undefinedness, which is the situation for propositions which cannot receive either truth value due to their presuppositions being false. Using our example, if the presupposition that Junior used to chase squirrels in the past is proven untrue, then it is not possible to say whether 29b is true or false, and the truth value of the corresponding proposition is said to be undefined. Therefore in this theory a presupposition is understood as a semantic relation between different propositions.

However, there is also a pragmatic account of presupposition, pioneered by Stalnaker (1974). As we saw when we reviewed Stalnaker's model of communication (1978) in the previous chapter, the propositions corresponding to the background assumptions mutually shared by all interlocutors in a conversation are understood as presuppositions. According to this line of thought then, a presupposition is no longer a semantic relation between propositions or a property of specific expressions but is explained in terms of the beliefs of interlocutors. As noted before, utterances serve to update the context by increasing the common ground (add more propositions to the stock of presuppositions) and thus decrease the context set (the set of possible worlds which could correspond to the actual one). This means that an utterance that carries certain presuppositions will be accepted as a valid conversational move only if these presuppositions are already satisfied, i.e. they are already part of the common ground. If they are not, but are presented as such by the speaker, the context turns out to be 'defective' and communication may break down.

Of course, this is just an idealised description of how things work, as we can always expect shortcomings in the interaction between human agents, such as a mismatch of background beliefs. However, assuming cooperative behaviour between different agents,

[18] Of course, presuppositions do not always project indiscriminately; if they did, accounting for them would be straightforward. Following Karttunen (1973), in certain environments presuppositions always project (holes), in others they are blocked (plugs) and in others they may do either depending on the circumstances (filters). An example where a presupposition fails to project is when the antecedent of a conditional satisfies the presupposition of the consequent:
(30) If Junior ever chased squirrels, then he stopped doing so.
Therefore it is the combination of projection and cancellability that makes accounting for them so challenging. However, as this section concerns the general property of projection, the specific conditions under which it does not take place need not concern us here, despite being essential for a proper account of presupposition.

it is in their interest for communication breakdown to be avoided at all costs. As a result, we can expect that interlocutors may be willing to admit into the common ground certain assumptions made by the speaker as if they were presuppositions, provided those are not too controversial. This process is known as 'accommodation' (Lewis 1979) and can be understood as a corrective strategy interlocutors resort to in order to prevent communication from failing altogether. In our example, the utterance 'Junior didn't stop chasing squirrels' would be felicitous in a context where it is part of the common ground that Junior used to chase squirrels. However, we can expect that if it is uttered in context where this piece of information is not commonly shared, interlocutors would more likely be willing to accommodate it, so that worlds in which he stopped chasing squirrels can be eliminated from the context set, rather than halt the discourse and protest to the speaker that this was not commonly shared, in which case speakers would deny the utterance any effect on the context set. Understandably, accommodation is a rescue rather than a routine strategy because in a cooperative conversation we expect speakers to contribute novel information by asserting it as new, and not trying to sneak it in by masking it as old.[19]

3.1.2. Projection beyond presuppositionhood

However, although projection seems to be regarded as a necessary condition for presuppositions, there is no consensus that it is a sufficient condition, as some theorists have proposed that other kinds of inferences can project, but they cannot be classified as presuppositions. An early noted case was that of non-restrictive relative clauses (NRRCs), which seem to project across the operators of the FoS test but feel much more like backgrounded assertions rather than presuppositions (Chierchia & McConnell-Ginet 1990). More recently, Potts (2005) re-adopted the old Gricean term of 'conventional implicatures' (CIs) to identify a class of meanings which although projective, are claimed to be distinct from presuppositions and completely independent from the asserted content, namely as functioning in an altogether distinct dimension of meaning. According to Potts, expressions which trigger CIs are supplements, such as appositives and parentheticals, and expressives, such as epithets and honorifics. Although the entire argument and system by Potts will be reviewed below in detail, what is relevant for the discussion here is to emphasise that projection can no longer be considered a distinctive characteristic of presuppositions.

[19] According to another approach (van der Sandt 1992), presupposition can be accounted for as a type of anaphora; this means that presupposition projection can be explained in terms of anaphora resolution. The idea in a nutshell is the following: we know that an anaphor is felicitous only when it has an antecedent in the discourse. But one important difference between presupposition triggers and pronouns is that unlike the former, the latter are mere variables, devoid of descriptive content of their own. Exactly due to their descriptive content, it is claimed that if presuppositions do not have an antecedent, they are able to force one into the discourse, which produces the projection effect. I will not be making use of this approach to presupposition as anaphora in this dissertation, and as my analysis does not depend on a particular theory of presupposition hopefully this is not of any concern.

3.1.3. Projection as not-at-issueness

According to Simons, Tonhauser, Beaver and Roberts (2010), projection is best explained in terms of irrelevance to the QUD (Question Under Discussion, Roberts 1996), which corresponds to the semantic question of a certain discourse, i.e. what the occurrence of the specific discourse aims to answer. Assuming that an utterance can correspond to several propositions, we can distinguish between that one which addresses the QUD and which is called at-issue, and those that do not address it as their content is more peripheral, which are called not-at-issue. According to this line of thought, semantic operators target the at-issue proposition but not the not-at-issue ones, which results in them projecting.

Defining thus projection in discourse terms, namely as not directly addressing the QUD, this proposal aspires to provide a unified account of projection, something that it claims has not been possible by other theories as those are more tailor-made either for classical presuppositions, or for Pottsian conventional implicatures. Specifically, the accusation against common ground approaches to projection is that they work for classical presuppositions which are already part of the common ground, but are unable to systematically account for informative implications such as expressives, appositives and non-restrictive relative clauses, all of which have been analysed as CIs by Potts (2005). Even though these approaches could invoke accommodation as a strategy for these projective meanings, this would defeat the point of accommodation as it would require making it from an exceptional strategy to align the common ground when needed to a default and necessary one for this class of expressions. Let's show this with an example.

Context 1: It is well known that Greg used to smoke heavily.

(31) Since Greg quit smoking recently, let's finish our cigarettes before we invite him over.

Context 2: Nothing is known about Greg's smoking habits.

(32) Since Greg quit smoking recently, let's finish our cigarettes before we invite him over.

The presupposition that Greg used to smoke is already part of the common ground in context 1, while in cases such as context 2, where it is not, we expect that if this utterance does not cause a communication failure it is because interlocutors are able to accommodate it. Let's try to compare this with an example involving a NRRC:

Context 3: It is well known that Dimitris loves meat.

(33) # Since Dimitris, who really likes meat, is joining us I suggest we go to a Brazilian barbecue restaurant.

Context 4: Nothing is known about Dimitris' eating preferences.

(34) Since Dimitris, who really likes meat, is joining us I suggest we go to a Brazilian barbecue restaurant.

In theory, we could appeal to accommodation in this case as well to explain why there is no communication breakdown in context 4. But then we need to explain the awkwardness (if not infelicity) noted in context 3. To highlight the contrast, classical presupposition triggers such as 'stop' only appeal to accommodation as a rescue strategy where necessary (context 2, but not 1), whereas for systematically informative inferences such as NRRCs we will need to assume it as an automatic default even when no rescue is needed (context 3), resulting in awkwardness.

Contrarily, multidimensional approaches (Potts 2005) which explain projection in terms of the independence between different dimensions can account for the more informative types but not for classical presuppositions. Specifically, the notion of independence lying at the heart of such accounts is fatally problematic for classical presuppositions (the so-called 'binding problem', Karttunen & Peters 1979), but is in fact a 'virtue' (Potts 2005: 79) for CIs since they are posited as totally independent from the asserted content.

Nevertheless, despite the alleged superiority of this theory that explains projection in terms of not-at-issueness, it seems that it has problems of its own. Although the argument that there is a connection between projection and not being at-issue is extremely convincing, this conclusion seems more relevant to a theory of discourse than semantics. As the authors say (2010: 322), they 'posit that at least some constructions or lexical items conventionally mark their content as not-at-issue', while at the same time admitting that they 'remain agnostic as to which expressions conventionally mark at-issue status'. In other words, it seems that we still need another theory to account for the conventional character of these inferences. Of course, the same could be said for the common ground approach we saw above, but the fact that it was rejected as inferior to the discourse one in terms of failing to account for specific classes of expressions seems problematic now, since the discourse approach itself turns out not to have anything to say about specific classes of expressions either. This is made more evident in subsequent work by Tonhauser, Beaver, Roberts and Simons (2013), in which they posit a 'taxonomy of projective content' composed of four subclasses, three of which cover some classical presupposition triggers and one which includes Pottsian CIs. The subclasses are distinguished based on two different properties, which they call 'Strong Contextual Felicity' (whether the context should already entail the inference) and 'Obligatory Local Effect' (whether the inference can be contained by an operator). Obviously, these properties are directly related to backgrounding, typical of presuppositions, and scopelessness, typical of CIs. Since the authors say nothing about the actual ontology of meaning, namely what a presupposition or a CI is and how they differ (if at all), their taxonomy reads mostly like a list useful for quickly referencing the similarities and differences between presuppositions and CIs which are already known, but devoid of real explanatory value as no elucidation of the conventional properties that make these expressions projective is provided.

3.2. Expressive meaning

As mentioned above, expressive meaning is a kind of meaning that is projective but is posited to constitute a different theoretical entity than presuppositions. This section will review existing accounts and formal models of expressive meaning.

3.2.1. The philosophical beginnings: Kaplan

Expressive meaning, just like presuppositions and other linguistic notions, was first noted by philosophers. Specifically, Kaplan (1999) first advocated that some expressions are better analysed in terms of what they display rather than what they describe; he proposes thus a distinction between 'descriptives', i.e. expressions whose meaning can be captured in terms of truth, and 'expressives', i.e. expressions whose meaning can be understood in terms of rules of use. According to Kaplan, some expressions which fall into this category are epithets, diminutives, ethnic slur terms, honorifics, and interjections. Since the meaning of expressives like the above is conventional, and does not depend on the intention of interlocutors, he advocates that it should be accounted for by a semantic theory and that this is in fact possible, allowing semantic theory in these cases to speak of conditions of use in the place of truth conditions.

3.2.2. Expressive meaning in the spotlight: Potts

However, despite the philosophical origins, it was the pioneering work of the linguist Chris Potts (2005) that brought expressives into the spotlight of research in linguistics, while it also fed back into philosophy by sparking off a resurgence of interest in the topic, especially in the domain of slurs (e.g. Richard 2008, Williamson 2009, Predelli 2010, 2013, Anderson & Lepore 2013, Camp 2013, etc). Specifically, Potts (2005) revived the old term 'conventional implicature' which was coined by Grice (1975/1989) to name a class of meanings which are conventional in nature (like assertions and unlike implicatures) but do not enter the truth conditions (unlike assertions and like implicatures). To give a very short history of the term, Grice himself did not say much about conventional implicatures and his brief treatment of them left more questions than answers. Subsequently, the term was adopted by Karttunen & Peters (1979) to distinguish semantic presuppositions from other inferences that were grouped into this category which they regarded as grossly heterogeneous; although their proposal for a multidimensional treatment was very influential, their naming suggestion went down completely ignored. Eventually, Bach (1999) attempted a final blow by dismissing the whole idea of conventional implicatures as a 'myth'; he claimed that assuming that a single sentence can correspond to more than one propositions, Grice's original conventional implicature triggers ('therefore', 'but', etc) can be analysed as contributing secondary assertions thus the notion of conventional implicatures can be completely discarded.

However, Potts (2005) argued that even if Grice's original examples don't fit the bill as Bach convincingly argued, there are in fact expressions that do. Specifically, Grice says that conventional implicatures are part of the 'conventional meaning of words' thus by using them the speaker 'certainly commits [himself] by virtue of the meaning of [his] words', although their content is not 'said (in the favoured sense)' though it is 'certainly indicated, and so implicated, that this is so' (1975/1989: 44-5). Potts argues that supplements and expressives are expressions that fit the description by Grice and provides a rigorous logic for CIs that is able to reflect their properties, which can be summarised as follows:

1. Scopelessness: CIs can never be under the scope of any operator, but always project out to claim widest scope. Of course, this means that they successfully pass the FoS test:

(35) a. Mike, who is an incredible dancer, is applying to med school.

b. It is not true that Mike, who is an incredible dancer, is applying to med school.

c. Is Mike, who is an incredible dancer, applying to med school?

d. Maybe Mike, who is an incredible dancer, will apply to med school.

e. If Mike, who is an incredible dancer, applies to med school, then he will have to study really hard.

However, unlike presuppositions which were showed to also pass this test, CIs project across the board. For example, CIs cannot be blocked by presupposition plugs such as non-factive predicates like 'believes'; contrarily, they escape and take wide scope, a characteristic called 'unpluggability' (Sauerland 2007):

(36) a. Brittany believes that her cat stopped smoking. But cats can't smoke, I don't know how she got that idea in the first place.

The presupposition that the cat used to smoke is contained inside the scope of 'believes', and is attributed to the attitude holder than the reporter, as shown by the felicity of the second sentence.

b. Brittany believes that her damn cat, who is a small dog, is the most adorable feline in the neighbourhood.

Both the expressive 'damn' and the supplement 'who is in fact a small dog' scope out of 'believes' and are understood as expressed by the reporter; if they didn't, then the sentence would feel awkward as it would attribute contradictory beliefs to the attitude holder, namely the beliefs that her pet is both a cat and a dog, and both contemptible and adorable.

2. Multidimensionality: This turns out to be a formal realisation of their independence property; since CIs are completely separate from the asserted content, they are represented in an altogether distinct dimension of meaning. The property of independence also distinguishes them from presuppositions in another respect: according to the classical treatment of presuppositions, they should be true in order for the propositions containing them to have a truth value (what Sudo 2012: fn 25 calls the 'Frege-Strawson view of presuppositions'), whereas CIs may certainly be false without rendering the truth value of the proposition undefined.

(37) Brittany's cat finally stopped smoking.

Unless the presupposition that Brittany's cat really smoked in the past is true, this sentence cannot be evaluated for truth or falsity.

(38) Brittany's cat, who is of Sumxu breed, is the smartest cat in the neighbourhood.

The CI that the cat is of Sumxu breed does not need to be true for the sentence to have a truth value. In fact, it is reported that this breed is extinct, if it ever existed.[20] However, the sentence can definitely be true if the specific cat is the smartest in the neighbourhood, and false otherwise.

(39) Brittany's damn cat ate all the dog food again.

This sentence can be judged true if the specific cat ate all the dog food and false otherwise, disregarding entirely the content of the CI which expresses a negative attitude.

3. Antibackgrounding requirement: Unlike presuppositions, CIs are not already part of the common ground; on the contrary, the information they contribute seems to be as new as that of normally asserted content.

(40) a. Rachel admitted that Finn was running late.

'Finn running late' is presupposed as part of the common ground; if it turns out not to be, it forcibly becomes so via accommodation.

b. Rachel notified us that Finn, who was running late, would arrive after noon.

'Finn running late' is offered as new information; this can be seen by the noticeable redundancy of discourses like the following, where the CI is already part of the common ground:

c. #Rachel admitted that Finn was running late. She notified us that Finn, who was running late, would arrive after noon.

4. Comments upon an at-issue core: As per their definition, the new proposition contributed by CIs is separable from the main assertion. However, CIs can act as side-comments to the asserted content, either directly or on the circumstances of utterance.

In order to defend the distinct theoretical status of CIs, Potts rigorously explains how they contrast with other kinds of meanings, namely at-issue entailments, conversational implicatures and presuppositions (2005: 26-36). Since the arguments for their distinction from assertions and conversational implicatures should follow from the description of the properties of CIs, we need not repeat them here. But at the cost of sounding repetitive, it will be useful for the discussion in the next chapter to take good note on their alleged differences from presuppositions, especially since it has been met with skepticism by some authors (e.g. Schlenker 2007, Sauerland 2007, Geurts 2007).

As we saw, projection is a common characteristic between presuppositions and CIs, as both pass the FoS test. But as shown above, unlike presuppositions, CIs have an anti-backgrounding requirement, are completely independent and are unpluggable. But according to Potts, an extra distinctive characteristic is that unlike presuppositions, CIs cannot be quantificationally bound:

[20] https://en.wikipedia.org/wiki/Sumxu

(41) a. 2 in 4 Americans regret voting for Donald Trump out of disgust with the status quo.

The presupposition that those individuals did indeed vote for DT works fine with the quantificational subject.

b. # 2 in 4 Americans, who were disgusted with the status quo, regret voting for Donald Trump.

The property of being disgusted, which is expressed in a supplement, cannot be predicated of the quantificational subject, resulting in an awkward construction.

3.2.2.1. Potts' (2005) formal system

Potts' logic for CIs, named $\mathcal{L}_{\text{CI}}$, is inspired by the system of Karttunen & Peters (1979), which was designed to translate natural language meaning in two distinct dimensions, one for the asserted content and one for presuppositions (although as mentioned above, they employed the term 'conventional implicature' for that). But in order to accommodate CIs, $\mathcal{L}_{\text{CI}}$ has the following distinctive characteristics:

1. It is not resource-sensitive: it is not the case that elements that participate in the composition process are used only once, contrary to common assumptions about the formalisation of natural language (Klein & Sag 1985). Specifically, the at-issue content participates twice in the composition: once on its own to yield the ordinary asserted meaning of the sentence, and once to serve as an argument for the CI function. But this is a reflection at the technical level of the facts that CIs are independent from the at-issue content but also act upon it as comments. Therefore although the lack of resource-sensitivity may feel a bit formally inept given the theoretical landscape in other formal systems in linguistics, it turns out to enable a faithful account of how CIs work.[21]

2. Either at-issue or CI (but not both): no lexical item contributes both an at-issue and a CI meaning. Given the postulation of total independence between the two dimensions, it should not come as a surprise that there are no single expressions endowed with both kinds of meanings. As we shall see in the description of the formal system, such entities only arise via composition, as the result of combining a CI functor to an at-issue argument. Indeed, the linguistic data Potts (2005) is concerned with do not contain any expression that could be said to contribute both kinds of content.[22]

3.2.2.2. Arguments for indirect interpretation

As mentioned before, CIs are said to form a new category of meaning, distinct from other semantic entailments such as at-issue content and presuppositions. Potts suggests that

[21] Although as Potts notes (2005: 86-7), citing p.c. with Asudeh, the lack of resource sensitivity is not really necessary, as the same effect could be achieved via a process of duplication.
[22] As we will see below, this was questioned with data by McCready (2010), which inevitably led to a reformulation of the formal system.

the crucial difference is to be found in the meaning language, where CIs carry distinct types from at-issue content. Although as he notes, this means that a translation language becomes essential, it seems to be the only viable option. Other possibilities would be either syntax or the model, but the first stumbles on the syntactic heterogeneity of CIs (supplements are constructions, expressives are lexical items) while the second has problems explaining the difference between the exact same clause used at-issue versus used as a supplement (CI).

3.2.2.3. Semantic types

Following the above, at-issue and CI content are distinguished by type:

a. e^a, t^a, and s^a are basic at-issue types for $\mathcal{L}_{CI}$.
b. e^c, t^c, and s^c are basic CI types for $\mathcal{L}_{CI}$.
c. If σ and τ are at-issue types for $\mathcal{L}_{CI}$, then $\langle\sigma, \tau\rangle$ is an at-issue type for $\mathcal{L}_{CI}$.
d. If σ is an at-issue type for $\mathcal{L}_{CI}$ and τ is a CI type for $\mathcal{L}_{CI}$, then $\langle\sigma, \tau\rangle$ is a CI type for $\mathcal{L}_{CI}$.
e. If σ and τ are at-issue types for $\mathcal{L}_{CI}$, then $\langle\sigma \times \tau\rangle$ is a product type for $\mathcal{L}_{CI}$, a subset of the set of at-issue types for $\mathcal{L}_{CI}$.
f. The full set of types for $\mathcal{L}_{CI}$ is the union of the at-issue and CI types for $\mathcal{L}_{CI}$.

Figure 3.1. Semantic types of $\mathcal{L}_{CI}$ (Potts 2005: 55)

Although only t^c will be actually used from the CI types (which makes sense since CIs are understood as propositional meanings), all types are defined for uniformity with the at-issue realm. Informally, the above definition of types stipulates the following:

1. At-issue content applies to at-issue content to yield at-issue content: this is just the usual process of functional application (Heim & Kratzer 1998). It is strictly determined that at-issue content cannot apply to anything else but homologous content.

2. CI content applies to at-issue content to yield CI content: this ensures that a CI can take an at-issue expression as its argument to yield a CI, but any other possibility is impossible; at-issue content simply cannot take a CI as argument, and a CI cannot take another CI as argument.

3. A CI type is only an output type: a CI type can take an at-issue type as its argument to yield a CI type again, but it can never be used as an argument itself (either of an at-issue or a CI type).

As Potts notes, the asymmetry between at-issue and CI types is a formal implementation of the idea that CIs function as comments on at-issue content.

3.2.2.4. Modes of composition

The rules of composition are type-driven and subject to certain semantic combinatoric rules ('tree-admissibility conditions') which stipulate that the nodes of the semantic parsetrees are at most binary, have at most one mother node and all share a single

common root. Potts introduces the metalogical symbol "•" (bullet) to mark the separation between independent dimensions of meaning.

3.2.2.5. Functional application tree structures

Given the two possible outcome types (at-issue or CI), there are two different tree structures for functional application, at-issue application and CI application (figures from Potts 2005: 62, 64[23]).

at-issue application

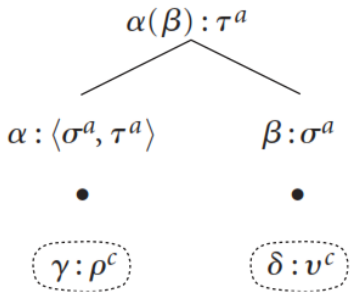


Figure 3.2. At-issue application

At-issue application works exactly like ordinary functional application, with the sole differences of having a superscript to mark the type of content, and the indication of optional CI material in the dotted boxes.

CI application

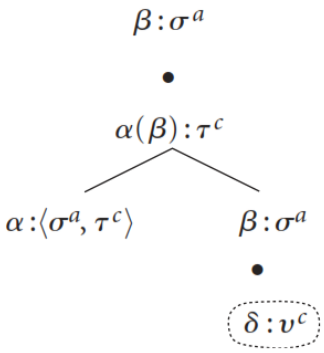


Figure 3.3. CI application

CI application constitutes a kind of functional application plus an identity function; it has two outputs, which are the following:

1) a new CI content (yielded by applying CI type to at-issue type), and

2) the at-issue content alone.

[23] Note that the figure of CI application, although taken from Potts (2005: 64), incorporates a correction made by Gutzmann (2015: 52), who noted that 'originally, the rule allowed for optional material not only at the node of the at-issue argument but also at the node of the CI expression functor'.

Because the at-issue content is used as an argument to yield the new CI type but is also returned intact (it is not consumed), the process is not resource-sensitive. As noted above, the lack of resource sensitivity enables us to account for several of the properties of CIs: their scopelessness, multidimensionality (independence), and their role as comments on at-issue content.

3.2.2.6. Application-free tree structures

Not all parsetrees in $\mathcal{L}_{\text{CI}}$ are formed via functional application, as some are already saturated or undergo type-shifting.

a. Isolated CIs

Some CI expressions already come saturated and as such do not require any argument:

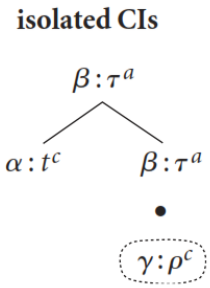


Figure 3.4. Isolated CIs (Potts 2005: 66)

Here we will need to tweak a bit the suggestion that CIs are comments upon an at-issue core; isolated CIs are better understood as comments on the circumstances, rather than a specific unit of at-issue meaning. Potts provides the following example (2005: 125):

(42) Luke has—and you'll never believe this—eaten fifty eggs.

In the example above, the so-called niched conjunction 'and you'll never believe this' literally interjects to express something about the whole sentence rather than any specific part of the at-issue meaning, namely that the sentence describes an unbelievable fact. This is even reflected graphically, as it is enclosed by dashes.

b. Feature semantics

It happens that a semantic contribution of syntactic features can turn at-issue content to CI content. Specifically in the case of supplements, the feature COMMA works as a type-shifting operator turning an at-issue function to a CI function, which can then partake in CI application:

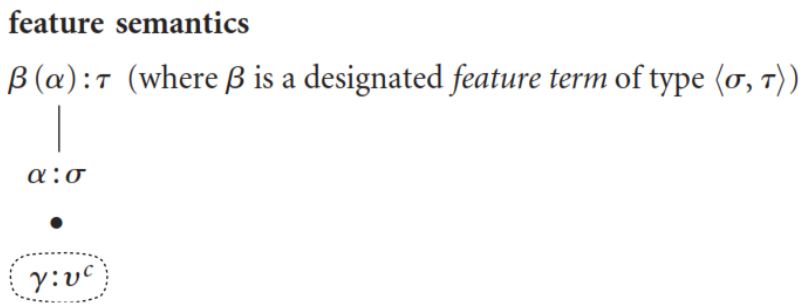


Figure 3.5. Feature semantics (Potts 2005: 66)

As can be shown in the example below, Potts assumes the feature COMMA as an operator present in the syntax, with implications on the semantics (resulting in a t^c type). By positing this feature, Potts is able to maintain the thesis that no linguistic item contributes both at-issue and CI meaning.

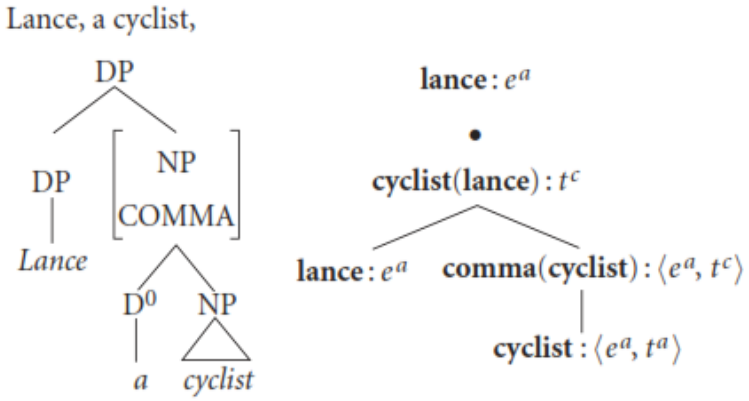


Figure 3.6. Potts' (2005) syntactic and semantic representation of a sentence with COMMA (Potts 2005: 66)

3.2.2.7. Interpretation

As it follows from the definition of CIs as separate propositions, it is expected that a single sentence may correspond to different propositions. To ensure multidimensional interpretation, Potts introduces the following rule for parsetree interpretation: the denotation of a sentence is given by the entire parsetree rather than just its root node, since there are two independent dimensions of meaning (formally separated by "•"):

parsetree interpretation

Let $\mathcal{T}$ be a semantic parsetree with the at-issue term $\alpha : \sigma^a$ on its root node, and distinct terms $\beta_1 : \langle s^a, t^c \rangle, \ldots, \beta_n : \langle s^a, t^c \rangle$ on nodes in it (extensionally, $\beta_1 : t^c, \ldots, \beta_n : t^c$). Then the interpretation of $\mathcal{T}$ is the tuple

$$\langle [\![\alpha : \sigma^a]\!]^{\mathcal{M}_i,g}, \{ [\![\beta_1 : \langle s^a, t^c \rangle]\!]^{\mathcal{M}_i,g}, \ldots, [\![\beta_n : \langle s^a, t^c \rangle]\!]^{\mathcal{M}_i,g} \} \rangle$$

where $[\![\cdot]\!]^{\mathcal{M}_i,g}$ is the interpretation function, taking formulae of the meaning language to the interpreted structure $\mathcal{M}_i$, relative to a variable assignment g.

Figure 3.7. Parsetree interpretation (taken from Potts 2005: 68)

The proclamation of this new category of meaning that is definitely projective but seems to differ significantly from presuppositions has caused serious debates across theorists about the status of these items. Some strongly resisted idea that these meanings justified the introduction of a new category (e.g. Geurts 2007), some saw such a move as 'complicating our ontology of meaning' (Macià 2002: 510) and advocating that so-called CIs can be explained as presuppositions with some theoretical adjustments, i.e. as a kind different than classical ones (Schlenker 2007, Sauerland 2007, Lasersohn 2007). Overall, however, the postulation of CIs as a specialised category was met with enthusiasm and adopted by many theorists (e.g. Abbott 2006, Tonhauser et al. 2013, McNally & Kennedy 2008, etc.), several producing a rich and lengthy literature based on Potts' seminal work and evolving it even further (Sawada 2011, 2014, Feng 2010, Liu 2012, Beltrama & Lee 2015, McCready 2008, 2009 and most notably 2010, Gutzmann 2013, 2015, etc).

3.2.2.8. An expressives-only dimension

In subsequent work (2007), Potts takes a different turn specifically with regards to expressives, in complete separation from supplements. He abandons the idea that expressives are propositional, which means that they are no longer conventional implicatures.[24] Instead, he treats them as operators that change the context, in terms of the emotional attitudes expressed by the speaker. As a result, the distinction between at-issue vs. CI is replaced by the distinction expressive vs. descriptive meaning, and it is no longer just type-theoretic but also model-theoretic.

According to Potts (2007), expressive meaning has the following distinctive properties:

a. Independence: The dimension of meaning contributed by expressive meaning is entirely separate from that of descriptive content.

[24] This is as long as one takes CIs to amount to propositions, as Potts (2005) does. In theory it could be possible to retain that expressives are still CIs but under the condition that one felt comfortable to say that CIs are non-propositional; in that case one would either have to exclude supplements from CIs or show that supplements are also non-propositional.

b. Nondisplaceability: Expressives predicate something of the utterance situation, and cannot be shifted from it.

c. Perspective dependence: Expressive content is tied to a perspective, which is usually the speaker's, but it is possible to differ in some cases.

d. Descriptive ineffability: It is not possible to accurately paraphrase expressive meaning using descriptive means, as the result is not equivalent.

e. Immediacy: Expressives behave like performatives in that their mere utterance seems to inflict their content rather than contribute it to conversation.

f. Repeatability: The repetition of an expressive results in strengthening its emotive content rather than sounding redundant.

These revisions result in the proposal of a formal system elaborated specifically for expressives. Some basic ideas remain the same as $\mathcal{L}_{\text{CI}}$: the system is truly multidimensional and as such there is a distinction at the level of types; expressives act on descriptive content but return it unaltered. However, several important alterations are introduced to account for the above defined properties of expressive meaning. Skipping the full technical details, these are the following:

Expressive indices: These provide the denotations of expressives and correspond to structured entities which encode the relationship between two entities in a numerical fashion by the intervals between [-1, 1]. For example, the negative relationship expressed by the use of 'damn' by a about b (where a and b $\in D_e$) can be said to correspond to <a [-0.5, 0] b>. The notion of expressive indices effectively accounts for the property of descriptive ineffability, since there is an ontological difference between the denotations of expressives and descriptive items.

Finer-grained contexts: In order to accommodate expressives, the notion of a context as a tuple is enriched with two more parameters:

a) the set of expressive indices of the context c_ε: c_ε is what expressives affect, the idea being that uttering an expressive alters the context of utterance and yields a new one, differing from the first in that c_ε has been changed (in a downward monotonic fashion). This is in fact a kind of context shifting. In Potts' own words (2007: 181):

> To utter an expressive is to alter the current context of interpretation c by inserting a new expressive index into c_ε or replacing one of the expressive indices $i \in c_\varepsilon$ with an index i' that is just like i except that its interval (middle) component is at least as narrow as that of i.

Building the idea of expressive indices as a contextual parameter helps explain their performative character, reflected by the property of immediacy.

b) the judge of the context c_j: Already in his 2005 Potts had noted that CIs are speaker-oriented, but here he moves a step further. In order to account for the perspective dependence of expressives, Potts adopts the notion of a contextual judge from

Lasersohn (2005)[25]. The idea is that expressives are necessarily tied to a perspective, which is usually the speaker's as a pragmatic default, but could also be someone else's under the right contextual conditions. This enables Potts to account for some known counterexamples to the speaker-orientation of expressives:

(44) My father screamed that he would never allow me to marry that bastard Webster. (Kratzer 1999)

(45) I am not prejudiced against Caucasians. But John, who is, thinks/claims that you are the worst honky he knows. (Schlenker 2003)

In these examples the attitudes conveyed by the expressive items are those of the respective attitude holders, rather than the speaker who acts as an attitude reporter. I will extensivenly discuss such examples in chapter 5.

3.2.2.9. Some implications of Potts' expressive turn

Although not of our main concern here, in concluding this section it is interesting to note some consequences of Potts' revisions. As we saw, Potts abandoned the idea that expressives are CIs and claimed that they differ model-theoretically from descriptive items. This has two implications: first, the category of CIs is now much poorer, consisting solely of supplements. This renders it much more vulnerable either to the original reduction of CIs to secondary assertions by Bach (1999) or to newer sophisticated accounts that put more burden on the syntax and assume a conjunctive semantics (Schlenker 2010)[26]. Second, we saw that Potts (2005) advocated the necessity of a meaning language in order to reflect the difference between CIs and at-issue content. Potts rejected syntax as a candidate for the distinction because of the heterogeneity between supplements and expressives, which comprised the category of CIs: as these two classes of expressions are syntactically dissimilar, then their similarity ought to be sought somewhere else, such as the meaning language.[27] With expressives

[25] To be precise, Potts adopts an alternative that was rejected by Lasersohn (2005) as inadequate for predicates of personal taste. As Lasersohn points out, if the judge becomes a contextual parameter, then we will need to concede that there is a change of context to account for commonplace examples like the following:
(43) John thinks that roller coasters are fun, but Mary thinks that roller coasters are not fun.
Moreover, such an analysis would fail to capture the intuition of disagreement, since we'd end up dealing with different contents. Because of this, Lasersohn proposes that relativisation to a judge must not take place at the level of context (which would result in a different content), but at the level of circumstances of evaluation (resulting in a difference of truth value, which feels intuitively correct for cases of so-called faultless disagreement). However, on the basis that unlike predicates of personal taste expressives do not seem to shift perspective mid-computation, Potts adopts c_j as a contextual parameter.

[26] In fact, as will be seen in chapter 5 where I advocate for a new analysis of expressives, I do agree with Potts (2007) that there is something fundamentally different about expressive meaning which is not shared by supplements. For this reason, I find non-multidimensional analyses of supplements, such as Schlenker's (2010) conjunctive proposal, as more fitting than Potts' (2005) CI one where supplements are grouped with expressives.

[27] We should note here that another alternative was the models, but Potts rejected it on the basis that 'the models for conventionally-implicated phrases are the same as those for at-issue phrases' (2005: 51). Interestingly, here again the sameness in the models seems to be based largely on

out of it, it seems possible to suggest that the difference between supplements and at-issue content does indeed lie solely in the syntax, rendering the idea of a meaning language no longer necessary and allowing thus direct interpretation. After all, as noted previously the so-called COMMA operator was posited in the syntax, so it can function as a distinctive characteristic of supplements, without the need to appeal to a dedicate type t^c – a good ole t would do.

3.2.3. Extensions to $\mathcal{L}_{CI}$

3.2.3.1. McCready's $\mathcal{L}^{+}_{CI}$

Soon after Potts inaugurated the study of expressive meaning as a popular topic, it was discovered that there were data that challenged some tenets of his theory, such as expressions which contribute both at-issue and expressive meaning ('mixed content'), as well as expressions which function as the main content of an utterance but possess only expressive meaning. Pejoratives such as ethnic slurs and honorifics are items of the first category, while several particles and evidentials fall in the second one. As Potts (2005) operated on the assumption that no such items are possible, $\mathcal{L}_{CI}$ could not accommodate them. The relevant adjustments were proposed by McCready (2010) in a revised logic called $\mathcal{L}^{+}_{CI}$. It is important to note here that McCready is following Potts (2005) and not (2007) in considering expressives to be conventional implicatures, just like supplements, hence the name of his article ('Varieties of Conventional Implicature'). Nevertheless, he adopts the more specific term 'CIE' content (conventionally implicated or expressive) for expressives.

Specifically, the problem is the following: in the cases of expressives which function as the main content and mixed content that is predicative in both dimensions (e.g. honorifics) we don't want any at-issue content returned after the CI application is complete. For this reason, McCready proposes resource-sensitive applications[28], a move which also necessitates adding the respective types (called 'shunting types' and marked by superscript s) so that the system can know where to apply what:

$$\frac{\alpha : \langle \sigma^a, \tau^s \rangle, \beta : \sigma^a}{\alpha(\beta) : \tau^s}$$

Figure 3.8. McCready's (2010) proposed application (taken from McCready 2010: 18)

This resource-sensitive application allows expressions with so-called shunting types to be saturated by an at-issue argument by consuming it fully, thus giving nothing back. As a result, it can account for cases where there is no asserted content at all, the main

supplements, since one and the same phrase can act either as a supplement or as at-issue. If one considers expressives to be different beasts, as Potts (2007) ended up doing, then this might be a viable alternative. (And as I will show when I make my own proposal, in chapter 5, it is worth pursuing).

[28] Unlike Potts who used trees, McCready presents the rules of her system in the style of proofs, which she takes as notational variants for her purposes despite their superficial differences.

content of the utterance being expressive. McCready (2010: 37) gives the following example with the Japanese adverbial 'yokumo':

(46) Yokumo koko ni kita (na)!

(yokumo) here to came (sentence-final particle)

'You have a lot of guts to come here!'

yokumo: $\langle t^a, t^s \rangle$, you-came-here: t^a
yokumo(you-came-here): t^s

Moreover, there is an application for mixed expressions, endowed with meaning in both dimensions. The two different contents of such expressions are represented as separated by a "♦" (diamond):

$$\frac{\alpha \blacklozenge \beta : \langle \sigma^a, \tau^a \rangle \times \langle \sigma^a, \upsilon^s \rangle, \gamma : \sigma^a}{\alpha(\gamma) \blacklozenge \beta(\gamma) : \tau^a \times \upsilon^s}$$

Figure 3.9. Application for mixed expressions (taken from McCready 2010: 20)

This application, which is also resource-sensitive, can account for mixed contents where one and the same argument is used in both dimensions of an expression to yield two propositions, an at-issue and a CI one.

(47) Sonoo sensei ga oideninarimashita.

Teacher Sonoo came-HON.

'Teacher Sonoo came.'

come ♦ be-honourable: $\langle e^a,t^a \rangle \times \langle e^a,t^s \rangle$, teacher Sonoo: e^a
come (teacher Sonoo) ♦ be-honourable (teacher Sonoo): $t^a \times t^s$

As soon as the above derivation is complete and there are two distinct propositions, the familiar Pottsian operator "•" (bullet) takes over to isolate them and to mark that the expression can move to the interpretation stage:

$$\frac{\alpha \blacklozenge \beta : \sigma^a \times t^s}{\alpha : \sigma^a \bullet \beta : t^s}$$

Figure 3.10. Isolation via bullet operator (taken from McCready 2010: 20)

come (teacher Sonoo) ♦ be-honourable (teacher Sonoo): $t^a \times t^s$
come (teacher Sonoo): t^a • be-honourable (teacher Sonoo): t^s

At-issue proposition: Teacher Sonoo came.

CI proposition: The speaker honours Teacher Sonoo.

3.2.3.2. Gutzmann's $\mathcal{L}_{TU}$

3.2.3.2.1. Conceptual remarks

A subsequent upgrade of McCready's $\mathcal{L}^{+}_{CI}$ was proposed by Gutzmann (2015) to enable the system to account for more data. But apart from the technical adjustments made, there are some conceptual novelties as well.

Specifically, Gutzmann breaks free with the tradition of treating expressives as conventional implicatures on a par with supplements, i.e. as regular propositions that happen to be independent from the at-issue meaning. In this regard, his point of departure is closer to Potts (2007) but eventually, it is to be found back in Kaplan (1999). Gutzmann follows Kaplan's insights faithfully by conceptualising the difference between expressive and descriptive meanings as use-conditions VS. truth-conditions. Effectively, Gutzmann puts forward a system for use-conditional meaning, of which expressive meaning is merely a subset; specifically, expressive meaning is defined as use-conditional meaning that conveys an evaluative attitude. Following Kaplan, Gutzmann posits that the same formal tools devised for truth-conditional meaning can be applied to use-conditional meaning, by making the necessary assumptions: substituting truth-conditions by conditions of use, and possible worlds where a sentence is true by contexts in which an expression is felicitously used:

Truth conditions	Use conditions
"Snow is white"	"Oops"
is true	is felicitously used
iff snow is white.	iff the speaker observed a minor mishap.

t-proposition "Snow is white": the set of possible worlds in which snow is white.
u-proposition "Oops": the set of contexts in which the speaker observed a minor mishap.

Again following Kaplan's (1999) proposal that an expression's 'semantic information' includes all the conventional information carried by the expression and not just those analysable in truth conditions, Gutzmann advocates the notion of 'hybrid semantics', which he defines as 'a semantic framework that employs these two dimensions in order to analyse the meaning of natural language expressions' (Gutzmann 2015: 8). As he tells us, the idea is that 'the overall interpretation of an expression hence delivers two dimensions (rendered as a tuple), one corresponding to its t-content and one to its u-content' (2015: 21). According to a hybrid semantic approach then, the meaning of a sentence can be given in two dimensions as follows:

If S is a sentence, then
a. S is true in a possible world w and a context c, if $w \in ||S||^{t}$
b. S is felicitous in a possible world w and a context c, if $c \in ||S||^{u}$

3.2.3.2.2. Types of UCIs

Before going into the formal system, it will be useful here to present Gutzmann's (2013) terminology for different kinds of use-conditional items (UCIs).

	f	2d	r-s	e.g.
isolated expletive	−	−		'ouch', 'oops'
isolated mixed	−	+		ethnic slurs, e.g. 'Kraut'
functional expletive	+	−	−	'damn'
functional shunting	+	−	+	Japanese adverbial 'yokumo'
functional mixed	+	+		Japanese honorifics verbs, e.g. 'irassharu' (go_{HON})

Table 3.1. Types of UCIs

f: functionality; whether the expression needs an argument
2d: 2-dimensionality; whether the expression has meaning in both dimensions (i.e. it's mixed)
r-s: resource-sensitivity; whether the expression fully consumes its argument[29]

3.2.3.2.3. The system: $\mathcal{L}_{TU}$

According to Gutzmann, despite the innovations proposed by McCready (2010), the resulting system $\mathcal{L}^{+}_{CI}$ still has some problems and limitations. These are the following:

1. Inability to handle use-conditional modification (e.g. 'bloody bastard'): as both expressions are UCIs they are combined with shunting application, which as we saw is a resource-sensitive process, meaning that no truth-conditional argument is returned to combine with the sentence predicate where it is needed, causing the derivation to crash.

2. Problems with abstraction: due to the maintained separation between different dimensions, in the case of a quantifier in object position (e.g. 'Andrew is gawking at every girl') the variable introduced by quantifier raising cannot be bound by the λ-operator because it is isolated in the UCI dimension.

3. Problems with quantification: though there's a rule for applying a one-dimensional argument to a two-dimensional functor (mixed application), there is no rule for applying a two-dimensional argument to a one-dimensional functor, which is necessary to account for constructions with subject quantification such as 'Everybody is gawking at her'.

4. Compositionality problems: This is not a particular issue of $\mathcal{L}^{+}_{CI}$ but it is inherited by its predecessor $\mathcal{L}_{CI}$. As we saw previously, during the interpretation process the whole parsetree is considered rather than merely the root node (Potts' 2005 'parsetree interpretation'). This was a necessity given that UCI meaning is completely independent, yet it needs to be interpreted. Nevertheless, this process does not respect the principle of compositionality (Frege 1892, Partee 1984), that the meaning of a complex expression depends on the meaning of its (immediate) parts and the way they are put

[29] Note that resource-sensitivity is only specified for functional expletive and functional shunting UCIs, in fact it is what distinguishes one from the other. It does not make sense to talk about resource-sensitivity for items that are [–f] (isolated expletive and isolated mixed UCIs) because in this case there is no argument at all, nor for items that are [+2d] (functional mixed) because in this case the argument is available in both dimensions anyway.

together[30]. By definition, in parsetree interpretation meaning is derived by looking into more than the immediate parts and the way they are combined, as the information is isolated in the second dimension and needs to be collected from there.

According to Gutzmann, all of the above problems can be solved by taking a step further and assuming multidimensionality not just at the level of interpretation (as in $\mathcal{L}_{CI}$ and $\mathcal{L}^{+}_{CI}$) but throughout the entire composition ('compositional multidimensionality'). This means that now there are 3 different dimensions at every step during the computation, rather than just at the interpretation stage:

$$\{\ \underbrace{t\text{-dimension}}_{t\text{-content}},\ \underbrace{s\text{-dimension}}_{u\text{-active content}},\ \underbrace{u\text{-dimension}}_{\text{saturated } u\text{-content}}\ \}$$

Figure 3.11. Gutzmann's (2015) 3 different dimensions (taken from Gutzmann 2015: 118)

t-dimension: for at-issue types, same as in $\mathcal{L}^{+}_{CI}$ (the set of possible worlds where the t-proposition is true)

s-dimension: for shunting types, it works as storage for UCI content that is still active and usable as an argument by another UCI expression

u-dimension: for saturated UCI content, same as in $\mathcal{L}^{+}_{CI}$ (the set of contexts where the u-proposition is felicitously used)

In terms of types, the only addition to $\mathcal{L}^{+}_{CI}$ needed is that use-conditional types can also be used as input, which will allow for expressive modification.

3.2.3.2.5. Multidimensional application

Gutzmann introduces a single application rule that covers all of $\mathcal{L}^{+}_{CI}$'s 4 different application rules: ordinary functional application (Heim & Kratzer 1998), expressive application (Potts 2005), mixed application and shunting application (McCready 2010), while also accommodating use-conditional modification and solving the problems of abstraction and quantification; this is possible because all three dimensions of meaning are available for every expression at every step of the derivation. Specifically:

- Expressive application (Potts 2005) becomes an instance of mixed application (McCready 2010) by the assumption that in the former instead of an empty t-dimension we have an identity function; in the case of a mixed UCI the t-dimension has meaningful content, in the case of a UCI without any truth-conditional content we just have the identity function in the t-dimension. This idea can account in a uniform way for the two different outcomes: in the case of mixed UCIs, the argument is used in both dimensions, in UCIs without a t-

[30] To be precise, as it is stated this is the principle of 'strictly local' compositionality – there is also a 'non-local' counterpart (Partee 1984, Larson & Segal 1995, Pelletier 2012) which in theory could allow for parsetree interpretation. This issue will be elaborated in chapter 4, where I will critically assess the whole enterprise initiated by McCready (2010) and further developed by Gutzmann (2015) will be critically assessed.

dimension of their own it appears as if it is returned intact due to the identity function.

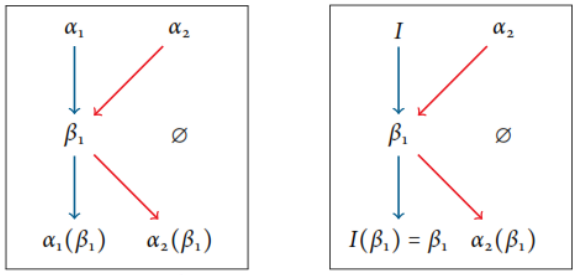


Fig. 3.12. Expressive application as an instance of mixed application with an identity function (taken from Gutzmann 2015: 120)[31]

- Ordinary functional application (Heim & Kratzer 1998) becomes an instance of expressive application (Potts 2005), which itself was reduced to mixed application (McCready 2010) above, by placing a copy of the t-dimension in the s-dimension of expressions in which it is empty (such as non-UCIs); this unifies the two applications because ordinary functional application can be seen as a case of expressive application where the function's s-dimension is a copy of its t-dimension.

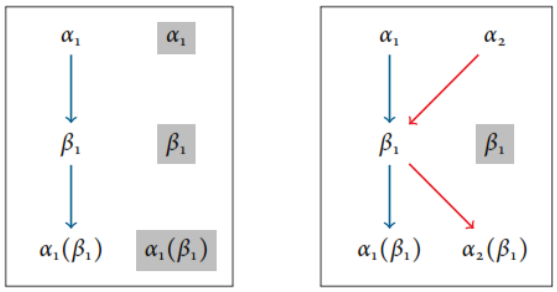


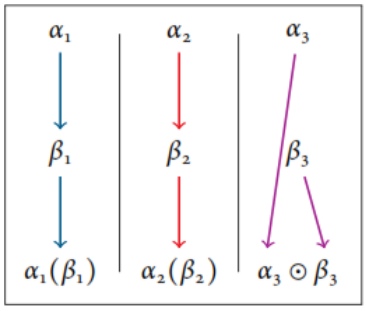
Fig. 3.13. Functional application as expressive application by copying the t-dimension to the s-dimension (taken from Gutzmann 2015: 121). The boxes appearing in grey are considered as compositionally inactive.

- In all cases the s-dimension of the argument is inactive in the composition, and it contains a copy of the function's t-dimension; because of this, applying the function's s-dimension (α_2) to the argument's t-dimension (β_1) is equivalent to applying the function's s-dimension (α_2) to the argument's s-dimension (since it also contains β_1). This enables the application to happen intra-dimensionally (within the same dimension) rather than trans-dimensionally (across dimensions).

[31] In this figure as well as in the next one (Fig 3.13), for reasons of simplicity only the t-dimension and s-dimension are presented, as the u-dimension is considered as always composed in the same way, namely by gathering all use-conditional meaning (its denotation is the intersection of sets of contexts).

- Shunting application (McCready 2010) can also be subsumed assuming that UCIs that fully consume their truth-conditional argument can equivalently be seen as returning it, but that is truth-conditionally trivial.

All the above amount to the following all-purpose rule of multidimensional application (the operator ⊙ merges use-conditional expressions in the logical language):

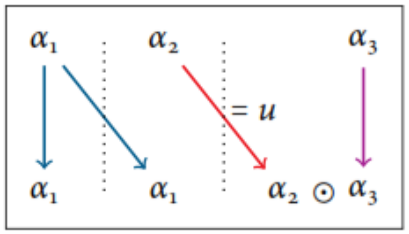


Multidimensional application

$$\frac{\alpha_1 : \langle \sigma, \tau \rangle \blacklozenge \alpha_2 : \langle \rho, \nu \rangle \bullet \alpha_3 : u \qquad \beta_1 : \sigma \blacklozenge \beta_2 : \rho \bullet \beta_3 : u}{\alpha_1(\beta_1) : \tau \blacklozenge \alpha_2(\beta_2) : \nu \bullet \alpha_3 \odot \beta_3 : u}$$

Fig. 3,14. Multidimensional application in arrow diagram and formula (taken from Gutzmann 2015: 122)

We also need an elimination rule to move content from the s-dimension to the u-dimension as soon as the composition is complete. Following our previous assumption that if an expression has an empty s-dimension the content of its t-dimension is copied there (needed so that we could reduce ordinary functional application to expressive application), we get the following:



Unary use-conditional elimination

$$\frac{\alpha_1 : \sigma \blacklozenge \alpha_2 : u \bullet \alpha_3 : u}{\alpha_1 : \sigma \blacklozenge \alpha_1 : \sigma \bullet \alpha_3 \odot \alpha_2 : u}$$

Fig. 3.15. Unary use-conditional elimination in arrow diagram and formula (taken from Gutzmann 2015: 128, 135)

As the above improvements secure multidimensionality throughout the composition, all the previously mentioned problems are resolved: use-conditional modifiers are supplemented by content in all dimensions, which allows them to combine with other UCIs via multidimensional application; there are no problems with abstraction and quantification, given that all expressions are equally multidimensional; and the system fully respects compositionality, since the root node, where interpretation takes place, is fully multidimensional itself.

3.2.3.2.6. Lexical extension rules (LERs)

Following compositional multidimensionality, composition rules act only on 3-dimensional objects. However, most often content is 1-dimensional (truth-conditional or use-conditional) and it can be at most 2-dimensional (both truth-conditional and use-conditional). Gutzmann proposes as the most parsimonious assumption that natural

language expressions are translated to lexical expressions of $\mathcal{L}_{TU}$ which are only specified for those dimensions that cannot be inferred from what is given, and the rest is derived. The rules that fill in the unspecified dimensions into the logic and render the necessary 3-dimensional format are called lexical extension rules (LERs). The table below is to show which dimensions each kind of expression comes specified for (maximum up to 2) and which are to be derived by the LERs.

Kind of expression	t-dimension	s-dimension	u-dimension
truth-conditional	☑	⬚	⬚
functional expletive UCI	⬚	☑	⬚
isolated expletive UCI	⬚	⬚	☑
isolated mixed UCI	☑	⬚	☑
functional mixed UCI	☑	☑	⬚
Functional shunting UCI	⬚	⬚	☑

Table 3.2. Dimensions of each UCI type

Skipping the exact formal details (see Gutzmann 2015: 138-9), LERs fill in the empty dimensions as follows:

Kind of expression	t-dimension	s-dimension	u-dimension
truth-conditional	☑	t-copy	U
functional expletive UCI	I	☑	U
isolated expletive UCI	T	t-copy	☑
isolated mixed UCI	☑	t-copy	☑
functional mixed UCI	☑	☑	U
functional shunting UCI	$\dot{F}_{\mapsto T}$	t-copy	☑

Table 3.3. LERs filling empty dimensions

I: identity function that returns the t-dimension of the argument
t-copy: copy of the t-dimension of the specific expression
U: neutral use-conditional object
T: trivial truth-conditional object
$\dot{F}_{\mapsto T}$: dummy function that maps its argument onto T

Apart from enabling an economical explanation of how all lexical items engage compositionally in a 3-dimensional fashion without postulating that the lexicon is loaded with 3 distinct dimensions for every single expression, the notion of LERs also helps to account for cross-linguistic variation and allows accounting for the possibility that some languages may lack certain kinds of expressives.

3.3. Expressives as necessarily de se

As we saw before, Potts (2005) already noted the speaker-oriented character of expressives, which he subsequently strengthened by positing the characteristic of perspective dependence (Potts 2007). This section aims to argue that this can be safely taken a step further to the explicit claim that expressive meaning is essentially de se; in other words, that by using an expressive the speaker self-ascribes a property, following the terminology of Lewis (1979) which we saw in the previous chapter.

In order to show this, we will cross-check some expressives in de se situations, i.e. situations in which the speaker self-ascribes the relevant attitude, versus non-de se ones. As the attitudes conveyed by expressives are emotive attitudes, it is assumed that the attitude holder cannot have them without being aware of it, due to the so-called first-person authority of emotional states (Neta 2011)[32]. Moreover in all the situations, it will be assumed that the speaker's linguistic competence includes knowledge of the meaning of the words s/he uses. Here 'meaning' is used in its folk pre-theoretical sense and covers both truth conditions and conditions of use (as in that the speaker knows both what the words he uses refer to and under what circumstances they are used).

In all the examples below, it will be assumed that the speaker is watching a recording of her/himself speaking, but fails to recognise him/herself in the video due to intoxication.

Yanis is watching a video of himself criticising in graphic terms the fact that German people are likely to elect Angela Merkel for another term. Now that he is drunk, apart from failing to recognise himself in the video, he doesn't really care about German politics at all and has no intense feelings about Merkel's possible re-election. His friend Alexis walks in the room and asks him what the video he is watching is saying.

(48) The Germans will vote for Merkel again.

(49) # The Krauts will vote for Merkel again.

In this specific context, (48) is felicitous but (49) is not, despite having identical propositional content. Importantly, it's not just the fact that the truth-conditions are identical (since 'German' and 'Kraut' are extensionally equivalent), but it's also the case that the speaker, Yanis, does indeed harbour a negative attitude against Germans in general. However, the crucial difference which makes (49) infelicitous is that when drunk, Yanis is not aware of his negative attitude nor does he intend to express it. So although (49) is true, it is infelicitous because of the lack of self-awareness necessary for the de se.[33]

Hiroki is watching a video of himself talking about how grateful he is to his teacher Sonoo for helping students like him, and how much he respects her. Now that he is drunk, apart from failing to recognise himself in the video, he doesn't really care about Sonoo's contribution. His boss happens to walk in the room and asks him what the video he is watching is saying.

(50) Sonoo-san ga gakusei-tachi-o tasuketa.

[32] It should be noted that this has been doubted by some experimental results which have suggested that people are not necessarily better informed about their own emotional states than those close to them (Gertler 2010). Given the focus of this dissertation, I will take first-person authority of emotional states as a given, recognising that it remains far from a resolved issue in the philosophy of mind.

[33] Note that example (49) would be predicted as felicitous by a system along the lines of Potts (2005), since the 'propositional content' of the expressive accidentally happens to be true here. However, since Potts (2005) excluded the possibility of mixed expressives such as 'Kraut' here, this observation mostly holds for McCready (2010), who adopted Potts (2005) and supplemented the technical provisions for mixed expressives.

Sonoo students helped

'Sonoo helped the students.'

(51) # Sonoo-san ga gakusei-tachi-o o-tasuke-ni-narimashita.

Sonoo students helped$_{HON}$

'Sonoo helped$_{HON}$ the students.'

Again, despite the fact that both sentences are true, (51) is infelicitous because Hiroki is not aware of having a respectful attitude towards teacher Sonoo at the moment of utterance.

Wilton is watching a video of himself complaining about how often he receives visits from Mormons trying to proselytise him. Now that he is drunk, apart from failing to recognise himself in the video, he doesn't really mind Mormons. His brother Alexandre walks in the room and asks him what the video he is watching is saying.

(52) The Mormons knock on people's doors every day.

(53) # The damn Mormons knock on people's doors every day.

Although both sentences are true, (53) is infelicitous because Wilton is not aware of having a negative attitude towards Mormons at the moment of utterance.

Willian is watching a video of himself complaining about the habit of some people visiting friends' homes uninvited. Now that he is drunk, apart from failing to recognise himself in the video, he doesn't really mind the idea of uninvited visits. Suddenly his friend Zeca shows up uninvited and sits down next to him.

(54) #Yokumo koko ni kita (na)![34]

(yokumo) here to came (sentence-final particle)

'You have a lot of guts to come here!'

Although Willian generally has a negative attitude about people showing up uninvited, he is not aware of it at the moment of utterance, which makes (54) infelicitous.

[34] Note that in this case, there is no neutral counterpart to this utterance, exactly because as noted before 'yokumo' is a so-called functional shunting UCI, i.e. it takes a truth-conditional argument and 'shunts' it to the expressive dimension, leaving nothing behind. As a result, it is not possible to construct a neutral version of an utterance with 'yokumo'; a construction involving only the truth-conditional component that functions as the argument of 'yokumo' here ('You came') would be one with a totally different discourse function, capable of updating the common ground, whereas when the truth-conditional part is consumed by 'yokumo' it is not informative but merely serves to express an emotive attitude. For detailed discussion see McCready (2010: 38-9). The example used here (54) is taken from McCready (2010: 37).

As can be seen from the examples above, the use of expressives is infelicitous in cases where the speaker does not self-ascribe the relevant attitude. Moreover, it is rather impossible, as in these cases a competent speaker would not select these expressions at all. As a result, the self-ascription of the relevant attitude seems to be a necessary condition for the use of expressives – though not a sufficient one, as the speaker may not decide to employ an expressive for other reasons (e.g. she may not wish to disclose her attitude). That is to say, expressives are shown to be essentially de se.

The idea that expressives might be necessarily de se is not new: it is briefly examined by McCready & Wechsler (2012) but only to be rejected. Specifically, it is claimed that it is possible to produce an utterance containing an expressive such as "Where's my damn wallet?" without self-ascribing an emotive attitude. On the contrary, they say that it is possible that by the production of the utterance the subject comes to self-ascribe that attitude: 'I come to know my excited state by my (instinctive) use of damn' (McCready & Wechsler 2012: slide 57). However, as the examples above showed, the use of expressives by the speaker in the absence of self-awareness is infelicitous, as well as extremely unlikely.

3.3.1. Multidimensional de se

But if expressives are necessarily de se, then this has to be accounted for by any formal system of expressive meaning. This section is going to review how this could be accommodated in the proposed frameworks. Although all the different formal proposals are continuations of each other (Gutzmann's (2015) $\mathcal{L}_{TU}$ is an improvement of McCready's (2010) $\mathcal{L}^{+}_{CI}$, itself an improvement of Potts' (2005) $\mathcal{L}_{CI}$), there are significant differences with regards to the denotations between Potts (2005) and (2007) and between Potts (2005, 2007)/McCready (2010) and Gutzmann (2015). Even though Gutzmann's (2015) $\mathcal{L}_{TU}$ constitutes the latest (and still uncontested) proposal for a formal system for expressives, Potts' work remains very influential, if not identified with the whole field of expressive meaning, so it's worth examining the issue for each framework separately.[35]

3.3.1.1. $\mathcal{L}_{CI}$

As we saw, for Potts (2005) expressives were CIs, just like supplements, so their denotations were regular propositions. However, one defining characteristic of CIs was their speaker-orientation, i.e. they invariably constitute commitments of the speaker of the utterance. Assuming that speakers always have access to their own beliefs, this is a way to say that they are essentially de se. Just as we would expect that a person who uses 'I' is aware that he is referring to himself (Perry 1979), we can equally expect that a speaker who uses a speaker-oriented expression is aware that he himself endorses its content. The speaker-orientation of CIs is not just posited but also captured in Potts' $\mathcal{L}_{CI}$

[35] McCready's (2010) $\mathcal{L}^{+}_{CI}$ constitutes an authentic extension of Potts' $\mathcal{L}_{CI}$ in the sense that it only complements it with more composition rules and adjusts the interpretation function accordingly, but it does not review any of the fundamental assumptions of the framework (bar its prohibition of mixed content, of course). As a result, it will be assumed that the discussion about $\mathcal{L}_{CI}$ equally applies to $\mathcal{L}^{+}_{CI}$ as well.

framework as follows. According to the rule of parsetree interpretation, the interpretation function is relativised to an intensional model, as well as an assignment function:

parsetree interpretation

Let $\mathcal{T}$ be a semantic parsetree with the at-issue term $\alpha : \sigma^a$ on its root node, and distinct terms $\beta_1 : \langle s^a, t^c \rangle, \ldots, \beta_n : \langle s^a, t^c \rangle$ on nodes in it (extensionally, $\beta_1 : t^c, \ldots, \beta_n : t^c$). Then the interpretation of $\mathcal{T}$ is the tuple

$$\langle [\![\alpha : \sigma^a]\!]^{\mathcal{M}_i,g}, \{[\![\beta_1 : \langle s^a, t^c \rangle]\!]^{\mathcal{M}_i,g}, \ldots, [\![\beta_n : \langle s^a, t^c \rangle]\!]^{\mathcal{M}_i,g}\} \rangle$$

where $[\![\cdot]\!]^{\mathcal{M}_i,g}$ is the interpretation function, taking formulae of the meaning language to the interpreted structure $\mathcal{M}_i$, relative to a variable assignment g.

Figure 3.16. Parsetree interpretation (taken from Potts 2005: 68)

Roughly, intensional models can be understood as world-views, and we assume one for each interlocutor. Specifically, this means that if Fede utters a sentence, its content should be interpreted in the intensional model $\mathcal{M}_{\text{Fede}}$, capturing thus speaker-orientation.

Nevertheless, there is a fundamental problem with this idea; as it was demonstrated by Wang, Reese and McCready (2005), Karttunen & Zaenen (2005) and Amaral, Roberts & Smith (2007), speaker-orientation is not inviolable. This fact was subsequently recognised by Harris & Potts (2009). Therefore the idea of relativisation to each speaker's own intensional model cannot be maintained, and the system needs to be amended so as to allow for perspective shifting. As Harris & Potts (2009) point out, there are two ways to do that: the semantic and the pragmatic one. The first one consists in semantic binding by attitude predicates, as it was suggested by Schlenker (2003, 2007) and Sauerland (2007). This idea is based on examples such as the following:

(55) I am not prejudiced against Caucasians. But John, who is, thinks/claims that you are the worst honky he knows. (Schlenker 2003)

However, this would be a good strategy only if perspective shifting occurred solely with attitude predicates. According to examples by Potts (2007) and Harris & Potts (2009), this is not the case; we can have perspective shifting outside attitude predicates, provided that the context allows it (e.g. if the perspective of an individual other than the speaker is very salient in the discourse):

(56) A CPJ report on Venezuela tells us how problems have 'escalated' in Venezuela under Chavez, i.e. the physical attacks against journalists under previous presidents have 'escalated' to Chavez calling the opposition, which includes the media, names. This is very, very serious, but I don't think another coup attempt is called for until Chavez resorts to dramatic irony or sarcasm. But if that vicious bastard uses litotes, then there's no other rational choice than an immediate invasion. (Potts 2007: 175-6)

(57) My brother Ken said that his math tutor has been in a great mood lately. The jerk is always nicer when he's paid in advance. (Harris & Potts 2009: 32)

This suggests that perhaps the correct way to capture this phenomenon is pragmatic.

3.3.1.2. Pragmatic accounts of perspective dependence

3.3.1.2.1. Potts (2007)

Again, there are different alternatives about how a pragmatic account of perspective shifting might go through. One such proposal is made by Potts himself (2007), but in his revised framework dedicated for expressives. This time instead of speaker-orientation, Potts speaks of perspective-dependence. The perspective is generally the speaker's but can deviate under the right conditions. As we saw above, technically this is done by adding a parameter for a judge in the context tuple, a move rejected by Lasersohn (2005) for predicates of personal taste because of two worries: it imposed a single perspective per context and resulted in different contents at the mouth of each speaker, failing to capture the intuition of disagreement. It was mentioned that Potts (2007) adopted it despite the first worry as he posited that given their independence, there can be no perspective shifting of expressives mid-computation, unlike predicates of personal taste. However, there have been counterexamples to that:

(58) I am not prejudiced against Caucasians. But Pierre, who is, has repeatedly made the claim that you are the worst honky that the frog's mother knows. (Schlenker 2011)

Although Harris & Potts (2009) recognise that there can be two different perspectives in play, no problem is raised for Potts (2007) as their article aims to defend a pragmatic solution to perspective dependence in general, and not favour any specific one.

With regards to the second worry about different content for different speakers, it could be said that this is not a problem for Potts (2007), as disagreement does not seem to be possible with expressives in the same way as with predicates of personal taste.

(59) A: This natto is delicious.

B: No, it's not. (the negation denies the PPT 'delicious')

(60) A: This music is fucking unbearable.

B: # No, it's not. (# in the sense that the negation does not target the expressive 'fucking').

However, in some cases expressives can show a disagreement pattern similar to PPTs, which suggests that sometimes expressives can contribute propositional content, as seen in the below example by Lasersohn (2007: 226):[36]

[36] I will have more to say about the contrast between expressives that can appear in disagreement structures ('bastard') and those that cannot ('fucking') in chapter 5, where I will lay out my own account of expressives.

(61) John: Bill is a damn bastard!
Mary: No he isn't!

To summarise, Potts (2007) proposed a pragmatic implementation of perspective dependence after his semantics-based idea of speaker-orientation (2005) was shown to be untenable. However, this move inevitably inherits the problems Lasersohn (2005) warned about when he examined and rejected it for predicates of personal taste, namely perspective shifting and the unaccountability of disagreement, which at least in some examples seem to concern expressives as well.

3.3.1.2.2. A proposal parallel to predicates of personal taste?

As Potts (2007) seems to have inherited the problems Lasersohn (2005) warned about by adopting his discarded solution, the most obvious alternative is to adopt the same solution Lasersohn proposed for predicates of personal taste. In fact, Potts (2007) himself seems to have pondered over this possibility, but to have hesitated to adopt it because 'it would so fully integrate the judge into the semantics that it would be another semantic parameter, i.e., a part of semantic content in the same way that worlds and times are' (2007: 174). It is not clear why this would be such a problem, since it was demonstrated by Lasersohn to be a viable strategy for predicates of personal taste, but given that Potts did not believe that the side-effects of Lasersohn's rejected solution applied to expressives, he seems to have preferred to avoid it. But since it turns out that those side-effects do apply to expressives as well, let's see what happens if we employ Lasersohn's proposal for PPTs to expressives.

Following Lasersohn's suggestion to wire the judge into the evaluation index, the denotations of sentences are no longer sets of worlds but sets of world-individual pairs (ignoring times). Although this move seems to make the denotation appropriate for capturing the de se, given that we know from the previous chapter that a pair of an individual and a world essentially amounts to a centered world (Lewis 1979), in this system there is no way to ensure that the judge is identical to the speaker of the context, which renders this proposal inadequate for expressives, despite its success with PPTs. In a way it makes sense that these two phenomena cannot be captured by exactly the same mechanism, because although they share perspective-dependence, they differ in that unlike PPT, expressives are de se, non-displaceable and do not display the same (faultless) disagreement pattern.

3.3.1.3. $\mathcal{L}_{TU}$

Although no explicit link between expressives and de se is made by Gutzmann (2015), it so happens that the formal system already reflects the de se character of expressives.

We know from our discussion in the previous chapter that an attitude de se is captured by the idea of centered possible worlds; namely, as an attitude not only towards the way the world is or could be (possible world), but also referring to the position of the subject in it (centered possible world). Now recall that in Gutzmann's framework, a sentence's

u-proposition is said to correspond to a set of contexts in which the sentence is felicitous. But as we know, a context can only have one speaker; in this sense, it can be understood as a centered possible world (Schlenker 2011, Schroeter 2017). Now let's recall the definition of truth conditions and use conditions:

Truth conditions	Use conditions
"Snow is white"	"Oops"
is true	is felicitously used
iff snow is white.	iff the speaker observed a minor mishap.

t-proposition "Snow is white": the set of worlds in which snow is white.
u-proposition "Oops": the set of contexts in which the speaker observed a minor mishap.

As seen above, in the u-proposition there is explicit mention not only of the context but also of the speaker, while in the t-proposition we only talk about worlds, without any mention of a centre. As a result, Gutzmann's system $\mathcal{L}_{\mathrm{TU}}$ already captures the necessarily de se character of expressives.

However, since now there is explicit mention of the speaker, perspective shifting as it was noted in the examples by Kratzer (1999) and Schlenker (2003, 2011) is a problem. One possible solution to that would be to postulate some kind of context shifting operator to account for examples with different perspectives in a single sentence (e.g. the one by Schlenker 2011). I will not elaborate on this possibility because as I will show in the next chapter, Gutzmann's (2015) framework has some fundamental flaws that prevent it from getting off the ground.

Chapter 4. Disentangling expressive meaning

4.1. Uncluttering the landscape: conceptual problems of the existing accounts of expressive meaning

4.1.1. McCready 2010

As we saw in the previous chapter, Potts' (2005) tenet that there are no expressions that have both at-issue and CI content was quickly refuted by the existence of lexical items such as racial slurs and Japanese honorifics[37]. As Potts' $\mathcal{L}_{CI}$ was designed based on the assumption that there are no mixed items, it was unable to account for them, which was McCready's motivation to evolve the system in order to increase its applicability to more data. As McCready notes (2010: 8, 28), it is not entirely impossible for $\mathcal{L}_{CI}$ to account for mixed items, however, this will require several extra assumptions:

> Potts's core logic is not able to handle examples of this sort of mixed content (due to limitations imposed by the type system) without additional, costly assumptions about semantic decomposition. […] There is no obvious motivation for decomposition, other than the limitations imposed by the analytical resources made available in $\mathcal{L}_{CI}$. Without independent motivation, it seems much more natural just to analyze them as mixed content bearers. At the very least, one would not want to be forced to a decompositional analysis by the type system underlying the work.

In fact, McCready does attempt a decompositional analysis herself, only to show how troublesome and thus less preferred it would be to an extension of the formal system to cover for these cases. Although it is indeed convincingly shown that a decompositional approach would seriously overcomplicate things, the formal extensions proposed by McCready carry an even greater risk: they threaten the core argument of the whole enterprise of $\mathcal{L}_{CI}$ and its followers, i.e. its usefulness in handling meanings which are ontologically (at least as far as the meaning language is concerned) different from all other kinds.

One of the most arduously argued points of Potts (2005) is that CIs are completely independent from at-issue content, a fact that clearly distinguishes them from

[37] Potts (2005) and more specifically Potts & Kawahara (2004) consider Japanese honorifics, but they only look at cases where honorification is morphologically added to the verb root (e.g. *o-warai-ni-natta*), and do not discuss or mention forms which are purely honorific (e.g. *ossharu*, *irassharu*). Of course, it is possible that these forms are approached as suppletive, i.e. as underlyingly consisting of a verb root and an honorific affix but pronounced as unanalysable forms, which would allow them to fit in the aforementioned analyses – although McCready's argument for semantic decomposition mentioned right above would still apply.

presuppositions, which he understands as preconditions on whether the at-issue content can have a truth value or not. As noted before, Potts claims that while his system is based on Karttunen & Peters' (1979) logic for presuppositions, even though the Binding Problem was fatal for their system of presuppositions, it constitutes a 'virtue' for $\mathcal{L}_{CI}$ (2005: 79), given the total independence of CIs of the at-issue content. This was indeed a very strong conceptual argument for why we should recognise a new dimension of meaning and why it deserves its own dedicated formal system. But now McCready's extension of the system to account for mixed items raises an important conceptual worry: perhaps the limitations imposed by the technical system were there to signal the category boundaries, reflecting the inherent characteristics of the class of CIs; this would mean that the concessions made on independence could have the consequence that the clear divide between presuppositions and CIs is lost, so perhaps there is no essential difference between the two and no CI-dedicated system is needed.

One way to diagnose this would be to see whether McCready's (2010) amendments render the system vulnerable to the Binding Problem that turned out to be deadly for the system of Karttunen & Peters (1979). We should note here that when Potts characterises it as a virtue, he does mention a possible counterexample (2005: 20):

(62) Every Democrat advocating [a proposal for reform] says the stupid thing is worthwhile.

Potts says that in this example 'the relationship between the quantifier 'every' and the expressive content of 'stupid thing' is not one of binding. Rather, what we seek for the expressive meaning in quantified cases is a generic quantification over the restriction on the at-issue quantifier'[38]. Although this explanation could perhaps feel a bit ad hoc, it eliminates the possibility of the Binding Problem applying to $\mathcal{L}_{CI}$. Another example involving quantification with expressives is given in subsequent work by Potts (2007: 171, example by Florian Schwarz):

(63) Whenever I pour wine, the damn bottle drips.

As explained in Potts (2007), an analysis in terms of quantificational binding here would mean that the speaker is not frustrated at the moment of utterance, but only in the situations he is quantifying over, i.e. whenever he pours wine. However, this is contrary to the intuition that the speaker is expressing that he is experiencing exasperation at this very moment. As this explanation seems to fare well with intuitions, again the threat of the Binding Problem is neutralised.

Nevertheless, the permission of mixed expressions by McCready (2010) seems to dangerously open the backdoor:

(64) Someone is gawking at Sandy.

[38] As we noted in the previous chapter, Potts (2005: 127) explicitly says that CIs cannot be quantificationally bound, which he takes to be another confirmation of the inherent difference between CIs and presuppositions. But conveniently, apart from this example the non-quantifiability of CIs is directly exemplified mostly by supplements, and thus indirectly extended to expressives. This is an extra motivation for Potts to explain (62) in terms of generic quantification as well as to say that an analysis in terms of binding across dimensions would be one that 'badly mishandled the data' (Potts 2005: 20), without giving much explanation.

At-issue: ∃x look(S, x)
CI: rude/stupid(x)

(65) Dareka irasshaimashita.
someone arrived$_{HON}$
'Someone has arrived.'
At-issue: ∃x arrive(x)
CI: be-honourable(x)

In the examples above, it is possible for the person who is acting as the argument of the asserted proposition to be different from that of the CI one; the formal system has no way of ensuring sameness of reference. Specifically, in (64) the asserted proposition is true iff there is someone who is looking at Sandy, while the CI proposition is true iff someone is rude, and obviously both propositions could be rendered true by different individuals. Similarly in (65), the asserted proposition is true iff someone has arrived, and the CI one is true iff someone is worthy of honour, without any necessary connection between the two individuals. As this is identical to the problem that brought down Karttunen & Peters (1979), McCready's well-disposed improvements of Potts' system seem to threaten its very foundations.

Another issue with McCready (2010) is a certain lack of clarity on the exact class the system applies to. She offers the following terminological clarification for his use of the cover term 'conventionally implicated and expressive' content (2010: fn 2):

> 'CIE content' is intended as a neutral term for conventionally implicated and expressive content. In this paper, the assumption is made that, to a first approximation, both conventional implicatures and expressives make use of roughly the same combinatoric system. Where the distinction matters, I will not use the cover term.

From the above it is obvious that McCready (2010) is following Potts (2005) system, rather than the one in Potts (2007), even though she is aware of it as she cites it. Although this might be a minor point, one would expect clearer argumentation of why she chose to evolve the older system rather than the new one, and what of the new system does not work. To be fair, McCready elsewhere (2009) does mention the following reasons for choosing the framework of Potts (2005) rather than (2007): the complexity of Potts (2007), plus the fact that Potts (2007) can only represent attitudes towards individuals, due to the definition of expressive indices. The second reason feels more legitimate than the first, as there seems to be a considerable conceptual difference between Potts (2005) and (2007) so that the degree of complexity cannot stand as a criterion entirely on its own. A clearly voiced disagreement of McCready (2010) with Potts (2007) is the explicit rejection of the 'descriptive ineffability' property posited by Potts (2007), as he even uses descriptive paraphrasing as a way to apply the binding test with conditionals. Nevertheless, given that Potts himself intended his 2007 system as more apt for expressives than his 2005 one, a more detailed explanation by McCready about her choice of formal systems might have placed her own contribution on firmer grounds.

4.1.2. Gutzmann 2015

As noted in the previous chapter, Gutzmann (2015) evolved McCready's (2010) framework even further, proposing significant changes both in the conceptual as well as the formal components of the system. As we recall from the previous chapter, Gutzmann noted that despite the extensions proposed to account for mixed expressions, McCready's $\mathcal{L}^{+}_{CI}$ was still unable to accommodate some cases, such as use-conditional modification (e.g. 'fucking bastard'), abstraction ('Hans is gawking at every girl') and quantification ('Everybody is gawking at Sandy'). Moreover, he found that the process of parsetree interpretation did not respect compositionality, due to the requirement that meanings were collected from deeper layers of the structure.

Interestingly, the problems with abstraction and quantification which Gutzmann says McCready's $\mathcal{L}^{+}_{CI}$ cannot account for reveal the amount of departure from Potts' (2005) initial idea, since Potts had ruled out the possibility of quantifiability of CIs. That is to say, being a follower of $\mathcal{L}^{+}_{CI}$, Gutzmann (2015) seems to inherit its problem of blurring the divide from allegedly ontologically different meanings (e.g. presuppositions), which casts doubts over the necessity of a specialised logic.

However, as noted before, one significant difference of Gutzmann's $\mathcal{L}_{TU}$ from McCready's $\mathcal{L}^{+}_{CI}$ is that the distinction is now between truth-conditional and use-conditional meaning, rather than at-issue and not-at-issue content, as he explicitly separates the two: 'the difference between truth-conditional and use-conditional content cannot plausibly be reduced to a distinction between contributing at-issue or side issue content' (2015: 272). But the problem with this new distinction is that it is much less easily defined than the previous one; although there is a precise definition of at-issueness as well as specific tests for diagnosing it (e.g. Simons et al 2010, Tonhauser et al 2013), what is exactly meant by 'conditions of use' is assumed rather than precisely explained by Gutzmann (2015), which renders the notion somehow vague and the distinction between truth and use conditions slippery to define. Although in Kaplan (1999), which constitutes the theoretical arsenal behind Gutzmann's $\mathcal{L}_{TU}$, the fact that this distinction is clear-cut is presented as a given, it turns out that it is not necessarily so.[39] Continuing

[39] After all, Kaplan (1999) was a talk defending the possibility of a formal account of expressions that had been discarded as untreatable by formal semantic methods, as well as sketching the general directions, rather than a fully worked out theory of the exact differences between the different kinds of meaning. Kaplan (1999) claims that some expressions could be formally treated 'by attending to rules of use – the right sort of rules of use' (1999: 3), but without precisely saying what would count as 'the right of rules of use' in each case, and as he admits, his definitions of the terms 'descriptive' and 'expressive' are given 'roughly' (1999: 5). For this reason, it is perhaps a bit shaky to base a formal semantic enterprise without further theoretical elaboration of the otherwise extremely insightful points made by Kaplan (1999). Although as Kaplan (1999) says his seminal 'Demonstratives' (1989) was also based on the idea that the meaning of some expressions is better given in 'rules of use', that one was a thorough research project where he explained exactly what he took to be the rules of use for the expressions he termed 'pure indexicals' and 'true demonstratives', in which case the distinction with all other

the tradition of the late Wittgenstein, Travis (1985, 1996, 1997) has repeatedly pointed out problems with the idea of meaning as truth conditions even in the most supposedly uncontroversial cases (1997: 89):

> Pia's Japanese maple is full of russet leaves. Believing that green is the colour of leaves, she paints them. Returning, she reports, 'That's better. The leaves are green now.' She speaks truth. A botanist friend then phones, seeking green leaves for a study of green-leaf chemistry. 'The leaves (on my tree) are green,' Pia says. 'You can have those.' But now Pia speaks falsehood.[40]

Without needing to go into the details of the debates between minimalism and contextualism,[41] we can take the moral of this story simply to be that the idea that certain expressions are susceptible to a truth-conditional treatment is not uncontroversial and thus the distinction between truth conditions and use conditions cannot be taken for granted but should be argued for. Even though this task might sound more relevant to the philosopher than the linguist, designing a formal linguistic framework on the presumption that there is a fundamental difference between truth conditions and use conditions seems to require a somehow precise delimitation of this difference. After all, it is always possible for a follower of so-called 'meaning eliminativism' (Recanati's 2003 term) or else 'occasionalism' (Borg's 2009 term) to argue for the adoption of conditions of use across the board, eliminating the distinction that lies at the very heart of $\mathcal{L}_{TU}$. As Borg (2009: 102) tells us, 'according to occasionalism there is simply no such thing as determinate content outside a context, for it is only in use that words and concepts come to have particular conditions of application'. Therefore, in order to justify a formal framework designed for use-conditional meaning, having a clear definition of what use-conditional meaning is and how it differs from truth-conditional one seems like a theoretical prerequisite.

Another possible problem for $\mathcal{L}_{TU}$ lies at the way denotations are defined. We recall from the previous chapter that the T-schema is replicated in the use-conditional dimension by substituting the notion of truth by that of felicity:

Truth conditions	Use conditions
"Snow is white"	"Oops"
is true	is felicitously used
iff snow is white.	iff the speaker observed a minor mishap.

The problem here is that unless the speaker observed a minor mishap simpliciter, 'oops' is not felicitous. Now note that there is a difference between the speaker observing a minor mishap, and the speaker believing that they did, but ultimately discovering that

---

expressions (non-indexicals, non-demonstratives) was clear. This seems much less the case in his 1999 work, as it has not been developed further than a talk.

[40] As noted by Bianchi (2010), both Searle (1979, 1980, 1992) and Travis (1975, 1981, 1985, 1996, 1997) have come up with many such examples, involving expressions that seem straightforwardly defined in truth conditions, such as 'duck', 'milk', 'red', 'open', 'cut', etc.

[41] For example, see Recanati (2003), Borg (2009).

they were wrong (e.g. the supposed mishap was in fact part of a movie script)[42]. As Gutzmann's (2015) condition stands, the use of 'oops' should be infelicitous in the second case, but this seems to clash with the de se character of expressives, which was demonstrated in the previous chapter. That is to say, given the necessarily de se nature of these words, the mere utterance of an expressive guarantees that the speaker thinks it is an instance of felicitous use.[43] This means that while attributing the conventional meaning of expressives in terms of speaker attitudes makes intuitive sense, phrasing it in terms of 'iff' seems problematic because it implies some sort of evaluation process. Even if one could invoke here Kaplan's (1999: 12) example of using 'oops' when observing the collapse of a building as 'expressively incorrect' to defend the idea that there is some kind of evaluation, if indeed the conventional meaning of 'oops' is something akin to the speaker having observed a minor mishap, then the example of the building can at best show that the speaker is cruel enough to consider such an event as a minor mishap, rather than that he has made a mistake – after all, who could effectively contest whether this was really a minor mishap to the speaker or not? We can imagine a villain who performs some horrible violent act and just utters 'oops' to show how minor it was to him[44]. That is to say, given the speaker's first person authority on his mental states, despite any objections that perhaps the use of an expressive was deviant compared to established practices, or even sounded like a 'macabre joke' (Kaplan 1999: 12), it seems very difficult to label it as 'incorrect'.

But there is an even more serious conceptual worry here. The use of the logical biconditional is problematic for the definition of use conditions due to its double directionality. In the case of truth conditions we can use 'iff' because the proposition corresponding to the sentence 'Snow is white' is true if snow is white, and also if snow is

[42] Kaplan (1999) examines such an example, and says that in a case like this 'oops' would be expressively incorrect: 'the moral of this story, in contrast to that of *ouch*, is that in this case, you, the oops-er, do not have privileged access to the state of affairs which must obtain in order for the expressive to be expressively correct.' (1999: 13). Kaplan's claim is based on his distinction between subjective and objective expressives, with the former expressing merely a state of the speaker and the latter more than that, such as facts. I do not agree with this distinction as it stands, because it goes against the de se characteristic of expressives, which I advocated for previously. But it is not clear whether Gutzmann (2015) agrees with it either, as he does not provide a different analysis for so-called objective expressives such as 'oops' and subjective ones such as 'ouch'; in fact, he does not even mention the distinction. So even if his proposed felicity schema could potentially work for objective expressives, it will not work for subjective ones.

[43] Kaplan (1999: 16) mentions the possibility of 'insincere use' of expressives such as when someone uses honorifics with a person they do not actually respect. This is relevant to the point about objective expressives mentioned in the previous footnote, which I do not agree with. Even more in the case of honorifics, whether a certain use is sincere or not is not easily contestable because of the speaker's first person authority on his mental states. Either if we analyse them as subjective expressives (conveying something along the lines of 'I respect you') or as objective expressives ('You are worthy of respect'), it essentially boils down to the speaker whether s/he is in the mental state encoded by the former or is in agreement with the latter. As a result, even though it is definitely possible for a speaker to use honorific language at the absence of genuine respect, it is not clear how one could verify any such case as an instance of insincere use, as it can only be retrospectively admitted by the speaker himself.

[44] Some musical examples from rather opposing genres: Postmortem's 'Oops! ... I Killed Again' (https://www.youtube.com/watch?v=Tdnea8c3qF0) and Britney Spears' 'Oops, I did it again (I played with your heart)' (https://www.youtube.com/watch?v=CduA0TULnow), if one considers causing heartbreak more than a minor mishap, of course.

white then 'Snow is white' is true.[45] However in the case of use conditions, although one can say that the expression 'oops' is felicitous if the speaker has observed a minor mishap, it feels less certain to say that if the speaker has observed a minor mishap then the expression 'oops' is felicitous. The reason is because in this case we are not dealing with a proposition, which is understood as an independent entity, with an absolute truth value, corresponding to a set of possible worlds, etc. Rather we are dealing with an utterance, and utterances are not independent entities but intentional actions of agents; this means, for example, that a speaker could genuinely believe to have observed a minor mishap, but consider that the use of 'oops' is not felicitous, perhaps because of the discourse context, which could be too formal, too strict, too focused on a specific topic to allow deviances, etc. As a result, although definitions in 'iff' terms are appropriate for truth-apt entities, they do not seem to be analogously extendable to non-truth-apt ones, contrary to what Gutzmann (2015) suggests.[46] All in all, it seems highly dubious that one could advocate for a correspondence theory of felicitous use in the exact same way as for a correspondence theory of truth, especially if one ascribes to the idea that there is an external reality which exists independently from human minds, which comes in contrast with the notion of felicitous use which seems to be a subjective notion rather than a collection of objective facts external to human minds.

The above perhaps more philosophical worries about the definition of use conditional meaning are also reflected in practical issues about delimitating the theory's domain of application. According to Gutzmann & McCready (2016: 2), the class of expressions that contribute use conditional meaning 'goes beyond the stereotypical characterisation of expressives in a strict sense. That is, besides the standard examples of expressive adjectives, honorifics, or ethnic slurs, there are also expressions like particles in German (Gutzmann 2015) or Japanese (McCready & Takahashi 2013) or even syntactic constructions (Frey 2010) that fall under the scope of the framework provided by the formal systems we will discuss in this paper' (which is essentially $\mathcal{L}_{\mathrm{TU}}$). Following the authors' work, we find that other expressions they place in this class are numerical classifiers (McCready 2009), phi-features (McCready 2010), exclamatives (Castroviejo Miró 2008, cited by Gutzmann 2015), intonation, syntax (e.g. topicalisation) and morphology (e.g. diminutives) (Gutzmann 2013), appositives and referential uses of definite descriptions (Gutzmann & McCready 2014) and even predicates of personal taste (Gutzmann 2016).[47] What seems to be the problem here is that the class of UCIs

[45] Of course we need to consider a world of evaluation, but I skip this here for simplicity of exposition.

[46] It should be noted that Gutzmann himself emphasises on the difference between truth and felicity, when he says 'truth is independent of actual use and depends on possible states of affairs, while felicitous use obviously is tied to a use case' (2015: 17); nevertheless, he still believes that it is possible to represent felicity conditions in terms analogous to the T-schema.

[47] To be fair, just as in the quote by Gutzmann & McCready (2016), Gutzmann (2013, 2015) also makes a distinction between 'expressives in the narrow sense', which he defines as 'expressions that express some emotional and evaluative attitude with a high degree of affect' and the rest, such as 'particles and non-lexical UCIs' (2015: 27). However, this distinction does not seem definitional or explanatory in any sense, as it's not clear whether the criterion is content (whether the word expresses 'some emotional and evaluative attitude with a high degree of affect') or morphosyntax (whether it is a particle or not a lexical item at all). For example, where would expressions such as 'delicious' or 'lovely', or diminutives lie? Because 'delicious' and 'lovely' are not particles but words, and one could say that they express an

starts to feel like a wastebasket for anything that resists a (straightforward) truth-conditional treatment, which means that the category itself is defined in negative terms (what it is not) rather than what it really is. This would be much less of a problem if it was openly clarified, as is in Sudo (2013: 275), where 'UCI' is used as a cover term: 'the non-truth-conditional (or 'use-conditional') aspects of the meanings of declarative sentences, e.g. presuppositions, conventional implicatures, etc'. But no such concession is made by Gutzmann (2015), who still marks a distinction between presuppositional and use-conditional meaning (as amounting to 'hierarchical' versus 'parallel' multidimensionality, 2015: 196). Given the vagueness of the term 'use-conditional meaning' that was diagnosed above, it is feared that the category is comprised by largely heterogeneous items which simply have not yet been accounted for by other means – or in some cases such as taste predicates or phi-features they have been, but still a new proposal in use-conditional terms is made perhaps simply because the formal framework allows it. This situation is reminiscent of the grouping of several heterogeneous phenomena under the rubric of 'presupposition', leading Karttunen & Peters (1977, 1979) to decry the use of the term and call for its abolition.

Moreover, the lack of terminological clarity that was noted above for McCready (2010) is also reflected in Gutzmann & McCready (2016: 1), where they say that they 'take the disapprobatory content of pejoratives to be expressive in nature (or possibly conventionally implicated)'. However in that article they adopt Gutzmann's (2015) approach to expressive meaning as a subgroup of use-conditional meaning rather than as a conventional implicature, so recognising the possibility that expressive meaning could be conventionally implicated is rather confusing, since under such an assumption the denotations would be propositional (unless there is an extra assumption that CIs are no longer propositions). As Gutzmann clearly notes (2015: 272), 'the difference between truth-conditional and use-conditional content cannot plausibly be reduced to a distinction between contributing at-issue or side issue content', so the statement that expressives could possibly be conventional implicatures is puzzling.

As we saw above, the recognition of mixed items rendered McCready's $\mathcal{L}^{+}_{\mathrm{CI}}$ vulnerable to the Binding Problem, and as expected this is inherited by Gutzmann's $\mathcal{L}_{\mathrm{TU}}$ since it also covers such expressions. Gutzmann anticipates this worry but says that 'this is not a problem specific to $\mathcal{L}_{\mathrm{TU}}$, but holds generally for multidimensional approaches. However, since quantification is not a major topic of this book, I cannot provide a solution for the binding problem here' (2015: 161). Nevertheless, quantification *is* the specific topic of Gutzmann & McCready (2016, 'Quantification with Pejoratives'), but no solution to the Binding Problem is proposed there either; conveniently, the analysis of quantification proposed in that paper is exemplified using universal rather than existential quantification so the issue is not touched upon.

---

emotional/evaluative/affective attitude, they risk being indistinguishable from classical expressives such as 'damn' or 'fucking'. As for diminutives, because they are not single expressions but morphemes, they would not be classified as expressives in the narrow sense, although one could say that they carry emotional/affective content much more than the aforementioned predicates of personal taste. All in all, the proposed distinction does not seem rigorous enough.

We also saw that one of Gutzmann's (2015) motivations to propose multidimensionality throughout the composition was because Potts' (2005) and McCready's (2010) parsetree interpretation seemed to be a threat for compositionality. Nevertheless, it seems that if we accept a distinction between 'strictly local' compositionality, where only the immediate parts are considered, and 'non-local', where the subconstituents of the immediate parts are also considered (Partee 1984, Larson & Segal 1995, Pelletier 2012), perhaps we could say that parsetree interpretation obeys non-local compositionality, although not the strictly local variety. But even if this answer is not satisfactory, it would also be possible to adopt a solution briefly mentioned by Potts (2005: 68), but dispreferred over parsetree interpretation: that of a CI store. Considering these two options, the accusation of non-compositionality against Potts (2005) would be much less serious and one could resist resorting to 3-dimensional objects across the board as is advocated in $\mathcal{L}_{\text{TU.}}$ Although we should be clear that Gutzmann (2015) did not introduce compositional multidimensionality only to solve the compositionality problem, but also to account for use-conditional modification, abstraction and quantification, which were not covered by either $\mathcal{L}_{\text{CI}}$ or $\mathcal{L}^{+}_{\text{CI}}$. However, given the high degree of complexity that ensued (the need for LERs, the '☆' operator, etc.) one could perhaps doubt whether a move as radical as positing multidimensionality throughout the entire composition was worth it. Moreover given the persistence of the Binding Problem, it seems quantification was not successfully accounted for either, which discounts a bit more from the value of this significant theoretical assumption.

4.1.2.1. General criticism against Gutzmann (2015)

One has the feeling at this point that the theoretical background is not stable enough to support the major assumptions (e.g. compositional multidimensionality) or the rather high degree of technical convolution that ensues. Conceptually, the framework of hybrid semantics is entirely based on Kaplan (1999), which as noted before was an informal talk sketching some possibilities rather than a fully worked out theory. Of course it is quite possible that all of Kaplan's insights are perfectly defendable but it feels that they should be developed further and more rigidly argued for. In essence, Gutzmann's (2015) framework extends way beyond a technical extension of Potts (2005) and McCready (2010): Potts (2005) called for the recognition of a distinct category of meanings (CIs), which were still propositional but needed a separate logic to be represented, while McCready (2010) complemented the system with necessary tweaks to accommodate more data. However, besides the technical novelties regarding semantic composition, Gutzmann's (2015) project seems to have important philosophical and metasemantic implications because of the proposed change in denotations. As a result, one would expect more in-depth discussion and detailed argumentation of the theory adopted, rather than merely a summary of it as it was informally advocated by Kaplan (1999), who repeatedly admits that it is given in 'rough' terms. In fact, Crespo (2015: 71-2) voices the following worry about Gutzmann's work:

> Following Kaplan (1999), who squibs an idea of how to specify felicity conditions for typical examples of interjections like *ouch* and *oops*, Gutzmann[48] is optimistic about how these two layers can be integrated. Gutzmann proposes what he calls a hybrid semantics in which truth and use conditions can work side by side. [...] But […] expressivism creates a division between descriptive and prescriptive sentences, between kinds of meaning, or between kinds of contents that sentences express, and then has to face the problem of the union and resemblance. How can these two kinds of sentences, meanings, or mental contents interact? If they are so similar, then what exactly is the decisive difference? In other words, if truth- and felicity-conditions are as similar as Kaplan (1999) suggested, then should not we try to find a way to conceptualise meaning in such a way that we put these two together instead of making a division?

4.1.2.2. Breaking down the different categories

After the critical discussion of the conceptual issues in Gutzmann (2015), the next destination is the data. This section will critically review the typology of UCIs proposed by Gutzmann (2013, 2015). Let's recall the different types of UCIs, which we saw in the previous chapter:

4.1.2.2.1. Types of UCIs

	f	2d	r-s	e.g.
isolated expletive	−	−		'ouch', 'oops'
isolated mixed	−	+		ethnic slurs, e.g. 'Kraut'
functional expletive	+	−	−	'damn'
functional shunting	+	−	+	Japanese adverbial 'yokumo'
functional mixed	+	+		Japanese honorifics verbs, e.g. 'irassharu' (go_{HON})

Table 4.1. Types of UCIs

f: functionality; whether the expression needs an argument
2d: 2-dimensionality; whether the expression has meaning in both dimensions (mixed)
r-s: resource-sensitivity; whether the expression fully consumes its argument

As we can see, Gutzmann (2015) proposed the above typology based on two properties: dimensionality and functionality. However, using functionality as a criterion to distinguish between different subcategories of UCIs seems in itself problematic, given Gutzmann's objectives. Even though he is clear that he defends 'a conventionalist view of semantics' (2015: 5), he states that one of the goals of his book is to 'develop a compositional, multidimensional formal framework for hybrid semantics' (2015: 9). Therefore if the mission is to provide an account of semantic composition, one would expect the property of interacting with other constituents (i.e. functionality) to be taken for granted, instead of allowing it to vary between plus and minus. After all, if some meanings always come saturated and cannot be arguments themselves either (due to

[48] Crespo (2015) is here referring to Gutzmann's (2016) 'If expressivism is fun, go for it', which she cites as Gutzmann (2014) as it was not published at the time she was writing.

being 'isolated'), one wonders what would count as good reasons for including them in the compositional process.[49, 50] For Gutzmann, such a reason is conventionality.

As he notes, there are two possible criteria in order to delineate the domain of what one means by 'semantics': truth-conditionality and conventionality. This is in fact Kaplan's (1999) point, and it has also been repeated by other philosophers, such as Kripke (2011) and Predelli (2013). However, it is not an accident that this point has been made by philosophers rather than formal semanticists; convention is a key notion in the discussion for a theory of meaning (ever since Plato's Cratylus) but its inclusion in compositional semantics should not be automatically justified until it is clear that it participates in semantic composition. For example, apart from the obvious point of phonetic differences (such as accent), information about the speaker's origins can be conveyed in a conventional manner by lexical differences across dialects (e.g. the use of 'pop by' and 'mobile' is conventionally perceived as British by US English speakers[51]). But it is not clear such information, albeit conventional, interacts with any other constituent in a compositionally relevant manner to yield meaning. On the contrary, it feels much more as a side-comment, or one that is added on top of everything else. Specifically for the examples just mentioned, if we represented the expressions' dialectal profile in the semantic composition, we would be making the wrong predictions since whether something is perceived as dialectal or not usually depends on pragmatic considerations such as the speaker's affiliation, the circumstances of discourse etc. (i.e. 'flat' and 'mobile' sound totally unmarked to ordinary British English speakers, although they could potentially perceive them as British-like in a communicative instance where another variety of English is spoken). Therefore it seems that conventionality cannot be a criterion for a definition of compositional semantics on its own, but needs to be constrained by something else to prevent our semantic representation from overflooding by too much information or becoming too pragmatic in nature. As noted above, one obvious constraint could be whether an expression takes an argument or can be used as an argument itself. If neither holds, it feels that there is no adequate justification for overcomplicating the semantics for representing such information there. Of course, this does not in any way mean that such information does not need to be represented; exactly because of its conventional character, expressions

[49] To be clear, it is the fact that such expressions are *both* saturated *and* not argument material either that points towards a treatment outside the compositional process. Just being saturated obviously won't do, as the whole idea of Frege's conjecture is to combine a saturated (such as a name or sentence) with an unsaturated component (Heim & Kratzer 1998). But we should also note that Gutzmann takes it that sometimes isolated UCIs can function as arguments, e.g. in constructions such as 'Ouch, man!' (2015: 92). I will discuss and try to refute this below.

[50] As Potts (2005: 44) first noted when he defined the notion of isolated CIs, 'when they [isolated CIs] arise, the 'side-comment' nature of CIs is pragmatic: in a sense, the speaker is commenting on the circumstances of utterance, rather than expressing an aside about the at-issue comment'. However despite acknowledging their pragmatic nature, Potts (2005) opts for a purely semantic treatment of isolated CIs. For Amaral, Roberts & Smith (2008) this leads to inadequate formal treatment: 'it seems that the isolated CI rule is a catch-all for phenomena which cannot be correctly accounted for by using CI application or COMMA' (2008: 723-4). For this reason, the present criticism against Gutzmann (2015) that isolated (marked as –functional) need not be represented in the grammar also extends to Potts (2005).

[51] https://www.youtube.com/watch?v=SlCApAgyayI

carrying such meaning deserve their own semantic profile[52]. Thus to clarify, the disagreement with Gutzmann (2015) is not whether conventionality is an adequate criterion for a theory of meaning, but whether it is also one for a theory of semantic composition. The immediate consequence here is that Gutzmann's categories of isolated expletive and isolated mixed UCIs can no longer figure in a typology of UCIs, provided that the field is that of compositional semantics. Let's see what implications these considerations have on the categories of isolated mixed and isolated expletive of Gutzmann's UCI typology.

4.1.2.3. Banishing isolated UCIs

4.1.2.3.1. Isolated mixed UCIs

As we saw before, these are expressions which have meaning in both dimensions (+2d) but do not require an argument at the use-conditional dimension (−f). The most typical examples are pejoratives such as ethnic slurs (e.g. 'Kraut' and others). Employing linguistic tests to show that the expressive part is scopelesss and unaffected by denial, McCready (2010) notes that 'pejoratives introduce mixed content: but only part of this content, the at-issue portion, is predicative [...] the CIE [conventionally implicated or expressive] content is propositional' (2010: 16). He represents this in $\mathcal{L}^{+}_{CI}$ terms as follows:

(66) ⟦Kraut⟧ = λx. German(x) ♦ bad (∩German)[53]

That is to say, the expressive component of slurs encodes a negative attitude against the general kind denoted by the at-issue content, not against the specific individual the slur is used about; that part is only inferred. In McCready's words, 'since the speaker asserts that Juan is German, and expresses a negative attitude toward German people in general, it is natural to conclude that the speaker holds a negative attitude toward Juan as well. It is also natural to conclude that the speaker intends, as part of the reason for his utterance, to indicate this attitude. The content that Juan is bad, then, is communicated, probably intentionally, but is not, strictly speaking, a part of the semantic content of the sentence' (2010: 22). We can complement the above representation by including the conversational implicature (symbolised as '+>') as follows:

(67) ⟦Kraut⟧ = λx. German(x) ♦ bad (∩German) +> bad(x)

Of course, we cannot deny that a conventional element is involved in this inference; competent speakers know that 'Kraut' is an offensive word, usually employed to express a negative attitude towards German people. But as per the discussion above, conventionality by itself is not sufficient for inclusion in the compositional process; if a certain element does not interact compositionally with anything else, i.e. it does not require an argument as it comes pre-saturated and cannot be used as an argument

[52] My employment of the word 'semantic' is as clarified by Predelli (2013: 69): ''Semantic', of course, in the sense of a theory of meaning, rather than a theory of truth-conditions.'
[53] The operator '∩' is used to derive a kind out of a property (cf. Chierchia 1998)

either, then there is little reason to represent it in the composition, as it does not participate in it.

Apart from the above conceptual argument, there seems to be some empirical evidence as well: the fact that the offensive component of slurs is defeasible in some environments testifies against considering it as part of semantic composition, and suggests that the adequate analysis for such terms should include some pragmatic considerations. Specifically in the case of slurs, a phenomenon where defeasibility is observed is that of appropriation (Anderson & Lepore 2011, Bianchi 2014). This refers to the practice of adoption of the offensive terms by the very groups they are used to offend. Some of the most well-known examples of appropriated slurs are that of 'nigger' by the African American community and that of 'queer' by the homosexual community – especially the latter one has even entered academic parlance. This means that members of the targeted groups use such terms to refer to themselves and each other, expressing thus camaraderie. However, despite this, their use can remain highly offensive when used by members of the out-group, which has the quite astonishing result that one and the same expression can be used in the exact same conversational context by two different speakers and have diametrically opposed effects: fellowship versus offence. Which effect will ensue seems to be entirely dependent on whether the speaker belongs to the in-group or the out-group; this is so persistent that it cannot even be contained by quotation, which is the ultimate linguistic device for distancing oneself and attributing commitment solely to the original speaker.[54] The phenomenon of appropriation of slurs serves as evidence for their significantly contextual nature and the high volatility of the non-truth conditional part of their meaning, which might further suggest that this meaning does not participate in the compositional mechanism, as has been proposed above in (67).

It is thus argued that pace Gutzmann (2015), the expressive or use-conditional meaning of so-called isolated mixed UCIs such as slurs[55] does not form part of compositional semantics[56].

---

[54] This can be easily demonstrated by a simple online search about whether the quotation of racial slurs is perceived as offensive or not. Some examples: https://www.theguardian.com/media/mind-your-language/2014/may/08/mind-your-language-n-word , https://www.mediaite.com/tv/discussion-when-quoting-racists-should-tv-reporters-repeat-horrible-slurs-on-air/ , https://patrickiber.org/2016/01/31/on-the-dilemma-of-quoting-racial-slurs-in-lecture/ , and many others.

[55] Following McCready's (2010) observation that the expressive component is propositional rather than predicative, the distinctive characteristic of expressions of this category is that the attitude conveyed (usually negative) is against a group, rather than the individual denoted. It thus includes slurs based on race, ethnicity, religion, sexual orientation, gender, but also disability (e.g. 'retard'). In recent years, due to society having become more sensitive to issues of gender equality and body positivity, we have witnessed some expressions which were previously regarded as swearwords, insults or merely describing something socially undesirable being perceived as/becoming slurs, such as 'slut' , 'bitch' and 'fat'. Some examples of such intuitions: https://www.theodysseyonline.com/slut-sexist-slur , https://finallyfeminism101.wordpress.com/2008/05/16/feminism-friday-on-bitch-and-other-misogynist-language/ , https://www.psychologytoday.com/blog/eating-mindfully/201409/is-fat-the-new-f-word , http://www.huffingtonpost.com/2015/06/29/other-f-

Apart from slurs targeting specific groups, other expressions in this category are pejoratives such as 'cur' (labelled 'coloured' by Frege 1897) and by extension general register variants such as 'umbilicus' for 'navel', 'bunny' for 'rabbit', etc. (examples from Predelli 2013).

However, we should note some differences between slurs and pejoratives. As we saw, slurs seem to display a negative attitude towards a group in general, rather than the specific individual referred to, which is done indirectly. This is why the use of 'Kraut' to refer to someone is felt to be offensive not solely to that person, but to all Germans. It does not seem to work quite exactly the same with pejoratives such as 'cur', which feel more towards the specific entity rather than the entire class. For example, there seems to be a difference of felicity in the two examples below:

(68) # This nigger won't stop leaving his afro hair care products in the communal shower every time. My adorable black boyfriend would never do that.

(69) This cur won't stop barking at 4 am every night. My own lovely doggie is so well-behaved that never barks during the night.

Even though the examples are not exactly parallel, the idea is that in the first one the use of 'nigger' signals that the negative attitude is against the entire race, whereas in the second case the use of 'cur' feels more like expressing a negative attitude towards the specific dog, rather than dogs in general (hence the felicity of expressing a positive attitude towards another member of the set immediately after, contrary to the first example). There are two ways one can account for this difference. One is to say that unlike slurs which are isolated mixed UCIs, pejoratives like 'cur' are functional mixed UCIs, meaning that they take an argument in the expressive dimension as follows:

(70) $[\![\text{cur}]\!] = \lambda x.\ \text{dog}(x) \blacklozenge \text{bad}(x)$

Another way would be to say that pejoratives like 'cur' are isolated mixed UCIs like slurs, and the above difference is merely one of degree, to be attributed to sociological factors: namely, people are much more sensitive with generic terms used for groups of other humans such as minorities rather than those used for animals, objects, etc.[57] As a result,

---

word_n_7671762.html . I will elaborate on the difference between slurs and swearwords in the section on functional expletive UCIs.

[56] To be clear, at the cost of sounding repetitive: I am not claiming that the expressive meaning of these expressions is not part of their semantic profile; I am merely saying that so long as Gutzmann's (2015) goal is to 'develop a compositional, multidimensional formal framework' (2015: 9), such items need not—and in fact should not—be accounted for in such a framework. Contrarily, I am in complete agreement with Predelli (2013) when he identifies his project concerns a 'semantic theory, at least in the everyday sense of 'semantics' as the study of meaning' (2013: 63). There can be no denying that a semantic theory defined in such terms ought to account for expressive meaning irrespective of its compositional relevance, or even all kinds of vocalisations that are conventional as well as intentional (to exclude 'verbal discharges' as Predelli calls them, such as a scream or a hiccup, for example). More on Predelli's (2013) theory in chapter 5.

[57] A possible exception to this claim is that there are some generic terms denoting particular groups of humans which do not feel like slurs, such as 'hack' (in the sense of a bad writer), 'ham' (bad actor), 'quack' (bad doctor), etc. But interestingly, although these expressions denote groups of people, unlike slurs they do not pejorate the larger groups in question (identified in terms of

a term expressing a negative attitude towards the set of people of a specific race has quite a shocking effect because of society's sensitisation to the problem of racism, which is not a linguistic problem hence its gravity for society, but is reflected by linguistic practices (i.e. the use of racial slurs). Contrarily, a term expressing a negative attitude towards the set of animals of some kind is not associated with any issue so big and is thus more likely to be perceived as offending the specific entity referred to in the speech act than the entire set, merely because of general knowledge and presumptions about society. If this is the case, then we could expect that a shift in the values of society would affect a pejorative's domain of offence. As we noted above, this seems to be confirmed: some pejorative terms against women, such as 'bitch' and 'cunt' are now widely regarded as misogynistic slurs offensive to the totality of women, rather than targeting the specific individuals they are used for. Another supporting example is that of the Greek equivalent of 'cur' ('κοπρόσκυλο'); although it is generally felt as used to express a negative attitude towards the specific dog referred to, similarly to the example above, it is in fact perceived as a slur by groups of people dedicated to animal welfare.[58] In other words, the claim here is that any single expression which denotes a class and encodes a derogatory attitude towards it is prone to becoming a slur if the linguistic community in question becomes sensitive about the welfare of the class in question; the mere fact that there is a dedicated expression which denotes and discredits a group at the same time can become perceived as an instance of prejudice.

Therefore, it seems that the difference of effect between slurs such as 'Kraut' and pejoratives such as 'cur' can be attributed to extra-linguistic factors, and they can both be classified as belonging to the same category, that of isolated mixed UCIs.

(71) $[\![\text{Kraut}]\!] = \lambda x.\ \text{German}(x) \blacklozenge \text{bad}\ (^{\cap}\text{German}) +\!> \text{bad}(x)$

(72) $[\![\text{cur}]\!] = \lambda x.\ \text{dog}(x) \blacklozenge \text{bad}\ (^{\cap}\text{dog}) +\!> \text{bad}(x)$

The same goes for register variants such as 'bunny' and 'umbilicus'[59]; just as with our analysis of slurs, by saying that these expressions are isolated mixed UCIs what is meant is that apart from their truth-conditional meaning, they carry some connotations which can be reduced to world knowledge about the circumstances in which such expressions are generally used. And because of this just like slurs, they do not require any argument in their use-conditional dimension, which makes them isolated.

---

an inherent characteristic such as race, sexual orientation, gender identity, etc.) but simply to denote a subset of the larger set, specifically those who are not very skilled at the craft. As a result, this allows a paraphrase in the sense of 'bad' + class denoting noun. On the contrary, slurs do not denote a subset but the entire set, and their meanings are more aptly paraphrased as 'bad by virtue of being part of' + class denoting noun. So the distinction between words like 'hack' and slurs should be clear. However, I would like to claim that there is also a distinction between words such as 'cur' and those like 'hack', which is that the former are pejorative whereas the latter are not. Of course, this does not mean that words like 'hack' cannot be used to insult or offend, but merely that they are not derogatory terms. Predelli (2013: 109) notes the distinction as follows: 'derogation, at least in the sense relevant here, is a primarily linguistic dimension, related to, but importantly distinct from the variety of meaning-independent considerations appropriate for the discussion of phenomena such as insulting or offending.'

[58] https://www.thespro.gr/2016/12/blog-post_68.html, https://tinyurl.com/yb22tjso

[59] Examples taken from Predelli (2013).

(73) ⟦bunny⟧ = λx. rabbit(x) ♦ cute (∩rabbit) +> child-directed speech-proper context

(74) ⟦umbilicus⟧ = λx. navel(x) ♦ formal (∩navel) +> formal context (e.g. among physicians)

And while we are in the discussion of register, another class of expressions that we could classify as isolated mixed UCIs is what Harada (1976) called 'performative' honorifics in Japanese, such as the verb ending '-masu' and the polite verbal form 'desu' (verb 'to be'). These are not directed towards a specific argument, but are akin to polite speech and thus convey the idea that in employing them, the speaker respects the addressee. Although 'desu' is a verb in itself, as '-masu' is a verbal suffix it might not seem to fit in the category of isolated mixed UCIs; but since it never appears on its own but forms an inextricable part of verbs, we can treat the verbal expressions that end in '-masu' as isolated mixed UCIs, e.g. 'tsukare-masu' ('get tired'), with the neutral, non-UCI form being 'tsukare-ru'. Potts & Kawahara (2004) treat both performative honorifics and argument-oriented honorifics (which Harada 1976 calls 'propositional', more of which below) as expressives and analyse them similarly, while in Potts (2005) although they are both regarded as expressives, performative honorifics are treated like utterance-modifying adverbs such as 'frankly' thus handled at the utterance level, while propositional honorifics are handled like expressives such as 'damn'. Performative honorifics are not discussed at all by either McCready (2010) or Gutzmann (2015). I would like to propose that performative honorifics could be classified as isolated mixed UCIs, with a representation as follows:

(75) ⟦tsukare-masu⟧ = λx. get-tired(x) ♦ politely(λx. get-tired(x)) +> the speaker respects the addressee

I believe that as advocated by Potts (2005), another possible analysis could be one as an utterance modifier (e.g. as if 'politely' could be used in the same way as 'frankly'), in which case we would need a more pragmatically informed account, as I will argue below for expressions that seem to interact with utterances. But since in the above analysis the inference of respect is analysed as a conversational implicature, the result would not be very different. In either case, the honorification in question does not seem like something we would be justified to represent in the compositional semantics[60].

Finally, another class that seems to pertain to the category of isolated mixed UCIs is that of verbal pejoratives, which amount to verbal expressions that encode a negative attitude, such as 'to goggle', 'to gawk' and 'to guzzle' in English, 'φαρμακώνω' ('to eat') , 'ξεραίνομαι' ('to sleep') 'ψοφάω' ('to die') in Greek, etc. In Gutzmann (2015: 89-90) such expressions are called 'coloured verbs' and are categorised as functional mixed UCIs.

[60] Another possible analysis specifically for '-masu' is that of a functional expletive UCI, as unlike 'desu', it does not appear independently but as a verb suffix which takes a verb root as its argument and yields a polite verb form. Conveniently, whichever of the two analyses is chosen makes no difference to this thesis' arguments against Gutzmann's (2015) UCI typology because as will be shown later, the category of functional expletive UCIs will not be contested. However even if both analyses of '-masu' are possible, I want to defend the isolated one mentioned just above, for the following reason. As will be shown below, the emotive/expressive attitude of functional expletive UCIs like the Japanese suffix '-yagaru' targets the event expressed by the verb they affix to, whereas '-masu' merely adds a note of politeness to the entire discourse context, without directly targeting either any entity or event as functional expletive UCIs normally do.

The example used there is the German transitive verb 'begaffen' ('to gawk at') and its meaning is given as follows:

(76) ⟦begaffen⟧ = λx.λy. look-at(x)(y) ♦ λx.λy. bad(look-at(x)(y))

Informally, its truth-conditional meaning is that the subject is looking at the object, while its use-conditional meaning is that 'the speaker is annoyed that (or by the way) the subject looks at the object' (Gutzmann 2015: 90). However, this representation is significantly different from that of other functional mixed UCIs, the prime example of which is Japanese honorifics:

(77) ⟦irassharu⟧ = λx. come(x) ♦ λx. be-honourable(x)

In the above representation, the expression is predicative in both dimensions (McCready 2010). This is different from the proposal for 'begaffen' in which the argument is not an entity but the entire content of the truth-conditional tier. Under a 'begaffen'-like representation, the Japanese honorific verb 'irassharu' would be represented as follows:

(78) ⟦irassharu⟧ = λx. come(x) ♦ λx. be-honourable(come(x))

However, this is not the correct representation for 'irassharu', simply because the target of honorification is clearly the subject of the verb, rather than the whole proposition or event encoded by the truth-conditional dimension. This can be demonstrated by the felicity of examples where the event described is regarded as the opposite of honourable by the speaker:

(79) Shachou-ga saisho-no kaigi-ni osoku irasshaimashita-no-ha rippana koui deha-arimasendeshita.
CEO-NOM to first meeting late arrived$_{HON}$ fact honourable conduct be-NEG-PAST
That the CEO arrived$_{HON}$ late at the first meeting was not very honourable conduct.

Therefore it seems that the difference between honorifics and verbal pejoratives is not merely that they are at the opposite places of the spectrum (positive vs. negative), but that the former express an attitude about the individual involved in the action whereas the latter express an attitude about the action itself. As a result, verbal pejoratives look a lot more like their nominal counterparts:

(71) ⟦Kraut⟧ = λx. German(x) ♦ bad (∩German) +> bad(x)

(72) ⟦cur⟧ = λx. dog(x) ♦ bad (∩dog) +> bad(x)

(82) ⟦begaffen⟧ = λx.λy. look-at(x)(y) ♦ λx.λy. bad(look-at(x)(y)) +> bad(x)

Gutzmann & McCready (2016) note the difference between honorifics and verbal pejoratives (in which they subsume the Japanese so-called anti-honorific verbal suffixes '–yagaru' and '–chimau') in that the former can be characterised in terms of social distance, formality etc. whereas the latter seem to be truly emotive.[61] This argument is

[61] That this distinction is true can also be shown by the way the different items are taught to L2 learners; the correct use of Japanese honorifics is usually taught in terms of hierarchy in Japanese

proven correct by contrasting the corresponding representations ('irassharu' vs. 'begaffen') as done above. Therefore, we may conclude that verbal pejoratives should be categorised as isolated mixed UCIs and not functional mixed UCIs (pace Gutzmann 2015).[62]

To summarise, it is argued that so-called isolated mixed UCIs do not form part of the compositional semantics, as their use-conditional meaning does not interact with any semantic component. The expressions identified as isolated mixed UCIs are as follows: slurs (e.g. 'Kraut', 'fag', etc.), register variants (e.g. 'umbilicus'), Japanese performative honorifics ('desu', '-masu'), as well as nominal (e.g. 'cur') and verbal pejoratives (e.g. 'to guzzle').

4.1.2.3.2. Isolated expletive UCIs

These expressions carry meaning only in the non-truth-conditional dimension (−2d) and they do not require an argument either (−f). This category includes all kinds of interjections (e.g. Kaplan's 'ouch' and 'oops', but also 'alas', 'argh', 'hurray', 'yikes', etc.), as well as items such as greetings (e.g. 'hello', 'goodbye', which is mentioned by Kaplan) and we should also add onomatopoeic elements which have not evolved into full-blown lexical items (e.g. 'kerching' or 'da-bum tss', but not items like 'fizzy' or 'to purr').

As Gutzmann (2013: 30) explains, isolated expletive UCIs 'do not interact with the rest of the sentence in any non-trivial way. When an interjection like 'ouch' is used together with a sentence, the truth-conditional content of the sentence is not affected by the presence of that interjection, whereas its use-conditional content is given by the emotion conveyed by 'ouch' itself.' But as noted above, given that these expressions do not take any argument but are not used as arguments themselves either, it is not clear why we should represent them in semantic composition. In fact, exactly because of their lack of interaction with anything else, they can only appear in peripheral positions (Gutzmann 2013: 6):

(83) Oops, I've lost my keys.
(84) *I've lost my oops keys.

They do not appear mid-sentence unless in quotation, which is of course a case of mere mention rather than use:

(85) You've made such a terrible mess and all you had to say was 'oops'.

However, one could say that there are cases in which isolated expletive UCIs do in fact interact with other components, such as the sentence-final expression 'man' (McCready 2010: 35):

---

society, whereas pretty much in any language typical expressives such as cursewords are usually explained in terms of how the speaker feels when using them.

[62] However, pace Gutzmann & McCready (2016), the Japanese verbal suffixes '-chimau' and '-yagaru' (as well as '-shimau' and '-chau') are not verbal pejoratives like 'begaffen' which would make them isolated mixed UCIs, but functional expletive UCIs, exactly because they are not verbs but suffixes. They will be discussed further below in the category of functional expletive UCIs.

(86) Ouch, man!
(87) Man, ouch!

In these cases 'ouch' appears to be modified by 'man', so this would be an argument that isolated expletive UCIs can interact compositionally with other expressions in some cases. But as noted before for isolated expletive UCIs in general, in both the examples 'ouch' appears in a peripheral position: either in the beginning, or in the end of the utterance. Because of this, it is possible to say that no modification is taking place, and this is merely a speech act comprising two utterances which do not interact compositionally but just appear consecutively, such as the following:

(88) Oh, hello!
(89) Hey, good morning!

So if the particle 'man' constitutes indeed a separate part, something like an utterance of its own, we expect it to be able to appear after almost any kind of (intentional) verbalisation, which is in fact the case:

(90) Hurray, man![63]
(91) Argh, man.[64]
(92) Shh, man.[65]
(93) Pff, man.[66]
(94) Boo hoo, man.[67]

McCready (2010) examines and rejects the idea that the relation in such cases is mere adjacency on the basis that 'Ouch, man' is felicitous in a scenario where the speaker has suffered a slight misfortune caused by the interlocutor (e.g. he stepped on her foot) but not in a scenario where the interlocutor has nothing to do with it (e.g. she tripped and fell on her own), while 'Ouch' alone is felicitous in both. For McCready, the contrast between these two scenarios makes it 'clear that 'man' is in fact doing something to the meaning of 'ouch' […] and so some kind of composition is at work' (2010: 36). However, I am not sure the argument is convincing enough, as mere adjacency is also able to explain the infelicity. For example, let's contrast two scenarios:

The speaker meets a baby girl in the park:
(95) Hello!
(96) Hello, princess!

The speaker meets an old and distinguished male professor in a book presentation:
(97) Hello!
(98) # Hello, princess!

There is clearly a felicity contrast, but it is not clear how this constitutes an argument for compositional interaction, as the greeting 'hello' on its own is felicitous in both scenarios and the infelicity is only caused by the addition of the term of address in the second case.

[63] https://www.reddit.com/r/Voltron/comments/5pk3k2/my_attempt_at_the_classic_pose/
[64] https://twitter.com/MannyNorte/status/19364186425397248
[65] https://archive.org/details/jamendo-033919
[66] https://social.shorthand.com/Village_DORP/j2Xoc1RYnn/solartaxi
[67] http://gawker.com/dave-grohl-accidentally-berates-grieving-son-for-crying-1725145495

As a result, adjacency is able to account for the infelicity and there is no sufficient evidence that composition is at play.

If the above discussion is on the right track, then 'man' is shown to be an isolated expletive UCI, just like 'ouch' and others, and not a functional shunting UCI as claimed by both McCready (2010: 34) and Gutzmann (2015: 84)[68]. The reason it was claimed to be a functional shunting UCI was because of its alleged interaction with a proposition p to display that 'the speaker has some kind of emotional reaction toward p (that it is good or bad)' (McCready 2010: 34). Under this analysis of 'man' as a functional shunting UCI, the idea is that it takes a proposition as an argument and consumes it entirely, leaving nothing at the truth-conditional dimension. As McCready observes, 'this analysis disallows the assertion of p itself, as desired' (2010: 34). In fact the reason this is 'desired' because of the following example:

(99) What's the weather like outside?
Man!

McCready claims that in this example the interlocutor can only infer that the speaker finds the weather extreme in some way, and the proposition modified by 'man' (which would clarify whether it's too hot or too cold) is not recoverable in any way, therefore it is not asserted.[69] The alternative analysis would be that the speaker only communicated that he is in an excited state, and that this is somehow related to the weather would follow merely from inference. Although McCready says that there seems to be no empirical way to choose between these alternatives, she finds that both of them support 'man' to be of a shunting type. However, it is not clear why this is so; if we take the second alternative that the speaker merely used an interjection to express his excited state (which was 'man' in this case, but could also have been 'ouch'), then there would be no argument (−functional), which would make the expression a typical isolated expletive UCI. As shown with the examples above, this analysis is even compatible with cases where 'man' appears alongside other expressions as in above examples, where co-occurrence was shown to be mere adjacency.

Therefore it seems that so-called isolated expletive UCIs are not really modified even though it might seem so, which strengthens the argument against representing them in the compositional semantics. The same goes for examples such as the following (from Predelli 2013: 75):

(100) Tomorrow, alas, it will rain.

---

[68] Although in earlier work Gutzmann (2013) classifies 'man' in a different category. Specifically he says that 'an example of a functional mixed UCI is given by the interjectional use of English *man* when combined with a sentence that contains a gradable adjective. On the truth-conditional layer, *man* intensifies the property predicated in the propositional content, i.e., *Man, it's hot!* means that it is very hot (McCready, 2009). On the use-conditional layer, man expresses that the speaker is somehow affected by that high degree of heat.' (2013: 31).

[69] Actually McCready admits that this criterion is not always reliable: 'the question of how extensively we should take particle meanings to be analyzable in terms of shunting types is left for another occasion; it turns on the empirical question of whether or not the propositional content of sentences modified by particles can serve as answers to questions. In many cases it is clear that they can, in others, perhaps not.' (2010: 34).

Although here 'alas' appears mid-sentence, it literally interjects: its position is not related to taking any argument, as it can appear in any position between different phrases just not inside them, which confirms its peripheral character:

(101) Alas, tomorrow it will rain.
(102) Tomorrow it will rain, alas.

(103)The rain, alas, has ruined the party.
(104) *The, alas, rain has ruined the party.

Of course, there is some kind of interaction taking place between 'alas' and the rest, as well as 'man' and what precedes or follows it. But what we should note here is what the interjection or particle interacts with is not a proposition, but an utterance. The criticism by Amaral, Roberts & Smith (2008) against Potts' (2005) treatment of adverbials like 'frankly' in an upper layer of meaning language ($\mathcal{L}_U$) is illuminating in this case:

> 'We might say that $\mathcal{L}_U$ simply introduces yet another dimension of meaning, that in which utterances, as opposed to propositions or other truth conditional entities, may be modified. But in what sense is this dimension merely a reflection of compositional semantics? What is modified is not the sentence or its compositionally derived parsetree in the meaning language, but the utterance, and the latter is a fundamentally pragmatic entity, as everyone from Bar-Hillel (1971) on has agreed. […]Their treatment arguably requires a theory of the relationship between compositional semantic interpretation and the pragmatics of speech acts.' (2008: 727, 729).

That is to say, although there can be interaction between so-called isolated expletive UCIs and other units of meaning, since these are essentially utterances, this interaction cannot be captured merely in terms of compositional semantics, even if that would be hybrid semantics; it urgently needs to be complemented by pragmatic considerations of how discourse works, or as Amaral, Roberts & Smith (2008) note, be accounted for by a dynamic theory, as such approaches 'are designed to model the interaction of syntactically-driven compositional semantics with contextually-enriched pragmatics in the course of interpretation.' (2008: 745).

4.1.2.4. Elucidating functional UCIs

4.1.2.4.1. Functional shunting UCIs

This category of UCIs is defined as carrying meaning only in the use-conditional dimension (−2d) and as requiring an argument (+f). However, their functionality has the following distinctive characteristic: resource-sensitivity. In other words, they entirely consume their truth-conditional argument instead of returning it, or else 'shunt' it in the use-conditional dimension. Apart from 'man' (McCready 2010) which was discussed above, other expressions said to be in this category are the Japanese adverbial 'yokumo' (McCready 2010), as well as exclamatives or unexpectedness intonation (Gutzmann 2015), e.g. 'How crazy Eleni is!'.

Following the discussion just above, it does not seem that 'man' can be classified as a functional shunting UCI merely because the claim that an utterance containing 'man' cannot be used as an answer to a question, while in the case of an utterance consisting only of 'man' we don't have to posit any shunting but just treat 'man' as an isolated expletive UCI. So, to repeat, 'man' has not been convincingly shown to be a functional shunting UCI and the evidence points more towards classifying it as an isolated expletive UCI.

However, the situation is different with 'yokumo' as well as with exclamatives; as shown by McCready (2010) and Gutzmann (2015) respectively, these expressions cannot be used to make an assertion that would satisfy a question, i.e. they do not seem to contribute novel information. Does this mean that we can justify a category of functional shunting UCIs based on these expressions?

I believe that the answer is no, and the reason for that is because as noted above about the interaction of isolated expletive UCIs such as 'alas', the interaction in question cannot be adequately captured by a solely semantic account. It was noted above that what 'alas' seems to interact with is an utterance rather than a sentence. In this sense, it could perhaps be argued that this interaction is not mere adjacency but modification, but it is crucial to note that this would amount to utterance modification, rather than of a sentence or proposition. It seems to me that the discussion above and the comments by Amaral, Roberts & Smith (2008) about the urgent need for a pragmatically informed or a dynamic account for expressions that interact with utterances rather than sentences conveniently apply in the discussion of functional shunting UCIs as well.[70] The fact that one can easily devise constructions that Gutzmann (2015) would classify as functional shunting UCIs seems to argue against a strictly semantic account of such cases. One example that comes to mind is the Greek expression 'αν είναι δυνατόν να…' plus a finite construction, which roughly means something like 'I can't believe that…'. Similarly to 'yokumo', utterances formed by this Greek expression and a sentence cannot be used as assertions:

Who did Austin marry?

(105) #Yokumo Dallas to kekkon shita na!
'He did an amazingly stupid and shocking thing by marrying Dallas! (McCready 2010: 39)

(106) # Αν είναι δυνατόν να παντρεύτηκε τον Ντάλας!
'I can't believe he married Dallas!'

In fact, one can easily think of other such constructions, such as the Greek phrases 'και τι κατάλαβε που…' (roughly 's/he didn't achieve anything noteworthy by…'), 'άκου εκεί να…' (roughly 'it's unbelievable that…'), etc. Therefore it seems that the example above shows that instead of proposing a functional shunting UCI-like semantic profile for such expressions, we should merely acknowledge that certain expressions interact with units higher than semantic content, such as 'the speech act performed by the utterance of the

[70] In fact, the first formal account of 'yokumo' that was given by McCready (2004) herself was in terms of dynamic semantics.

argument constituent' or 'the speech situation in which the argument takes place' (Amaral, Roberts & Smith 2008: 743). As suggested by Amaral, Roberts & Smith (2008), a plausible account of cases such as the above example, where the speaker is expressing some kind of disapproval, is expected to make connections to speech act theory. For example, a category of Searle's (1976) expressive illocutionary acts that might be fitting for this example is that of 'deploring'. Although this is merely a suggestion that a possible theory of these expressions could be one linked to speech act theory, what seems to be sure is that the quest for an exclusively semantic account of such phenomena is in vain.

It is therefore argued that so-called functional shunting UCIs do not constitute a semantic category and thus should not figure in a typology of compositionally relevant use-conditional or expressive meanings.[71]

4.1.2.4.2. Functional mixed UCIs

This category includes expressions which have meaning in both dimensions (+2d) and also require an argument (+f). The prime example in this category is that of Japanese honorifics, specifically those that are argument-oriented (called 'propositional' by Harada 1976). As McCready (2010) explains, such expressions are predicative in both dimensions, and they can be represented as follows (McCready 2010, Gutzmann 2015):

(77) $[\![\text{irassharu}]\!] = \lambda x.\ \text{come}(x) \blacklozenge \lambda x.\ \text{be-honourable}(x)$

However, there is strong evidence that Japanese honorifics are actually presuppositional in nature. First of all, as noted at the start of the chapter, an account of honorifics in terms of two different dimensions stumbles on the presuppositional (when treated by a two-dimensional framework) woe known as the Binding Problem[72]:

(65) Dareka irasshaimashita.
someone $\text{arrived}_{\text{HON}}$
'Someone has arrived.'
At-issue: ∃x arrive(x)
CI: honourable(x)

As explained before, although Potts' (2005) initial project of expressives as CIs was immune to the Binding Problem, given the predicated total independence, as soon as truly mixed meanings such as Japanese honorifics are included in the data, the old problem simply resurfaces. At a first glance, the fact that these meanings seem to be

[71] However, one characteristic of 'yokumo' that could testify for a semantic treatment is that it conveys a strong factive presupposition, in the sense that it feels awkward to utter a sentence with 'yokumo' when the factivity of the complement has not been established (Yasu Sudo, p.c.). Nevertheless, as this also holds for the Greek expressions mentioned above as corresponding to 'yokumo', it seems quite likely that the factivity is coming from the exclamative force of the utterances involved, rather than from any specific expression.

[72] Potts & Kawahara (2004) contend that honorifics cannot appear in quantificational constructions, but as shall be explained below, the apparent incompatibility is in fact due to the unnaturalness of their examples.

susceptible from a presupposition-targeting malady is an indication that they could be presuppositional in nature.[73]

Indications aside, however, a presuppositional analysis of Japanese honorifics has already been proposed. According to Sudo (2012), unlike bona fide expressives Japanese honorifics are neither scopeless, as they can be embedded under operators, nor independent, since they can be quantified in. Let's see how Sudo (2012) shows this with data:

(107) Sono ronbun-no chousha-ga yuumei-na kyouju nara, LSA-ni irassharu hazu da.
the article-GEN author-NOM famous professor if, LSA to come$_{HON}$ definitely
'If the author of that article is a famous professor, he will definitely come$_{HON}$ to LSA.'

(108) Demo gakusei-nara, ko-nai kamoshirenai.
but student if, come-NEG come
'But if he is a student, he might not come.'

In the above example, the speaker is not referring to any specific person he has an attitude towards, but embedding honorification under a conditional operator, essentially entertaining the possibility that the author of the article in question might be a professor and saying that under this condition, they are worthy of respect. As we can see in the immediately following sentence, under a different condition in which that person is a student, no display of respect is needed as shown by the neutral counterpart of the verb. But if honorification is accounted as expressive or else use-conditional meaning, then it should be on a different dimension and as such it should be unaffected by operators such as conditionalisation:

(109) # If Conner is a bastard, then that bastard Conner was promoted. (Potts 2005: 157)

(110) # If that bastard Kresge arrives on time, he should be fired for being so mean. (Potts 2007: 170)

As unlike the above, the example with the Japanese honorific verb 'irassharu' is acceptable, it seems that it pairs up more with presuppositions than expressives:

(111) If Aretousa ever smoked, then she stopped smoking.

In the above example, the presupposition that Aretousa used to smoke is filtered by the conditional construction and does not project out, just like in the example with 'irassharu' the inference that speaker honours the author of the article does not project out but is contained by the conditional.

However, it should be noted here that Sudo (2012) mentions that the following example is infelicitous, which could potentially be seen as a complication:

(112) #Sono gakusei-ga yuumei-na kyoujyuu dat-tara, LSA-ni irassharou darou.
that student-NOM famous professor be if, LSA to come$_{HON}$ probably
'If that student were a famous professor, he would come$_{HON}$ to LSA.'

[73] To be clear, this is not a problem for presuppositions, but a problem for a two-dimensional theory of presuppositions.

However, he observes that this is due to a preference of honorifics for de re readings, which can only be circumvented under the right pragmatic conditions; these conditions are not met in this example as the speaker is acquainted with the individual talked about, but seem to be in the example where the speaker does not know who that individual might be, as in (107). Yanovich (2010) has observed the same preference for de re in gender features, which have been shown to be presuppositional in nature (Sudo 2012):

(113) (We know that Jesse is a man.) #If Jesse$_i$ was a woman, I would marry her$_i$.
(114) (We don't know if Jesse is a man or woman) If Jesse$_i$ is a woman, I will marry her$_i$.

As a result, examples such as (112) do not constitute counterexamples. I will come back to such cases in chapter 5 and put forward an explanation of why they prefer de re readings.

Another example against the scopelessness of honorifics provided by Sudo (2012: 72) is one where the honorific is under the scope of an attitude verb:

(115) Taro-wa [watashi-no shiranai dareka-ga nani-o meshiagat-ta ka] kiitekita.
Taro-TOP [I-GEN not-know someone-NOM what-ACC eat$_{HON}$-PAST Q] asked me
'Taro asked me about someone I do not know what he ate$_{HON}$'.

In the above example, the speaker is not acquainted with the person Taro inquired about, thus the honorification can only be attributed to Taro than the speaker. Yet this example is felicitous, contrary to what is predicted by a theory in which honorifics are in a different dimension altogether. Again, this behaviour is rather reminiscent of presuppositions:

(116) Luca believes that Fede has realised that cats can grasp complex mathematical concepts.

Moreover, Sudo (2012) has presented data that have shown that honorifics can be bound by quantifiers, which should not be possible for expressives given their alleged independence:

(117) Onnano hito-ga hitori/zenin irasshatta.
female person-NOM one/all come$_{HON}$-PAST
'One lady/All of the ladies came$_{HON}$.' (the speaker respects her/them)

(118) Dono senshu-mo irassharanakatta.
which player-NPI come$_{HON}$-NEG-PAST
'None of the players came$_{HON}$.' (the speaker respects all players)

(119) Kankyaku-no hotondo-ga o-warai-ni-natta.
audience-GEN most-NOM laugh$_{HON}$-PAST
'Most of the audience laughed$_{HON}$.' (the speaker respects them)

The claim that Japanese honorifics specifically cannot be quantificationally bound was made in Potts & Kawahara (2004: 4)[74]:

[74] Remember that we also mentioned in the beginning of this chapter that Potts (2005) discussed an example where an expressive was in a quantificational construction but dismissed an

(120) ??Hotondo-no kyoujyu-ga [*pro* sono kurasu-wo o-shie-ni-natta to] omotteiru.
most facutly-NOM [*pro* that class-ACC teach$_{HON}$-PAST C] believe
'Most of the professors believe that they taught$_{HON}$ that class'.

However, Sudo (2012) notes that the infelicity of this sentence is not due to the impossibility of the honorific to be quantificationally bound, but to the mismatch between the honorific form of 'teach' and the neutral form of 'believe'. According to Sudo (2012), if we use the honorific form for the second verb as well, the sentence becomes perfectly acceptable:

(121) Hotondo-no kyoujyu-ga [*pro* sono kurasu-wo o-shie-ni-natta to] omotteirassharu.
most facutly-NOM [*pro* that class-ACC teach$_{HON}$-PAST C] believe$_{HON}$
'Most of the professors believe$_{HON}$ that they taught$_{HON}$ that class'.

Given the above, it seems that the claim that honorifics cannot be quantificationally bound cannot be sustained. Again, the possibility for quantificational binding is reminiscent of presuppositions:

(122) Most of the band members have stopped using hallucinogenic drugs before going on stage.

Based on the above data and discussion, Japanese honorifics do not confirm core characteristics of expressives or else UCIs, but behave much more akin to standard presuppositions. As Japanese (argument-oriented) honorifics are not just the prime but actually the sole example of functional mixed UCIs provided by Gutzmann (2015), it is proposed that there is no category of functional mixed UCIs at all.

4.1.2.4.3. Functional expletive UCIs

The category of functional expletive UCIs includes items which contribute meaning only in one dimension (−2d) and also require an argument (+f). Here we find the primordial expressives introduced by Potts (2005), such as 'damn'[75], 'bastard', 'fucking', etc. As these expressions take another expression as their argument, it cannot be doubted that they should be represented in the compositional process. However, one should note that some of these items in which they appear to also have content in the truth-conditional dimension (thus to be +2d), as they can appear in predicative positions:

(123) That bastard headmaster has expelled me again.
(124) The headmaster is a bastard.

(125) That jerk Junior is bringing moribund country mice again.
(126) Junior is a jerk.

---

explanation in terms of binding as inappropriate for the data (see footnote 37). However as the current discussion is advocating that honorifics are not expressives or UCIs (if expressives/UCIs are understood in terms of being in a different dimension), we will stick to honorifics and not make use of that example in the current discussion.

[75] We should note that 'damn' can also be used on its own, as an interjection. Gutzmann calls such UCIs that belong to more than one category as 'multifunctional' (2015: 32).

(127) That bitch Katia is treating everyone badly.
(128) Katia is a bitch.

However, not all of them can:

(129) We don't know how to solve this fucking problem.
(130) *This problem is fucking.

(131) As soon as I heard this album I just knew it was a damn masterpiece.
(132) *This masterpiece is damn.

Zimmermann (2007) notes the possibility of predicative, thus descriptive-like uses and sees two ways to account for it: either a type-shift operation, or lexical ambiguity. However, there are problems with both views. First, if we go for the type-shift view, we will have trouble explaining why this is not possible for all expressives, and any explanation of selective type-shifting might end up being no less idiosyncratic than lexical ambiguity. If we opt for lexical ambiguity instead, Zimmerman notes that this will fail to account for the fact that whether an expression functions as expressive or descriptive seems to depend on its syntactic position. However, it is possible that the objections to the second view are not as fatal; if we adopt a lexical ambiguity view, we can interpret the dependence on syntactic position as a consequence rather than a cause of the expression's profile in each case (expressive or descriptive). According to such a view, expressives can only appear in attributive positions, and those that seem to appear in predicative ones are actually homophonous descriptive expressions. In this case perhaps one worry will be the intuition that the discourse effect is the same, and descriptive-like uses of these expressions feel as emotionally charged as their expressive counterparts. However, it is possible that this is due to pragmatic factors. Let's see the examples below, one from Greek and one from English:

(133) Αυτή η πουτάνα η Κάτια φέρεται άσχημα σε όλους.
this prostitute the Katia treats badly to everyone
'This bitch Katia is treating everyone badly.'
Katia is treating everyone badly. • I don't like Katia.
('Katia is a prostitute' NOT entailed)

(134) Η Κάτια είναι πουτάνα.
the Katia is prostitute
'Katia is a bitch.' OR 'Katia is a prostitute.'

(135) That bastard Kresge is famous.
Kresge is famous. • I don't like Kresge.
('Kresge was born out of wedlock' NOT entailed)

(136) Kresge is a bastard.
I don't like Kresge. OR Kresge was born out of wedlock.

I think that what these examples show is that if we adopt a lexical ambiguity view then we could say that it is possible for a descriptive item (usually one related to a taboo subject as in the examples) to give rise to an expressive counterpart, and in this case they could both co-exist in the lexicon. When this happens, although predicative uses involve the descriptive rather than the expressive item, we can expect that even the

descriptive uses will be endowed with an emotive attitude, but this will be because of the taboo subject. In these cases the predicative use can be either literal (prostitution, being born out of wedlock) or figurative, and in the figurative cases the offence will be similar to that of the expressive counterparts but it should be attributed to a pragmatic inference.

Moreover, as the items labelled as functional expletive take an argument, it is clear that the attitude expressed is targeted towards their argument rather than the general group the entity talked about belongs to. This is in contrast with so-called isolated mixed UCIs which were discussed previously, as those targeted the entire group and the offence to the entity was merely inferred. Let's show the contrast with an example:

(137) I don't want to live in a honky neighbourhood.

The speaker is expressing a racist attitude against white people; the reason for the negative attitude is clearly related to race.

(138) I don't want to live in a damn white neighbourhood.

The speaker is expressing a negative attitude against a neighbourhood of predominantly white residents, or against the idea of him living there. It is not clear whether the reason for the negative attitude is related to race or anything else not inherently related to race; for example, it could be that the speaker is upset because he thinks that there are certain services he will have trouble finding in a white neighbourhood, such as dreadlock salons.

Therefore, it seems that there is a stark difference between swearwords and slurs, although as noted in the discussion of slurs it is possible for expressions to migrate from the first category to the second, depending on shifts in social attitudes.

Furthermore, I would like to argue against Gutzmann & McCready (2016), who seem to classify Japanese so-called anti-honorifics as functional mixed UCIs, and propose that they should be classified as functional expletive UCIs instead, essentially revindicating the original proposal by Potts & Kawahara (2004) and Potts (2005) who treated them not differently than 'damn'. However, I also argue that the term 'anti-honorifics' is unfortunate as they are inherently different from honorifics.[76] According to Potts & Kawahara (2004), the Japanese expressions '-chimau' and '-yagaru' are classified as anti-honorifics, which signal that the speaker has a negative attitude towards the proposition of the sentence in which it appears. Here is an example from Potts & Kawahara (2004: 2):

(139) John-wa [Mary-ga nesugoshi-chimat-ta] koto-wo shitteiru.
John-TOP Mary-NOM oversleep$_{\text{ANTIHON}}$-PAST fact know
'John knows that Mary overslept.' • 'It sucks that Mary overslept.'

[76] To be clear, when I use the term 'honorifics' in this section I am specifically referring to Harada's (1976) propositional honorifics, which are argument-oriented, such as 'irassharu' ('come$_{\text{HON}}$'), and not to performative honorifics such as '-masu'. This is done for brevity and because these are what most people refer to as Japanese honorifics, but I hope it won't be a confusing abbreviation.

However, Gutzmann (2015) as well as Gutzmann & McCready (2016) treat anti-honorifics akin to verbal pejoratives such as 'to guzzle', which as mentioned before they analysed as functional mixed UCIs. They also differentiate between '-chimau' and '-yagaru', saying that the former targets the proposition while the latter targets the subject (Gutzmann & McCready 2016: 12):

(140) Taro-ga Jiro-wo nagut-chimat-ta.
Taro-NOM Jiro-ACC $\text{hit}_{\text{ANTIHON}}$-PAST
'Taro punched Jiro.' • 'That was bad.'

(141) Sam-ga warai-yagat-ta.
Sam-NOM $\text{laugh}_{\text{ANTIHON}}$-PAST
'Sam laughed.' • 'I view Sam negatively.'

However, the view that '-yagaru' targets the subject cannot be correct, as its use is felicitous in sentences where there is no individual (example offered by Yasu Sudo, p.c., the expressive component is given along the lines of Gutzmann & McCready 2016):

(142) Kuraku natte ki-yagatta.
dark-ADV $\text{become}_{\text{ANTIHON}}$-PAST
'It got dark.' • 'That was bad'.

In fact, the example above can be recreated not just with '-chimau', but also its more neutral/formal counterpart '-shimau', as well as its more colloquial one '-chau'. Although all the below have the same truth-conditional content, their difference can be explained in terms of speech level:

(143) Kuraku natte kite-shimatta.
dark-ADV $\text{become}_{\text{ANTIHON}}$-PAST
'It got dark.' • 'That was unfortunate.'

(144) Kuraku natte ki-chimatta.
dark-ADV $\text{become}_{\text{ANTIHON}}$-PAST
'It got dark.' • 'That sucked'.

(145) Kuraku natte ki-chatta.
dark-ADV $\text{become}_{\text{ANTIHON}}$-PAST
'It got dark.' • 'That sucked big time'.

(146) Kuraku natte ki-yagatta.
dark-ADV $\text{become}_{\text{ANTIHON}}$-PAST
'It got dark.' • 'That fucking sucked'.

I have included the original example with '-yagaru' to show that all these expressions can be placed in a continuum as far as the register or level of formality is concerned. The translations are not intended to be exact, but just to show that different shades are possible.[77] As in all these examples there is no individual involved, it is obvious that the expressive attitude is towards the event described, falsifying the claim of Gutzmann &

[77] Apart from the possibility that these specific renditions are not faithful enough, following McCready (2014) perhaps any translation will in a way fall short of the original expressive meaning due to the 'genuine incommensurability' of expressives of different languages.

McCready (2016) that '-yagaru' targets the subject. The examples show us that all these expressions are used in a similar way, and they only differ among them in terms of intensity, which can be attributed to register. As stipulated before that register connotations are not part of the compositional process, the only compositionally relevant expressive component of these verbal endings is that the speaker is unhappy about the event described by the verb in which they are attached. Since we saw that this is possible with events which do not involve individuals, this means that when there is an individual involved, any inference that the speaker has a negative attitude towards them should be explained in pragmatic terms, i.e. attributed to the low register of the expression chosen:

(147) Sam-ga warat-chimat-ta.
Sam-NOM laugh$_{\text{ANTIHON}}$-PAST
'Sam laughed.' • 'It sucked that Sam laughed.' +/> 'I view Sam negatively.'

(148) Sam-ga warai-yagat-ta.
Sam-NOM laugh$_{\text{ANTIHON}}$-PAST
'Sam laughed.' • 'It fucking sucked that Sam laughed.' +> 'I view Sam very negatively.'

After we have identified the Japanese verbal endings '-shimau', '-chimau', '-chau' and '-yagaru' as functional expletive UCIs that express an attitude about the event, it should be clear that the term 'anti-honorifics' is not suitable for two reasons. First of all, as Gutzmann & McCready (2016) say, they do not have the opposite effects in the same domain but rather pertain to different domains altogether: honorifics having to do with formality, social hierarchy, etc., while these expressions seem more emotive than anything. But second, we see that the difference in what they take as argument (honorifics target an entity, these expressions an event) means that any attitude these expressions may convey about entities is due to pragmatic rather than semantic reasons. Moreover, since these expressions are functional expletive UCIs, they are also distinct from verbal pejoratives (e.g. 'to guzzle'), which were shown to be isolated mixed UCIs; this is pace Gutzmann & McCready (2016) who group '-chimau' and '-yagaru' together with verbal pejoratives (which are functional mixed UCIs to them). All in all, what is clear is that the term 'anti-honorific' is not fitting for these expressions used by Gutzmann & McCready to justify this category, since it is not the case that they function similarly to honorifics (i.e. taking an entity as an argument) and achieve the opposite result (i.e. insult/depreciation instead of honorification).

4.1.2.5. Verdict on the typology of UCIs

After analysing the five different categories of UCIs proposed by Gutzmann (2013, 2015), it was shown that isolated UCIs (which subdivide in mixed and expletive) should not figure in the typology if the domain of interest is compositional semantics. Moreover, it was argued that due to their interaction with utterances which are pragmatic entities, functional shunting UCIs cannot be adequately accounted for in strictly semantic terms (or at least static semantic terms). Furthermore, it was exemplified that functional mixed UCIs behave more like standard presuppositions than expressives or UCIs. Eventually, this means that from Gutzmann's (2013, 2015) five different categories of

UCIs, only functional expletive UCIs seem to instantiate the idea of expressive meaning as a different kind of meaning from existing categories (at-issue, presuppositions). Interestingly, the properties of this category (−2d, +f) are exactly those originally advocated by Potts (2005, 2007) before the extensions proposed by McCready (2010) and Gutzmann (2015), which in a way feels that we have taken a long detour to be back to square one. Perhaps this is not so surprising if one considers the first criticisms of these extensions noted at the beginning of this chapter: it was observed that McCready's (2010) amendments intended to cover mixed expressions essentially breached independence, a core property of expressives as pertaining to a different category of meaning altogether as proposed by Potts (2005)[78], as well as that Gutzmann's (2015) radical revision of the conceptual background and the resulting reconstruction of the entire formal framework was essentially unconstrained, in a way that almost anything could be classified and treated as a UCI.

	f	2d	r-s	e.g.	Expressive meaning (non-presuppositional & compositionally relevant)
isolated expletive	−	−		'ouch'	×
isolated mixed	−	+		'Kraut'	×
functional expletive	+	−	−	'damn'	✓
functional shunting	+	−	+	'yokumo'	×
functional mixed	+	+		'irassharu'	×

Table 4.2. Types of UCIs as expressive meaning

f: functionality; whether the expression needs an argument
2d: 2-dimensionality; whether the expression has meaning in both dimensions (mixed)
r-s: resource-sensitivity; whether the expression fully consumes its argument

4.2. Presuppositional alternatives

As mentioned previously, due to their projective behaviour expressives look a lot like presuppositions, which made a lot of authors support the idea that they are indeed a kind of presupposition (Macià 2002, 2006, Sauerland 2007, Lasersohn 2007, Schlenker 2003, 2007). The most concrete proposal from the presuppositional field comes from Schlenker (2007) and is presented below.

4.2.1. Expressives as indexical presuppositions: Schlenker (2007)

In a brief commentary to Potts (2007), Schlenker (2007) offers an elaborate sketch of how expressives could be accounted for in the existing ontology of meaning as 'expressive presuppositions'. The gist of the proposal is that 'expressives are lexical

[78] Hence the conversion of the Binding Problem from a virtue back to a deadlock, as was noted.

items that carry a presupposition of a particular sort, namely one which is indexical (it is evaluated with respect to a context)[79], attitudinal (it predicates something of the mental state of the agent of that context), and sometimes shiftable (the context of evaluation need not be the context of the actual utterance).' (2007: 237). Based on these assumptions, the expressive presuppositions of the racial slur 'honky' and the French familiar pronoun 'tu' are represented as follows:

(148) ⟦honky⟧(c)(w)≠ # iff[80] the agent of c believes in the world of c that white people are despicable. If ≠ #, ⟦honky⟧(c)(w) = ⟦white⟧(c)(w)

(149) ⟦tu⟧(c)(w)≠ # iff the agent of c believes in the world of c that he stands in a familiar relation to the addressee of c. If ≠ #, ⟦tu⟧(c)(w) = the addressee of c

Schlenker's (2007) claim is that the distinctive characteristics of expressive meaning as proposed by Potts (2007) follow naturally from the properties of indexicality and attitudinality advocated by his approach. First, independence follows from presuppositionhood, as it is a fact for all presuppositions that are separate from regular at-issue meaning. Next, nondisplaceability directly follows from indexicality, while perspective dependence can be derived by an account of shifted indexicality, which is not ad hoc but has been independently motivated to account for the cross-linguistic variation of indexicals (Schlenker 2003). Then, immediacy follows from the combination of indexicality and attitudinality, which are responsible for the performative character of expressives and render them 'self-fulfulling presuppositions', as will be explained below. The properties of descriptive ineffability and repeatability are not accounted for, but Schlenker raises reasonable doubts on their validity. With regards to descriptive ineffability, he notes that it could be due to the mismatch between presupposition and assertion, but also alludes to Geurts (2007), who notes that descriptive ineffability is a property of many lexical items, such as colour terms. As for repeatability, he notes that when an expressive is used repeatedly (e.g. 'I left my damn keys in my damn car', Potts 2007), it modifies a different constituent each time, hence the lack of redundancy.

Elaborating on the notion of self-fulfilling presuppositions, Schlenker (2007) observes that this could account for the anti-backgrounding character of expressives, which is significantly different from standard presuppositions. First, he reminds us that informative presuppositions are a phenomenon that has already been noticed independently from expressive meaning (Stalnaker 2002, von Fintel 2006). To show how this works, let's assume that in the context of each one of the examples below, there is only one president who has the characteristic mentioned (examples from Schlenker 2007: 240):

(150) #The brown-haired/tall president will take us out of this quagmire.
(151) The stupid/fantastic president will take us out of this quagmire.

[79] It should be noted that indexicality was proposed as a means to account for expressive meaning also by Sauerland (2007), Lasersohn (2007) as well as Amaral, Roberts and Smith (2008).
[80] Note that the use of 'iff' here is not problematic as it was for Gutzmann's (2015) schema as explained above, exactly because Schlenker's (2007) representation is merely presuppositional so still within the realm of truth conditions, whereas Gutzmann's (2015) was advocated as use-conditional.

Based on what we know about presuppositions, the definite descriptions (which include adjectives) trigger the presupposition that there exists exactly one individual who has that characteristic. But (150) is infelicitous because we already know that there is only a single individual who is a president and brown-haired/tall, which means that the adjectives are not doing any semantic work, flouting thus the Gricean principle of manner. However (151) is acceptable, because the adjectives do have a semantic mission: to provide information on the speaker's attitude towards the president. This is precisely an instance of an informative presupposition.

Schlenker observes that there are two ways of explaining why expressives should be analysed as informative presuppositions. One explanation would be that accommodation (Lewis 1979) is easy with presuppositions about the speaker's attitudes, given that a one has authority over one's mental states. But since accommodation is usually seen as a repair strategy and one might wish to avoid any idea of repair, another explanation is one based on Stalnaker's (2002) proposal that if it is common belief that the speaker believes that something is common belief, and if the interlocutors also believe that same thing, then it becomes common belief indeed. In the discussion about presuppositions which express attitudes, if the speaker believes that s/he has the attitude conveyed by the expressive and also that the interlocutors know that s/he does, as soon as the addressee hears the speaker utter the expressive presupposition s/he will come to believe it as it predicates something of the speaker's mental states (by being indexical and attitudinal) and the addressee has no reason or means to contest it, even though it was not previously in the common ground. In this way, expressive presuppositions are systematically informative, but also 'self-fulfilling', in the sense that mere signalling from the speaker that s/he believes to be harbouring the attitude encoded by the expressive presupposition (and that the interlocutors also believe it) suffices for its acceptance in the common ground.

Summarising, Schlenker (2007) provides convincing evidence that expressives constitute a kind of presupposition that is indexical and attitudinal. This move is theoretically parsimonious in that no new category of meaning is introduced, while the introduction of a subcategory of presuppositions which are indexical is not radical and is perhaps even expected given that we know that there is at-issue meaning which is indexical and there are no theoretical reasons to preclude indexicality at the presuppositional dimension (Sauerland 2007). Overall, the presuppositional analysis of expressives seems theoretically appealing as well as convincing, however Schlenker (2016) himself mentions the following caveats. First, there is a considerable difference of pragmatic effect between ordinary and so-called expressive presuppositions (2016: 726):

(152) Everybody knows that I hate Caucasians. Are you one?

(153) Are you a honky?

As it should be evident to the reader, even though both examples carry the presupposition that the speaker is prejudiced against Caucasians, the level of offence of (153) is much greater. The second significant difference is that unlike ordinary

presuppositions, expressives display a more unconstrained projective behaviour, which can be seen in filtering environments (Schlenker 2016: 726):

(154) Nobody stole my car, or it was John who did.

(155) I am not prejudiced against Caucasians, or you are the worst honky I know.

As the above examples show, the presupposition of (154) is filtered out and thus does not project, while it strongly feels that the expressive presupposition of (155) projects out. This seems like a significant difference between run-of-the-mill presuppositions and expressives, but as Schlenker (2016) notes, a plausible explanation could be the triggering of ignorance implicatures, as it is rather odd for the speaker to display ignorance of his/her own mental state. I will elaborate on this immediately below.

4.2.2. Objections against the presuppositional analysis

The task of scrutinising the claim that expressives are presuppositions is undertaken by Thommen (2017), who carefully cross-examins them side to side with ordinary presuppositions. He finds that even when controlling for the possible confounds of ignorance implicatures and intensionality, expressives do project more robustly than ordinary presuppositions, which strongly suggests that they are not a kind of presupposition.

As mentioned right above, a possible reason why expressives project out of disjunction filters could be the triggering of ignorance implicatures. Thommen (2017) proposes that in order to circumvent this confound, one could try subjunctive conditional filters instead, as follows (2017: 189-90):

(156) If France was a monarchy, the monarch of France would be bald.

(157) If I were Germanophobic, then my colleagues would be Germanophobic too.

(158) ! If I were Germanophobic, then John would be a boche.

While the construction is the same in all examples, we see projection only in (158), where the slur *boche* is used. But as Thommen notes, this is explained by the indexical character of expressive presuppositions, as they were proposed by Schlenker (2007). Specifically, because of indexicality, the consequent of the conditional is tied to the actual world rather than any epistemically accessible worlds introduced by the modal operator, which accounts for the projective behaviour:

(159) $\forall w \in R(w^*)$ (($[\![\text{I am Germanophobic}]\!]^{(c)(w)} = 1$) → (the speaker of c is Germanophobic in w))

(160) $\forall w \in R(w^*)$ (($[\![\text{I am Germanophobic}]\!]^{(c)(w)} = 1$) → (the speaker of c is Germanophobic in w*))

Based on the above, Thommen (2017) contends that a legitimate comparison between ordinary presuppositions and expressives ought to control for both confounds of

ignorance implicatures as well as intensionality. To this end, Thommen (2017) introduces a game show where players deliberately try to give as little information as possible for other players to guess their identity, which means that no ignorance implicatures should arise in disjunctive constructions. And because disjunctive constructions are extensional, the confound of intensionality does not arise either. Thommen provides the following examples (2017: 193):

(161) Either I don't hate Caucasians, or I hate Caucasians and my daughter married a Caucasian.

(162) ! Either I don't hate Caucasians, or my daughter married a honky.

Evidently, there is a difference in the projective behaviour between (161) and (162), with the former providing no clues or commitments about the speaker's ethnic prejudices, and the latter conveying a certain level of offence. Because this environment controls for both previously suspected confounds, Thommen (2017) concludes that the contrast between (161) and (162) testifies against the idea that expressives are a kind of presupposition.

Chapter 5. A new account of expressives: presuppositions plus associative meaning

This chapter will advocate a novel account of expressives as presuppositions (pace Potts 2005, 2007, McCready 2010, Gutzmann 2015, and in agreement with Schlenker 2007), which are not indexical but ordinary (pace Schlenker 2007), while at the same time endowed with additional use-conditional meaning (modelled on Predelli 2013, rather than Gutzmann 2015) which shall be called 'associative'.

The first step in presenting the new account is defending the presuppositional view against Thommen's (2017) objections.

5.1. Assessment of Thommen's (2017) claims against the presuppositional view

The first observation in assessing Thommen's (2017) claims is that slurs are not particularly good examples when it comes to figuring out what expressive meaning is. First, they are quite different from classical expressives such as 'damn', 'bastard' etc. because as was noted before they do not participate in semantic composition in the same way as the latter do (they are 'propositional', or 'isolated', i.e. they come saturated, unlike 'predicative' classical expressives, in McCready's 2010 words). This means that whatever conclusions we draw from observing the behaviour of slurs will not be automatically applicable to classical expressives. But secondly and most importantly, much more than typical expressives, slurs are extremely sensitive words with potentially detrimental social effects, which could be more adequately explained based on the idea of taboo or prohibition (Anderson & Lepore 2013). As was mentioned in the

previous chapter, often the offence of slurs cannot be contained even by quotation, so it should not come as a surprise that slurs cannot be filtered but project across the board. The fact that slurs project more robustly than presuppositions, therefore, does not seem as sound evidence that expressives are not presuppositions, as slurs are different from classical expressives. So we need to see whether Thommen's tests produce the same results using classical expressives in order to assess if there is indeed strong evidence against the presuppositional account.

To remind the reader, Thommen's general point was that expressives are not filterable like ordinary presuppositions are, and the task now is to see whether this is confirmed by classical expressives as it was with slurs. But it turns out that there are examples of (classical) expressives with conditionalisation, a fact which suggests filterability:

(163) I consider John a saint. But if he ever screws me over, I'll crush the bastard like a bug. (Lasersohn 2007: 12)

(164) I love my computer, I've had it for almost 7 years and it's still working fine. But if it breaks down during the interview, I swear I'll smash the damn thing.

(107) Sono ronbun-no chousha-ga yuumei-na kyouju nara, LSA-ni irassharu hazu da. Demo gakusei-nara, ko-nai kamoshirenai. (Sudo 2012)[81]
the article-GEN author-NOM famous professor if, LSA to come$_{HON}$ definitely. but student if, come-NEG come
'If the author of that article is a famous professor, he will definitely come$_{HON}$ to LSA. But if he is a student, he might not come.'

Because of the conditional constructions involved, in the above examples it feels that no offence (or honouring) is actually taking place, as the speaker is merely stating the conditions under which it *would* take place. As Lasersohn (2005: 227) notes for his own example:

> A speaker of (12) [in this dissertation: 163, R.C.] might be accused of being volatile, but not incoherent. This discourse does not commit the speaker to the position that John is a bastard; on the contrary, the speaker makes clear that he or she considers John to be a saint. The description of John as a bastard is conditionalised on the (unexpected) event of his "screwing over" the speaker, so that the expressive content does not project up to the sentence as a whole.

Lasersohn's explanation directly applies to (164), as this example was constructed based on his own. And as was previously noted for (107), the speaker only finds the referent worthy of respect/honouring under the condition that they are a professor, not

[81] This example uses Japanese honorifics, which as discussed previously, have been shown to be much more like presuppositions than Pottsian expressives by Sudo (2012). As a result, it might seem that it should not figure here, since the current discussion is about classical expressives. However, I want to remind the reader that the claim in the previous chapter was that Japanese honorifics are not a special kind of meaning which is both non-presuppositional (and not at-issue, of course) and compositionally relevant, as advocated by Gutzmann (2015), but an instance of good ole presupposition. I have decided to include Japanese honorifics as an example in this discussion exactly because it is a discussion defending the idea that classical expressives are also presuppositions – just like Japanese honorifics. Any reader who strongly feels that this example shouldn't have been included is invited to disregard it and consider only the other examples.

unreservedly. So we have evidence for filterability for the classical expressives 'bastard' and 'damn', very similarly to Japanese honorifics which have already been shown to be presuppositions (Sudo 2012).

But even if we acknowledge that there is some filtering, we may also grant that there is *some* kind of projection which is not 100% filtered. That this is happening can be diagnosed by interlocutors' possible reactions/challenges to the expressives (*bastard*, *damn*) or to honorifics; if the filtering was 100% successful, then no challenge would have been possible, as in the following case with an ordinary presupposition:

(165) If France is a monarchy, the King of France is bald.

Hey wait a minute! I didn't know France was a monarchy.

But interestingly, the examples involving classical expressives and honorifics are not felicitously challenged by the classical 'HWAM I didn't know…' test for presuppositions (Shannon 1976, von Fintel 2004) but rather metalinguistically, by employing a metalinguistic variant of the test which targets the term in question, such as 'HWAM you shouldn't use that word/why are you using that word' (Camp 2013)[82,83]:

(166) I consider John a saint. But if he ever screws me over, I'll crush the bastard like a bug.

Hey wait a minute! I didn't know you considered John a bastard!

Hey wait a minute! You shouldn't use that word (to talk about John). It's offensive.

(167) I love my computer, I've had it for almost 7 years and it's still working fine. But if it breaks down during the interview, I swear I'll smash the damn thing.

Hey wait a minute! I didn't know you thought badly of your computer.

Hey wait a minute! You shouldn't use that word. It's offensive.

(168) Sono ronbun-no chousha-ga yuumei-na kyouju nara, LSA-ni irassharu hazu da. Demo gakusei-nara, ko-nai kamoshirenai.

---

[82] In a paper about slurs, Camp (2013: 342) contends the following:

> 'the most natural form of challenge to a slur is typically metalinguistic, as in:
> Hey wait a minute! You shouldn't use that word to talk about Hispanic people – it's offensive and demeaning.
> Notably, this is also the form of challenge that is most natural for other perspectival expressions, such as 'tu'/'vous', slang expressions for parents, or thick pejorative terms like 'snitch' '.

[83] A similar metalinguistic way of targeting the expressive content is Predelli's (2013) 'Cautious Assent Test', i.e. 'yes, but…' as a way to agree to the character but challenge what he calls the 'bias' component of meaning (more on this below).

'If the author of that article is a famous professor, he will definitely come$_{HON}$ to LSA. But if he is a student, he might not come.'

Chotto matte! Anata-ga kyouju-wo sonkei shite iru towa shira-na-katta-yo.
Hey, wait a minute! I didn't know you respected professors.

Chotto matte! Sono sonkeigo nandayo? Chotto yarisugi janai?
Hey, wait a minute! Why an honorific? Isn't it a bit too much?

Considering the above, the verdict on Thommen's (2017) claims is not absolute. Putting aside slurs for the reasons mentioned above, contrary to his claims classical expressives do display filterability. But at the same time, this filterability is not as impervious as in the case of ordinary presuppositions, since something projects out nonetheless (as seen by the possibility to challenge the utterances by a metalinguistic variant of HWAM). So it seems that expressives manifest an ambivalent behaviour by being both presupposition-like in one sense (part of them is filtered) and non-presupposition-like in another (part of them projects out).

I will argue that an adequate explanation of the ambivalent behaviour of the data is that expressives are in fact presuppositions, but which are also endowed with some extra properties which are responsible for their non-presuppositional behaviour. These extra properties can be understood in terms of additional meaning which is represented in terms of conditions of use, as preliminarily advocated by Kaplan (1999) and eventually developed in a full-fledged framework by Predelli (2013), to which I turn next.

5.2. Predelli's Theory of Bias (2013)

A major philosophical project on meaning that cannot be captured in terms of truth conditions has been offered by Predelli (2010, 2013, 2017). Following Kaplan's (1989) distinction between the 'vagaries of actions' and 'the verities of meanings' (1989: 585), Predelli (2010) notes that in some cases expressions may impose certain constraints on how they can be used, which cannot be captured in terms of truth-conditions. In some cases, such use constraints may have to do with purely extra-semantic facts, but in others they are dictated by the conventional meaning of expressions. For example, an utterance such as 'There is an instance of English produced now' intuitively feels as guaranteed to be saying something true, while an utterance such as 'Hurray, I'm sorry to hear that' strikes as us rather anomalous. But obviously, one cannot say that the former is a logical truth, because it is a purely contingent matter that this is the case, and one cannot express the anomaly of the latter in terms of falsity, but more like in terms of incorrect (if not insensitive) use. To this end, Predelli's (2013) proposal is that such intuitions are best captured in terms of conditions of use, or more specifically in terms of classes of contexts in which certain expressions are felicitously (appropriately, non-defectively, etc) used.

To clarify, Predelli's (2013) plan is not to reject truth-conditional semantics by wholly ascribing to the 'meaning is use' slogan, but to delineate its boundaries as well as complement it with an additional framework that extends beyond these boundaries.

This clear delineation allows for a solution to certain challenges posed for truth-conditional semantics, such as the seemingly guaranteed truth of 'I am here now' (Kaplan 1989) and the so-called 'addressing puzzle' (Tsohatzidis 1992, Zimmerman 1997), among others. According to Predelli (2013), such problems result from confounding use-imposed constraints with truth conditions, or as he calls it, when a 'Fallacy of Misplaced Character' takes place. In order to account for these puzzles, Predelli (2013) introduces the notion of settlement according to a theory of use. The notion of settlement can encompass that of truth, but can also naturally extend beyond it. Specifically, when a sentence is true by virtue of character alone, then we can say that it is generally settled, as it will be true in all contexts of use. As character has to do with truth-conditionally relevant aspects of meaning, this means that all logical truths are generally settled. However, there is also the possibility for a sentence to be merely generally settled, or else, generally settled not by virtue of character alone but due to extra-semantic constraints that follow from the theory of use in question. This elegantly allows Predelli to account for intuitions of guaranteed truth by making a distinction between what follows from truth conditions and what not, which makes the aforementioned puzzles dissolve.

More specifically, with regards to the first puzzle, in order to eschew the problematic claim that a contingent matter such as that expressed by 'I am here now' is a logical truth, Kaplan (1989) accounted for the intuition of guaranteed truth by restricting it to so-called 'proper contexts', namely contexts in which the agent of the context exists in the world of the context at the time of the context. This allowed him to say that while 'I am here now' is not true in any context (as a logical truth would be), it is true in any *proper* context. But Predelli's (2013) proposal enables him to account for the intuition of guaranteed truth without the dubious distinction between proper and improper contexts: settlement depends on the theory of use in question. When it comes to face-to-face communication, obviously 'I am here now' is merely generally settled, as in such circumstances the speaker of the context is at the world as well as time of the context. But when it comes to the notorious cases of messages left on answering machines or notes written for posterity, it follows naturally that this sentence will not be settled. As far as the so-called addressing puzzle is concerned, for Tsohatzidis (1992) the perceived anomaly of sentences such as 'Je ne te tutoie jamais, mon cher Paul' ('I never address you (*tu*) as "tu", my dear Paul') shows that the whole enterprise of accounting for meaning in terms of truth conditions is doomed to failure. But Predelli's (2013) proposal shields the truth-conditional approach from such false accusations by marking a clear distinction between settlement in terms of truth-conditionally relevant versus irrelevant aspects of meaning, and thus accounting for the anomaly of sentences such as this one in a level outside that of truth-conditions.

As noted before, Predelli's (2013) proposal is that the constraints encoded in a theory of use can be represented in terms of the class of contexts in which a specific expression can be felicitously used, and in some cases these constraints have to do with extra-semantic topics, while in others they form part of the conventional profile of an expression. From this follows that the meaning of an expression is not always exhausted by its truth-conditional endowment (its character) but may also encode information about how this expression is felicitously used, which Predelli (2013) terms 'bias'. This

amounts to saying that the meaning of an expression can be represented as a character-bias pair (Predelli 2013: 66):

(169) *meaning*(hurray) = <*char*(hurray), *bias*(hurray)>

Specifically, the bias component represents those aspects of meaning which, albeit truth-conditionally irrelevant, impose restrictions on the conditions under which an expression may be felicitously used, which are represented as sets of contexts. Below are two such examples (Predelli 2013: 65, 82):

(170) c ∈ CU (hurray) only if c_a approves of something in c

(171) c ∈ CU (fuck) only if c_a is a participant in register coarse in c

Informally, this amounts to saying that a context c belongs to the class of contexts of use for 'hurray'/'fuck' only if the agent of c approves of something/is a participant in register coarse in c. It is crucial to note here that the bias meaning of an expression encodes necessary conditions ('only if…') for membership to the class of contexts of felicitous use, not necessary and sufficient (iff) conditions. Although Predelli himself does not dwell on this issue at all, I find that it is important to emphasise it after the criticisms against Gutzmann (2015) in chapter 4. To remind the reader, it was noted there that while Gutzmann's (2015) proposal amounted to a faithful replication of the T-schema in the use-conditional dimension, it is highly dubious that one could advocate for a correspondence theory of felicity in the exact same way as for a correspondence theory of truth; assuming that there is an external reality which exists independently from human minds, where something either is the case or not, the notion of truth seems to be quite different from that of felicity, which is essentially a subjective notion rather than a collection of objective facts. Moreover, this also highlights that despite the apparent similarity between the frameworks of Predelli (2013) and Gutzmann (2015)[84],

[84] As noted by Gutzmann himself (2015: 22, fn.8): 'See also Predelli (2010, 2013), who presents an analysis similar to mine, but (i) does not develop a compositional framework as I will do in the second part of this book, and (ii) only addresses expressive validity and not Kaplan's "validity-plus" '. Having already argued in chapter 4 against (i) (i.e. that the framework needed should not be compositional), a few words are in order about (ii), which if intended as a criticism (given 'only'), is a rather misguided and unfair one, since Predelli explains the difference between his expressive validity and 'validity-plus' so it should be clear there were clear theoretical motivations for addressing the former rather than the latter. The difference is as follows: Kaplan's 'validity-plus' corresponds to the idea of 'truth-plus-preserving', which basically simply adds something on top of the notion of truth, an 'attitude' (Kaplan 1999: 7). But Predelli's expressive validity is an independent notion based on the idea of 'usability preservation', and not a mere adjunct to truth. Recognising an independent notion enables us to capture the difference between ordinary (truth-based) validity and expressive validity, which cannot be captured by validity-plus given that the latter is an integrative notion which entails ordinary validity. As Predelli notes by offering the following contrast (2013: 119-20), 'some expressively valid arguments are not character-guaranteed': (|≠ stands for ordinary invalidity, |e stands for expressive validity)

alas Fa **|e** *unfav(i, that Fa)*
Alas, Trump was elected. **|e** The speaker is unfavourable towards the fact that Trump was elected.

alas Fa **|≠** *unfav (i, that Fa)*
Alas, Trump was elected. **|≠** The speaker is unfavourable towards the fact that Trump was elected.

as they both essentially propose to encode conditions of use in terms of sets of contexts, there is a critical difference between the two at a core level of the representation. As noted in chapter 4, I find this a fatal problem which threatens to bring down Gutzmann's (2015) whole enterprise of hybrid semantics, while Predelli's (2013) proposal seems adequate for representing conditions of use. In the same way seen just above for the meaning of interjections and register variants, Predelli (2013) shows how the theory of bias accounts in the same way for the non-truth-conditional meaning of other expressions, such as slurs, honorifics, diminutive suffixes, nicknames, etc.

5.2.1. Criticisms against the theory of bias

In this section I will present and attempt to address a criticism of Predelli (2013) by Sherman (2014). Specifically, Sherman (2014) raises the worry that the theory of bias, as it stands, could be employed to represent not only non-truth-conditional meaning as intended by Predelli (2013), but also truth-conditional meaning such as presuppositions. In Sherman's (2014) own words:

> If we can account for both truth-conditional and non-truth-conditional meaning in terms of bias, then why bother distinguishing them? Why not account for all meaning in terms of bias? This seems like a mistake; we can call it the Fallacy of Misplaced Bias. But what's the mistake? What's special about truth-conditional meaning that makes it worth singling out? Truth-conditional meaning is special because representation is a central task of language use. A framework for non-truth-conditional meaning ideally ought to provide resources useful for the task performed. What task does a given kind of non-truth-conditional meaning perform? It depends. A use of an expression might function to impose an information state on the context, to bring about a change in preferences, to introduce a question under discussion. Like the example of presupposition, all of these tasks can be accounted for in terms of bias. But that only highlights the fact that bias is not tied to any one task. The framework, in other words, is so general that much of the work needed to account for a particular type of non-truth-conditional meaning will need to be reconstructed from within the framework.

In fact, Sherman's (2014) criticism against Predelli (2013) is very similar in nature to Crespo's (2015) against Gutzmann (2015) and Kaplan (1999), which was presented in chapter 4. The worry expressed by both Sherman (2014) and Crespo (2015) essentially

---

This means that if we make up a symbol for validity-plus (e.g. |+, its negation being |+≠), the following holds:

alas Fa **|+≠** *unfav (i, that Fa)*
Alas, Trump was elected. **|+≠** The speaker is unfavourable towards the fact that Trump was elected.

The result is that while all valid-plus arguments are expressively valid (given the fact that validity-plus is a hybrid notion of truth enhanced by attitude), as shown above there are expressively valid arguments which are *not* valid-plus. It should thus be clear why Gutzmann's (ii) does not count neither as a legitimate criticism of Predelli or as an asset of his own framework over Predelli's.

amounts to asking what is the distinctive difference between character and bias, or else between truth-conditional and use-conditional meaning, since the proposed framework for the latter element of these pairs seems to be usable for accounting for the former one as well. This is a legitimate worry, which I shared in chapter 4 after assessing Gutzmann's (2015) hybrid semantics framework. Nevertheless, I think that Sherman's (2014) similar worry for Predelli (2013) can be responded, though perhaps only partially or perhaps temporarily. The reason the defence is partial is because it will only apply to those expressions whose bias component displays the distinctive characteristic of non-displaceability (Potts 2007), such as interjections, register variants and slurs. For each of these categories of expressions that he examines, Predelli (2013) applies tests which involve syntactic embedding to diagnose whether this property is manifested or not before declaring a certain component of meaning as bias. This means that there is indeed a distinctive difference between character and bias, that of non-displaceability. Thus said, Sherman's (2014) worry can be answered as follows: no Fallacy of Misplaced Bias may be committed as long as one makes sure to apply tests for the diagnosis of non-displaceability before characterising a certain meaning as bias. However, it should be clear that this response hinges on non-displaceability acting as a safeguarding mechanism; for any expressions that do not exhibit this property, the distinction between character and bias feels much harder to define. This is what makes this defence temporary, because as will be discussed in the appendix of chapter 7, there might be valid reasons to abandon the distinction altogether.

5.3. A new account of expressives as presuppositions plus associative meaning

5.3.1. Associative meaning

This section will argue that the ambivalent behaviour of expressives can be adequately explained by an account which conciliates the presuppositional view with the idea that there is something ontologically distinct about so-called expressive meaning.

First of all, this account agrees with the assumption by Potts (2007) that there is indeed an ontologically different dimension of meaning, endowed with distinctive characteristics (i.e. independence, non-displaceability, perspective dependence, descriptive ineffability, immediacy and repeatability). However, it is argued that this meaning lies outside semantic composition, thus it does not participate in it (pace Potts and his followers). I propose to call this meaning 'associative', adopting a term by Leech (1981)[85]. I argue that associative meaning essentially corresponds to Predelli's (2013) bias, so following his model it can be represented as the class of contexts where the expression in question is felicitously used, as explained above.

A few words are in order about what exactly it means that associative meaning is ontologically distinct. When Potts (2005, 2007) made this claim for his expressive

[85] Leech (1981) defined associative meaning as the counterpart of 'conceptual meaning', roughly equivalent to truth-conditional meaning. For more details about Leech's (1981) notion see specifically his pages 12-16.

meaning, he merely meant to say that is different from at-issue and presuppositional meaning. However, for the present account the essential difference lies in the contrast between truth-conditions and conditions of use. More specifically, this contrast entails that associative meaning is devoid of a Fregean sense, if such a notion is to be understood as a kind of abstract meaning which is shared by speakers but is independent from their minds (cf. Frege's anti-psychologism argument). Associative meaning is understood as a collection of uses, which entails not only that there is a user involved but also that his/her conscious decision to use this meaning in such-and-such circumstances can define it. Of course, this is essentially a psychologistic claim since it says that associative meaning lies in the mind (or more accurately in a collection of minds), unlike somewhere like a mind-independent Third Realm where Fregean senses are said to be.

This ontological difference is also reflected in the way associative meaning is learned by speakers: it is argued that this is done in a token-reflexive way, by observing the collection of instances of use. This means that every use is a legitimate use, as it just gets added to the collection and has the ability to change the meaning, by changing the association. Of course, this does not happen overnight which signals that there are some constraints, such as the assumptions that the speaker is a competent user of the language, that there was no miscommunication or misunderstanding of the discourse, etc. However, such a change has the potential to happen relatively quickly compared to a change in the meaning of words with an abstract meaning (Fregean sense). Examples of this relatively quick change of association is the re-appropriation of slurs of all kinds (Anderson & Lepore 2011, Bianchi 2014), the adoption of Japanese male personal pronouns by female speakers not to signal identification as male but to convey traditionally masculine attitudes such as assertiveness, competitiveness, or rejection of social norms (Miyazaki 2004, Jugaku 1979, Reynolds 1998), and the use of swearwords such as 'bastard' (Potts 2007) or the Greek 'μαλάκας' ('wanker' or 'asshole') to signal affection, solidarity and familiarity (Vergis & Terkourafi 2015).

5.3.2. Implications of the associative meaning hypothesis for the question of whether expressives are presuppositions or not

As seen above, this account says that associative meaning lies outside semantic composition, and does not participate in it. However, it should be recognised that some expressions which are compositionally relevant can also carry associative meaning (which is compositionally irrelevant). It is argued that in this case even if these two different dimensions do not interact within the compositional process, their co-existence in a single lexical item creates certain effects which can result in what is perceived as ambivalent behaviour.

Specifically, the proposal is that expressives such as *damn*, *bastard*, *fucking* and Japanese honorifics are indeed presuppositions – this is how they form part of semantic composition, *qua* presuppositions – but also carry associative meaning. The idea is that their presuppositional nature explains the filterability observed, and their associative meaning explains the projectivity observed. Importantly, their presuppositions are

understood as ordinary presuppositions rather than indexical presuppositions (pace Schlenker 2007), but the associative meaning itself is indexical, as it is modelled as a set of contexts. So this proposal agrees with Schlenker (2007) that all the distinctive characteristics proposed by Potts (2007) can be explained on the basis that these inferences are indexical, however pace Schlenker (2007) it identifies the observed indexicality not as a property of the presupposition, but of the associative meaning which co-inhabits the same lexical item. The result is that we have (ordinary) presuppositions which are indexical (because they also have associative meaning, which is indexical), but we do not have indexical presuppositions.[86] Let's show how this works with an example.

We start with the assumption that *bastard* carries the lexical presupposition that the referent is not nice/vile/bad (or something along these lines), but also has the associative meaning that the speaker is angry at the referent, which is modelled as the set of contexts in which the speaker is angry at the referent.

(172) John believes that that bastard Kresge is rich.

Potts (2005, 2007) would say: *bastard* is not filtered by the attitude predicate to signal that John finds Kresge a bastard, but it projects out to mean that it is the speaker who thinks Kresge is a bastard. So *bastard* is not a presupposition because it cannot be filtered.

Schlenker (2007) would say: since *bastard* is an indexical presupposition, its projection across the attitude predicate can be naturally explained by indexicality. That is to say, because the expressive presupposition is tied to the actual context of utterance, the attitude expressed by *bastard* pertains to the speaker of the context.

The current account says: Assuming that *bastard* means something like not nice/vile/bad and it is an (ordinary) presupposition, it can be read either *de dicto* or *de re*. But because we assume that *bastard* also has associative meaning, which is indexical (tied to the speaker of the current context), there is a strong preference for the *de re* reading over the *de dicto* one. The reason is because if the presupposition is read *de dicto* (i.e. according to John's beliefs), it is hard to explain why the speaker himself is angry at the referent, which is what the associative meaning encodes. In such a case, it would have been better conversational strategy for the speaker to choose a term without associative meaning, because s/he would not risk being mistaken as the one angry when his/her intention was simply to report John's belief.[87,88] So because of their

[86] I should clarify that the claim here is that expressives are not instances of indexical presuppositions, not that there is no such thing as indexical presuppositions. This analysis of expressives as ordinary rather than indexical presuppositions does not in any way preclude that there might be other linguistic phenomena which are actually instances of indexical presuppositions. After all, given that we know that assertive meaning can be indexical, there is no theoretical justification to rule out the possibility that presuppositional meaning can also be indexical.

[87] This explanation is quite similar to the one given by Lasersohn (2007: 228) for a very similar example:

> '(14) Sue believes that that bastard Kresge should be fired.
> Because expressives are so emotionally charged, and because their use can carry a significant social risk, I suspect that speakers are especially cautious about using them in

associative meaning, which is tied to the speaker, such presuppositions have a strong preference for the *de re* reading.[89] Note that this explanation is essentially the same as the one given by Sudo (2012) for why gender features are not indexical presuppositions (as per Cooper 1983, Yanovich 2010), but as ordinary presuppositions with strong preference for global accommodation and *de re* construal.

5.3.3. Lexical semantics of expressives[90]

At this point it is important to address a possible complication for the present account, i.e. the seeming lack of projection observed when expressives are in non-veridical predicative positions, as shown in the contrast below:

(172) John believes that that bastard Kresge is rich.

(173) Kresge is not a bastard.

As mentioned right above, the projection in (172) signals that it's the speaker who finds Kresge a bastard, and not John. However, in (173) where we have a denial of the expressive, there seems to be no projection of the associative meaning: it's not just that no inference that the speaker finds Kresge a bastard projects out, in fact it is explicitly said that this is not the case. What are the options here to rescue the associative meaning hypothesis from a potentially devastating counterexample? As mentioned in chatper 4, Zimmerman (2007) suggested two possible solutions for Potts (2007), who faced a similar problem by predicative uses of expressives: type-shifting or lexical ambiguity. Given the compositional irrelevance of associative meaning, the first is unavailable for the present account. So it seems that we're stuck with the second one, which is not necessarily untenable (it was worked out and argued in chapter 4), yet it comes at a cost as we will be forced to concede that expressions that feel intuitively the

---

> embedded contexts where there is a chance of their content ''leaking'' – except, of course, if the speaker does agree with the content of the expressive and is willing to make this agreement public. But if this caution means that speakers systematically refrain from using (14) unless they are willing to publicly agree with the description of Kresge as a bastard, (14) will come to imply a speaker commitment to this description – in other words, the presupposition will project, despite the usual classification of *believe* as a plug. This, I suspect, is the correct explanation for why expressive content normally projects, rather than any theoretical distinction between presupposed and expressive content.'

I think that although this explanation is towards the right direction, it doesn't go far enough to be explanatory. What does it mean exactly that 'expressives are so emotionally charged'? How is this property defined and how does it manifest? Unless the property of being 'emotionally charged' is clarified, one might claim that it is precisely this property which imposes a theoretical distinction between presupposed and expressive content, contrary to what Lasersohn (2007) sets out to argue. In fact, it seems to me that this was the main claim in Potts (2007). I think that as the current account successfully pins down the distinct property of expressives as their associative meaning component, it allows Lasersohn's (2007) explanation to go through.

[88] Of course, there is also the possibility that this is a case of mixed quotation (Kuno 1988, Maier 2009), i.e. that the speaker said *John believes that the "bastard" Kresge is rich.* Such a case can be straightforwardly explained, but we assume that it is not always the case.

[89] Sauerland (2007: 233) was the first to note the 'obligatory de re effect' of expressives, and pointed it out as a problem for Potts (2007) as his proposal could not account for it.

[90] I thank Yasutada Sudo for pointing out this complication and for stimulating and constructive discussion.

same are distinct, something that feels uncomfortable especially in examples such as the above, where (173) could in fact be an immediate response to (172).

But it turns out that there is a way to explain the difference noted above without resorting to lexical ambiguity: it has to do with the precise lexical content of the associative component of expressives. Conveniently, by clarifying it we will also be able to pin down an important difference between different expressives: as noted by Zimmerman (2007), some expressives can appear in predicative positions (such as 'bastard' above), but others cannot (e.g. 'damn' or 'fucking').

The explanation consists in proposing that there are two different kinds of expressives: one which combines a compositionally relevant component together with an associative one, such as 'bastard', 'motherfucker', Japanese honorifics etc, and one which consists only of associative meaning, such as 'damn', 'fucking', etc. Merely for ease of discussion, let's label the first category 'complex' and the latter 'simplex' expressives. The proposal is that complex expressives have the following two properties: first because of their compositionally relevant component (which may be as generic as 'person' or even 'sentient entity', given that they can be used to talk about animals too) they can appear in predicative positions, and second their associative meaning is specific, in the sense that the attitude expressed by the associative meaning is directed specifically towards a specific entity. In particular, this means that the associative meaning of a complex expressive such as 'bastard' is that the speaker is angry at the referent of the expression, if a referent exists.

On the other hand, simplex expressives have the opposite two properties: they cannot appear in predicative positions (at the absence of a compositionally relevant component, predication is impossible), and the attitude expressed is not tied to any specific referent. The latter means that they can be employed to express the speaker's anger either towards a specific referent, if s/he so wishes, or in general, as seen in the examples below:

(174) The fucking kids scratched my car.

(175) They scratched my fucking car.

In (174), the speaker is angry at the kids for having scratched his car. But in (175), the intuition is not that the speaker is angry at his own car (even though 'fucking' appears right before 'car') but at the fact that his car got scratched. So specifically, the idea is that the associative meaning of a simplex expressive such as 'fucking' is that the speaker is angry, but the target of this attitude is left unspecified.

So the problem about the lack of projection in (173) is to be explained as follows: because complex expressives express an attitude towards a specific entity, in non-veridical contexts such as (173) it so happens that there is *no* specific entity in question; as a result, the attitude encoded in their associative meaning is not consummated and thus it does not project. Apart from negation which was seen in (173), the same phenomenon is observed in other non-veridical contexts such as the antecedent of a conditional:

(176) If he really is a bastard, he won't offer to give us a helping hand.

An interesting implication of the different lexical semantics of simplex and complex expressives is that they can explain the so-called 'repeatability' property (Potts 2007). According to Potts (2007: 182), repeatability is a distinctive property of expressives:

(177) Damn, I left my damn keys in the damn car.

(178) # I'm angry! I forgot my keys. I'm angry! They are in the car. I'm angry!

Although Potts' (2007) intuition is towards the right direction, I would like to claim that the repeatability property is mostly manifested by simplex expressives because of the non-specificity they afford:

(179) # Bastard, I left my bastard keys in the bastard car.

(180) Damn, I left my damn keys in the damn car.

5.3.4. Application to data

Given the above assumptions, this account explains the data as follows.

(163) I consider John a saint. But if he ever screws me over, I'll crush the bastard like a bug. (Lasersohn 2007: 12)

Assumptions about *bastard*:

Presupposition: the referent is not nice/vile/bad

Associative meaning: the speaker is angry at the referent (association: this expression is used by the speaker when they are angry at the referent)

The presupposition is conditionalised: the speaker doesn't think that John is not nice/vile/bad, on the contrary he makes clear how highly he thinks of him. He is just saying that in case John does something bad to him, he will think badly of him (i.e. as not nice/vile/bad).

The associative meaning projects out: the speaker is angry at this very moment, hence his choice of this expression, just thinking about the possibility that John might screw him over.

(164) I love my computer, I've had it for almost 7 years and it's still working fine. But if it breaks down during the interview, I swear I'll smash the damn thing.

Assumptions about *damn*:

Presupposition: the referent is not nice/vile/bad

Associative meaning: the speaker is angry (association: this expression is used by the speaker when they are angry)

The presupposition is conditionalised: the speaker doesn't think that her computer is not nice/vile/bad, on the contrary she makes clear how much she likes it. She is just saying that in case it breaks down during a critical moment such as an interview, she will think badly of it (i.e. as not nice/vile/bad).

The associative meaning projects out: the speaker is angry at this very moment, hence her choice of this expression, just thinking about the possibility that the computer might break down at such a critical moment such as an important interview.

(107) Sono ronbun-no chousha-ga yuumei-na kyouju nara, LSA-ni irassharu hazu da. Demo gakusei-nara, ko-nai kamoshirenai. (Sudo 2012)
the article-GEN author-NOM famous professor if, LSA to come$_{HON}$ definitely. but student if, come-NEG come
'If the author of that article is a famous professor, he will definitely come$_{HON}$ to LSA. But if he is a student, he might not come.'

Assumptions about *irassharu* (come$_{HON}$):

Presupposition: the subject is worthy of respect/honouring

Associative meaning: the speaker recognises the referent as socially superior (association: this expression is used by the speaker about someone when they find the referent worthy of honour)

The presupposition is conditionalised: the subject is worthy of respect/honouring only under the condition that they are a professor. This does not hold under the assumption that they are a student.

The associative meaning projects out: the speaker is signalling that he respects professors in general, hence his use of honorifics in this case for the (hypothetical) professor. This holds irrespective of whether the individual talked about turns out to be a professor or not.

(63) Whenever I pour wine, the damn bottle drips. (Schwarz, mentioned in Potts 2007)

Assumptions about *damn*:

Presupposition: the referent is not nice/vile/bad

Associative meaning: the speaker is angry (association: this expression is used by the speaker when they are angry)

The presupposition is quantified in: the speaker could be saying this after he just had a dripping accident, but could also be saying it in a discussion about his general wine-pouring skills. Because of the quantification here, the *de dicto* construal prevails; the

speaker isn't explicitly saying that a specific bottle is not nice/vile/bad, he is saying that in all situations in which he tries to pour wine, the bottle drips and in those situations he finds the bottle in question not nice/vile/bad. (However, the de re construal is also perfectly possible, in a scenario in which the speaker has repeatedly tried to pour wine from a specific bottle, which has dripped in every single attempt).

The associative meaning projects out: the speaker is angry at this very moment, hence his choice of this expression, just thinking about the cases in which he has poured wine and the bottle dripped.

5.3.5. Examples with attitude predicates

Examples with attitude predicates are somehow controversial, as there are different intuitions about whether there is projection or not. Some say that there is projection (Potts 2007, Predelli 2013), others that there is not (Kratzer 1999, Schlenker 2003, 2007). As the present account posits a strong preference for de re but does not preclude the possibility of de dicto reading, it is compatible with both these intuitions, although it predicts that one should be much more prominent. More specifically, the prediction is that the more loaded the associative meaning of the expression in question, the stronger the feeling of projection/the harder the filtering, as it is assumed that in such cases *de re* construal will be strongly preferred over *de dicto* due to the influence of the associative meaning, for the pragmatic reasons mentioned above. So for example we expect *motherfucker* to project more robustly/be less filterable than *bastard*.

(44) My father screamed that he would never allow me to marry that bastard Webster. (Kratzer 1999: 6)

Assumptions about *bastard*:

Presupposition: the referent is not nice/vile/bad

Associative meaning: the speaker is angry at the referent (association: this expression is used by the speaker when they are angry at the referent)

The presupposition is filtered: the speaker doesn't think that Webster is not nice/vile/bad, on the contrary she must like him a lot since she wants to marry him, contrary to her father's wishes. He who finds Webster not nice/vile/bad is the speaker's father.

In this example it seems that the associative meaning does not project out, or at least not as much. The explanation is that because the associative meaning is not extremely strong (while we take it that *bastard* is used when the speaker is angry at the referent, it is a relatively mild expression to show anger towards someone, compared to other choices), *de dicto* construal of the presuppositional content (*not nice/vile/bad*) is possible, just as it would be for a case with an ordinary presupposition such as *My father*

screamed that he knew I was a Brexit supporter, where filtering is possible. One could perhaps think that this example is a counterexample to the associative meaning hypothesis, since we don't see projection. But note that the explanation is not that there is zero projection, but that it is much weaker because the associative meaning of *bastard* isn't as loaded, which allows for the *de dicto* construal of the presupposition and results in a sentence which can be taken to report the speaker's father's attitude towards Webster, and not her own. As evidence that there is some minor projection I suggest the fact that a speaker would rather not utter this sentence comfortably in a very high register discourse where uttering *bastard* might result in the speaker being perceived as vulgar. It is expected that in such a case she would either make it explicitly clear that she is quoting her father's use of *bastard* (a case of mixed quotation), or she would use a safer synonym without any associative meaning (e.g. *bad guy*).

(181) My father screamed that he would never allow me to marry that motherfucker Webster. (adapted from Kratzer 1999)

Assumptions about *motherfucker*:

Presupposition: the referent is not nice/vile/bad

Associative meaning: the speaker is very angry at the referent (association: this expression is used by the speaker when they are very angry at the referent)

Here we have a much stronger sense of projection, which the current account attributes to the strong associative meaning of *motherfucker*.

Theoretically, the presupposition could be filtered: the speaker doesn't think that Webster is not nice/vile/bad, on the contrary she must like him a lot since she wants to marry him, contrary to her father's wishes. He who finds Webster not nice/vile/bad is the speaker's father.

But because the associative meaning of *motherfucker* is so loaded, the projection is really hard to contain. The idea is that because *motherfucker* has much stronger associative meaning than *bastard*, it is predicted that the *de dicto* construal of the presupposition (*not nice/vile/bad*) while not impossible, is too risky or simply not good conversational strategy for the speaker. Therefore the associative meaning forces the *de re* construal, which explains the strong sense of projection. But if one assumes that the speaker were to utter this in a very low/vulgar register environment, the associative meaning would be less strong and filtering could be possible.

(182) John said that whenever he pours wine, the damn bottle drips.

Here we see two operators at play, an attitude predicate plus quantification; this example is essentially the one by Schwarz (mentioned in Potts 2007) embedded under an attitude verb. As before, we take it that the presupposition of 'damn' is that the referent is not nice/vile/bad (though I'll use just 'bad' here for ease of exposition). In theory, there should be two possible readings per operator, so four in total, as follows:

182i) De re + de re: John said that there is a specific bottle such that whenever he pours wine from it, it drips, and the speaker finds that bottle bad.

(182ii) De re + de dicto: John said that there is a specific bottle such that whenever he pours wine from it, it drips, and John finds that bottle bad.

(182iii) De dicto + de re: John said that whichever bottle he uses to pour wine from, it drips, and the speaker finds such bottles bad.

(182iv) De dicto + de dicto: John said that whichever bottle he uses to pour wine from, it drips, and John finds such bottles bad.

As mentioned before, the present account predicts that whenever a presupposition is also endowed with associative meaning, a de re reading will be strongly preferred to a de dicto one. For this reason, the de re + de re reading (182i) is predicted as a much preferred one: the speaker is angry at a specific bottle which John said keeps dripping. However, even though it might initially be expected that the de dicto + de re (182iii) should also be predicted as fine, since also in this case the presupposition of 'damn' is read de re, this does not follow from the present account due to a pragmatic clash. As the NP 'the bottle' is read de dicto here, according to John's beliefs, it would make very little sense for the speaker to be angry at a hypothetical bottle s/he is not acquainted with in any way. As a result, this reading is predicted as odd and rather impossible to get, due to the clash between the de dicto NP and the indexicality of the associative meaning, and this prediction is borne out. As for the readings in which the presupposition of the expressive is read de dicto (182ii and 182iv), the present account does not disallow them, but merely predicts that they should be less prominent than (182i), given the preference for de re reading, which I believe is confirmed.[91]

Yet again as above, here too we can test the prediction that the de dicto reading becomes increasingly harder to get when the associative meaning in question is more loaded. To see this, let's repeat the examples above by again replacing 'damn' by its stronger counterpart 'motherfucking':

(183) John said that whenever he pours wine, the motherfucking bottle drips.

(183i) De re + de re: John said that there is a specific bottle such that whenever he pours wine from it, it drips, and the speaker finds that bottle bad.

(183ii) De re + de dicto: John said that there is a specific bottle such that whenever he pours wine from it, it drips, and John finds that bottle bad.

(183iii) De dicto + de re: (predicted as not possible)

(183iv) De dicto + de dicto: John said that whichever bottle he uses to pour wine from, it drips, and John finds such bottle bad.

The intuition here is that the most prominent reading for (183) is (183i), and that (183ii) and (183iv) are really hard to get – if not impossible, then definitely harder than

[91] Note that the present account does not make any prediction of preference between (182ii) and (182iv), as their difference falls outside its scope.

(182ii) and (182iv) were as readings of (182). This is exactly what the present account predicts.

5.3.6. Schlenker's examples with slurs

This section will discuss some well-known examples by Schlenker (2003, 2007, 2012) which involve racial slurs. As mentioned above, I do not take slurs to be really uncontroversially comparable to classical expressives like *damn* or *bastard*, as their behaviour in discourse is akin to that of taboo words. In terms of the present account, this means that they are endowed with extremely loaded associative meaning; so loaded that sometimes it can hardly be contained even by quotation, as seen in the examples in footnote 54 (page 70). As a result, it is predicted that the associative meaning of these terms will project unconditionally.

Based on this, I take it that in any example where there is intuition of lack of projection, it should be attributed to mixed quotation (and even in this case, it is not guaranteed that there will be zero projection, given the taboo). This can be shown by the following contrast, provided by Schlenker (2003)[92]:

(184) # I am not prejudiced against Caucasians. But if I were, you would be the worst honky I know. (Schlenker 2003: 43)

(45) I am not prejudiced against Caucasians. But John, who is, thinks/claims that you are the worst honky he knows. (Schlenker 2003: 43)

In (184) the slur projects out, but not so much in (45), and the difference between them can be explained as that (45) allows for mixed quotation due to the attitude predicates *thinks/claims* while (184) does not have this possibility. So according to this analysis, the unproblematic reading of (45) in fact corresponds to (45i)[93]:

---

[92] To be clear, Schlenker (2003) did not provide this contrast to advocate for mixed quotation, but for his account of shifted indexicality which hinges on the presence of an attitude verb. In that paper, he convincingly shows that shifted indexicality can successfully account for the behaviour of Amharic indexicals, while an account based on mixed quotation is untenable – and it is after establishing shifted indexicality based on Amharic data that he proposes that it can also apply to expressives in English. However, just because the mixed quotation account was shown not to work for Amharic indexicals, it doesn't mean that it should be excluded for expressives, which is the stance adopted here. Nevertheless, I need to acknowledge that the main motivation for my resorting to mixed quotation is the detrimental effect of slurs, which seems to surpass that of other expressives; as mentioned above, the idea is that slurs are like taboo words whose mere presence in the discourse is risqué, obviously when used but sometimes even when quoted (including cases where the speaker mentions the slur to clearly condemn its use by someone else, as seen in the examples in footnote 54). Given their highly risqué nature, I therefore take it that the apparent lack of projection in certain cases can only be explained as a case of successful (mixed) quotation in some environments. However, I admit that this explanation based on mixed quotation is a kind of reductio ad absurdum just because at this point I cannot see another way to account for the ability of slurs to sometimes offend and others not, and further evidence is needed to support it, which I have to leave for future research.

[93] As noted before, it is quite possible that the mere quoting of a slur can offend. In these specific examples I assume that it does not, just to highlight the contrast between use and mention. But

(45i) I am not prejudiced against Caucasians. But John, who is, thinks/claims that you are the worst "honky" he knows.

Based on the associative meaning hypothesis, the prediction is that the more loaded the associative meaning of the slur[94], the clearer the intuition of mixed quotation, otherwise the speaker seems as expressing contradictory attitudes. I believe that this prediction is borne out (I use '#' to show that one perceives some oddness resulting from seemingly contradictory attitudes, and '##' to signal that it's much higher):

(45ii) ## I am not prejudiced against black people. But John, who is, thinks/claims that you are the worst nigger he knows. (adapted from Schlenker 2003)

(45iii) I am not prejudiced against black people. But John, who is, thinks/claims that you are the worst "nigger" he knows. (adapted from Schlenker 2003)

(58) # I am not prejudiced against Caucasians. But Pierre, who is, has repeatedly made the claim that you are the worst honky that the frog's mother knows[95]. (Schlenker 2011: 1588)

(58i) I am not prejudiced against Caucasians. But Pierre, who is, has repeatedly made the claim that you are the worst "honky" that the frog's mother knows.

(58ii) ## I am not prejudiced against black people. But Pierre, who is, has repeatedly made the claim that you are the worst nigger that the frog's mother knows. (adapted from Schlenker 2011)

(58iii) I am not prejudiced against black people. But Pierre, who is, has repeatedly made the claim that you are the worst "nigger" that the frog's mother knows. (adapted from Schlenker 2011)

5.3.7. Conclusion

There is indeed a different dimension of meaning, that of associative meaning, which is completely independent from compositional semantics and ought to be modelled differently, i.e. as a class of contexts of felicitous use:

(185) $c \in CU$ (damn) only if c_a is angry/annoyed etc.

(186) $c \in CU$ (fucking) only if c_a is very angry/annoyed etc.

(187) $c \in CU$ (bastard) only if c_a is angry/annoyed etc. at the referent

---

even if the reader strongly feels that mere mention is also offensive, s/he should still feel a considerable difference in offence between use and mention.

[94] I assume here that although slurs of all kinds are very offensive, there can be gradation among them, depending on the social position of the group which is being denigrated (e.g. chink VS redneck).

[95] Note that in this example the slur 'frog's mother' is expected to project (which is why example 58i is not predicted as odd). The motivation for Schlenker (2011) to construct this example was to show that expressives violate 'Shift Together' (Anand & Nevins 2004), or else that it is possible for two expressives within the same sentence to be attributed to different perspectives.

(188) c ∈ CU (motherfucker) only if c_a is very angry/annoyed etc. at the referent

With regards to the debate about the theoretical standing of expressives, it is proposed that classic expressives such as *damn*, *fucking*, *bastard*, and Japanese honorifics are indeed ordinary presuppositions, but also endowed with associative meaning, which results in a strong preference for *de re* construal for pragmatic reasons. As for racial and other slurs, despite their similarities with expressives I contend that due to extralinguistic reasons their associative meaning is so loaded that they essentially behave like taboo words, defying even quotation. In this sense, it might be wise not to use them as representative items of the class of expressives, as their behaviour is largely controlled by reasons that fall out of the interests of linguistic theory.

Finally, this account calls for some terminological remarks. This account recognises an ontologically distinct class of meaning which has the characteristic properties proposed by Potts (2007), but unlike him it argues that this kind of meaning is not compositional. To highlight the difference from Potts (2007), who calls such meaning 'expressive', the present account calls this distinct class of meaning 'associative'. However, I will still employ the term 'expressives' for the class of items which combine compositionally relevant meaning with associative meaning.

Chapter 6. Self-reference in Japanese: a description

This chapter will describe the landscape of how speakers refer to themselves in Japanese. The relevance of this topic to the previous chapters on de se and expressives is that so-called personal pronouns in Japanese are endowed with rich associative meaning. An interesting observation that will emerge from the description of self-reference in Japanese is that unlike languages such as English, Japanese allows proper names and nouns to be used in the same way as personal pronouns. This will be examined in chapter 7, where I will propose a fully-fledged framework of the semantics of expressions used for self-reference in Japanese, i.e. personal pronouns, nouns and proper names.

6.1. Introduction

Japanese is one of the several languages of Southeast and East Asia which have been reported to have paradigms of several pronominal forms for the first person.[96] As we shall see below, while all the forms of the paradigm are used to identify oneself as the speaker in the given context, they also encode other pertinent information about the speaker, such as their gender, age, position related to the addressee, appropriate register, etc.

Pre-theoretically, there are at least four different ways speakers can refer to themselves in Japanese: by personal pronouns, nouns denoting a kinship or professional role, names, or without any overt form (ellipsis). Let's see those one by one.

The topic of personal pronouns in Japanese is rather thorny, with strong disagreements about the exact status of this class of expressions. This debate will be analysed further below, but for now for purposes of exposition I will use the term '1st person pronoun' quite liberally and without any theoretical commitment yet, to mean expressions which are used primarily to denote the speaker. Although most descriptive works in Japanese linguistics mention the existence of various 1st as well as 2nd person pronouns, they confine themselves to listing a few and mentioning that there are more (e.g. Harada 1976, Hinds 1986, Martin 1988, and several others). To my knowledge, the only source that provides a detailed list of Japanese personal pronouns is Tsujimura (1968), who lists 51 forms for the first person and 81 for the second diachronically, while also splitting them into six different historical periods. Although these forms span across hundreds of years, the number is still significantly large; out of the 51 first-person forms, Tsujimura marks 22 that were in use contemporarily to the time he was writing (marked below in boldface).

		Antiquity			Modernity		
		Early Antiquity	Late Antiquity	Early Middle Ages	Late Middle Ages	Early Modernity	Modernity
a	あ						
are	あれ						
onore	おのれ						
maro	まろ						
mi	み						
yatsugare	やつがれ						
wa	わ						
wake	わけ						
wanu	わぬ						

[96] Others being Thai, Vietnamese, Burmese (Cooke 1968, Siewierska 2004), Korean (Song 2006), Indonesian (Flannery 2010), Javanese (Uhlenbeck 1968) and Khmer (Haiman 2011).

ware	われ						
waro	わろ						
onora	おのら						
koko	ここ						
kokomoto	ここもと						
nanigashi	なにがし						
mitzukara	みづから						
warera	われら						
wareware	われわれ						
shousei	小生						
boku	僕						
yo	余 (予)						
ore	おれ						
konata	こなた						
kore	これ						
soregashi	それがし						
wagami	わがみ						
waraha	わらは						
chin	朕						
ora	おら						
kochi	こち						
konokata	この方						
temahe	てまへ						
midomo	みども						
washi	わし						
watakushi	わたくし						
sessha	拙者						
oira	おいら						
kochito	こちと						
kochitora	こちとら						
kochira	こちら						
kotchi	こっち						
watai	わたい						
watashi	わたし						
wachiki	わちき						
watchi	わっち						
wareware	われわれ						
gesetsu	下拙						
atai	あたい						
atakushi	あたくし						
atashi	あたし						
wagahai	わが輩						

Table 6.1. Tsujimura's list of Japanese 1st person pronouns (1968: 389-392)

As Tsujimura (1968) was a historical study on honorifics written in Japanese, it provided minimal information about use and only about a few specific forms (e.g. used by the warrior class, by prostitutes etc.). However, sources written in English usually provide further information about the meaning of these pronouns and how they differ among them, which is mostly defined in different combinations of gender and formality. Below we see two such tables provided by Shibatani (1990) and Shibamoto-Smith (2004).

Gender distinction in pronominal forms

	Formal ⟵			⟶ Informal
1st person				
Male speaker	***watakusi***	***watasi***	***boku***	***ore***
Female speaker	***watakusi***	***watasi***		***atasi***

Figure 6.1. Shibatani's (1990) inventory of Japanese 1st person pronouns (taken from Shibatani 1990: 371)

First-person pronominal forms

	Context			
	Formal ⟵			⟶ *Informal*
Women	*watakushi* (*atakushi*)	*watashi*		*atashi* (*atai*)
Men	*watakushi* (*jibun*)	*watashi* (*washi*)	*boku*	*ore*

Figure 6.2. Shibamoto-Smith's (2004) inventory of Japanese 1st person pronouns (2004: 120)

Nevertheless, as Shibamoto-Smith (2004) stresses, such classificatory tables typically reflect normative usage, rather than necessarily actual. According to sociolinguistic research, the use of 1st person pronouns by Japanese speakers can deliberately transcend conventional gender barriers (Okamoto 1995), for numerous reasons and with different effects. For example, pronoun choice can be a conscious assertion of gender identity for homosexual and transgender people (Lunsing and Maree 2004, Abe 2004, 2010). However, apart from the expression of gender identity, the non-canonical use of gendered pronouns may also have to do the association of traits that are stereotypically seen as male or female; for example, female speakers, especially young ones, may use male pronouns in contexts where they want to compete with boys (Jugaku 1979), accentuate stereotypically masculine virtues such as strength or

assertiveness (Miyazaki 2004), or simply to break with traditional femininity stereotypes without necessarily denying their gender identity (Reynolds 1998). That is to say that the use of gendered-marked expressions does not straightforwardly designate the gender identity of the speaker, but it may communicate it indirectly by implying it. Therefore, the availability of different 1st person pronouns for Japanese speakers to choose from enables them to convey or signal additional meanings that largely exceed mere self-reference. Below is a blog entry by a male native speaker of Japanese, cited in Toff (2009: 124):

> 'When you use *ore*, people think you sound arrogant or unpleasant but when you use *boku*, they think you sound like a nerd or that you're too serious. I personally don't like the sound of *boku*. I don't want to use the b sound here. But at the same time, I feel uncomfortable about using *watashi* in any situation. In English there's only one 'I' that can be used regardless of your position or sex. I envy that. I wish Japanese was like that.'

Besides the plurality of 1st person pronouns to choose from, Japanese speakers consistently use nouns or even their own names with the intention to refer to themselves (Suzuki 1978). For example, when the speaker is talking to a family member that is lower in the family hierarchy, it is customary to refer to oneself using the kin title s/he holds relative to that family member (or the youngest member present in the conversational context, if any) rather than any 1st person pronoun, e.g. 'okaasan' for mother, 'ojisan' for uncle, etc. Sometimes not necessarily the participation, but the mere presence of a junior member in a conversation may trigger such usage. Similarly, self-reference based on title is also possible for specific professions, e.g. 'sensei' ('teacher') (Suzuki 1978, Morita 2009). Nevertheless, both kin and profession titles are used for self-reference only when the speaker is in a hierarchically superior position to the addressee, and not in the opposite situation or when the relationship is symmetrical. That is to say, a daughter or a younger sister cannot refer to themselves as 'musume' ('daughter') or 'imouto' ('young sister') respectively, nor a pupil or student as 'seito' ('pupil') or 'gakusei' ('student') (Suzuki 1978, Shibatani 1990). In slightly older varieties of the language, there were noun-like self-referring forms used by specific professional groups, such as 'honkan' (lit. 'this official') and 'toushoku' (lit. 'this post') employed by government officials and civil servants or, 'gusou' (lit. 'foolish monk') used by monks, etc. (now mostly obsolete but occurring in historical drama and animated series). Interestingly, these forms are listed in Japanese dictionaries as 'personal pronouns employed by policemen/officials/monks' (see e.g. Daijirin dictionary, Kadokawa dictionary).

Moreover, one's name can also be used for self-reference (Lunsing and Maree 2004). This is a practice common with young children (Clancy 1985) and also exploited by younger women for the child-like and cuteness connotations of this form of self-reference (Maynard 2016). All in all, Japanese affords itself a significant array of 1st person pronouns, as well as routinely allow the use of nouns and proper names for self-reference. As was pointed out by Morita (2009: 177), all these self-referring strategies can be exploited by the same speaker as engaging in different discourse instances:

> For example, a mother may refer to herself as *okaasan* 'mom' or *mama* when talking to her child, but may use *obachan* 'aunty' when she is talking to her child's friend. A few minutes later, she may refer to herself as *oneechan* 'elder sister' when she is talking to her younger sister. If the same speaker is an elementary school teacher, she may use *sensei* 'teacher' when speaking to her students, as in *sensei no ato ni tsuite itte kudasai* 'repeat after teacher (me),' but *watashi* or *watakushi* when conversing with the schoolmaster. In addition, she may refer to herself as *atashi* to a female friend, but as *watashi* to an interlocutor who is just an acquaintance. There are dialectal variations as well. [...] Each term for 'I' can index (among other factors) the speaker's gender, age, status, and regional origin; the formality of the context; the intimacy of the relationship between the interlocutors; and the position of the speaker vis-à-vis her immediate interlocutor as well as the larger community.

Despite all these possible choices for self-reference in Japanese, it so happens that very often none is chosen. Japanese is a pro-drop language, and naturally personal pronouns are omitted as often as possible, to the extent that frequent use of 1st person pronouns is considered egocentric and unnatural (Hasegawa 2014), while the use of a 2nd person pronoun can be seen as contemptuous or even rude towards the addressee (Azuma 2000). Since all forms to refer to oneself or to address the hearer are associated with some connotation or other, speakers may deem it safer to avoid forms that would commit them to certain implications. As Ishiyama (2008: 218) notes, 'the socio-pragmatic implicature is responsible for both proliferation and avoidance of personal pronouns in Japanese'.

As this preliminary exposition aims to show, in Japanese discourse a speaker can refer to him/herself via diverse means: 1st person pronouns, nouns, names, or implicitly without using any overt form.[97] Unlike languages with verbal agreement for person, in Japanese this shift among different forms for self-reference happens at no cost of shifting the grammatical person. Even in English, a language with considerably impoverished morphology for marking person agreement on the verb, only the 1st person pronoun ('I') licenses 1st person verbal agreement, whereas any other form of self-reference, such as a noun to describe or a name to denote oneself, obligatorily licenses 3rd person agreement, giving the impression that the person is speaking about him/herself as if s/he were someone else (what is known in literary studies as 'illeism'). In such languages illeism is seen as an extremely marked option to talk about oneself and is often interpreted as odd, or even as arrogant and narcissistic.[98] However as we

[97] Of course, Japanese employs a number of other mechanisms to facilitate the disambiguation of reference given the combination of dropping the subject together with the lack of person agreement. Some of these are the sophisticated honorifics system which allows the speaker to distinguish between in-group and out-group, the grammatical marking of 'territory of information' (Kamio 1995) which encodes whether evidence for psychological events such as sensations and emotions has been acquired directly or indirectly, etc. As noted by Shibatani (1990: 364) 'in Japanese [...] a fair number of phenomena, rather than a single phenomenon such as agreement, conspire to bring about a situation in which missing nominals, or pros, can be identified'. For a detailed study about how Japanese compensates for ellipsis and a proposed algorithm of how reference is tracked, see Nariyama (2003).

[98] A typical and recent example is the public's astonishment and mockery of public figures such as Donald Trump and Kanye West talking about themselves in the third person, and its

saw in Japanese, speakers shift between different forms of self-reference routinely and much less markedly.

6.2. The data

As noted above, the number of forms used for self-reference in Japanese can be quite large. With regards to 1st person pronouns, despite the 51 diachronically attested in Tsujimura (1968), it is possible for this number to change depending on the criteria we adopt for what counts as a pronoun, as we shall see below in the grammatical analysis. Certainly, the number of nominal expressions (such as kinship and professional terms) which can be used for self-reference cannot be defined exhaustively, and obviously the same goes for names. This is mentioned to clarify that this study will propose an analysis for each category of expressions used for self-reference, using specific expressions from each category as examples but meaning to be applicable to other expressions as well. That is to say, I aim to offer a general framework of how self-reference works in Japanese on the basis of representative examples, hopefully powerful enough to extend to expressions I will not explicitly mention.

The dialectal variety chosen is that of standard Tokyo Japanese, and the expressions selected are based primarily on the ability of speakers of this variety to comprehend and secondarily to produce them; this caveat is introduced so as to allow us to study some interesting items which have largely fallen out of active use by speakers, but are still understood by them when encountered in dramas, animated series or comics. The information about the meaning and use of these expressions comes from both academic (e.g. sociolinguistic or historical studies) and non-academic sources. The latter refer to searches in Google and in the online corpus Kotonoha Shonagon[99], as well as to solicited and unsolicited native speaker intuitions. Most of the latter have been taken from online discussions on language use at the Japanese version of Yahoo Answers (Yahoo Chiebukuro), where many native speakers engage in lively discussions over the meaning and use of 1st person pronouns and other expressions used for self-reference.

The 1st person pronoun-like forms are described first. Although not the same amount of information is available for all the terms, given that some are more common or versatile than others, for reasons of structural clarity the information available is presented in three parts: gender and register, written form, and usage. Next are nouns, such as kinship and professional terms, for which I describe their stand-alone meaning and usage. Lastly I describe the use of names as devices of self-reference.

6.2.1. 1st person pronouns

---

association of this phenomenon to extreme arrogance and narcissism: (http://www.bbc.co.uk/news/magazine-33943762 , http://www.ballerstatus.com/2009/02/17/kanye-west-i-have-the-right-to-talk-in-the-third-person-im-a-walking-brand/). Nevertheless despite laypeople opinion, scientific research has found no correlation between illeism and narcissism (Raskin & Shaw 1988).

[99] http://www.kotonoha.gr.jp/shonagon/search_form

1) ***watashi***

Gender and register: Canonically used by both male and female speakers, but with a difference in register; it is perceived as formal when used by a man while it can be either formal or informal/casual when used by a woman (Shibatani 1990, Hasegawa 2014). It is the most common and most unmarked 1st person pronoun in Japanese.

Kanji (Chinese character): *Watashi* originates from *watakushi* (see below) and it is still written with the same Chinese character as *watakushi* (私, which means 'private'). As such, it is only distinguished from *watakushi* in oral discourse or when written in the hiragana syllabary.

Usage: Its use in casual discourse is associated with femininity, while its use by a man gives the impression of formality. As a result, it cannot be said that *watashi* is truly unmarked in the sense of the English 'I' which is used indistinctively by speakers of any characteristics. Some men consciously use *watashi* outside formal contexts in order to avoid the particular connotations of the male-specific pronouns (see below). In general *watashi* is the most neutral way to refer to oneself compared to all the other choices available, only because it has fewer connotations than the other forms and not because it comes without any connotations of its own.

2) ***watakushi***

Gender and register: Gender neutral and extremely formal (Hasegawa 2015).

Kanji (Chinese character): Its Chinese character (私) literally means 'private'.

Usage: It is the standard form used in the interaction with hierarchically superior interlocutors and it conveys humility and deference (Daijirin dictionary). However, it may also be used when the speaker is superior to the addressee in order to signal extreme formality, which is why it can also convey the idea that the speaker is of a high social class; it is the pronoun used in public by the current emperor (Shillony 1999).

3) ***boku***

Gender and register: Male, relatively informal to casual (Hasegawa 2015).

Kanji (Chinese character): Its Chinese character (僕) literally means 'manservant'.

Usage: Although casual, it does not sound impolite in formal contexts. It is associated with a somehow soft and boyish attitude rather than a manly one, and it can sound a bit childish and not masculine enough most likely because of the competition with the other male pronoun, *ore*, which is more strongly masculine (see below). Interestingly, because it is the 1st person pronoun stereotypically used by little boys, it can also used as a 2nd person pronoun to address a little boy (Ishikawa et al 1981, Clancy 1985). Its use by a girl usually gives a strong tomboyish impression, although the reasons why a female speaker may use it vary from expressing a male gender identity (Asada 1998),

exhibiting stereotypically masculine virtues such as strength or assertiveness (Miyazaki 2004), especially in contexts of competition with boys (Jugaku 1979), or simply to break with traditional femininity stereotypes without necessarily denying their gender identity (Reynolds 1998).

4) ***ore***

Gender and register: Male, extremely informal (Hasegawa 2015)

Kanji (Chinese character): Its Chinese character (俺) means 'I'.

Usage: It is used with peers of equal status or lower (Kadokawa dictionary), which makes its effect depend on the occasion of use; it can sound rude or even vulgar on occasions when the speaker is not already quite intimate with the addressee, making the speaker sound rough, macho or even aggressive (Kadokawa dictionary), but when there is a certain familiarity it can signal that the speaker is on equal terms and feels quite intimate with the hearer(s), e.g. being friends or in a close relationship (it is often used by men when speaking to their friends or partners, Tsujimura 1996). If it is used by a female speaker, it sounds extremely rough.

5) ***atashi***

Gender and formality: Female, relatively informal and casual (Hasegawa 2015).

Kanji (Chinese character): None; it originates from *watashi*, and it does not have a Chinese character itself.

Usage: It is strongly associated with femininity. Because of this, it is one of the pronouns of choice of gay men (Lunsing and Maree 2004, Abe 2004, Toff 2009).

6) ***washi***

Gender and formality: Male.

Kanji (Chinese character): Associated with two different Chinese characters which both mean 'I' (儂 or 私, the second is the same as for *watakushi/watashi*).

Usage: Although now fallen out of active use, it was typically used by men of approximately 50 years of age and older with intimates and addressees of lower status (Shibamoto 1985). As such, it can convey an idea of superiority or patronising attitude, especially if the addressee is younger.

7) ***jibun***

Gender and register: Male, formal.

Kanji (Chinese character): 自分 (lit. 'self-part'). Note that *jibun* is a reflexive pronoun, meaning precisely 'oneself', but may also be used as a 1st person pronoun.

Usage: Nowadays it is mostly used by male athletes (Abe 2004, Hasegawa and Hirose 2005). Traditionally it was the standard way of self-reference for soldiers (Toff 2009), which gives it a hierarchical and old-fashioned flavour when used nowadays, but its use was officially discouraged after WW2 both in the army and public services by the government-issued *Korekara no keigo* ('Honorific expressions in the future', 1952). Although it has been traditionally used by male speakers, it may be perceived as gender neutral because of its literal meaning of 'self' and be used by speakers who wish to avoid gender connotations, e.g. women (Toff 2009), lesbians or transgender men (Abe 2010).

8) ***atai***

Gender and register: Female pronoun, very informal; a more informal variant of *atashi* (Kadokawa dictionary).

Kanji (Chinese character): None; it is essentially a corrupted form of *atashi* (see above).

Usage: Originally used by women and girls in the Tokyo red light district (Daijirin dictionary). It is not commonly used nowadays, but speakers understand and perceive it as associated with a 'bad girl' image, e.g. delinquent female characters in comics and films. It is also one of the pronouns used by gay men (Abe 2004). Probably because it is somehow archaic now, it has also been reported as having a 'hillbilly' connotation in urban areas (Miyazaki 2004: 271).

9) ***uchi***

Gender and register: Female, casual or even colloquial.

Kanji (Chinese character): It literally means 'inside' (内) or 'home' (家) (Miyazaki 2004). Despite its origin, it is mostly written in the hiragana syllabary than with a kanji.

Usage: Originally from the dialect of Western Japan (Kansai area), but in the past few years its use has spread to standard Japanese and it has gained popularity with young girls in urban Japan (Toff 2009). Although it is used mostly by female speakers, it is perceived as less overtly feminine than *atashi* and more casual to use with friends (Miyazaki 2004).

10) ***boku-chan***

Gender and register: male, casual to colloquial level.

Etymology: It is a combination of the male 1st person pronoun *boku* with the title suffix –*chan*, which is used to 'express intimacy and familiarity between close friends and for children' (Bowe et al 2014: 152). However, in general title suffixes in Japanese are not felicitously used for oneself but only for non-speaker participants, as is the case with all honorifics (Brown 1996). As noted above, *boku* is sometimes used as a 2nd person pronoun to address a little boy, and *boku-chan* as a mix of a 1st person pronoun with a title suffix seems to be an overt product of this process.

Usage: Used by a 'male infant' (Hasegawa 2015). Just like *boku*, it can also used as a 2nd person pronoun to address a little boy (Suzuki 1978).

11) ***ore-sama***

Gender and register: Male, extremely informal.

Etymology: A combination of *ore* (see above) with the title suffix *–sama*, which is typically used to address people of much higher rank and to display deference (Bowe et al 2014). Miyazaki (2004: 260) renders it to English as 'Mr. I' or 'the honourable I'.

Usage: Arrogant and self-elevating (Daijirin dictionary). It sounds 'boastful and roguish' (Takeuchi 1999: 64) but because of the mix of the 1st person pronoun with an honorific title suffix it may feel like a 'jocular possibility' (Whitman 1999: 377) used in order to achieve certain effects, e.g. in comics and animated series (Maynard 2016).

6.2.2. Nouns

6.2.2.1. Kinship terms

okaasan, otoosan, mama, papa, obasan, ojisan, obaasan, ojiisan, oneesan, oniisan, etc.

('mother', 'father', 'mummy', 'daddy', 'aunt', 'uncle', 'grandmother', 'grandfather', 'elder sister', 'elder brother')

Meaning: The literal meanings of these kinship nouns are given above, but it should be clarified that apart from *mama* and *papa*, all the rest are respectful forms; if we break them down, they have the honorific prefix *o–* in the beginning and the honorific title suffix *–san* in the end. The former roughly contributes the meaning 'honourable' to the words it is attached (Nariyama 2003) while the latter is added to terms referring to other people, never to oneself, to convey 'a greater degree of formality' (Bowe et al 2014: 152)[100].

Usage: All of these kinship terms are used for the speaker to refer to him/herself when talking to a junior member of the family with whom s/he bears the relation attested in the self-referring expression; in such situations, the use of personal pronouns to refer to

[100] All these expressions in their exact form are used to refer to an addressee with whom the speaker bears the specific kinship relation (e.g. one refers to one's mother as *okaasan*, etc).

oneself is infelicitous (Loveday 1986, Alpatov 2006). However, in some conversations with family members a speaker may use a self-referring term that denotes his/her relation to the most junior member of the family even if that member is not the addressee or even present in the conversation. This means that a wife and a husband might refer to themselves as *okaasan* ('mother') and *otoosan* ('father') respectively even between them when their children are not present (Ishikawa et al 1981, Alpatov 2006), or a grandmother or grandfather might refer to themselves as *obaasan* ('grandmother') and *ojiisan* ('grandfather') when talking not just to their grandchildren, but to their children as well (Suzuki 1978). Moreover, it is possible for a speaker to use a kinship term such as *obasan* ('aunt') or *ojisan* ('uncle') to refer to him/herself in situations where no such kinship relation to the addressee applies, as long as the addressee is a child or quite junior in age (Suzuki 1978).

Interestingly, all the kinship terms presented here which can be used by the speaker to refer to him/herself reflect superior positions in the family hierarchy. We do not find terms such as *musume* ('daughter'), *musuko* ('son'), *mago* ('grandchild'), etc. The reason why this is so will be evident after the explanation provided below.

An explanation of how kin role-based self-reference works is what Suzuki (1978) calls 'empathetic identification'. According to this notion, a speaker chooses a term of self-reference based on the perspective of someone else, with whom s/he empathises. In the context of the family, the perspective that is chosen is that of the youngest member in the family. As Suzuki explains (1978: 122):

> In the Japanese family, an older member may address a younger member, using the kinship term designating the latter's position from the viewpoint of the youngest member of the family. When an older family member, in a dialogue with a younger member, mentions a third member who is older than the latter, he does not view this person linguistically from his own standpoint, but from the standpoint of the hearer, that is, the youngest member.

Therefore, an older member refers to him/herself using the exact same term the youngest member of the family would use to address him/her, not even a stylistic variant of the same form.[101,102] According to Whitman (1999), the use of titles (including kin terms) is obligatory in Japanese and several other Southeast Asian languages when addressing social superiors, while 2nd person pronouns are infelicitous and inappropriate as they are used primarily to address equals and inferiors.[103] As such, the

[101] As Suzuki (1978: 109) clarifies, 'if a particular child addresses his father as 'papa', the latter calls himself 'papa' too, and not 'otoosan', when talking to the child'.

[102] However, there is some indication that this pattern does not apply as much nowadays in the relationship between siblings; although younger members might call their older siblings as *oniichan* ('elder brother') or *oneechan* ('elder sister'), the older siblings use (informal) pronouns to refer to themselves when talking to their younger siblings (Yasutada Sudo, p.c.). I think this has to do with a change of attitudes about family hierarchy in Japanese society, and the relationship between siblings seems easier to affect due to being the least hierarchical one within a family (e.g. compared to that between parents and offspring).

[103] The reason why titles are used to address superiors and pronouns to address inferiors can be easily explained by the fact that the main function of titles is deference (even in English, 'Mr. Jones' is a more respectful way to address someone than merely 'Jones') and they are often associated with prestigious roles (e.g. English *Dr.*, Spanish *Don*, etc), while that of pronouns is

only way the youngest member of the family can address an older member is using the respective kinship term, and according to the idea of empathetic identification it is the only way the older member can refer to him/herself in the context of communication with the youngest member. However as it was mentioned above, a husband and wife might refer to themselves as *otoosan* ('father') and *okaasan* ('mother') respectively even when their children are not present, which means that the idea of empathetic identification may persist even if the youngest member is not the addressee but is merely present in the conversation, or even if s/he is absent.

The idea of empathetic identification also accounts for why the 1st person pronoun *boku* and its variant *boku-chan* can be used as terms of address for a young boy: the (older) speaker emphasises with the addressee (the young boy) to such a point, that s/he calls him the way the young boy refers to himself. Below is an example of a mother addressing her son as *boku* (Clancy 1985: 454):

(189) Boku, Puu-san mi-te.

I Pooh HON look IMP

Lit. 'I, look at Pooh'

'Look at Pooh.'

Whitman (1999) mentions that the female 1st person pronoun *atashi* can also be used in the same way to address a little girl, but that no other 1st person pronouns can be felicitously used in this manner (however, I have been told that the use of 'watashi' is also possible in this case of addressing a girl, Yasutada Sudo, p.c.) This follows from the idea of empathetic identification, as *boku*, *boku-chan* and *atashi* are all pronouns typically used by more junior/younger people. It is interesting to note here that although Suzuki (1978) originally introduced empathetic identification as a uniquely Japanese phenomenon, it also applies to languages other than Japanese. One example is the use of the 1st person pronoun to address a young child in French, noted by Aoki (1989). Shibasaki (2014: 136) gives the following French example:

(190) Je reste assis. Je ne bouge pas. Je suis un bon garçon.

I stay sitting. I do not move. I am a good boy.

'You must remain seated. You must not move. You must be a good boy.'

Another phenomenon that seems to be very similar is what has come to be known in English as 'nurse we' (Collins & Postal 2012), though it has also been noted in Italian (Servidio 2014), German (Roehrs 2006) and Greek (Holton et al. 2012). The following example is from Collins and Postal (2012: 34):

---

merely reference. Of course there are respectful variants of 2nd person pronouns in several languages, such as the Spanish *usted*, the French *vous*, etc.; interestingly, although such forms are used for the 2nd person singular, grammatically they correspond to either the 3rd person or a plural version of the 2nd person, which can be explained in terms of the association of obliqueness and indirectness with politeness (Wales 1996). Nevertheless, in Japanese there is practically no 2nd person pronoun that is more polite than the use of a title reflecting the addressee's social role, especially when that is superior to the speaker or when the speaker wants to be polite (Hori 1995), and as such pronouns are used only for equals or inferiors.

(191) Are we feeling better today? (nurse to single patient)

6.2.2. 2. Professional terms

1) ***sensei***

Meaning: The combination of its Chinese characters (先生) literally means 'the one who comes before'. The word itself means 'teacher', but it is a common way to address not just teachers but also other prestigious professionals as well such as medical doctors, politicians, lawyers, etc. (Ishikawa et al 1981). Because it can be used as an honorific title (also together with someone's surname, e.g. a teacher whose surname is Sonoo can be called 'Sonoo sensei'), it is not combined with any honorific title suffix such as *–san* or *–sama*.

Usage: It can be used by a teacher of any level to refer to him/herself but only when the addressee is his/her pupil.[104]

There are a few other professional terms that can be used by the speaker to refer to him/herself, but those are only used when the addressee is a child. These are the following:

2) ***oishasan***, 'doctor'

3) ***kangofusan***, 'nurse'

4) ***omawarisan***, 'policeman'

5) ***obousan***, 'priest'

It is noteworthy that just like in the case of kinship terms, also the above expressions contain honorific elements: the honorific prefix *o–* in the beginning (with the exception of *kangofusan*) and the honorific title suffix *–san* in the end. It seems that what is at play also here is the idea of empathetic identification (Suzuki 1978) that was used above for kinship terms, as all of the professional terms that can be used for self-reference denote a superior hierarchical position, and the speaker can only employ them when the addressee is junior.

6.2.3. Names

[104] It should be noted that it's uncommon for a university professor to call themselves 'sensei' or 'kyouju' ('professor'), while they are usually called such (university students usually call them the latter). The relevant factor might be that this type of self-reference sounds patronizing, and given that university students are considered adults rather than children, it is probably somehow inappropriate to use 'sensei' or 'kyooju' for self-reference in such circumstances. For the same reason a university professor wouldn't refer to themselves by either of these terms when talking to university staff, although university staff call them such.

As reported by Clancy (1985), Japanese children use their names to refer to themselves, as they acquire personal pronouns later than children of Indo-European languages. Citing Horiguchi (1979, 1981), Clancy notes that the most common forms of self-reference used by children is a nickname plus the diminutive suffix *–chan* (which belongs to the same paradigm as the honorific title suffixes *–san* and *–sama* noted above, and just like them *–chan* is not canonically used for oneself). She suggests that this practice is related to the high level of complexity of the Japanese addressing system: as explained above, family members do not use personal pronouns to refer to themselves when talking to children but kinship terms, which means that the child has limited exposure to 1st person pronouns. Moreover, the sociolinguistic complexity of the Japanese pronominal system is also a great difficulty for children. As such, the easiest and most common way for the child to refer to him/herself is the way s/he is called by adults, i.e. by name or nickname, with or without the diminutive suffix *–chan*[105].

Although the practice of using one's name for self-reference by children is motivated by practical issues, it happens that it is also adopted by some adult speakers, specifically by young women (Maynard 2016). According to Maynard, speakers give the following two reasons for this practice: first, to avoid the formality associated with *watashi* and second, to pose as a child in order to appear cute and attractive particularly towards men[106]. To summarise, the use of one's name for self-reference in Japanese seems to be confined to children and young women, and it is not observed in men.

6.3. Current analyses of Japanese personal pronouns

As we have seen above, Japanese has a multitude of forms that resemble 1st person pronouns; after all, they are used by speakers to refer to themselves, their most straightforward and common translation to any foreign language is the 1st person pronoun of that language, and Japanese dictionaries define them as 1st person pronouns (e.g. Daijirin dictionary, Kadokawa dictionary, etc). However, there is a very big debate as to whether these 1st person pronoun-looking expressions are really personal pronouns, or nouns in disguise used for specific purposes. It is to this that we now turn.

6.3.1. General views on personal pronouns

First, I will examine some prominent views about what personal pronouns are, and see what the implications are for the Japanese class of expressions. Second, I will discuss specific studies on Japanese personal pronouns and assess their claims.

Traditionally, personal pronouns are grammatical items whose function is 'to indicate the two principal speech roles, namely that of 'being the speaker' and 'being the addressee' respectively' (Bhat 2004: 6). It is important to note the distinction between denoting the individuals who undertake the roles of speaker and the addressee in a

[105] Children usually switch from names to 1st person pronouns as soon as they enter nursery school, when they start socialising with other children (Clancy 1985, Ide 1990).
[106] The practice of an adult woman pretending to be a child is called *burikko*, which literally means 'fake child' (Miller 2004). It ranges from baby talk (e.g. high pitch) to the way one dresses or even behaves (e.g. appearing embarrassed, innocent, helpless, etc).

given discourse and denoting the discourse roles of speaker and addressee; the former can be done using any noun whereas the latter is considered to be the primary function of personal pronouns. This function is possible due to the ability of personal pronouns to act as 'shifters' (Jespersen 1923), meaning to vary their reference continuously during the flow of discourse so as to track the speech participants. Generally, personal pronouns are said to have the following characteristics: unlike nouns, they do not carry lexical meaning but only denote person deixis, they are not usually combined with modifiers or determiners, and they are considered a closed class, that is they are limited in number and they are fairly stable diachronically (Heine and Song 2011). These characteristics of personal pronouns enable them to have the function they do; first, if they carried lexical meaning on the basis of which their referents could be identified, then their ability to function as shifters would be severely limited to specific participants. Similarly, as modifiers and determiners are used to facilitate the identification of referents while personal pronouns are assigned with denoting discourse roles than referents in particular; it is expected that the modifiers and determiners will not be used together with personal pronouns (Bhat 2004). Finally, it would be 'dysfunctional' (Siewierska 2004: 2) for languages to have a wide array of forms to denote speech roles, as these are crucial for following the discourse.

Despite Benveniste's (1971) claim that one cannot imagine a language without the expression of person, under these general characteristics of what personal pronouns are, things already start to look dire for the Japanese expressions. First, several of them seem to carry meanings on the basis of which their referents can be if not identified at least severely narrowed down, e.g. the use of 'atashi' points that the speaker is very likely to be female, or at least perceive oneself as such. Moreover, they can be used with modifiers or determiners, e.g. 'bikkuri shita watashi' ('surprised me'), 'kono watashi' ('this me'), while they also form the plural in exactly the same way as nouns, e.g. 'watashi-tachi' (I-plural, 'we') is similar to 'tomodachi-tachi' (friend-plural, 'friends') (Siewierska 2004). Finally, given their large number and numerous variations, it is very hard to consider them a closed class. The fact that Japanese personal pronouns, as well as those of other Southeast and East Asian languages such as Thai, do not confirm these characteristics has led several researchers to suggest that these languages do not have real personal pronouns and that these expressions are actually nouns (e.g. Suzuki 1978, Mühlhäusler and Harré 1990, Longobardi 2008). Discussing such doubts, Bhat (2004) proposes that answering the following questions could settle the debate: whether a description is needed to disambiguate the referents of these expressions in narration (cf. the English 'I' in narration does not allow the reader to understand who is speaking, and the narrator needs to clarify that), whether they can take complements and modifiers which would enable them to identify their referents, and whether appository phrases are needed after these expressions in oath-taking contexts so as to identify the referent, as in English.

However, it turns out that even answering these specific questions for Japanese still does not bring us closer to a definite solution. With regards to the first question about whether specifying the speech participants using names or other descriptive expressions is necessary in narration, this can be the case in Japanese, but not always. That is to say, if the two speech participants are of the same hierarchical position and

gender, then they could be using the same 1st person pronoun and therefore a name or description would be needed so as to distinguish them, when one narrates the dialogue. However, if there were differences in gender or hierarchy, then they would most likely be using distinct expressions, thus no further specifications would be necessary. As far as the second question goes, modifiers and complements seem to be tolerated before person terms in Japanese (Noguchi 1997), which should aid in differentiating between speech participants. And lastly, the name of the oath-taker is required to come after the self-referring expression during an oath, because although the choice of 1st person pronoun can provide some information about the speaker (e.g. gender), it absolutely does not identify the referent up to the binding standards required by an oath.[107] Therefore again we find that Japanese person terms converge with Indo-European personal pronouns in some aspects, such as not being able to establish the referents without a description in narration and the name in an oath-taking context, while in others they diverge from them, such as the fact that they can accept modifiers which can facilitate identifying the referent.

As Siewierska (2004) notes, the answer to whether the personal pronoun-like expressions of these languages are really pronouns or not depends on one's definition of what personal pronouns are, which may vary in different approaches. In the functional literature, although pronouns are distinguished from nouns, it has been proposed that the distinction between personal pronouns is not absolute but gradual, and the personal pronouns of different languages may be in different points of a 'pronominality scale' (Sugamoto 1989).

(192) Sugamoto's 'pronominality scale' (1989), cited in Siewierska (2004: 9)

i. closed class membership
ii. lack of morphological constancy
iii. lack of specific semantic content
iv. lack of stylistic and sociolinguistic implicative properties
v. expression of grammatical person
vi. inability to take modifiers
vii. restrictions on reference interpretation

Applying these pronominal criteria to Japanese, and given the discussion before, it turns out that it doesn't confirm most of them. If we use this scale for cross-linguistic comparisons, we will be able to place different languages in a noun-pronoun continuum. Siewierska (2004) specifically does this for English, Polish, Japanese and Thai to find out that albeit the vicinity between the first two and last two, all these languages differ among them and they are in different points on the pronominality scale:

[107] Example from a website of a wedding chapel giving instructions about what to say during the ceremony (http://crystalwd.wix.com/crystalwedding#!about-2/ca7f):
'Watashi, ___ (shinrou no namae)-ha, anata-ni no sono yasashii kokoro –wo itsumademo taisetsu-ni shi shougai-wo kakete shiawase-ni suru koto-wo chikaimasu'
'I, ___ (groom's name), vow with all my heart to cherish you and make you happy for the rest of my life'

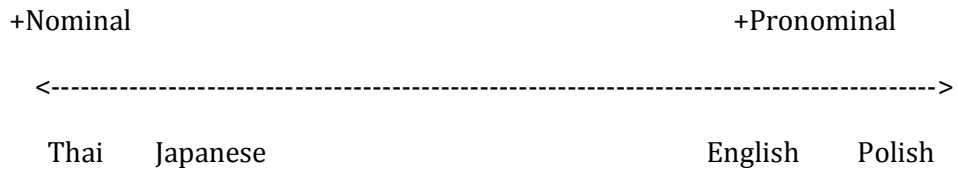


Figure 6.3. Sugamoto's (1989) Pronominality scale (Siewierska 2004: 9)

The idea of the pronominality scale helps to graphically illustrate the fact that Japanese personal pronouns have more nominal characteristics than English pronouns, but they still manifest several pronominal characteristics as well, and in fact more than other languages such as Thai. Although this is an unquestionable virtue of the pronominality scale, some issues should be pointed out. As Ishiyama (2008) notes, the characteristics in the list are mentioned without any hierarchical order, while clearly some are more important than others. With regards to i (closed class membership), Ishiyama argues that the reason the lists of personal pronouns in Japanese (e.g. Tsujimura 1968) appear so inflated is because they include demonstrative forms which may be used to denote the speaker or the addressee but which do not denote the discourse *roles* of speaker and addressee. Specifically, Ishiyama (2008: 144) argues that 19% of the 1st person and 25% of the 2nd person forms in Tsujimura's (1968) list should not be analysed as personal pronouns but as demonstratives (although he does not specify which).

However, even after removing the demonstrative-based items Tsujimura's list remains significantly large, especially from an Indo-European perspective. But this also has an explanation: according to Whitman (1999), personalisation (the grammatical process by which an expression which did not previously denote a discourse role comes to do so) is much more frequent in languages which lack person agreement, such as Japanese. This is because complete personalisation requires a change in agreement pattern so that the non-personal origins of the new personal forms are no longer recognisable, and given that this step is unnecessary in languages without person agreement the process is much easier in those. That is to say, this typological characteristic of Japanese has made it much easier not only for nouns to evolve into pronouns, but even to import pronouns from foreign languages: for example *boku* is one of the several Chinese pronouns that came into Japanese[108,109] and *mii* ('me') is a novel import from English (Morita 2009).

[108] It is interesting to note that older varieties of Japanese had small and closed paradigms of personal pronouns: pre-Old Japanese (before 700 AD) had only one form for the 1st person (*a*), while Old Japanese (700-800 AD) had 2 short (*wa*, *a*) and 2 long forms (*ware*, *are*). Moreover, 1st and 2nd person pronouns were used frequently in Old Japanese, much more than in subsequent stages of the language (Frellesvig 2010). This seems to have changed during the period of Early Middle Japanese (800-1200 AD), during which 'there is no longer a system of personal pronouns as such, but an inventory of terms of address and of self-reference' (Frellesvig 2010: 245). This development has been dubbed 'depronominalisation' by Vovin (2003: 95), and it should be stressed that it took place during a period when Japan was under heavy cultural influence from China, which was also reflected on the extensive import of Chinese elements into the Japanese language, a process known as 'sinification' (Frellesvig 2010).

[109] Another Chinese import is the notorious 'chin', a personal pronoun for the exclusive use of the emperor (Keene 2013), whose peculiar nature has been mentioned by Nunberg (1993) and Abbott (2010). Nowadays 'chin' is entirely obsolete, as the current emperor has adopted

Furthermore, Ishiyama observes that ii (lack of morphological constancy) is typologically biased, as it is a characteristic of Japanese to show case via particles instead of having suppletive forms. He also finds that iv (lack of stylistic and sociolinguistic implicative properties) is problematic, as it would exclude polite forms in languages such as French, Spanish and German which are clearly pronominal. Moreover, he mentions that the tolerance of modifiers and determiners by Japanese pronouns in actual discourse is rare (as shown in an empirical study by Shibasaki 2005). As such, he concludes that features iii (lack of specific semantic content) and v (expression of grammatical person) are the most crucial ones for the notion of a personal pronoun: the role of personal pronouns is to denote discourse roles and in order to do so successfully they need to be devoid of semantic content.

In the generative approach, personal pronouns have been identified not as a separate syntactic category but as a kind of determiner which is composed by a set of φ-features such as person, number and gender (Postal 1969). That is to say, the phrases headed by personal pronouns are not pronoun phrases, but determiner phrases (DPs). Following this, Noguchi (1997) proposes that personal pronouns constitute a feature which distinguishes between different sets of lexical items, that of referential dependency. This feature means that the identification of the referents of pronouns is determined by the extra-linguistic context, which contrasts them both with common nouns and proper names; common nouns are semantic predicates able to identify a referent only via a determiner, and proper names are able to identify the referent by themselves (they are inherently referential). This proposal allows Noguchi to capture the difference between English-like and Japanese-like personal pronouns by identifying the former as syntactically determiners (D-pronouns), while the latter are syntactically nouns (N-pronouns).[110] This means that although there is a difference in categorial status between the two, it is not a difference of pronominal status as pronouns do not form a syntactic category anyway (there is no such phrase as a 'pronoun phrase'). Thus English-like personal pronouns head DPs, while Japanese-like personal pronouns head NPs. Noguchi's proposal successfully accounts for the intuition that Japanese has words that function like personal pronouns, since they display referential dependency in the same way as their English counterparts, but moves the valid observation about their noun-like characteristics to the level of syntactic structure. Importantly, this account successfully explains why Japanese personal pronouns lack a prime characteristic of pronouns, namely the possibility to be construed as bound variables, a fact noted by several researchers such as Nakai (1976), Kitagawa (1981), Nakayama (1982), Saito & Hoji (1983) and Hoji (1990, 1991).

A similar account is that of Longobardi (2008), who suggests that Japanese personal pronouns do not have the φ-feature of Person in their syntactic representation but only in their conceptual or pragmatic representations. This follows from Longobardi's proposal that D is where the φ-feature of Person is assigned (based on observations of

'watakushi' (Shillony 1999), but it has left its permanent trace in the Japanese rendition of the French phrase of Louis XIV 'L'état, c'est moi' ('I am the state'), which is 'Chin-ha kokka nari'.

[110] This also enables Noguchi to account for the fact that Japanese 3rd person pronouns cannot be construed as bound variables like their English counterparts, positing that only functional items (D-pronouns) can be bound, whereas lexical ones (N-pronouns) cannot.

the distribution of pronouns, proper names and common nouns in English and Italian), which combined with Kuroda's (1988) observation that Japanese has no person agreement, leads him to conclude that Japanese lacks the category D altogether. The predictions of this hypothesis are confirmed in Japanese, of which the most relevant for our discussion is the fact that pronouns, proper names and common nouns have the same distribution. This conclusion is similar to Noguchi (1997) in the sense that it doesn't dispute the ability of these expressions to denote discourse roles, but it postulates that this will not be dictated by the syntactic structure but by other factors such as pragmatic constraints.

To summarise, Japanese so-called personal pronouns are quite different from their counterparts in Indo-European languages. They resemble nouns in a number of ways: they have the same distribution in the sentence, they take the same morphological markers for plural, they may tolerate modifiers and determiners. Moreover, their significant number (which persists even if we exclude demonstrative forms) makes it difficult to consider them a grammatical paradigm, and they encode meanings (such as gender, age, etc) that may potentially constrain their ability to function as shifters. Nevertheless, it is undeniable that these expressions encode the speaker role; that is to say, although Japanese has a variety of ways to refer to oneself, including ellipsis, in case it needs to be made clear who the speaker is the only certain way to disambiguate is using one of these so-called 1st person pronouns. In the example of taking an oath (footnote 96), in order for the groom to be bound by it, he needs not only to state his name but also a 1st person pronoun (*watashi* in our example of a relatively formal context) in order to unambiguously state that he himself is taking a binding oath. Therefore, although there is a plethora of valid reasons to say that Japanese so-called pronouns are not really syntactically represented as pronouns but as nouns (as per Longobardi), it is important to acknowledge that they exhibit the feature of referential dependency in the same way as their counterparts in other languages (as per Noguchi). However, specifically with regards to personal pronouns, the characteristic of referential dependency needs to be further constrained by the notion of person deixis, in order to be able to distinguish the former from demonstrative pronouns or adverbs (as per Ishiyama). As such, because Japanese have referentially dependent expressions whose purpose is to denote the speech roles of speaker and addressee, disregarding the element of syntactic distinction, I take it that Japanese has a category of expressions which can be called personal pronouns.

6.4. Summary of data

After presenting the data, it seems that pre-theoretically they can be distinguished in three preliminary categories: personal pronouns, nouns (including both kinship and professional terms), and names. From the remaining items in our preliminary category of 1st person pronouns, it is argued that two of them should not be classified as distinct forms but as variants. These are *boku-chan* and *ore-sama*, which as mentioned in the data description are combinations of the existing 1st person pronouns *boku* and *ore* with

the title suffixes –*chan* and –*sama* respectively. As it should be clear from the description above, these complex forms retain all the connotations of the pronouns they come from, but the addition of the title suffixes adds something extra. For this reason, it is argued that these two forms should be classified as variants of the pronouns they come from, rather than independent.

Disregarding the variants, we are left with the following inventory of 1st person pronouns: *watashi*, *watakushi*, *boku*, *ore*, *atashi*, *washi*, *jibun*, *atai*, *uchi*. It is interesting to note that all of these forms appear in Shibamoto-Smith's (2004) inventory (shown in figure 6.2 above), apart from *uchi*[111]. As mentioned in the data description, *uchi* is a novel form for standard Japanese. Therefore, I take it that the selection of the data for the 1st person pronoun category is well-motivated, as it largely coincides with Shibamoto-Smith's (2004) inventory apart from three additions which have been independently motivated.

To summarise the data, the items that were previously described have been categorised as follows.

1) So-called 1st person pronouns

watashi, *watakushi*, *boku*, *ore*, *atashi*, *washi*, *jibun*, *atai*, *uchi*

Variants: *boku-chan*, *ore-sama*

2) Nouns

Kinship terms: *okaasan* ('mother'), *otoosan* ('father'), *mama* ('mummy'), *papa* ('daddy'), *obasan* ('aunt'), *ojisan* ('uncle'), *obaasan* ('grandmother'), *ojiisan* ('grandfather'), *oneesan* ('elder sister'), *oniisan* ('elder brother'), etc.

Professional terms: *sensei* ('teacher'), *oishasan* ('doctor'), *kangofusan* ('nurse'), *omawarisan* ('policeman'), *obousan* ('priest')

3) Names

6.5. The projective meaning of Japanese 1st person pronouns

This section aims to pinpoint the projective nature of the meanings carried by Japanese 1st person pronouns. I will use specific tests to establish their projective character, and

[111] Shibamoto-Smith (2004) also mentions *atakushi* in her inventory as a female formal to semi-formal form, but I have not included it in my data list due to lack of specific enough data.

for ease of exposition, I will select a number of the so-called 1st person pronouns that were mentioned above in order to apply the tests. However, it will be assumed that the results apply to all the forms in the category of 1st person pronouns, and this assumption rests on the insights of sociolinguistic literature (see references above) and confirmation from native speakers. The selected forms to which tests will be applied are: *watashi*, *watakushi*, *boku*, *ore*, *atashi* and *washi*. The first five forms have been chosen because they are used more frequently and the sixth (*washi*) is somehow less common nowadays but native speakers still have relatively clear intuitions about its implications.

6.5.1. Projecting or not?

In order to confirm that these implications project as non-at-issue meaning we can use the Family of Sentences (FoS) test (Chierchia and McConnell-Ginet 1990) which consists in testing whether the projective behaviour attested in the affirmative form of a sentence can be embedded under the scope of negation, question, modality, and the antecedent of a conditional.

Family of Sentences test

(193) *watashi*

(193a) Watashi no sekinin datta.

At-issue: 'It was my responsibility.'

Non-at-issue: The speaker is a woman / a man in a relatively formal occasion.

(193b) Watashi no sekinin dehanakatta.

At-issue: 'It wasn't my responsibility.'

Non-at-issue: The speaker is a woman / a man in a relatively formal occasion.

(193c) Tabun watashi no sekinin datta.

At-issue: 'Maybe it was my responsibility.'

Non-at-issue: The speaker is a woman / a man in a relatively formal occasion.

(193d) Watashi no sekinin datta ka?

At-issue: 'Was it my responsibility?'

Non-at-issue: The speaker is a woman / a man in a relatively formal occasion.

(193e) Moshi watashi no sekinin datta baai, ima sugu machigai wo naosu.

At-issue: 'If it was my responsibility, I'll correct the mistake right now.'

Non-at-issue: The speaker is a woman / a man in a relatively formal occasion.

(194) *watakushi*

(194a) Watakushi no sekinin deshita.

At-issue: 'It was my responsibility.'

Non-at-issue: This is a very formal occasion.

(194b) Watakushi no sekinin deha arimasen deshita.

At-issue: 'It wasn't my responsibility.'

Non-at-issue: This is a very formal occasion.

(194c) Tabun watakushi no sekinin deshita.

At-issue: 'Maybe it was my responsibility.'

Non-at-issue: This is a very formal occasion.

(194d) Watakushi no sekinin deshitaka?

At-issue: 'Was it my responsibility?'

Non-at-issue: This is a very formal occasion.

(194e) Moshi watakushi no sekinin datta baai, ima sugu machigai wo naoshimasu.

At-issue: 'If it was my responsibility, I'll correct the mistake right now.'

Non-at-issue: This is a very formal occasion.

(195) *boku*

(195a) Boku no sekinin datta.

At-issue: 'It was my responsibility'

Non-at-issue: The speaker is a young male and the register is relatively informal.

(195c) Boku no sekinin dehanakatta.

At-issue: 'It wasn't my responsibility.'

Non-at-issue: The speaker is a young male and the register is relatively informal.

(195c) Tabun boku no sekinin datta.

At-issue: 'Maybe it was my responsibility.'

Non-at-issue: The speaker is a young male and the register is relatively informal.

(195d) Boku no sekinin datta ka?

At-issue: ‘Was it my responsibility?’

Non-at-issue: The speaker is a young male and the register is relatively informal.

(195e) Moshi boku no sekinin datta baai, ima sugu machigai wo naosu.

At-issue: ‘If it was my responsibility, I’ll correct the mistake right now.’

Non-at-issue: The speaker is a young male and the register is relatively informal.

(196) *Ore*

(196a) Ore no sekinin datta.

At-issue: ‘It was my responsibility.’

Non-at-issue: The speaker is male and the register is very informal.

(196b) Ore no sekinin dehanakatta.

At-issue: ‘It wasn’t my responsibility.’

Non-at-issue: The speaker is male and the register is very informal.

(196c) Tabun ore no sekinin datta.

At-issue: ‘Maybe it was my responsibility.’

Non-at-issue: The speaker is male and the register is very informal.

(196d) Ore no sekinin datta ka?

At-issue: ‘Was it my responsibility?’

Non-at-issue: The speaker is male and the register is very informal.

(196e) Moshi ore no sekinin datta baai, ima sugu machigai wo naosu.

At-issue: ‘If it was my responsibility, I’ll correct the mistake right now.’

Non-at-issue: The speaker is male and the register is very informal.

(197) *atashi*

(197a) Atashi no sekinin datta.

At-issue: ‘It was my responsibility.’

Non-at-issue: The speaker is a young female and the register is relatively informal.

(197b) Atashi no sekinin dehanakatta.

At-issue: ‘It wasn’t my responsibility.’

Non-at-issue: The speaker is a young female and the register is relatively informal.

(197c) Tabun atashi no sekinin datta.

At-issue: ‘Maybe it was my responsibility.’

Non-at-issue: The speaker is a young female and the register is relatively informal.

(197d) Atashi no sekinin datta ka?

At-issue: ‘Was it my responsibility?’

Non-at-issue: The speaker is a young female and the register is relatively informal.

(197e) Moshi atashi no sekinin datta baai, ima sugu machigai wo naosu.

At-issue: ‘If it was my responsibility, I’ll correct the mistake right now.’

Non-at-issue: The speaker is a young female and the register is relatively informal.

(198) *washi*

(198a) Washi no sekinin datta.

At-issue: ‘It was my responsibility.’

Non-at-issue: The speaker is an old man.

(198b) Washi no sekinin dehanakatta.

At-issue: ‘It wasn’t my responsibility.’

Non-at-issue: The speaker is an old man.

(198c) Tabun washi no sekinin datta.

At-issue: ‘Maybe it was my responsibility.’

Non-at-issue: The speaker is an old man.

(198d) Washi no sekinin datta ka?

At-issue: ‘Was it my responsibility?’

Non-at-issue: The speaker is an old man.

(198e) Moshi washi no sekinin datta baai, ima sugu machigai wo naosu.

At-issue: ‘If it was my responsibility, I’ll correct the mistake right now.’

Non-at-issue: The speaker is an old man.

The FoS test confirms that Japanese 1st person pronouns have content that projects.[112] But since projection is also a characteristic of ordinary presuppositions, what needs to be shown is non-displaceability, i.e. inability to be filtered. This will be done by syntactically embedding the terms in question under a belief predicate and see whether their not-at-issue meanings project beyond the scope of the belief predicate. In the examples below, the 1st person pronoun that triggers the implications described above is embedded under the attitude predicate *omotte-iru/imasu* ('thinks') together with a phrase denying the implication[113].

(199) Kanai-ha watashi-ga shitsurei-na hito da to omotte-imasu.

'My wife thinks that $I_{\text{male formal}}$ am a rude person.'

(200) Tsuma-ha watakushi-ga shitsurei-na/gouman-na hito da to omotte-imasu.

'My wife thinks that $I_{\text{very formal}}$ am a rude/arrogant person.'

(201) Sensei-ha boku-ga josei da to omotte-iru.

'The teacher thinks that I_{male} am female.'

(202) Sensei-ha ore-ga josei da to omotte-iru.

'The teacher thinks I_{male} am female.'

(203) Sensei-ha atashi-ga dansei da to omotte-iru.

'The teacher thinks I_{female} am male.'

(204) Douryou-tachi-ha washi-ga amari toshi dehanai to omotte-iru.

'My colleagues think that $I_{\text{old male}}$ am not old/aged.'

The above examples have been specifically constructed to express a contradiction between the at-issue (what someone else thinks of the speaker) and the projective meanings (how the speaker sees/presents him/herself). The projective contents of the forms (gender, formality, age) that are targeted for contrast follow the descriptive observations seen above. As these sentences do not present the attitude holder(s) as holding contradictory beliefs, but the projecting implication is understood as made by the speaker, it is concluded that the implications of the 1st person pronouns cannot be filtered or plugged and project to the global context. Moreover, it is also possible to construct examples which would be felicitous if the relevant meaning were filterable, such as the following:

[112] We should mention here that another popular test for projection is the 'Hey, wait a minute!' test (Shanon 1976, von Fintel 2001), but as was advocated in chapter 5, this test is apt for presuppositions, while the projected meaning of expressives is better challenged metalinguistically (Camp 2013).

[113] Note that I only mention the critical bit of meaning in subscript (e.g. register or gender), but not necessarily all existing meanings when it does not make a difference to the argument, in the interests of simplicity.

(205) # Watakushi-no kanai-ha ore-ga shitsurei-na hito da to omotte-imasu.[114]

'My$_{\text{very formal}}$ wife thinks that I$_{\text{male very informal}}$ am a rude person.'

(206) # Boku-no sensei-ha atashi-ga josei da to omotte-iru.

'My$_{\text{male informal}}$ teacher thinks that I$_{\text{female}}$ am female.'

(207) # Atashi-no sensei-ha boku-ga dansei da to omotte-iru.

'My$_{\text{female informal}}$ teacher thinks I$_{\text{male informal}}$ am male.'

(208) # Boku no douryou-tachi-ha washi-ga amari toshi dehanai to omotte-iru.

'My $_{\text{male informal}}$ colleagues think that I$_{\text{old male}}$ am not old/aged.'

Thus they are shown to confirm both scopelessness and non-displaceability, which justify labelling their projective meaning as associative, following the insights of chapter 5.

However, even if 1st person pronouns carry meanings which behave like expressives, one should not forget that they also have at-issue meaning, i.e. they denote the speaker:

(209) Watashi ha Nihon-jin da.

'I$_{\text{female informal/male formal}}$ am Japanese'

(210) Watakushi ha Nihon-jin desu.

'I$_{\text{very formal}}$ am Japanese'

(211) Boku ha Nihon-jin da.

'I$_{\text{male informal}}$ am Japanese'

(212) Ore ha Nihon-jin da.

'I$_{\text{male very informal}}$ am Japanese'

(213) Atashi ha Nihon-jin da.

'I$_{\text{female very informal}}$ am Japanese'

(214) Washi wa Nihon-jin da.

'I$_{\text{old male}}$ am Japanese'

Common asserted content: 'I am Japanese', i.e. the speaker is Japanese.

The verdict from these tests is that Japanese 1st person pronouns act just like expressives, whose projective behaviour was explained in terms of associative meaning in chapter 5. Having identified and explained that component of meaning, in the next

[114] Note that this is still infelicitous even if the speaker does indeed refer to himself as 'ore' when talking to his wife.

chapter I will address the question of how to represent their at-issue meaning, as well as that of nouns and names when employed by the speaker for her/his self-reference.

Chapter 7. Names, nouns and so-called 1st person pronouns: the semantics of self-reference in Japanese

7.1. Preliminaries

As we saw in chapter 2, both names and indexicals clearly defy the Fregean picture; both are predicated as directly referential, thus devoid of a Fregean Sinn. Nevertheless, if we follow the orthodox traditions for each (Kripke 1980 and Kaplan 1989 respectively, which will henceforth be called using the collective term 'Kripke-Kaplan account'), direct reference is as far as their similarities go, as they are distinguished by some stark differences. Having discussed the details in chapter 2, let's just remember the most important bits here: according to Kripke (1980), the sole function of names is to contribute the individual they stand for to a proposition, which is made possible by means of a pre-existing relationship of causal transmission. Put simply, names entirely lack any sort of linguistic meaning, and their relevance for reference and thus linguistic communication goes no further than them acting as 'tags' (Marcus 1961). Contrarily, indexicals are undeniably endowed with some kind of linguistic meaning, what Kaplan (1989) calls 'character'. So even though indexicals are also said to be directly referential like names, there is a difference in the mechanism between the two cases: in the case of names we have the agent's participation in the event of a causal chain, while in the case of indexicals we have the linguistic knowledge of the character meaning. Inevitably, this difference of how reference is instrumented has a consequence on the issue of cognitive significance: as we saw from Perry's (1979) 'essential indexical' story, the character meaning of indexicals enables a special kind of cognitive significance, which is totally absent from names, as well as descriptions. In terms of Kaplan's (1989) terminology of character and content meaning as functions, the differences between different kinds of expressions can be summarised as follows:

type of meaning / type of expression	**character**	**content**
Names	constant	constant
Indexicals	non-constant	constant
Descriptions	constant	non-constant

Table 7.1. Character and content functions for different kinds of expressions.[115]

Nevertheless, one could say that the stark contrast between names and other kinds of expressions is mostly based on examples from English, while it does not seem to be confirmed by other languages. For example, it has been observed that in languages such as Catalan and Greek names behave a lot like nouns by requiring definite articles throughout (Matushansky 2008), which may hint towards a treatment of names along the lines of descriptions – there will be more to say about this later on.[116] The specific claim I will make is that the data from Japanese about how names and indexicals are used do not confirm the strict distinction between the two kinds of terms, which follows from the Kripke-Kaplan account. To give a foretaste, I will argue that the difference of

[115] The row about descriptions is greyed out so as to highlight the distinction between names and indexicals, which is more pertinent to our present discussion.

[116] Of course, descriptive accounts of names have a very long tradition in philosophy, to which Kripke (1980) constituted a reaction. Here we are merely concentrating on how the distinction between names and other kinds of expressions can be challenged by cross-linguistic observations, rather than philosophical arguments, of which, of course, there are several.

cognitive significance between indexicals and names, as taught by Perry (1979), is not instantiated in Japanese as names can receive de se interpretations.

As we learned from the descriptive part of chapter 6, Japanese allows speakers to use their own name when they engage in self-reference, and it will be demonstrated below that these are in fact de se situations, rather than merely de re. Of course there are various explanations why this practice is observed, for example that children find it easy to refer to themselves using their name, which is expected to be unique in the specific context and this is how everyone else refers to them anyway. One reason why this is preferred is because the inventory of 1st person pronouns is long and complicated, requiring knowledge of sociolinguistic intricacies way beyond the ability of a small child. But this is not necessarily a Japanese-specific phenomenon, as it is known that even in languages where the pronominal system is much less complicated children tend to struggle with personal pronouns and call themselves via their name (e.g. in Dutch, De Houwer & Gillis 1998). In fact, it has been suggested that children's difficulty for personal pronouns might be cross-linguistically explicable in terms of a developing Theory of Mind (Wechsler 2010). However, while the phenomenon of self-reference via one's name is cross-linguistically common in child language, where there is a developing grammar, in Japanese it is not confined to that but is also observed in adult speakers, more specifically young women. Again, there is a possible sociolinguistic explanation for this, which could be that due to this being a child-like behaviour, it is adopted by young women to convey the impression of cuteness and innocence, which tend to be viewed as positive feminine characteristics in Japanese society.[117] But importantly, whatever the sociolinguistic or cognitive explanations for this practice will be, what is of interest to us here is not what underlyingly motivates it, but how to account for it in terms of a semantic account for names and indexicals.

But of course, in order for the phenomenon of self-reference via names to be in any way relevant for a semantic account, it will have to be convincingly argued that what is involved in these cases is indeed self-reference, and not merely reference to an entity that happens to be oneself. It is only if it can be proven that what is involved is genuine self-reference that the Kripke-Kaplan doctrine, proclaiming a stark distinction between names and indexicals, can be questioned. In other words, unless it is shown that such uses of Japanese names are truly de se, this phenomenon cannot be considered but a Japanese counterpart of speaking about oneself in the third person in a non-de se way, which is of course also possible and attested in English, as we saw in Perry's (1979) story. If that were the case, then the mere difference between Japanese and English would lie in the frequency of this phenomenon, as it seems to happen more often in Japanese than in English, which could be satisfactorily accounted for by pragmatics- or sociolinguistics-based factors.

[117] But even if one tries to play down the use of names in self-reference (de se) in Japanese by saying that it's only adopted by a specific group of speakers such as children and women acting like children, there is still the extensive use of names in the case of addressing the hearer (de te), which is observed almost throughout the social spectrum. Of course one can and in fact should attribute this to the lack of a neutral, unmarked 2nd person pronoun (McCready 2015), but we're still entitled to wonder about the implications of this absence on the semantics of the expressions which enable these functions we usually consider as exclusive to personal pronouns in English and other languages.

7.2. Names and nouns in de se positions

The first step in order to build the worthwhile contrast is to establish that names in English (and other languages which work similarly in the pronominal domain) cannot be used de se. This should be straightforward at this point, considering the theories we have reviewed so far: Kripkean semantics for names preclude any possibility for a meaning-based cognitive significance that could license a de se attitude. This is not to say that one is barred from having a de se attitude while referring to oneself by name, but merely that this is not reflected in the linguistic form. In other words, the use of a name for self-reference in English can never be unambiguously de se, and is perceived by default as a third-person way to refer to oneself (illeism). The most knock-down argument that such uses are third-person is that they license the respective 3rd person agreement; even if the speaker is perfectly aware that s/he is engaging in self-reference, there is no way s/he can use 1st person agreement with any other form that the 1st person pronoun:

(215) This researcher is tired of the constant lack of support from the department.
(216) * This researcher am tired of the constant lack of support from the department.

Although the syntactic constraint of 1st person agreement being licenseable only by 1st person pronouns and not any other expression is telling, it does not suffice to rule out the possibility of (215) being de se rather than merely de re. In order to do so, we will need to construct a scenario where the two readings can be contrasted. Here is an attempt:

Rodrigo is drunk, watching a video of himself complaining about the constant lack of support at the university department where he works as a researcher. Due to his drunkenness he fails to recognise himself in the video and he has even forgotten that he works as a researcher at a university (he thinks he's still at his old job as a musician). His friend Vlad walks in the room and asks him what the video he is watching is saying.

(215a) This researcher$_{\text{de re}}$ is tired of the constant lack of support from the department.

(215b) # This researcher$_{\text{de se}}$ is tired of the constant lack of support from the department.

(215c) # I am tired of the constant lack of support from the department.

Due to Rodrigo's lack of self-awareness in the scenario above, both (215b) and (215c) constitute infelicitous responses, due to the de se element attested in (215c) and hypothesised in (215b). Given the difference between overt pronouns and PRO (Chierchia 1989) seen in chapter 2, this results in a contrast in how we can report Rodrigo's utterance (215):

(216) Rodrigo communicated that this researcher is tired of the constant lack of support from the department.

(217) # Rodrigo communicated being tired of the constant lack of support from the department.[118]

As the above scenario should show, (215) is only felicitous when it is interpreted non-de se.

The use of 'imposters' (Collins & Postal 2012), i.e. third-person expressions for self-reference (or addressee-reference) in English (one of which is of course names) can be accounted in three ways: one is a discourse-driven one, in which case the phenomenon is not dependent on any specific linguistic forms (Siewierska 2004, Stirling & Huddleston 2002) – let's call this the notional-functional view. Another one is a syntactic one, according to which when expressions are used as imposters, there is a distinctive syntax lying underneath which licenses a 1st (or 2nd) person feature to non-pronominal forms (Collins & Postal 2012). The undisputed advantage of the syntactic over the notional-functional view is that it can explain some puzzling agreement phenomena, such as the fact that plural imposters can license 1st person reflexive pronouns (Collins & Postal 2012, Podobryaev 2014). The third one is a notional-semantic account (Podobryaev 2014), which retains the general thesis of the notional-functional view while explaining the peculiar cases of agreement by positing special semantic operators able to manipulate the assignment function and affect person features (assuming those are part of complex indices as per Sudo 2012). But whichever of the three views one decides to choose, it is clear that when non-pronominal expressions are used to express a de se (or de te) attitude, the burden of the explanation falls on things other than the expressions themselves (the discourse, syntax, or a combination of discourse and semantic operators, respectively).

As mentioned before, in English there seems to be no way we can be certain that a name (or noun) is used impostrously (de se) or as a third-person expression (de re), apart from the hindsight of explicitly confirming speaker intentions.[119] But in the case of Japanese, if it can be decisively shown that names can be used in positions which are unambiguously de se, then we could say that there is something about the expressions themselves which enables them to occupy such positions. Let's see some structures which could be considered possible candidates for unambiguously de se positions.

It has been observed that 'in Japanese, the speaker cannot report in a direct form the psychological state of anyone but himself. Thus, only the first-person subject is possible in a sentence that directly describes a psychological slate' (Shibatani 1990: 383). That is to say, the use of certain predicates in Japanese presupposes some kind of first-person authority; specifically, these are called 'experiencer predicates' (McCready 2007) or

[118] I used 'communicated' here because the more appropriate 'said' does not license a construction with PRO. An alternative that allows 'said' would be Castañeda's (1966) 'he*', in which case the contrast would be rendered as follows:
(218) Rodrigo said that he is tired of the constant lack of support from the department.
(219) # Rodrigo said that he* is tired of the constant lack of support from the department.

[119] This is not quite true about plural imposters, as those are able to license 1st person agreement in reflexive pronouns (Collins & Postal 2012). However as this is not possible for singular ones, it has to be due to the interaction with number and not just person features, which is a path worth exploring, but not for us as we are specifically interested in person here. For a detailed proposal about this see Podobryaev (2014).

'psychological predicates' (Hasegawa & Hirose 2005). In functional linguistics, this phenomenon is known as 'territory of information' (Kamio 1994, 1997). Some examples of such predicates are 'sabishii' ('lonely'), 'samui' ('cold'), 'atsui' ('hot'), 'itai' ('painful'), 'kanashii' ('sad'), 'tanoshii' ('pleasant'), 'kowai' ('fearsome'), the volitional construction verb stem + 'tai' ('want to…), as well as the verb 'omou' ('think').[120] In order to use these expressions in the 3rd person, one needs to add an evidential marker such as –*gatte iru* ('looks like') or *soo da* ('seems like'). The contrastive examples below are from Hasegawa & Hirose (2005: 227-230):

(220) a. Watashi wa samu-i.
I TOP feel.cold
'I feel cold.'

b. #Haha wa samu-i.
mother TOP feel.cold
'My mother feels cold.' [Intended]

c. Haha wa samu-gat-te i-ru.
Mother TOP feel.cold-EVID-CONJUNCTION be
Lit. 'My mother is showing signs of feeling cold.'

d. Haha wa samu-soo da.
Mother TOP feel.cold-EVID COPULA
'My mother appears to feel cold.'

(221) a. Watashi wa kanashi-i.
I TOP be.sad
'I'm sad.'

b. Haha wa kanashi-i.
mother TOP be.sad
'My mother makes me sad.' NOT 'My mother is sad.'

(222) a. Watashi wa koohii o nomi-ta-i.
I TOP coffee ACC drink-want
'I want to drink coffee.'

b. #Haha wa koohii o nomi-ta-i.
Mother TOP coffee ACC drink-want-NPST
'My mother wants to drink coffee.' [Intended]

c. Haha wa koohii o nomi-ta-gat-te i-ru.
Mother TOP coffee ACC drink-want-EVID-CONJ be
Lit. 'My mother is showing signs of wanting to drink coffee.'

(223) a. Watashi wa haha wa byooki da to omo-u.
I TOP mother TOP ill COP.NPST QUOT think
'I think my mother is ill.'

[120] Fascinatingly, this person restriction applies only in the present tense (also known as 'non-past' tense), and not in the past (form ending in –*ta*) or the continuous present (-*tte iru*). The fact that this person restriction is only observed in the simple present form might be explained in terms of indexicality, which would be responsible for the resulting non-displaceability. This possible connection is left for future research.

b. Haha wa byooki da to omo-u.
mother TOP ill COP.NPST QUOT think
'I think my mother is ill.' NOT 'My mother thinks she is ill.'

As we see from the examples above, the predicates 'samui' ('feel cold'), 'kanashii' ('sad'), 'nomitai' ('want to drink') and 'omou' ('think') can only be used with a first person subject, not with a third person one. Interestingly, as McCready (2007) notes, such predicates are also acceptable with a second person subject, provided that this is part of a question (examples from McCready 2007: 2-3):

(224) Watashi/*anata/*kare-wa samui desu.
I/you/he-TOP cold COP
'I'm/*You're/*He's cold.'

(225) ?Watashi/anata/*kare-wa samui desu ka?
I/you/he-TOP cold COP Q
'*Am I/are you/*is he cold?'

As McCready (2007) explains, the infelicity of such predicates with a first person subject in questions can be attributed to the fact that 'it is odd to ask someone else about one's own physical sensations' (2007: 3). So we can say that these predicates can only be used in the first person in assertions because of the speaker's first-person authority on his own mental states, while the fact that they can be used in the second person in questions does not constitute a counterexample as in these structures the speaker is essentially inquiring about the hearer's own privileged access. Again, we see the same interesting correspondence have observed before between de se and de te.

However, the notion of first person involved in these predicates should be clarified so that one does not jump to the premature conclusion that it strictly corresponds to the 1st person singular we have become accustomed to from the grammatical 1st person singular in English and other languages with similar person distinctions. The notion of first person as the subject of these Japanese constructions is one that includes rather than strictly delimitates the speaker; of course the idea of the speaker being included in the subject does not exclude the possibility of coinciding exactly with the speaker, but this in no way mandatory as in the case of the 1st person singular in English. Perhaps one fitting notion when characterising the subject in question is that of 'in-group' ('uchi' in Japanese), a concept of sociolinguistics and cultural studies which corresponds to a wider group the speaker identifies as belonging to, such as a kinship group, a company, a club, etc. (for the distinction between in- and out-group or else 'uchi' and 'soto', see Wetzel 1994, Bachnik 1994). Therefore, apart from the examples above where the subjects coincided exactly with the speaker, the following constructions are also acceptable:

(226) Watashitachi, futsuu no onnanoko ni modoritai.[121]
We ordinary girls want-to-go back to
'We want to return to being ordinary girls.'

(227) Heisha wa hatten shitai.[122]

[121] This is the statement by the idol group 'Candies' when they announced their break-up: https://en.wikipedia.org/wiki/Candies_(group)

Humble-company grow want-to
'Our company wants to grow.'

As per the examples above, we need to clarify the notion of first person involved in these constructions as follows: the subject necessarily includes the speaker, but may extend beyond him/her. This caveat, albeit essential to specify the exact notion of first person involved and to distinguish it from its grammatical counterpart as it has been conceptualised by observing languages such as English, does not have any implications on the proposal about using these constructions as decisively de se positions.

Therefore to show that names can appear in exclusively de se positions in Japanese, all we need to do is see whether substituting the 1st person pronouns by names in the examples (220)-(223) above leads to felicitous sentences. Let's assume that the speaker is called Ikuho, and uses her name to refer to herself:

(228) Ikuho wa samu-i.
Ikuho TOP feel.cold
'I feel cold.' (also: 'My name is Ikuho.')[123]

(229) Ikuho wa kanashi-i.
Ikuho TOP be.sad
'I'm sad.' (also: 'My name is Ikuho.')

(230) Ikuho wa koohii o nomi-ta-i.
Ikuho TOP coffee ACC drink-want
'I want to drink coffee.' (also: 'My name is Ikuho.')

(231) Ikuho wa haha wa byooki da to omo-u.
Ikuho TOP mother TOP ill COP QUOT think
'I think my mother is ill.' (also: 'My name is Ikuho.')

To reinforce the argument, let's assume that Ikuho has amnesia and has forgotten her own name. She has found her old diary (although she fails to recognise it as hers, since she doesn't recognise her own name) and is reporting the feelings and sensations the author is describing.

(232) # Ikuho wa samu-i.[124]
Ikuho TOP feel.cold
'Ikuho feels cold.'

(233) # Ikuho wa kanashi-i.
Ikuho TOP be.sad

---

[122] Example kindly translated by Miyuki Kamiya, p.c.
[123] Note that the English rendition makes it seem as if we have an additional sentence here, which might make the reader wonder what kind of an inference it is. However, I want to claim that given that this is a de se situation, i.e. the name is used as a self-referring term in Japanese, it belongs to the assertive core of the sentence.
[124] Since Ikuho here thinks she is talking about someone else, in standard Japanese at least she would be expected to use the honorific title '-san' after the proper name rather than employ it bare. However, this would make these utterances slightly different from their de se counterparts, potentially making the contrast less effective. For this reason, we will have to assume here that Ikuho is not being very polite, but given her amnesiac state of confusion such a pitfall may be overlooked.

‘Ikuho is sad.’

(234) # Ikuho wa koohii o nomi-ta-i.
Ikuho TOP coffee ACC drink-want
‘Ikuho wants to drink coffee.’

(235) # Ikuho wa haha wa byooki da to omo-u.
Ikuho TOP mother TOP ill COP QUOT think
‘Ikuho thinks her mother is ill.’

The contrast between when Ikuho is expressing her own feelings and sensations as her own, versus when she is reporting them as those of a person who happens to be herself, the assumption that the above constructions constitute de se positions seems to be correct. Therefore, we see that names in Japanese can indeed occupy de se positions, which are exclusively reserved for the 1st person pronoun in English:

(236) I am cold. (said by Natalia)
(237) *Natalia am cold.
(238) Natalia is cold. (even if uttered by Natalia herself, always ambiguous between de se, i.e. impostrous use, and de re, i.e. referring to oneself without knowing it)

Therefore it may be concluded that since names in Japanese can be used in de se positions, whereas in English they cannot, their semantics cannot be identical. In other words, there must be something in the semantics of names in Japanese that permits them to occupy de se positions. Since in English only the 1st person pronoun can occupy such positions[125], it could be said that the de se-ness is licensed by the 1st person singular feature. The quest is clear: we need to figure out what could be responsible for licensing the de se-ness in the case of names in Japanese.

Obviously, given the stark difference between names in English and Japanese shown by contrasting examples (228)-(231) with (237)-(238), we cannot maintain the same semantics for both, which suggests that the Kripkean theory of names is inadequate for Japanese. We will need to look for a proposal that is elaborate enough to be able to account for the fact that names in Japanese allow for both de se as well as de re readings, the latter just like their English counterparts. It will be argued that this can be undertaken by a variant of the camp of the theories of names as descriptions, and more specifically Matushansky’s (2008) proposal that names are predicates with quotative content (i.e. they mention the name) which are also equipped with a slot for the naming convention in force at the specific context of utterance.

7.3. Matushansky’s (2008) proposal

[125] To be precise, what happens is that while names in both languages could be used de se, in English they cannot occupy positions which are unambiguously de se (e.g. 1st person verbal agreement, as shown by examples above), but they can appear in constructions which are truly ambiguous between de se and de re (as also shown). So the contrast between English and Japanese amounts to the following: in Japanese names can occupy unambiguously de se positions, while in English they cannot.

Let's see how Matushansky's (2008) argument unfolds. First of all, let's remember from chapter 2 that there are two general camps when it comes to the semantics of names, one that treats them as descriptions (Fregean, also Russellian theories) and another that considers them directly referential (Kripke 1980). As far as direct reference is concerned, names are entirely simple both in terms of semantics as well as syntax, possessing no linguistic meaning or any internal structure whatsoever.[126] Contrarily, the descriptive alternative is compatible with the possibility of a more complex underlying structure, although it does not demand it.

Matushansky's (2008) crucial observation is that while there has been a long debate on names in argument positions, not enough attention has been paid to their use in predicative positions, such as the following:

(239) Whenever I meet Spanish speakers I introduce myself as Rosita.
(240) When we were in Crete, I used to call Ruoying Rena.
(241) If she has any children in the future, she'd like to name the boy Markos and the girl Aretousa.
(242) They're thinking of baptising their baby daughter Myrto.

Following the directly referential approach to names, predicative uses of names are to be accounted for as cases of mere mention rather than real use, i.e. by assuming that names appearing in predicative positions are nothing more than instances of unanalysable sounds. In this way, names are given a different semantics in argument than in predicative positions. However alternatively, if one opts for a descriptive approach to names, it turns out to be possible to propose a unified semantics for all positions, argument and predicative. But apart from the attractive parsimony of the second option, strong support for it comes from the syntax of the naming construction cross-linguistically, as Matushansky informs us.

Specifically, Matushansky's insight is that in a number of typologically unrelated languages the naming construction is compatible with a predicative analysis of proper names, while a metalinguistic explanation in terms of mere mention[127] shows to be inadequate; this suggests that names always enter the syntax as predicates, so in argument positions they are syntactically and semantically similar to NPs, albeit with a distinctive indexical slot which accounts for the differences in their behaviour compared to nouns. Matushansky backs her proposal with data from several different languages and carefully examines its implications for the syntax-semantics interface, but here it will suffice to present only the main arguments that justify it, before adopting and adjusting it to our specific purposes.

To begin with, assuming that naming constructions involve unanalysable quotes as the directly referential approach dictates, their underlying syntax should be that of

[126] However, it should be noted that this last claim that direct reference precludes structural complexity has been flagged as insubstantial by Predelli (2017) , who points out Kaplan's 'dthat'(1978) as a prominent counterexample.
[127] We will call theories which advocate that names are essentially unanalysable quotes as 'mere mention' theories; the adjectival qualification is needed in order to differentiate them from so-called quotation theories, which also advocate that (some kind of) mention is involved, but which retain that this forms part of a more complex semantic structure.

ditransitive verbs. Under such a view, both NPs (the individual named, and the name given) should be arguments of the verb. However, the choice of pronouns in the interrogative form testifies against such an analysis, as the pronoun chosen is not animate even if the argument it refers to happens to be:

(243) What/*who did you call your dog after all?
(244) What/*who did they decide to name their son?

While when the interrogative constructions involve names in argument positions, the pronouns are animates if the arguments themselves are:

(245) Who was barking all night again?
(246) Who was born today?

This difference already indicates that unlike in argument positions, names in the naming construction do not behave as referring devices. As this testifies against the assumption of a ditransitive structure as required by the direct reference view, it is worthy to consider other alternatives. Interestingly, the syntax of naming verbs happens to be very similar to that of change-of-state verbs, such as verbs of nomination; for example, the same verbs can be used both in naming and nomination constructions, such as *call* and *name*.

(247) In 332 BC, Alexander the Great was named Pharaoh of Egypt.
(248) The Pharaoh of Egypt in 332 BC was named Alexander the Great.

Now, it is known that verbs of nomination take a small clause complement, i.e. a minimal syntactic structure which contains a subject and a non-verbal predicate (Stowell 1989) – very similarly to the verb *make:*

(249) Spongebob Squarepants was named Employee of the Month.
(250) After Mr. Krabs retired, Carl was made manager of the business.

Given the similarity between nomination and naming constructions, we have good reasons to suspect that proper names in the naming construction are in fact predicates. Furthermore, it has been noted that cleft constructions are not possible with predicate NPs:[128]

(251) *It is Lebanese that Yves is.
(252) *It is a pastry chef that Vaishali wants to be.

And interestingly, cleft constructions are also impossible with names in the naming construction, which again points towards a predicative analysis.

[128] In fact Matushansky (2008) mentions that cleft constructions are not possible not only with predicate NPs but also PPs, e.g. *'It was in the States that Lucie was' (2008: 579). However, this does not seem to be supported by data, as one can easily find such examples online:
(253) It was in the States that The Moody Blues, who toured behind Canned Heat (where Brian Wilson would likely have seen them), really took off, and gained a popularity that eventually trickled back to England. (http://www.justinhayward.com/justin-hayward-comes-to-morristown-august-30/)
(254) It was in Europe that we built our first factory outside Japan, a huge undertaking for the time. (http://www.honda.co.uk/cars/world-of-honda/present/honda-in-europe.html)

(255) *It was Hikaru that they named him.

Even more support for the predicative hypothesis for the naming construction comes from the observation that nomination constructions can have a bare NP predicate only as long as it denotes a singleton set, whereas they require the presence of an article in case the set is not singleton:

(256) Mark was named champion for 2017.
(257) *Mark was named candidate for 2017.

While Stowell (1989) interprets this as evidence of correlation between the DP category and whether an NP is an argument or predicate, Matushansky proposes to use it as a diagnostic of a small clause structure as follows: if a certain NP is semantically definite, but shows up without an article in a construction where another NP can be considered as its subject, then the structure at hand is a small clause. This fits neatly with the observation that in languages where names usually come with a definite article (known as preproprial article) such as dialectal German, Pima (Uto-Aztecan), Tagalog, Catalan and others, names actually come bare in the naming construction, which can potentially be considered as extra support to the predicative hypothesis. Nevertheless, as Matushansky sensibly notes, article omission in this case could also be explained by appealing to lack of referentiality, rather than as an indication of predicativity. In fact, lack of referentiality is to be expected if one adopts a theory of mere mention, since in this case we are dealing with a sheer sequence of sounds, devoid of referential features.

However, mere mention analyses struggle to explain the appearance of the copula in the naming construction in languages such as Korean and Welsh, a phenomenon which is entirely unproblematic under the assumption that names in these cases are actually predicates. Moreover, even stronger support for the predicative hypothesis emerges from observing predicative case-marking in some languages. In particular, some languages mark predicates with a specific case, which we interestingly also notice in the naming construction in these languages; for Arabic the predicative case is the accusative, for Hungarian the dative while for Finnish it is the translative case, and these are also the cases of choice in the naming construction for these languages respectively. Moreover, in yet other languages we have the phenomenon of case agreement, i.e. the assignment of the same structural case between subject and predicate, which Matushansky proposes to use as yet another small clause diagnostic, and which is observed in languages such as Latin and Greek. Interestingly, Greek actually confirms both of the proposed diagnostics that the naming construction involves a small clause, as we note both the omission of the preproprial article as well as case agreement in the naming construction.

Apart from all the aforementioned arguments that have to do with syntax, there is also a decisive argument against mere mention theories that relates to interpretation. Specifically, although in cases where mere mention is indeed involved we can precede the quote by a characterisation such as ‘the word/expression’, this is impossible in the case of names:

(258) ‘Fthinóporo’ is difficult to pronounce for non-native speakers of Greek.
(259) The word ‘fthinóporo’ is difficult to pronounce for non-native speakers of Greek.

(260) They are considering naming their son Hiroyuki.
(261) *They are considering naming their son the name (of) Hiroyuki.

Summarising the above, there is robust crosslinguistic evidence that the naming construction should be analysed as one involving a predicate rather than unanalysable quote, which strongly supports descriptive over directly referential theories of proper names. The result of this is that if names enter the syntax as predicates, then in argument positions they should be analysed as consisting of a determiner plus a restrictor, just like common nouns. But of course, despite their similarities, proper names have certain characteristics that clearly set them apart from common nouns, and which should be explained deservedly.

Now if the above is on the right track and names are indeed predicates, an expected proposal of what their semantics might look like is something like the following:

(262) $[\![\text{Vera}]\!] = \lambda x.$ x is an entity named /'vʲerə/

Analyses such as the above have in fact been proposed by several authors, such as Kneale (1962), Bach (1981, 1987, 2002), Geurts (1997) and Recanati (1997). Such proposals have been called 'quotation theories' due to the mention of the phonological string in question, which is essentially the only component that distinguishes between different names. Although this analysis seems to be on the right track considering the discussion above, Matushansky makes the critical observation that it could be further decomposed, in order to incorporate a separate argument slot for the naming convention R, which was independently proposed by Recanati (1997). This would effectively make names 2-place predicates, in which case the representation would look as follows (Matushansky 2008: 592):

(263) $[\![\text{Vera}]\!] = \lambda x \in D_e . \lambda R_{<e, <n, t>>} . R (x) (/'v^jerə/)$
where n is a sort of the type e (a phonological string)

Matushansky's proposal is that this extra slot is saturated differently depending on whether the name is in a predicate position (in the naming construction) or in an argument one: when the name is a predicate of a small clause in the naming construction, the naming convention is provided by the verb involved, while when it is in argument position it is provided by the context. The addition of the naming convention in the lexical entry for names may seem like an extra complexity, yet it has two great advantages, one for each position in which a name may appear. I will present the implications for when the name is in an argument position further below, but as far as the naming construction is concerned, adding an argument for the naming convention allows for a uniform yet fine-grained enough analysis of different naming verbs, as the difference between them can be captured by allowing the naming convention R to vary accordingly:

(264) $[\![\text{Vera}]\!] = \lambda x \in D_e . R_1 (x) (/'v^jerə/)$, where R_1 = naming relation
(265) $[\![\text{Vera}]\!] = \lambda x \in D_e . R_2 (x) (/'v^jerə/)$, where R_2 = baptising relation
(266) $[\![\text{Vera}]\!] = \lambda x \in D_e . R_3 (x) (/'v^jerə/)$, where R_3 = nicknaming relation
Notably, this uniform analysis of different naming verbs is only made possible thanks to the naming convention R occupying its own distinct argument slot; in the coarser version was seen in (262), where the relation is primitively taken to be that of naming,

there is no way to represent the difference between distinct verbs of naming as their semantics looks identical.

It should be noted here that while Matushansky (2008) chooses to treat the context parameter as an argument, it can also be represented as a parameter in the interpretation function. The difference between the two representations has been shown to be critical in analyses of shifted indexicals, where the former option explains shifting in terms of quantification over context variables (Schlenker 2003) while the latter accounts for it in terms of modal operators which manipulate context parameters (Anand & Nevins 2004, Anand 2006). Given that I have treated the context as a parameter elsewhere in this dissertation, I will go for this latter solution, but given Matushansky's (2008) original treatment as an argument, it should be clear to the reader that nothing crucial hinges on this choice. Based on this, the above representation (263) is rendered as follows:

(267) $[\![\text{Vera}]\!]^c = \lambda x \in D_e \,.\, R_c\,(x)$ (/ˈvʲerə/)

One notable consequence of representing the context as a parameter is observed on the semantics of naming verbs. As mentioned above, what distinguishes the semantics of different naming verbs is the kind of naming relation they introduce, e.g. naming VS baptising VS nicknaming, etc. Now if the context is represented as a parameter and since the naming convention (R) is one of its coordinates, this means that naming verbs manipulate the naming convention coordinate of the context, i.e. they are so-called 'monsters' (Kaplan 1989). So, while before different naming verbs were distinguished based on how they saturated the λR argument of the name, this difference is now captured in terms of how they manipulate of the R coordinate of the context.

Turning now to composition, Matushansky's proposal for the compositional semantics of the naming consttruction is the following (2008: 610):

(268) $[\![\text{Alice is nicknamed Al}]\!]^w \approx \exists R$ [R is a nicknaming convention in w & R (Alice) (/æl/) in w]

Relativisation to a world parameter is added to account for the non-rigidity of names in this position, which is not captured by the lexical entry for names. As for the existential quantification, the idea is that it is introduced by the root of the naming verb. To show how this specific assumption is motivated, Matushansky starts by proposing a syntactic structure for naming verbs where they are decomposed into change-of-state (BECOME) and causative (CAUSE) components:

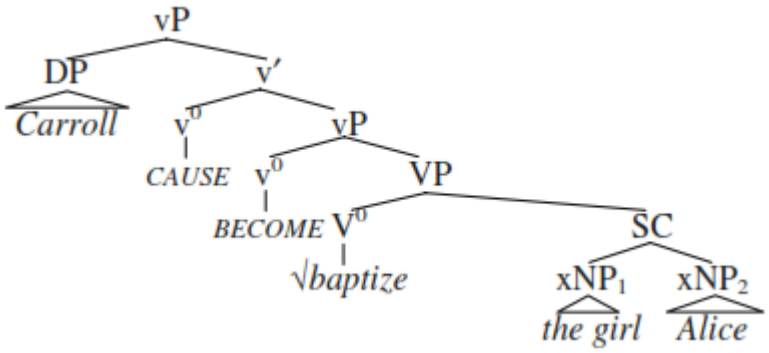


Figure 7.1. Syntactic representation of the structure of naming verbs (taken from Matushansky 2008: 610)

Note the separation of the verb root from the agent and the verb: the CAUSE component of the vP is responsible for assigning the theta-role of agent of the naming verb, while the BECOME component amounts to the event.[129] The proposal is that the verb root introduces existential quantification over naming conventions, whilst restricting them according to the specific lexical content of the verb involved. At the same time it enforces intensionalisation so that the naming convention can be evaluated in a given possible world. This results in the following meaning for the VP shown in the tree above:

(269) $[\![[_{VP} \sqrt{}\text{baptise } [_{SC} \text{ the girl Alice}]]]\!]^{c} = \lambda w. \ \exists R_{<e,<n,t>>} . \text{BAPTISM (w) (R)} \ \& \ [\![[_{SC}\text{the girl Alice}]]\!]^{c'}$

where c' is just like c except that the naming convention in c' is R

Turning now to the case of names in an argument position, the naming convention is provided by the context, i.e. it corresponds to the 'contextually salient naming convention in force between the speaker and the hearer, or more strictly speaking, the naming convention of the speaker that is presupposed to be shared by the hearer' (2008: 592). Above it was mentioned that the addition of the naming convention affords a great advantage for the representation of names in an argument position, which is the following: the extra coordinate R enables us to account for one of the most crucial as well as distinctive characteristics of proper names: their rigidity.

As explained in chapter 2, Kripke (1980) presented compelling arguments against the Fregean approach to proper names, essentially showing that it fails to account for their rigid character. Interestingly, although Matushansky's proposal essentially belongs to the descriptivist camp, the addition of the slot for the naming convention equips the proposal with a way to capture rigidity. Specifically, as previously seen the R argument is a free variable that is saturated by the context; in effect, this makes the representation indexical. And as also seen in chapter 2, rigidity follows from indexicality, given that in a double-indexed semantics indexicals depend on the context of utterance for their reference, which stays unaffected and fixed across different circumstances of evaluation.

[129] The BECOME component is motivated by data from Finnish, where the same case (translative) is assigned in small clauses involving change-of-state and naming verbs (see examples in Matushansky 2008: 611).

Therefore Matushansky's proposal is able to account for the rigidity of proper names, which is to be celebrated for at least three reasons: firstly, it complies with Kripke's insights of the admittedly rigid character of proper names. Secondly and most importantly, it manages to account for these philosophical insights in a way that makes sense for linguistic analysis; as we saw, the cross-linguistic data provides very strong indications against the idea that names are syntactically and semantically simplex, as it naturally follows from the idea that they are completely devoid of linguistic meaning and function merely as 'tags' (Marcus 1961). Thirdly, this proposal does not only combine philosophical insights with linguistic data, but even within the level of linguistic analysis it succeeds in reconciling the obvious similarity between proper names and nouns phrases in syntactic terms with their striking difference in semantic terms, as proper names are solely read de re, contrarily to noun phrases which also allow de dicto readings.

Although we have had to dedicate much space into presenting Matushansky's argument given its reliance on the syntax of the naming construction cross-linguistically, its implications for the semantics can be summarised quite briefly, despite its great significance for our purposes. The combination of the descriptive element with an indexical one adds a very powerful tool to our toolkit, because as noted before the Japanese data defy the absolute distinction between indexicals and names advocated by the Kripke-Kaplan account. As seen in chapter 6, Japanese self-reference employs an intermixture of personal pronouns, nouns and names. And interestingly, the boundary between the first two categories is so blurry that there are arguments for considering them a single category or two distinct ones depending on whether one's theoretical commitments lie in the functional or the generative camp respectively.

7.4. The semantics of self-reference in Japanese

7.4.1. The semantics of names in Japanese

As seen just above, Matushansky's addition of the naming convention as an argument of names turned them into indexicals, in other words, dependent on the context of utterance just like words like *I*, *here*, *now*. As seen in chapter 2, a context can be analysed as a tuple of different parameters, usually the speaker, hearer, time, location, world, but possibly more depending on our interests, as Potts (2007) added the judge which was mentioned in chapter 3, or just like Matushansky introduced the parameter of a naming convention in the context of utterance. Now although she explicitly added only this contextual parameter to her semantic representation of names, given that it is just one element of the tuple, we should be able to rewrite the semantics proposed in (267) by explicitly representing all the elements of the tuple, and just assume that the others are not actively relevant:

(270) $[\![\text{Vera}]\!]^{<s_c, h_c, t_c, l_c, w_c, R_c>} = \lambda x \in D_e \,.\, R_c\,(x)$ (/'vʲerə/)

In principle this representation is equivalent to the one in (267), if we assume that as far as names are concerned (at least following Matushansky), the only contextual coordinate that is relevant is the one of the naming convention R, and the others are

semantically idle for yielding the extension of a name (note how from all the coordinates, only R appears in the output of the function). So all we need to retain from this is that Matushansky's move to indexicalise the semantics of names by adding a single contextual coordinate opens the door to representing all of them there, in principle making them available to assume an active role given extra assumptions. But of course, just like Matushansky had to present arguments to justify the introduction of R to the semantics in the first place, we will also need to have good enough reasons to contend that any of the other coordinates becomes actively involved too.

Following the discussion of the use of names as terms for self-reference in Japanese, which were shown in examples (228)-(231), I would like to propose that this constitutes a good enough reason. As seen there, names in Japanese can occupy de se positions, which suggests that there must be something about their semantics that allows this. Contrarily, we saw that this is impossible for names in English, which prompted us to say that there must be an important semantic difference between the two. However this difference should also be observable intralinguistically, i.e. within Japanese, as names there can occupy de se positions but can also be read de re, just as their English counterparts. The proposal offered here is that this difference boils down to whether the contextual coordinate of the speaker is active or inactive in the representation. In other words, names are always context-dependent as Matushansky tells us, but in Japanese there is underspecification as to whether this dependency refers only to the coordinate of the naming convention R (cases of de re) or also to that of the speaker (cases of de se):

(271) De re: $[\![\text{Vera}]\!]^{<s_c, h_c, t_c, l_c, w_c, R_c>} = \lambda x \in D_e \,.\, R_c\,(x)$ (/'vʲerə/)

(272) De se: $[\![\text{Vera}]\!]^{<s_c, h_c, t_c, l_c, w_c, R_c>} = \lambda x \in D_e \,.\, s_c, = x, R_c\,(x)$ (/'vʲerə/)

At a first glance, this might seem ad hoc as we do not motivate when exactly the speaker parameter is active and when not. However, we should remember that the introduction of R as an argument was independently motivated by Matushansky (2008). And if R constitutes one of the elements of the context tuple, it feels natural to say that the rest also figure in the representation. Therefore the one thing that needs to be explained is how the activation of the speaker coordinate works in Japanese, especially since it is optional, so it cannot be conveniently attributed to a cross-linguistic difference.

I would like to propose here that although the speaker coordinate is an element of the context, which forms part of the semantic representation of names both in English and in Japanese, it is susceptible to activation only in Japanese because of the lack of a designated 1st person indexical. Following Kaplan (1989), the semantics of the 1st person pronoun *I* in English corresponds to the speaker coordinate of the context, as we saw in chapter 2:

(273) $[\![\text{I}]\!]^{c,w} = s_c$

However, as we saw in chapter 6, Japanese does not seem to have an expression exactly corresponding to that, but a multitude of expressions that can assume this role in discourse. The proposal thus is that because Japanese lacks an expression especially dedicated to undertake the role of referring to the speaker coordinate of the context, it

allows for this coordinate to potentially become activated in any position in which it appears, i.e. any context-sensitive expression.[130] And the reason why this is possible can be easily attributed to communicative efficiency: as we learned from the story of the essential indexical in chapter 2, self-reference can be crucial to thought and consequently its communication. Therefore if a language lacks the 1st person pronoun, given how useful such an expression is, it should be expected that it will try to compensate for it by allowing other expressions to fill in the gap. It should be mentioned here that one prediction of the current proposal is that there should be no language that both has an English-type 1st person pronoun as well as de se proper names. Although a definitive answer to this question deserves dedicated cross-linguistic research, preliminarily it can be said that the prediction is borne out by many Indo-European languages, if we apply the same tests used above for English. Of course, one caveat is that of child language, as it is known that toddlers frequently mix personal pronouns and names when referring to themselves (Budwig 1985, van der Meulen 2001, etc). Since this alternation between personal pronouns and names occurs at a young age and eventually eclipses as the grammar develops, it can be seen as supportive of the hypothesis that true 1st person pronouns (as soon as they are established as such) rule out de se names. However, yet another explanation of de se names in child English is that children might perceive personal pronouns as names (Clark 1978), something which would be very much in the spirit of the current proposal. Nevertheless, all this is merely speculative at this point, as serious empirical and cross-linguistic research is required to tackle this issue, which I leave for the future.

Summarising the proposal for the use of names in Japanese in de se positions, powered by Matushansky's (2008) indexicalisation of names, it was suggested that due to the lack of a designated indexical for the speaker of the context in Japanese, the speaker parameter can get activated in Japanese context-sensitive expressions as names are, enabling them thus to be used as de se expressions.

7.4.2. The semantics of so-called Japanese 1st person pronouns

As we saw in the descriptive part in chapter 6, Japanese has a very long array of expressions that function as 1st person pronouns, both diachronically but also synchronically, especially in comparison to languages such as English. It was explained there how in syntactic terms these expressions do not qualify as personal pronouns but

[130] Of course, this sounds like a claim grand enough to massively overgenerate; for example, what about context-sensitive expressions like *kochira* ('here') and *ima* ('now')? Interestingly, the first one can actually be used as a 1st person pronoun (see historic list in chapter 6), but the second one cannot. Lamentably, I have no rigorous solution to this problem of overgeneration. However, it could be that this is because in fact this is not really a problem; our account predicts that any context-sensitive expression could be used as a 1st person pronoun, and this might in principle be true, even though it doesn't mean it will necessarily happen in all cases. As the data presented in chapter 6 can show, the class of expressions that can be used as 1st person pronouns in Japanese is essentially open, and it seems that virtually any expression could come to assume this role if speakers feel like and of course agree to it. As we noted, our proposal is that this is possible because of the speaker parameter being susceptible to activation, but this should and need not be taken to mean that it always happens. This explanation will be further clarified on the section on Japanese 1st person pronouns below.

are indistinguishable from nouns, but how in functional terms their use resembles that of personal pronouns a lot. Given that these expressions are used by speakers as 1st person pronouns, it follows that they can occupy de se positions. And indeed, they comfortably pass the tests we used in the beginning of this chapter to diagnose that names in Japanese can appear in de se positions. This realisation brings back the same puzzle we had with names before: although these expressions look like plain nouns, or else definite descriptions, their behaviour shows that there must be something in their semantics that enables them to be read de se. We just saw how adopting Matushansky's (2008) analysis of names and proposing a case of underspecification of the active context parameters helped us account for the use of names as de se expressions in Japanese. Perhaps unsurprisingly, I will choose the tried path here (assuming it works of course), by adopting the same proposal for Japanese so-called 1st person pronouns, suggesting that they are *names for the speaker*.

Inevitably, the proposal needs clarification: what does 'names for the speaker' even mean? As mentioned in the beginning of this chapter and shown in table 7.1, the categories of names, indexicals and descriptions are understood as entirely distinct. As Kaplan (1989) taught us, there is an indexical that directly refers to the speaker ('I') but there are also descriptions such as 'the speaker' that can refer to him/her (non-rigidly, of course). As for names, following Kripke (1980), both their character as well as content meanings are said to be constant. Then we examined Matushansky's alternative proposal of names as descriptions endowed with an indexical component, which seemed to combine the best of both worlds (by explaining both predicative-like behaviour in their syntax as well as rigidity in their semantics). The claim here is that the variation of Matushansky's (2008) proposal in order to account for names in Japanese suffices to propose the category 'names for the speaker'. Specifically, it is claimed that so-called Japanese 1st person pronouns are essentially names that speakers adopt in order to notify other conversational participants that they are assuming the speaker role.[131] This allows us to understand the open class characteristic of Japanese so-called personal pronouns a bit more: the idea is that given the absence of syntactic restrictions, the language allows speakers to relatively easily introduce new terms to the category, i.e. start using nouns as 1st person pronouns. Although this may sound as dangerously liberal, it is in fact constrained by the pressure of communicative efficiency: frequent collective use of a certain expression as a 1st person pronoun has the effect that speakers recognise this expression as part of the class of words that speakers use to refer to themselves. This process essentially corresponds to establishing a convention in the community of speakers, which is why it closely reminds us of what Matushansky (2008) called the naming convention, borrowing Recanati's (1997) term. According to Recanati (1997: 140-1):

[131] Another supportive point for this analysis of Japanese so-called 1st person pronouns as names for the speaker comes from the forms 'boku-chan' and 'ore-sama' mentioned in chapter 6, which are essentially two items from the pronoun inventory combined with title suffixes, which are customarily used with proper names. Of course, the argument also extends to nouns such as 'okaasan' or 'omawari-san', which also contain a suffix title (note that the use of 'sensei' with a title suffix is excluded for what seem to be pragmatic reasons, given that 'sensei' may in itself be used as a title, and its use conveys particular respect).

> In this framework, a proper name refers by linguistic convention to whoever (or whatever) happens to be the bearer of that name; but who (what) is the bearer of the name is a contextual, non-linguistic matter, a matter of social convention. The reference of the name thus depends on a contextual factor, as the reference of an indexical expression does. […] Slightly more precisely, the linguistic meaning of an indexical expression refers the hearer to a relation R which is supposed to hold in context between the expression and the reference. Knowledge of this relation is what enables the hearer to identify the reference: the reference is the object (or an object) which contextually stands in that relation to the expression. For example, the reference of 'I' is the person who utters ' I '. Proper names are only a special case of this phenomenon, on the indexical view: the meaning of a proper name NN refers the hearer to a relation which holds in context between the name and its reference, namely the name-bearer relation. The reference of NN is the entity which is called NN in the context of utterance.

Specifically, the reason it is proposed that so-called Japanese 1st person pronouns are analysed as names for the speaker is because unlike 'I', they do not qualify as indexicals: their visible nominal characteristics and how easy a nominal expression can invade this category do not permit us to say that they have a non-constant character and a constant content, like bona fide indexicals do (see Fig. 7.1), since they retain some of their uses as nouns, as we saw in chapter 6. At the same time though, we cannot say that they are not different from expressions such as 'the speaker', given that at least in some examples, they demonstrate rigidity:

(274) Rainen hanashite-wa girishago-o hanasu hazuda.
Next year speaker-TOP Greek-ACC speak certainly
'Next year, the speaker will certainly speak Greek.'

(275) Rainen watashi-wa girishago-o hanasu hazuda.
Next year I-TOP Greek-ACC speak certainly
'Next year, I will certainly speak Greek.'

Given an adequate context, the term *hanashite* ('speaker') in (274) could refer to the specific person who is giving a speech in the Greek-Japanese Chamber of Commerce (de re), or to whoever who will happen to be assigned with the task of giving this speech next year (de dicto). However, *watashi* can only be understood as referring to the speaker of the current context, ore more accurately as used by him/her to self-refer (de se).

The above claim concerning rigidity should not be threatened by observations that Japanese personal pronouns are shiftable, like shown in Sudo (2012: 236):

(276) Mary-wa [John-ga watashi-o suki-da to] itta.
Mary-TOP John-NOM me-ACC like C said
'Mary said that John likes {me, her}.'

Importantly, what is claimed there is not that Japanese personal pronouns are equivalent to descriptions such as 'the speaker', but that they are shiftable indexicals. In other words, even if Sudo's claim holds, it only means that Japanese may allow the

context to shift, and not that Japanese personal pronouns are not tied to it in the way indexicals are, which makes the argument harmless to the current claim.

In a nutshell, the proposal is that the semantics of Japanese so-called 1st person pronouns are similar to those of names, and specifically that they should be understood as names that are used for the speaker, by the speaker:

(277) $[\![\text{boku}]\!]^{<s_c, h_c, t_c, l_c, w_c, R_c>} = \lambda x \in D_e . s_c = x, R_c (x)$ (/bo̤kɯ̤$^{\beta}$/)

As one can see, the above representation is the same as the one in (272), proposed for names in Japanese when used in de se positions. Given the sameness of the representation, one may ask whether the claim is that names in Japanese and expressions which are used as 1st person pronouns are essentially identical, which feels a bit problematic. But in fact, despite the fact that the two representations look identical, they should not be understood as such. First of all, let's remind ourselves that the claim was that the representation of names is underspecified as to whether the speaker parameter is relevant or not, which was necessary to be able to account for both de re and de se uses of names in Japanese. For so-called 1st person pronouns though, we do not require this kind of underspecification, and the proposed representation always entails de se. Secondly and most importantly, I would like to propose that there is a significant difference between (272) and (277), but which is not visible in the representation, and it lies in the meaning convention. Specifically, the claim is that R for proper names corresponds to the naming convention shared between speaker and hearer in the context of utterance, as advocated by Recanati (1997) and adopted by Matushansky (2008). In the case of expressions that are used as 1st person pronouns though, I would like to claim that R is somehow a bit more constrained, i.e. it involves a smaller pool of candidates. This does not mean that the class of these expressions is closed, because as mentioned previously it is not; it just means that the naming convention is a bit stricter than in the case of names, where in principle any string of sounds could constitute a name and can instantly be used as one. As Recanati (1997: 152) notes:

> Before the convention exists – as when a proper name is introduced for the first time to refer to a given object – the name by itself provides no way of identifying the reference; the latter is identified on a purely contextual basis. (However, even in this type of case, there is a convention in the offing, namely the convention which the speaker intends to be initiating by his use of the name.)

The proposal about expressions used like 1st person pronouns is that while they also carry a naming convention in their semantics like names do, it involves a group of terms which have been pre-established. Exactly because it requires some prior establishing, the number of items included in this domain is smaller as in the case of proper names. To reflect this, the representation should be slightly modified as follows:

(278) $[\![\text{boku}]\!]^{<s_c, h_c, t_c, l_c, w_c, R_c>} = \lambda x \in D_e . s_c = x, R_c (x)$ (/bo̤kɯ̤$^{\beta}$/)

Given the parallelism between the semantics for names in (272) and pronoun-like expressions in (278), but observing the difference that R_1 is a naming convention that involves a smaller set of items, namely terms that are commonly used for the speaker's

self-reference, this is how the characterisation of Japanese 1st person pronouns as names for the speaker is justified.[132]

Finally, to link this section with the previous chapter, it is argued that the compositional semantics of Japanese so-called 1st person pronouns are as in (278), but without denying that the different forms in this category differ not just in their phonology (as different proper names do), but also in what was called in chapter 5 associative meaning. Of course given that this kind of meaning is compositionally irrelevant, as it was extensively argued in chapters 4 and 5, it does not figure anywhere in the above representation, which is compositionally oriented. Therefore a complete representation of Japanese so-called 1st person pronouns must be given in two distinct dimensions, as follows:[133]

(279)

- $[\![\text{boku}]\!]^{<s_c, h_c, t_c, l_c, w_c, R_c>} = \lambda x \in D_e \,.\, s_c = x, R_c\,(x)$ (/bo̞kɯ̟$^{\beta}$/)
- $c \in CU$ (boku) only if c_a is a little boy/young male/tomboyish girl[134]

7.4.3. The semantics of nouns as terms of self-reference in Japanese

As mentioned in chapter 6, apart from so-called 1st person pronouns and names, several nouns in Japanese are also frequently used for self-reference, and in some cases they constitute the only felicitous choices. Unsurprisingly, they also pass the tests we previously used to show that names can occupy de se positions:

(280) Okaasan wa samu-i.
Mother TOP feel.cold
'I feel cold.' (also: 'I am your mother.')[135]

(281) Okaasan wa kanashi-i.
Mother TOP be.sad

---

[132] To give a very rough analogy from a totally different domain, we can understand the terms used in the game of 'Simon says' as names for the speaker in that context. Typically, the only such name is 'Simon', and the person who uses it does so by being aware of self-referring (and of course of giving commands, but that's secondary to our purposes). However, it is possible to admit more names into the list, such as themed ones like 'Cupid says' or 'Uncle Sam says'. In theory, on an occasion such as Valentine's Day the list of possible names for the speaker in this game should consist of at least two items, namely 'Simon' and 'Cupid', but potentially more if the people involved have previously played this game using any other name that does not depend on a specific occasion. For information about how the specific game works, see https://www.wikihow.com/Play-Simon-Says.

[133] Provided that such an analysis is towards the right path, it opens the question of whether it could also be relevant for other linguistic phenomena, such as gender (e.g. contemporary use of 'they' as a gender-neutral 3rd person singular pronoun). Due to limitations of space, I have to leave this for future research.

[134] As noted in the descriptive part in chapter 6, each form can be associated with several meanings.

[135] Note how the additional sentence of the English rendition does not only include some information about the speaker (that she is a mother), but also about the addressee (who is the other member in the two-place parental relationship). We will see below that this will indeed figure in the representation.

'I'm sad.' (also: 'I am your mother.')

(282) Okaasan wa koohii o nomi-ta-i.
Mother TOP coffee ACC drink-want
'I want to drink coffee.' (also: 'I am your mother.')

(283) Okaasan wa chichi wa byooki da to omo-u.
Mother TOP father TOP ill COP QUOT think
'I think my father is ill.' (also: 'I am your mother.')

Seeing how the above patterns exactly with names and pronouns as shown above, we're back to the question of what in their semantics allows them to do so. As expected, the answer here will be completely unoriginal, as we will essentially adopt the same formula proposed for names, and eventually 1st person pronoun-like forms: they have an argument slot that indexicalises them, enabling them to refer to the speaker parameter of the context. As mentioned previously, the idea motivating this is that in context-sensitive expressions in Japanese the speaker parameter is prone to activation due to the absence of an expression dedicated to denoting it. Their indexicality, and specifically the relevance of the speaker parameter, makes them similar to names and 1st person pronoun-like forms: this is exactly why all these forms can be used as de se in Japanese. However, there is a difference between nouns on the one hand and names and 1st person pronouns on the other, which is the role of the naming convention (R/R_1); because nouns are endowed with descriptive content that enables them to establish their reference by means of checking truth conditions, there is no need for a naming convention to do so. As a result, this context parameter is not active in the specific representation, which is proposed to be as follows:

(284) $[\![\text{okaasan}]\!]^{<s_c, h_c, t_c, l_c, w_c, R_c>} = \lambda x \in D_e . \lambda y \in D_e .$ x is mother to y, $s_c = x$, $h_c = y$

Notably, because kinship terms denote 2-place relations, their semantics denote this by having a slot for a second entity. Very significantly, this 2-place relation is one between the speaker and the hearer, which is why in the above formula both the speaker and the hearer parameter are active. Similarly to the argument of why s_c is prone to activation, I take it that we may also attribute the same ability of h_c to the lack of a dedicated indexical for this purpose in Japanese. Now, note that this is different in the case of professional terms, for which the following slightly modified semantics is proposed:

(285) $[\![\text{obousan}]\!]^{<s_c, h_c, t_c, l_c, w_c, R_c>} = \lambda x \in D_e .$ x is a priest, $s_c = x$

One could argue that potentially these terms could also be seen as 2-place relations (e.g. one is a priest to worshippers, a policeman to civilians, a doctor to patients, etc.). This is indeed a possibility for professional terms as well, and probably the best choice for 'sensei' (teacher), given that it is generally only used with a student as the addressee. However, it does not seem to be as required for the rest of professional terms, as they do

not show the exact relation between the speaker and the addressee as much as kinship terms and 'sensei', but mostly the speaker's professional role.[136]

7.4.4. Summarising

Summarising, the proposed semantics for Japanese so-called 1st person pronouns, as well as names and nouns used as devices of self-reference are as follows:

(279) So-called 1st person pronouns:

- $[\![\text{boku}]\!]^{<s_c, h_c, t_c, l_c, w_c, R_c>} = \lambda x \in D_e . s_c = x, R_c = R_1 (x)$ (/bo̤kɯ̤$^{\beta}$/)
- $c \in CU$ (boku) only if c_a is a little boy/young male/tomboyish girl

(272) Names: $[\![\text{Vera}]\!]^{<s_c, h_c, t_c, l_c, w_c, R_c>} = \lambda x \in D_e . s_c = x, R_c (x)$ (/'v^{j}erə/)

(284) Nouns (kinship terms): $[\![\text{okaasan}]\!]^{<s_c, h_c, t_c, l_c, w_c, R_c>} = \lambda x \in D_e . \lambda y \in D_e .$ x is mother to y & $s_c = x, h_c = y$

(285) Nouns (professional terms): $[\![\text{obousan}]\!]^{<s_c, h_c, t_c, l_c, w_c, R_c>} = \lambda x \in D_e .$ x is a priest & $s_c =$ x

Conveniently, the above semantic representations agree with the syntactic observations we saw in chapter 6: from a generative point of view, that Japanese does not have genuine personal pronouns and that its so-called personal pronouns are essentially nouns. And from a functional point of view, that these nouns are used in discourse just like 1st person pronouns, i.e. they denote the speaker. Moreover, thanks to Matushansky's (2008) revindication of a descriptive approach to names, there is no longer any essential difference between proper and common nouns. As a result, it can be said that there is a unifying account of self-reference in Japanese: syntactically speaking, all such expressions are noun phrases, and semantically, they are all sensitive to the speaker parameter of the context. Zooming in closer at this unifying proposal, we see the following semantic differences: unlike nouns, names and so-called pronouns are also sensitive to the parameter of the naming convention, though a different one (R for names and R_1 for pronouns). Within the group of nouns, there is a difference between kinship and professional terms, with the first ones also involving the hearer parameter.

7.5. Crosslinguistic implications

Inevitably, the claim that Japanese has a multitude of expressions whose semantics allows one to refer to oneself should be examined in crosslinguistic terms, as at a first glance this does not seem to be happening in English. The question here is, why is Japanese so different from English, or in more general terms, assuming that there are other languages like Japanese as mentioned in chapter 6, how do we explain this

[136] Note that this representation is powerful enough to apply to the case of de te uses of these nouns. Lamentably, I also have to leave this discussion for another occasion given to space limitations.

significant difference between Japanese-like and English-like languages? If this difference is not motivated adequately, one risks reducing the explanatory value of the proposal to a merely descriptive one.

The proposed answer is that in fact, despite appearances, there is no such significant difference between Japanese and English. Specifically, the claim is that the notion of an expression that refers to the first person (i.e. in a de se way) but is not a pure indexical (i.e. it is different from 'I') is not as foreign to English as it sounds. In fact, English has at least one such expression: 'yours truly'. This expression is a paramount example of what Collins & Postal (2012) called 'imposters', as mentioned above. To remind ourselves, imposters are expressions which though grammatically third person, notionally correspond to the first (or second) person. According to Collins & Postal (2012) who inaugurated work in this topic, other examples of such expressions are 'the (present) author', 'the undersigned', as well as names and kinship terms – sounding all too similar to our description of self-reference in Japanese. But interestingly, unlike all the terms given as examples of imposters, 'yours truly' is the only one that occurs solely as an imposter, i.e. it can only denote the speaker and not any other individual. In semantic terms, 'yours truly' is essentially a de se expression. But in syntactic terms, it is not a 1st person pronoun, as Collins & Postal (2012) have convincingly argued. So, in a way, it is very similar to Japanese so-called 1st person pronouns, which are necessarily de se (unlike names and nouns, in which case the s_c parameter is merely susceptible to activation, rather than permanently active), but are essentially nouns rather than pronouns. This neatly allows us to say that the phenomenon described is not unique to Japanese, relieving us from the burden of accounting for the stark crosslinguistic difference. And helpfully, just so that 'yours truly' does not seem like an exceptional expression we are relying too heavily on, we should notice that there is a French equivalent to it: 'bibi' (Collins & Postal 2012), and allegedly even a Spanish one: 'menda' (Batchelor 2006).

7.5.1. Some possible complications

Noticing the similarity between English so-called imposters and expressions that are used for self-reference in Japanese, we remind ourselves that there is a contrast between pronouns and imposters, i.e. between 'I' on the one hand and 'yours truly' as well as Japanese so-called 1st person pronouns on the other. Although Collins & Postal (2012) focus mostly on the syntactic differences, I will follow Podobryaev (2014), who employs semantic criteria to draw the distinction. Podobryaev (2014: 26) proposes to 'identify an expression as a Kaplanian indexical, if it refers to the speaker or the hearer coordinate of a given context, and if it cannot be ever interpreted as a variable, so, for example, it cannot ever be semantically bound', while he adopts Sudo's (2012) variable semantics in combination with enriched indices for personal pronouns. Using tests to show that imposters such as 'yours truly' cannot have bound interpretations, unlike regular pronouns as we know from the known case of so-called 'fake indexicals' (Partee 1989, Kratzer 2009), he contends that imposters correspond to Kaplanian indexicals, unlike pronouns.

Why is that such a problem for us? Well, if Japanese so-called 1st person pronouns are very similar to imposters such as 'yours truly', and the latter is confirmed to be a Kaplanian indexical, then this claim could be extended to these forms, bringing down the analysis of them as names for the speaker, and ultimately nouns. In theory, this could just mean that the particular analysis fails, and one could just say the semantics of all these expressions simply reduce to s_c, while the differences between them lie in the domain of associative meaning. However, there is an even graver consequence: we would find ourselves against the deadlock of having to explain how come a language allows the introduction of new indexicals so easily, given that the class of so-called 1st person pronouns in Japanese is rather open. But luckily, there is something that can be invoked in order to discount the threat coming from Podobryaev's (2014) claim: push the blame on semantic boundability from the variable vs. indexical pronoun to the influence of focus.

Specifically, let's take notice that virtually all constructions in which pronouns receive a bound interpretation involve focus:

(286) I am the only woman around here who could admit that I could be wrong. (Partee 1989)
(287) Only I got a question that I could answer. (Heim 1991)
(288) I'm the only one here who can take care of my children. (Kratzer 2009)

But obviously, Podobryaev (2014) did not make his claim about the non-boundability of imposters without supporting it with an example that involves focus. The contrasting pair is the following:

(289) Only yours truly talks to people who criticise his theory.
(290) Only yours truly talks to people who criticise my theory.

According to Podobryaev, (289) allows both for a strict as well as a bound reading, unlike (290) which disallows the bound one. The contrast between the two sentences is definitely indicative that we cannot simplistically say that focus is the sole culprit for bound readings, and I will not contend to have an answer to this intriguing puzzle. Nevertheless, there happens to be encouraging data in Roeper (1999, 2006) which suggest that perhaps bound readings are not exclusive to pronouns but given the right construction and context, they could even be observed beyond pronouns. Roeper (2006) offers examples from the following categories: proper names, indefinites, definite, point-of-view indexicals, and even VP gerunds:

(291) Only Fred still talks like Fred in front of royalty (everyone else puts on a phony accent).
(292) Only a very odd person still looks like a very odd person after using modern make-up techniques (ordinary people don't look ordinary).
(293) In these old pictures, only the living room still looks like the living room (no other room looks like itself).
(294) Only now can one appreciate now (then one could not appreciate then or then one could not appreciate now).
(295) Only ballet dancing looks like ballet dancing in a still photo (breakdance doesn't look like breakdance, squaredance doesn't look like squaredance... in a still photo).

Even these extremely interesting examples, though, should not be hastily interpreted as proof that focus alone is able to induce bound readings onto anything; in the same paper, Roeper (2006) notes that several constructions (such as dative indexicals, expletive expressives, and others) can act as intervenors, ultimately blocking bound readings even in focus environments. Undoubtedly, all these intriguing data deserve a serious and extensive study that can shed light onto the exact role of focus in giving rise to bound readings, which I sadly cannot undertake here. However, I will interpret the existence of data such as (291)-(195) as telling indications that bound readings are not unique to pronouns, which ultimately allows us to weaken Podobryaev's (2014) claim that non-boundability is a reliable criterion for identifying Kaplanian indexicals, salvaging thus the proposal for Japanese so-called 1st person pronouns, which seem much like imposters, as a species of names (and ultimately, nouns) rather than indexicals.

7.6. Chapter appendix: An alternative proposal of Kripke-Kaplan plus Predelli[137]

Let's remember here that the motivation for adopting Matushansky's (2008) proposal of names as descriptions endowed with an indexical element was to account for the de se character of names in Japanese, which cannot be explained by a Kripke-Kaplan account which posits a strict distinction between names and indexicals. This led to a developed version of Matushansky's (2008) account of names, which was adapted and also applied to nouns and so-called Japanese 1st person pronouns. Given that the whole analysis is based on Matushansky's (2008) account of proper names, it should be established whether it is indeed the best possible choice, as well as explain why this is so. The choice of a descriptive approach was deemed appropriate given the data of names being used as predicates, bearing articles, etc. Moreover, by positing an indexical element in the representation it was possible to circumvent the problems that inevitably come with a descriptivist stance, as famously noted by Kripke (1980) and as presented in chapter 2. Given all this, it might feel as if we have the best of both worlds.

Nor surprisingly, things are not as simple as this. According to Predelli (2009, 2017), an account that analyses the meaning of names in terms of 'the bearer of N', whether this component is seen as a description that yields the referent, or as the name's non-constant character, inevitably runs to the problem of wrongly guaranteeing some inferences as logically valid. Let's show this with an example.

(296) Sophia is laughing.

If the meaning of the proper name 'Sophia' is rendered as 'the bearer of /Sofía/ in this context' (along the lines of Matushansky's proposal), then the following inferences are logically valid, or as Predelli (2017) calls them, they are 'character-guaranteed':

(297) There exists a name.

Or even more far-fetched:

(298) There is a language.

[137] I thank Yasutada Sudo for extensive and fruitful discussion on this topic.

It feels bizarre that these inferences should be logical truths, as they seem to follow from the practicalities of language use, not from the meaning of the expressions themselves. In Predelli's (2013) terminology which was introduced in chapter 5, such inferences are 'merely generally settled', rather than character-guaranteed. This argument seems to threaten the claim that Matushansky's (2008) account has all the benefits of the Kripkean proposal thanks to the rigidity afforded by the indexical component that is added, as it turns out to lead to problems not faced by Kripke's account. This is indeed a serious issue that one has to address before opting for Matushansky (2008) over Kripke (1980), yet as mentioned previously one important advantage of Matushansky (2008) was that it can account for the observation that names can be de se in Japanese, something deemed impossible under a Kripkean view.

But if it were possible to account for the de se use of names in a Kripkean framework, then given the ensuing problems with faulty inferences mentioned above, it would hardly be justifiable to choose Matushansky's (2008) proposal. One such possibility can be afforded by accounting for the de se not at the level of character, since the Kripkean proposal advocates that proper names have an uninteresting one, but at the level of bias (Predelli 2013). Although Predelli (2017) does not discuss the exact idea of names having de se uses, his comments about how cognitive significance should be divorced from matters of character are quite fitting here (2017: 68):

> Appeals to cognitive value as the starting point for semantic inquiry are, at best, foolhardy. More fundamentally reckless is the attitude according to which cognitive value is primarily a property of expressions, rather than of uses and users—an attitude that makes it almost inevitable to look for character- and/or content-based peculiarities behind the intuitive discrepancy between, say, 'Cicero is Cicero' and 'Cicero is Tully'.

That is to say, if cognitive significance is to be sought outside the level of character, at the domain of uses, then it will be possible to maintain the Kripkean picture and account for de se at the level of bias. In fact, Predelli (2017) chooses this path in order to defend the Kripkean (or as he calls it, 'Millian') picture from alleged counterexamples such as uses of names as predicates, bound variables, etc. (mentioned above) by analysing such cases as phenomena related to the expressions' use rather than their character meaning (2017: 123):

> It seems inevitable that what is at issue are not the semantic properties of any of the expressions they involve: if certain outcomes evaporate even when these properties are kept constant, they must surely stem from the vagaries of articulation, rather than the verities of character.

At first glance, this seems like a very attractive path to take, menacing the entire account of Japanese self-reference that was designed based on Matushansky (2008). However, it might turn out that this path is itself rife with problems, as I will suggest below.

7.6.1. Problems for the combinatory approach of Kripke-Kaplan plus Predelli

It turns out that the problem of this combinatory account is related to Sherman's (2014) criticism against Predelli (2013), which was discussed in chapter 5: the lack of a clear distinction between bias and character essentially opens the door to the possibility that all meaning could be accounted for in terms of bias, and not just non-truth-conditional one as intended by Predelli (2013). It is reminded here that in chapter 5 non-displaceability was suggested as a possible criterion between what should be accounted for as bias and what not, but understandably this will only work for a handful of expressions, essentially bona fide expressives. Now that the theory of bias is made relevant to names, this problem resurfaces. Let's see how.

As noted above, Predelli (2013) advises against any attempts to locate cognitive significance at the level of character, as he says that it is related to issues of use instead. However, this opens the door to the following bizarre possibility: the existence of indexicals that are not de se (assuming that bias is not an obligatory component).[138] Given our learnings from the essential indexical story (Perry 1979), this is a critical conceptual problem for any account that banishes cognitive significance from character and pushes it over to bias instead, such as Predelli's (2017) and consequently for the combinatory account of Kripke-Kaplan plus Predelli.

What could one do against this conceptual problem? This is exactly where Sherman's (2014) criticism comes handy, or even inspirational: why not bite the bullet and abolish the difference between character and bias altogether, and use bias across the board. In any case, if we remember Predelli's (2013) framework from chapter 5, character-guaranteed inferences are those that are true in all contexts of use (generally settled), whereas those that follow from the pecularities of use are true only in those contexts covered by the theory of use in question (merely generally settled). One way to see this distinction is as quantitative, which means that the difference between inferences that are character-guaranteed and those that follow from questions of use is not an ontological one.

But of course, for this radical solution to work one has to make sure that character can indeed be subsumed by bias. As previously mentioned, the only expressions which are said to have an interesting (i.e. non-constant) character are indexicals; the original idea is that a character is a function from contexts to contents, so this is how indexicals get their reference. If we eradicate the notion of character then indexicals will lose their special status, but they could be accounted for as variables instead. To exemplify this with the indexical 'I', it would be analysed as a variable just like 3rd person pronouns, securing reference from the assignment function, while the additional information that it is used to refer to the speaker of the context, together with the cognitive significance of the de se, would be accounted for at the level of bias. In fact, this looks quite similar to an idea supported by a number of linguists, namely that 1st and 2nd pronouns are variables whose person features are presuppositions just like gender features (e.g. Heim & Kratzer 1998, Schlenker 2003, Sauerland 2008, and others).

[138] This side-effect was noticed and suggested to me by Yasutada Sudo (p.c.).

The result of this radical strategy will be that the inferences that follow from Matushansky's (2008) account will no longer be character-guaranteed but use-guaranteed, in total agreement with Predelli (2009, 2017). This means that the problem raised by Predelli (2009, 2017) against Matushansky (2008) does not emerge and her proposal is deemed a valid choice as an account for proper names. Of course, so is the combinatory account of Kripke-Predelli (now stripped of the Kaplanian component, with character out of the picture). In theory it would not make much of a difference which of the two one would choose, as the analysis proposed above for Japanese names, 1st person pronouns and nouns could be maintained as corresponding to the bias component in both cases.

The conclusion from this section is that if we push the problem of mistakenly character-guaranteed inferences to its logical end, then the only outcome is to admit Sherman's (2014) point about the lack of a meaningful distinction between character and bias, and eventually posit its dissolution. Predelli (2013) convincingly argued that character cannot do what bias can, yet it turns out that bias may be able to do everything character can. Of course, another alternative would be for one not to be as troubled by the issue of character-guaranteed inferences (assuming perhaps naively that what is at issue here is propositions as reflecting utterances, rather than as timeless entities completely independent from language), which would mean that Matushansky's (2008) account of names as descriptions involving an indexical element could be maintained, and preferred over a Kripke-Kaplan plus Predelli approach on the basis that it can account for the de se property of names in Japanese by locating it in the character, an option unavailable to its rival. I will not make a choice here about which of the two alternatives is preferable, and I leave it to the reader to pick her/his favourite one.

Chapter 8. Concluding remarks

The following may serve as a short summary of the major claims defended in this thesis:

1. **Expressives are necessarily de se**: Although this claim is not entirely novel, as it follows from Schlenker's (2007) analysis of expressives as 'indexical presuppositions' and it has been discussed (but rejected) by McCready & Wechsler (2012), it is corroborated in this thesis by specific examples. As a result, it is established that any framework of expressives should be able to reflect this property, and possible ways this is can be accounted for or not by the existing frameworks is examined.

2. **The category of 'expressive' or 'use-conditional' items has not been properly defined, and the extant frameworks are misguided**: Due to the lack of a rigorous delimitation of the category of expressives, or more recently UCIs, it can be (and has been) used as an all-purpose label for anything that is not part of assertive meaning, which deprives it of theoretical significance. The efforts to address so-called 'mixed' items that appeared as counterexamples to Potts (2005, 2007) by McCready (2010) and especially Gutzmann (2015) led to a complex and not restrictive enough logic which allows almost anything that lied outside the assertion to be classified and treated as a UCI. Moreover, there are some conceptual problems in the framework, such as the fact that felicity is treated on a par with truth while they are fundamentally different notions, and the unfortunate return of the Binding Problem. During the critical dissection of Gutzmann's (2015) framework in Chapter 4, the alleged five different categories of UCIs were examined and it was found that one of them is much more presupposition-like (functional mixed), two of them are not justified to be in the compositional mechanism (isolated expletive and isolated mixed) and one of them should be accounted for in discourse/pragmatic rather than strictly semantic terms (functional shunting). As a result, the only category left was that of functional expletives, which are exactly the main items Potts (2005, 2007) initially focused on.

3. **The debate over whether expressives are a kind of presupposition or a distinct category of their own is settled with a conciliatory response**: the ambivalent behaviour of expressives, i.e. the fact that a part of their meaning may be filtered (presupposition-like) but another part may unrestrictedly project (unlike presupposition) is explained by proposing that expressives are comprised by two components of meaning, one compositionally irrelevant, which is dubbed 'associative meaning' and modelled according to Predelli's (2013) Theory of Bias, and one compositionally relevant, which is presuppositional, as per the insights of Schlenker (2007) and other authors. However, expressives are defined as ordinary rather than indexical presuppositions (pace Schlenker 2007), but also endowed with indexicality via their 'associative' component. These claims are shown with detailed examples.

4. **A formal account of self-reference in Japanese is offered**: while it has been known that Japanese so-called personal pronouns are not genuine pronouns but nouns,

this is the first time that a full semantic treatment is offered not just about these expressions, but also nouns and names which are also used for the speaker's self-reference in Japanese.

5. **The absolute distinction between names and indexicals is denied**: With regards to the theoretical implications of the proposed account for the Japanese data, the major one is that it argues against the clear-cut distinction between indexicals (Kaplan 1989) and names (1980). More precisely, as names in Japanese can be used like first person indexicals, i.e. de se, they cannot be accommodated by an analysis that considers them entirely devoid of linguistic meaning, as per Kripke (1980). Instead, they are analysed following Matushansky's (2008) account which falls in the category of 'quotation theories' of proper names, and which is adapted for so-called Japanese personal pronouns and common nouns used as such, offering an integrative account of non-pronominal de se reference.

6. **The de se is more widely encoded in natural language than previously thought**: As it is shown that associative meaning is indexical and it lies in a different dimension, it is established that the de se can be encoded beyond the assertive dimension, and that any kind of expression can be de se, as long as it is endowed by associative meaning. Moreover, it is suggested that even in the assertive dimension, the de se can be encoded not just by pronominal expressions, which was already known (Chierchia 1989 Schlenker 2003, Anand 2006, Sudo 2012), but also by expressions which are not strictly pronouns, yet are used similarly, such as nouns and proper names.

References

Abbott, Barbara (2006) 'Where have some of the presuppositions gone?'In Betty J. Birner and Gregory Ward (eds.) *Drawing the Boundaries of Meaning: Neo-Gricean studies in pragmatics and semantics in honor of Laurence R. Horn*. Amsterdam: John Benjamins.

Abbott, Barbara (2010) *Reference*. Oxford: Oxford University Press.

Abe, H. (2004) 'Lesbian Bar Talk in Shinjuku, Tokyo' in Okamoto S. and Shibamoto J. (eds.) *Japanese Language, Gender, and Ideology: Cultural Models and Real People*. Oxford: OUP.

Abe, H. (2010) *Queer Japanese: Gender and Sexual Identities Through Linguistic Practices*. New York: Palgrave Macmillan.

Alpatov, V. (2006) 'Words of Kinship in Japanese' in Boikova, E. and Rybakov, R. (eds.) *Kinship in the Altaic World: Proceedings of the 48th Permanent International Altaistic Conference, Moscow 10-15 July, 2005*. Wiesbaden: Harrassowitz.

Amaral, Patricia, Craige Roberts and Allyn Smith (2007) 'Review of The Logic of Conventional Implicatures by Chris Potts'. Linguistics and Philosophy 30(6): 707–749.

Anand, Pranav (2006) *De De se*. Ph.D. dissertation, MIT Linguistics.

Anand, Pranav and Andrew Nevins (2004) 'Shifty operators in changing contexts'. In Kazuha Watanabe and Robert B. Young (eds.) *Proceedings of SALT 14*, 20–37. Ithaca, NY: CLC Publications.

Anderson, Luvell & Ernie Lepore (2013) 'What Did You Call Me? Slurs as Prohibited Words'. Analytic Philosophy 54.3, 350-363.

Aoki, S. (1989) 'Bunpoo no taishooteki kenkyuu: Furansugo to Nihongo' (A contrastive analysis of grammar in French and Japanese) in Yamaguchi, Y. (ed.) *Nihongo to Nihongokyooiku*, Vol. 5 (Japanese and Japanese Pedagogy 5), 290–311. Tokyo: Meiji-shoin.

Asada, S. (1998) 'Tachi, Mother, Woman, The Many Faces of Love', in Summerhawk, B. , McMahill C. and McDonald D. (eds.) *Queer Japan: Personal Stories of Japanese Lesbians, Gays, Transsexuals, and Bisexuals*. Norwich: New Victoria.

Au, Terry (1999) ‘Language and Thought’. In Robert A. Wilson and Frank C. Keil (eds.) *The MIT Encyclopedia of the Cognitive Sciences*. Cambridge, Massachusetts: MIT Press.

Azuma, S. (2000) 'Linguistic Strategy of Involvement: An Emergence of New Political Speech in Japan' in De Landtsheer, C. and Feldman, O. (eds.) *Beyond public speech and symbols: explorations in the rhetoric of politicians and the media*. Westport: Praeger.

Bach, K. (1981) ‘What’s in a name’. Australasian Journal of Philosophy, 59, 371–386.

Bach, K. (1987) ‘Thought and reference’. Oxford: Oxford University Press.

Bach, K. (2002) ‘Giorgione was so-called because of his name’. Philosophical Perspectives, 16, 73–103.

Bach, Kent (1999) ‘The myth of conventional implicature’. Linguistics and Philosophy 22(4): 367–421.

Bachnik, Jane (1994) ‘Introduction: Uchi/Soto: Challenging our concept of self, social order and language’ in Jane Bachnik and Charles Quinn (eds.) *Situated meanings: Inside and outside in Japan: Self, society and language*. Princeton, NJ: Princeton University Press.

Batchelor, Ronald E. (2006) *Using Spanish Synonyms*. Cambridge: CUP.

Beaver, David (1997) ‘Presupposition’. In: van Benthem, J., ter Meulen, A. (Eds.), Handbook of Logic and Language. Elsevier, Amsterdam.

Beaver, David, Craige Roberts, Mandy Simons, and Judith Tonhauser (2017) ‘Questions Under Discussion: Where Information Structure Meets Projective Content’. Annual Review of Linguistics Vol. 3:265-284.

Bekoff, Marc (2000) 'Animal Emotions: Exploring Passionate Natures: Current interdisciplinary research provides compelling evidence that many animals experience such emotions as joy, fear, love, despair, and grief—we are not alone' BioScience, Volume 50, Issue 10, 1: 861–870.

Beltrama, Andrea and Jackson Lee (2015) ‘Great pizzas, ghost negations: The emergence and persistence of mixed expressives’. Sinn und Bedeutung 19.

Benveniste, E. (1971) *Problems in General Linguistics*. Translated by Mary Elizabeth Meek. Cora Gables, FA: University of Miami Papers.

Bermúdez, José-Luis (1998) *The Paradox of self-consciousness*. Cambridge: MIT Press.

Bhat, D. N. S. (2004) *Pronouns*. Oxford: OUP.

Bianchi, Claudia (2010) 'Contextualism'. In Jan-Ola Östman and Jef Verschueren (eds.) Handbook of Pragmatics: 2010 Instalment. Amsterdam: John Benjamins.

Bianchi, Claudia (2014) 'Slurs and Appropriation: An Echoic Account', Journal of Pragmatics, Volume 66, Pages 1-162.

Borg, Emma (2009) 'Meaning and context: a survey of a contemporary debate'. In: D. Whiting (ed). The Later Wittgenstein on Language. Palgrave. 96-113.

Bowe, H., Martin, K. and Manns, H. (2014). *Communication Across Cultures: Mutual Understanding in a Global World*. Cambridge: CUP.

Brown, R. (1996) 'The Language of Social Relationship' in Slobin, D. et al (eds.) *Social Interaction, Social Contest, and Language*. Mahwah, NJ: Lawrence Erlbaum.

Budwig, N. (1985). I, Me, My and 'Name': Children's early systematizations of forms, meanings and functions in talk about the self. Papers and Reports on Child Language Development 24. 30–7.

Büring, Daniel (1999) 'Topic'. In Bosch, Peter & Rob van der Sandt (eds) *Focus: Linguistic, Cognitive, and Computational Perspectives*. Cambridge: CUP. 142-165.

Camp, Elizabeth (2013) 'Slurring Perspectives'. Analytic Philosophy Vol. 54 No. 3, 330–349.

Cappelen, Herman and Josh Dever (2013) *The Inessential Indexical: On the Philosophical Insignificance of Perspective and the First Person*. Oxford: Oxford University Press.

Castañeda, Hector-Neri (1966) ''He': A Study in the Logic of Self-Consciousness'. *Ratio*, 8: 130–157.

Castroviejo Miró, Elena (2008). 'An expressive answer: Some considerations on the semantics and pragmatics of wh-exclamatives', Proceedings of the Chicago Linguistic Society 44.2: 3–17.

Chierchia, Gennaro (1989) 'Anaphora and attitudes de se'. In R. Bartsch, J. van Benthem and B. van Emde Boas (eds). *Semantics and Contextual Expression*. Dordrecht: Foris. 1- 31.

Chierchia, Gennaro 1998. "Reference to Kinds across Language." Natural Language Semantics 6 (4): 339–405.

Chierchia, Gennaro and Sally McConnell-Ginet (1990) *Meaning and Grammar*. Cambridge: MIT Press.

Clancy, P. (1985) 'The Acquisition of Japanese' in Slobin D. (ed.) *The Crosslinguistic Study of Language Acquisition, Volume 1: The Data*. Hillsdale, N.J: L. Erlbaum Associates.

Clark, E. V. (1978). From gesture to word: On the natural history of deixis in language acquisition. In J. S. Bruner & A. Garton (eds), Human growth and development: Wolfson College lectures 1976, 85–120. Oxford: Oxford University Press.

Collins, C. and P. M. Postal (2012) *Imposters: A Study of Pronominal Agreement*. Cambridge, MA: MIT Press.

Cooke, J. R. (1968) *Pronominal Reference in Thai, Burmese, and Vietnamese*. Berkeley: University of California Press.

Cooper, R (1983) Quantification and Syntactic Theory. Studies in Linguistics and Philosophy, vol. 21. Reidel, Dordrecht.

Crespo, María Inés (2015) 'Affecting meaning: Subjectivity and evaluativity in gradable adjectives'. PhD dissertation, University of Amsterdam, Institute for Logic, Language and Computation.

Daijirin dictionary (スーパー大辞林 3.0)

De Houwer, A. & Gillis, S. (1998). Dutch child language: an overview. In S. Gillis & A. De Houwer (eds), The acquisition of Dutch. Amsterdam: Benjamins.

Dekker, Paul (2002) 'A Proper Architecture for Presupposition and Quantification'. MS, University of Amsterdam.

Egan, A., (2006), 'Secondary qualities and self-location', Philosophy and Phenomenological Research 72: 97-119.

Evans, Gareth (1981/1985) 'Understanding Demonstratives', in his Collected Papers, Oxford: Oxford University Press, 291–321.

Feng, Guangwu (2010) *A Theory of Conventional Implicature and Pragmatic Markers in Chinese*. Leiden: Brill.

Flannery, G. (2010) 'Open and Closed Systems of Self-References and Addressee-Reference in Indonesian and English: A Broad Typological Distinction' in De Busser, Rik and Treis Yvonne (eds). *Selected Papers from the 2009 Conference of the Australian Linguistics Society.* ANU.

Frege, Gottlob (1892/1948) 'On Sense and Reference'. Philosophical Review, 57: 209–30.

Frege, Gottlob (1897/1979). 'Logic', in Hans Hermes, Friedrich Kambartel, and Freidrich Kaulbach(eds),Posthumous Writings: Gottlob Frege, trans. Peter Long and Roger White. Oxford: Blackwell, 126–51.

Frege, Gottlob (1918/1956) 'The Thought: A Logical Inquiry'. Mind, 65: 289–311.

Frellesvig, B. (2010) *A History of the Japanese Language*. Cambridge: CUP.

Frey, Werner (2010) 'A'-movement and conventional implicatures: About the

grammatical encoding of emphasis in German', Lingua 120: 1416–35.

Gertler, Brie (2010) Self-knowledge. Routledge

Geurts, B. (1997). Good news about the description theory of names. Journal of Semantics, 14, 319–348.

Geurts, Bart (2007) 'Really fucking brilliant'. Theoretical Linguistics 33.2, 209-214.

Grice, Paul (1989) *Studies in the Way of Words*. Cambridge MA: Harvard University Press.

Gutzmann, Daniel & Elin McCready (2014) 'Using descriptions.' Submitted to Empirical Issues in Syntax and Semantics 10.

Gutzmann, Daniel (2013). 'Expressives and beyond: An introduction to varieties of use-conditional meaning', in Daniel Gutzmann and Hans-Martin Gärtner (eds) *Beyond Expressives: Explorations in Use-Conditional Meaning. Current Research in the Semantics Pragmatics-Interface 28*. Leiden: Brill, 1–58.

Gutzmann, Daniel (2015) *Use-conditional Meaning. Studies in Multidimensional Semantics*. Oxford Studies in Semantics and Pragmatics 6. Oxford: Oxford University Press.

Gutzmann, Daniel (2016) 'If expressivism is fun, go for it! Towards an expressive account of predicates of personal taste.' In: van Wijnbergen-Huitink, Janneke & Cécile Meier (eds.): Subjective Meaning. Alternatives to Relativism. Berlin: de Gruyter. 21–46.

Gutzmann, Daniel and Elin McCready (2016) 'Quantification with Pejoratives'. In: Finkbeiner, Rita & Jörg Meibauer & Heike Wiese (eds.). Pejoration. Amsterdam: John Benjamins.

Haiman, J. (2011) *Cambodian Khmer*. Amsterdam: John Benjamins Publishing Company.

Harada, S. I. (1976) 'Honorifics'. In M. Shibatani (ed.), Syntax and Semantics Vol. 5, Japanese Generative Grammar. New York: Academic Press, 499–561.

Harris, Jesse A. & Christopher Potts (2009) 'Perspective-shifting with appositives and expressive'. Linguistics and Philosophy 32(6):523-552.

Hasegawa, Y. (2014) *Japanese: A Linguistic Introduction.* Cambridge: CUP.

Hasegawa, Y. and Hirose, Y. (2005) 'What the Japanese Language Tells Us about the Alleged Japanese Relational Self', *Australian Journal of Linguistics* Vol. 25, No. 2, pp. 219-251.

Heim, Irene (1982) *The semantics of definite and indefinite noun phrases*. Ph.D. thesis, University of Massachusetts, Amherst.

Heim, Irene and Angelika Kratzer (1998) *Semantics in Generative Grammar*. Oxford: Blackwell.

Heim, Irene. 1991. Notes on the first person. Class notes.

Heine, B. and Song, K-A. (2011) *On the grammaticalisation of personal pronouns*. Journal of Linguistics 47.03, 587–660.

Hinds, John (1986) *Japanese*. London: Croom Helm.

Hoji, Hajime. (1990). On the so-called overt pronouns in Japanese and Korean. In Eun-Jin Baek (Ed.), Papers from the seventh International Conference on Korean Linguistics (pp. 61–78). Toronto: University of Toronto Press.

Hoji, Hajime (1991) Kare. In Carol Georgopoulos and Roberta Ishihara (Eds), Interdisciplinary approaches to language: Essays in honor of S.-Y. Kuroda. Dordrecht: Kluwer.

Holton, D., Mackridge, P., Spyropoulos, V. and Philippaki-Warburton, E. (2012) *Greek: A Comprehensive Grammar of the Modern Language.* London: Routledge.

Hori, M. (1995) 'Subjectlessness and Honorifics in Japanese: A case of Textual construal' in Hasan, R. and Fries, P. (eds.) *On Subject and Theme: A discourse functional perspective*. Amsterdam: John Benjamins.

Horiguchi, S. (1979) 'Nisaiji no hanashikotoba ni mitateru danjo no sai', Paper presented at the Sixth International Christian University Linguistics Symposium.

Horiguchi, S. (1981) 'Ichininshooshi, nininshooshi, koshooshi' in F. C. Peng (ed.) *Nihongo to danjosa: Male/female differences in Japanese*. Tokyo: The East-West Sign Language Association.

Horn, Larry (2013). 'I love me some datives: Expressive meaning, free datives, and F-implicature', in Daniel Gutzmann and Hans-Martin Gärtner (eds) *Beyond Expressives: Explorations in Use-Conditional Meaning. Current Research in the Semantics Pragmatics-Interface 28*. Leiden: Brill, 1–58.

Huang, C.T. James and C.-S. Luther Liu (2001) 'Logophoricity, attitudes and ziji at the interface'. In *Syntax and Semantics: Long distance reflexives*. New York: Academic Press.

Ide, S. (1990) 'How and why do women speak more politely in Japanese?' in Ide, S and McGloin, N. H. (eds.) *Aspects of Japanese women's language*, 63–79. Tokyo: Kuroshio.

Ishikawa, A., et al. (1981) 'Address terms in modern Japanese: A Sociolinguistic analysis' Sophia Linguistica, (8), 129-141.

Ishiyama, O. (2008) *Diachronic perspectives on personal pronouns in Japanese*. State University of New York at Buffalo dissertation. Ann Arbor: ProQuest.

Japanese Ministry of Education (1952) Korekara no keigo (Honorific expressions in the future) http://www.bunka.go.jp/kokugo_nihongo/sisaku/joho/joho/kakuki/01/tosin06/

Jespersen, Otto (1923) *Language: Its Nature, Development and Origin*. London: George Allen and Unwin.

Jugaku, A. (1979) *Nihongo to onna* (The Japanese language and women). Tokyo: Iwanami Shoten.

Kadokawa dictionary (角川類語新辞典)

Kamio, A. (1995) 'Territory of information in English and Japanese and psychological utterances', *Journal of Pragmatics* 24, 235-264.

Kamio, Akio (1997) Territory of information. John Benjamins, Amsterdam.

Kamio, Akio (1994) The theory of territory of information: The case of Japanese. Journal of Pragmatics 21: 67-100.

Kaplan, David (1968) 'Quantifying in'. Synthèse 19: 178-214.

Kaplan, David (1989) 'Demonstratives: An essay on the semantics, logic, metaphysics, and epistemology of demonstratives and other indexicals' in J. Almog, J. Perry and H. Wettstein (eds) *Themes from Kaplan*. New York: Oxford University Press, pp. 481-563.

Kaplan, David (1999) 'The Meaning of Ouch and Oops: Explorations in the theory of Meaning as Use'. Ms., UCLA.

Karttunen, Lauri & Annie Zaenen (2005) 'Veridicity'. In Graham Katz, James Pustejovsky & Frank Schilder (eds.) *Annotating, Extracting and Reasoning about Time and Events*. No. 05151 in Dagstuhl Seminar Proceedings, Dagstuhl, Germany: Internationales Begegnungs- und Forschungszentrum fur Informatik (IBFI), Schloss Dagstuhl, Germany.

Karttunen, Lauri & Stanley Peters. 1979. Conventional implicature. In Choon-Kyu Oh & David Dinneen (eds.), Syntax and Semantics 11: Presupposition, 1–56. New York: Academic Press.

Karttunen, Lauri (1973) 'Presuppositions and compound sentences'. Linguistic Inquiry 4(2): 169–193.

Karttunen, Lauri and Stalney Peters (1979) 'Conventional implicature'. In: Choon-Kyu Oh and David Dinneen (eds) *Syntax and Semantics, Vol. 11: Presupposition.* 1-56. New York: Academic Press.

Karttunen, Lauri and Stanley Peters (1977) 'Requiem for presupposition'. In Proceedings of the 3rd Annual Meeting of the Berkeley Linguistics Society, pp. 360-371.

Kitagawa, Chisato (1981) Anaphora in Japanese: kare and zibun. Coyote Papers 2.61-75. Tucson, AZ: University of Arizona Linguistic Circle.

Keene, Donald (2013) *Emperor of Japan: Meiji and His World, 1852-1912*. New York: Columbia University Press.

Kindermann, Dirk (2012) *Perspective in context: relative truth, knowledge and the first person*. PhD thesis, University of St. Andrews.

Klein, Ewan (1979) *On Sentences Which Report Beliefs, Desires and Other Mental Attitudes*. PhD Dissertation, Cambridge University.

Klein, Ewan and Ivan Sag (1985). 'Type-driven translation', Linguistics and Philosophy 8: 163–201.

Kneale, W. (1962). Modality de dicto and de re. In E. Nagel, P. Suppes, & A. Tarski (Eds.), Logic, Methodology and Philosophy of Science. Proceedings of the 1960 International Congress (pp. 622–633). Stanford: Stanford University Press.

Kotonoha Shonagon Online Corpus, http://www.kotonoha.gr.jp/shonagon/search_form

Kratzer, Angelika (1999). 'Beyond ouch and oops: How descriptive and expressive meaning interact', in Cornell Conference on Theories of Context Dependency. Ithaca, NY: Cornell University.

Kratzer, Angelika. 2009. Making a pronoun: fake indexicals as windows into the properties of pronouns. Linguistic Inquiry 40(2). 187–237.

Kripke, Saul (1980) *Naming and Necessity*. Harvard University Press.

Kuno, S. (1988) Blended quasi-direct discourse in Japanese. In W. Poser (Ed.), Papers from the Second International Workshop on Japanese Syntax, pp. 75–102. Stanford: CSLI.

Kuroda, S.-Y. (1965) Generative Grammatical Studies in the Japanese Language. Doctoral dissertation, MIT.

Kuroda, Shige-Yuki (1988) Whether we agree or not: A comparative syntax of English and Japanese, Linguisticae Investigationes, 12: 1-47.

Kusumoto, Kiyomi (1998) 'Tenses as logophoric pronouns'. MIT/UConn/UMass Semantics Workshop.

Lakoff, George (1972) 'Linguistics and Natural Logic', in Donald Davidson and Gilbert Harman (eds) *Semantics of Natural Language*. Dordrecht: Reidel, pp. 545–665.

Larson, Richard and Gabriel Segal (1995) *Knowledge of Meaning*. Cambridge: MIT Press.

Lasersohn, Peter (2005) 'Context dependence, disagreement, and predicates of personal taste'. Linguistics and Philosophy 28(6): 643–686.

Lasersohn, Peter (2007) 'Expressives, Perspective and Presupposition.' Theoretical Linguistics 33.2: 223-230.

Leech, Geoffrey (1981) Semantics: The Study of Meaning. Penguin.

Lewis, D. (1979) 'Score-keeping in a language game'. In Rainer Bauerle, Urs Egli and Arnim von Stechow (Eds.) *Semantics from a Different Point of View*. Berlin: Springer.

Lewis, David (1979) 'Attitudes de dicto and de se'. The Philosophical Review 88.4: 513-543.

Liu, Mingya (2012) *Multidimensional Semantics of Evaluative Adverbs*. Leiden: Brill.

Longobardi, G. (2008) 'Reference to individuals, person, and the variety of mapping parameters'. in A Klinge & H Müller (eds), *Essays on Nominal Determination: from morphology to discourse management*. John Benjamins, pp. 189-211.

Loveday, L. (1986) *Explorations in Japanese Sociolinguistics*. Amsterdam: John Benjamins.

Lunsing, W. and Maree, C. (2004) 'Shifting Speakers: Negotiating Reference in Relation to Sexuality and Gender' in Okamoto S. and Shibamoto J. (eds.) *Japanese Language, Gender, and Ideology: Cultural Models and Real People*. Oxford: OUP.

Macià, Josep (2002) 'Presuposicion y significado expresivo'. Theoria: Revista de Teoria, Historia y Fundamentos de la Ciencia. 3 (45): 499-513.

Magidor, Ofra (2015) 'The Myth of the de se', Philosophical Perspectives Volume 29, Issue 1, Pages 249–283.

Maier, E. (2009). Japanese reported speech: against a direct–indirect distinction. In H. Hattori, T. Kawamura, T. Ide, M. Yokoo, and Y. Murakami (Eds.), New Frontiers in Artificial Intelligence, Volume 5447 of Lecture Notes in Computer Science. Berlin/Heidelberg: Springer.

Marcus, Ruth Barcan (1961) 'Modalities and Intensional Languages,' *Synthèse*, 13: 303–322.

Martin, S. (1988) *A reference grammar of Japanese*. Rutland: Charles E. Tuttle Company.

Matushansky, Ora. 2008. On the linguistic complexity of proper names. Linguistics and Philosophy 31(5). 573–627.

Maynard, S. (2016) *Fluid orality in the discourse of Japanese popular culture.* Amsterdam: John Benjamins.

McCready, Elin (2007) Context shifting in questions and elsewhere. In E. Puig Waldmuller, ed., Proceedings of Sinn und Bedeutung 11, pages 433–447. Barcelona: Universitat Pompeu Fabra.

McCready, Elin (2008) 'Classification without assertion' in Matthew A. Tucker, Anie Thompson, Oliver Northrup, and Ryan Bennett (eds.) Proceedings of the Fifth Formal Approaches to Japanese Linguistics Conference, volume 64 of MIT Working Papers in Linguistics, 141–154.

McCready, Elin (2009) 'Conventional Implicature and Classifiers', Journal of Cognitive Science 10: 2, 195-208.

McCready, Elin (2009). 'What man does', Linguistics and Philosophy 31.6: 671–724.

McCready, Elin (2010) Pronominal features and conventional implicature. In Hiroshi Yoshiba, Kazuo Nakazawa, Shinichi Takeuchi, Shigeo Tonoike, A. Kawabata, H. Nomura, and S. Yamamoto (eds) Eigokenkyu no jisedai ni mukete: Current studies for the next generation of English linguistics and philology: a festschrift for Minoji Akimoto, 199–212. Hituzi Syobo.

McCready, Elin (2010). 'Varieties of conventional implicature', Semantics and Pragmatics 3.8: 1–57.

McCready, Elin (2014) 'Expressives and Expressivity', Open Linguistics Vol. 1 issue 1.

McCready, Elin (2015) Reliability in Pragmatics. Oxford: Oxford University Press.

McCready, Elin and Stephen Wechsler (2012) 'Meaning as use: Problems and prospects', NASSLLI Workshop on Indexicality, Expressives, and Self-References, University of Texas, Austin, TX. Available online at <http://nasslli2012.com/files/courses/workshop-final.pdf.>.

McCready, Elin and Yohei Takahashi (2013). 'Good reasons', in Daniel Gutzmann and Hans-Martin Gärtner (eds), Beyond Expressives: Explorations in Conventional Non-truth-conditional Meaning, Current Research in the Semantics/Pragmatics Interface 28. Leiden: Brill, 201–29.

McNally, Louise and Chris Kennedy (2008) *Adjectives and Adverbs: Syntax, Semantics and Discourse*. Oxford: Oxford University Press.

Miller, L. (2004) 'You Are Doing Burikko! Censoring/Scrutinizing Artificers of Cute Femininity in Japanese' in Okamoto S. and Shibamoto J. (eds.) *Japanese Language, Gender, and Ideology: Cultural Models and Real People*. Oxford: OUP.

Millikan, Ruth (1990), 'The myth of the essential indexical', Noûs 24: 723-34.

Miyazaki, A. (2004) 'Japanese Junior High School Girls' and Boys' First-Person Pronoun Use and Their Social World', in Okamoto and Shibamoto (eds) *Japanese Language, Gender, and Ideology: Cultural Models and Real People*. Oxford: OUP.

Morita, E (2009) 'Arbitrating community norms: The use of Me in Japanese Discourse' in Reyes A. and Lo A. (eds) *Beyond Yellow English: Toward a Linguistic Anthropology of Asian Pacific America*. Oxford: Oxford University Press, 175-192.

Mühlhäusler, P. and Harré, R. (1990) *Pronouns and people: The linguistic construction of social and personal identity*. Oxford: Blackwell.

Nakai, Satoru (1976) A study of anaphoric relations in Japanese. Amherst, MA: University of Massachusetts, MS.

Nakayama, S. (1982) On English and Japanese Pronouns. MA thesis, University of Tokyo.

Nariyama, S. (2003) *Ellipsis and reference tracking in Japanese*. Amsterdam: John Benjamins.

Neta, Ram (2011) 'The Nature and Reach of Privileged Access'. In Anthony Hatzimoysis (ed.) Self-knowledge. Oxford: Oxford University Press.

Ninan, Dilip (2010) 'De Se Attitudes: Ascription and Communication' Philosophical Compass 5:551-67.

Noguchi, Tohru (1997) 'Two Types of Pronouns and Variable Binding'. *Language*, Vol. 73, No. 4, pp. 770-797.

Nunberg, Geoffrey. 1993. Indexicality and deixis. Linguistics and Philosophy 16(1). 1–43.

Okamoto, S. (1995). ' "Tasteless" Japanese: Less "feminine" speech among young Japanese women' in Hall, K. and Bucholtz, M. (eds.), *Gender articulated: Language and the socially constructed self*, 297–325. New York: Routledge.

Partee, Barbara (1984) 'Compositionality'. In Frank Landman and Frank Veltman (eds.) *Varieties of Formal Semantics*, 281–312. Dordrecht: Foris.

Partee, Barbara. 1989. Binding implicit variables in quantified contexts. In Chicago Linguistic Society 25, 342–365.

Peacocke, Christopher (2010) 'Self-Consciousness', Revue de Métaphysique et de Morale 4: 521-551, issue entitled "Le Moi /The Self/ Le Soi" ed. B. Longuenesse.

Peacocke, Christopher (2014) *The Mirror of the World*. Oxford: Oxford University Press.

Pearson, Hazel. 2012. The Sense of Self: Topics in the Semantics of De Se Expressions: Harvard University Ph.D. dissertation.

Pelczar, M., & Rainsbury, J. (1998). The indexical character of names. Synthese, 114, 293–317.

Pelletier, Francis Jeffry (2012) 'Holism and Compositionality'. In Markus Werning, Wolfram Hinzen, and Edouard Machery (eds.) *The Oxford Handbook of Compositionality*. Oxford: Oxford University Press.

Percus, O., & Sauerland, U. (2003) 'On the LFs of attitude reports'. Proceedings of the Conference "Sub7 - Sinn Und Bedeutung", University of Konstanz. pp. 228-242.

Perry, J. (1977) 'Frege on demonstratives', Philosophical Review 86: 474-497.

Perry, John (1979) 'The Problem of the Essential Indexical'. Noûs 13:3–21.

Plato. Cratylus.

Podobryaev, Alexander. 2014. Persons, imposters, and monsters. PhD dissertation, MIT.

Postal, P. (1969) 'On so-called 'pronouns' in English' in Reibel, D. and Schane, S. (eds.) *Modern studies in English: Readings in transformational grammar*, 201-24. Englewood Cliffs, NJ: Prentice-Hall.

Potts, Christopher (2003) *The Logic of Conventional Implicatures*. PhD dissertation, University of California Santa Cruz.

Potts, Christopher (2005) *The Logic of Conventional Implicatures*. Oxford Studies in Theoretical Linguistics. Oxford: Oxford University Press.

Potts, Christopher (2007) 'The expressive dimension'. Theoretical Linguistics 33(2): 165–197.

Potts, Christopher (2007b) 'The centrality of expressive indicies'. Theoretical Linguistics 33(2): 165–197.

Potts, Christopher and Shigeto Kawahara (2004) 'Japanese honorifics as emotive defnite descriptions'. Proceedings of SALT 14, 235–254.

Predelli, Stefano (2009) Socrates and 'Socrates'. American Philosophical Quarterly 46.3: 203-212.

Predelli, Stefano (2010). 'From the expressive to the derogatory', in Sarah Sawyer (ed.), *New Waves in Philosophy of Language*. Houndmills and Basingstoke: Palgrave-Macmillan, 164–85.

Predelli, Stefano (2013) *Meaning without Truth*. Oxford: Oxford University Press.

Predelli, Stefano (2017) Proper Names, A Millian Account. Oxford: Oxford University Press.

Quine, W.V. (1956) 'Quantifiers and Propositional Attitudes'. Journal of Philosophy 53: 177–87.

Quine, Willard van Orman. (1969) 'Ontological Relativity'. In van Orman Quine, Willard (ed.), *Ontological Relativity and Other Essays*, 26-68. New York: Columbia University Press.

Raskin, R. and Shaw, R. (1988) 'Narcissism and the use of personal pronouns'. *Journal of Personality*, Volume 56, Issue 2, 393–404.

Recanati, F. (1997). Direct reference: From language to thought. Oxford: Blackwell.

Recanati, François (2005) 'Literalism and contextualism: Some varieties'. In: G. Preyer and G. Peter (eds). Contextualism in Philosophy: Knowledge, Meaning, and Truth. Oxford: Clarendon Press. 171-96.

Reynolds, K. A. (1998). 'Female Speakers of Japanese in Transition' in J. Coates (ed.), *Language and Gender: A Reader*. Oxford: Blackwell.

Richard, Mark (2008) *When Truth Gives Out*. Oxford: Oxford University Press.

Roberts, Craige (1996) 'Information structure in discourse: Toward an integrated formal theory of pragmatics'. In Jae Hak Yoon & Andreas Kathol (eds.), OSUWPL, vol. 49, 91–136. The Ohio State University, Department of Linguistics.

Roberts, Craige. 1996. Information structure in discourse: Towards an integrated formal theory of pragmatics. In Jae-Hak Yoon & Andreas Kathol (eds.), Papers in semantics, vol. 49 Ohio State University Working Papers in Linguistics, Department of Linguistics, The Ohio State University.

Roehrs, D. (2006) 'Pronominal noun phrases, number specifications, and null nouns' in Hartmann, J. and Molnárfi, L. (eds.) *Comparative Studies in Germanic Syntax: From Afrikaans to Zurich German*. Amsterdam: John Benjamins.

Roeper, T. (2006) Not Only I: Notes on the Syntax of Focus Binding. Manuscript, University of Massachusetts, Amherst.

Roeper, T., 1999. Notes on the Only I phenomenon. Manuscript, University of Massachusetts, Amherst.

Saito, Mamoru and Hajime Hoji (1983) Weak Crossover and Move a in Japanese. Natural Language & Linguistic Theory 1.

Sauerland, Uli (2007) 'Beyond unpluggability'. Theoretical Linguistics 33(2): 231–236.

Sauerland, Uli (2008) On the semantics markedness of phi-features. In Daniel Harbour, David Adger & Susana B/ejar (eds.), Phi Theory: Phi-Features across Modules and Interfaces, 57–82. Oxford: Oxford University Press.

Sawada, Osamu (2011) 'The meanings of positive polarity minimizers in Japanese: a unified approach'. Proceedings of SALT 20: 599–617.

Sawada, Osamu (2014) 'On the context-dependent pragmatic strategies of Japanese self-diminutive shift'. In Urtzi Etxeberria, Anamaria Fălăuş, Aritz Irurtzun, and Bryan Leferman (eds.), Proceedings of Sinn und Bedeutung 18, 377-395.

Schlenker, Philippe (1999) *Propositional Attitudes and Indexicality: a Cross-categorial Approach*. PhD dissertation, MIT.

Schlenker, Philippe (2003) 'A plea for monsters'. Linguistics and Philosophy 26(1): 29–120.

Schlenker, Philippe (2003) Indexicality, logophoricity, and plural pronouns. In Jacqueline Lecarme (ed.), Research in Afroasiatic Grammar II (Selected Papers from the Fifth Conference on Afroasiatic Languages, Paris, 2000), 409–428. Amsterdam: John Benjamins.

Schlenker, Philippe (2007) 'Expressive presuppositions'. Theoretical Linguistics 33(2): 237–245.

Schlenker, Philippe (2011) 'Indexicality and de se reports', forthcoming in Klaus von Heusinger, Claudia Maienborn, and Paul Portner (eds.), *Handbook of Semantics*, Mouton de Gruyter.

Schlenker, Philippe (2016) The Semantics/Pragmatics Interface. In: The *Cambridge Handbook of Semantics*, eds. Aloni and Dekker, Cambridge University Press.

Schlenker, Philippe: 2010, "Supplements Within a Unidimensional Semantics I: Scope". In Logic, Language and Meaning: 17th Amsterdam Colloquium, Amsterdam, The Netherlands, December 16-18, 2009, Revised Selected Papers, Springer (edited by Maria Aloni, Harald Bastiaanse, Tikitu de Jager, and Katrin Schulz)

Schroeter, Laura (2017) 'Two-Dimensional Semantics', in Edward N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy* (Summer 2017 Edition), , URL=<https://plato.stanford.edu/archives/sum2017/entries/two-dimensional-semantics/>.

Searle, John (1976) 'A classification of illocutionary acts', Language in society 5: 1-23.

Searle, John (1979) *Expression and Meaning*. Cambridge: Cambridge University Press.

Searle, John (1980) 'The Background of Meaning', in J. Searle, F. Kiefer & M. Bierwisch (eds.) *Speech Act Theory and Pragmatics*: 221–232. D. Reidel Publishing Company.

Searle, John (1992) *The Rediscovery of the Mind*. Cambridge: MIT Press.

Servidio, E. (2014) 'Imposters and Secondary Sources in Italian' in Collins, C. (ed.) *Cross-Linguistic Studies of Imposters and Pronominal Agreement*. Oxford: OUP.

Shannon, Benny (1976) 'On the two kinds of presuppositions in natural language'. Foundations of Language 14. 247–249.

Sherman, Brett (2014) Notre Dame Philosophical Review of Stefano Predelli, *Meaning without Truth*, Oxford University Press. https://ndpr.nd.edu/news/meaning-without-truth/

Shibamoto, J. (1985). *Japanese women's language*. New York: Academic Press.

Shibamoto-Smith, J. (2004) 'Language and Gender in the (Hetero)Romance: "Reading" the Ideal Hero/ine through Lovers' Dialogue in Japanese Romance Fiction' in Okamoto and Shibamoto (eds) *Japanese Language, Gender, and Ideology: Cultural Models and Real People*. Oxford: OUP.

Shibasaki, R. (2005) *Personal pronouns and argument structure: Discourse frequency, diachrony and typology*. PhD thesis, University of California, Santa Barbara.

Shibasaki, R. (2014) 'More thoughts on the grammaticalisation of personal pronouns' in *Grammaticalisation – Theory and Data*, eds. Sylvie Hancil and Ekkehard König. Amsterdam: John Benjamins.

Shibatani, M. (1990) *The languages of Japan*. Cambridge: CUP.

Shillony, B-A. (1999) *Collected Writings*. London: Routledge.

Shoemaker, S. (1968), 'Self-reference and self-awareness', Journal of Philosophy 65: 555-567.

Siewierska, A. (2004) *Person*. Cambridge: CUP.

Simons, Mandy, Judith Tonhauser, David Bearer and Craige Roberts (2010). 'What projects and why', Proceedings of SALT 20. Ithaca, CCL Publications, 309–27.

Song, J. J. (2006) *The Korean Language: Structure, Use and Context*. London: Routledge.

Stalnaker, Robert (1974) "Pragmatic presuppositions", in Munitz, M. and Unger, P. (eds.), *Semantics and Philosophy*, New York University Press, 197–214.

Stalnaker, Robert (1978). 'Assertion', in Peter Cole (ed.), *Pragmatics, Syntax and Semantics IX*. New York: Academic Press, 315–32.

Stalnaker, Robert (1981), 'Indexical belief', Synthèse 49: 129-151.

Stalnaker, Robert (2002) 'Common Ground', Linguistics and Philosophy 25, 701-721.

Stirling, Lesley, and Rodney Huddleston (2002) Deixis and anaphora. In Rodney Huddleston and Geoffrey K. Pullum (eds.), *The Cambridge grammar of the English language*, 1449–1564. Cambridge: Cambridge University Press.

Stowell, T. A. (1989). Subjects, specifiers and X-bar theory. In M. Baltin & A. Kroch (Eds.), Alternative conceptions of phrase structure. New York: Academic Press.

Strawson, P. (1950) 'On Referring'. Mind 59: 235/320-344.

Sudo, Yasutada (2012) *On the Semantics of Phi Features on Pronouns*. Ph.D. dissertation, MIT Linguistics.

Sudo, Yasutada (2013). 'Biased polar questions in English and Japanese', in Daniel Gutzmann and Hans-Martin Gärtner (eds) *Beyond Expressives: Explorations in Use-Conditional Meaning. Current Research in the Semantics Pragmatics-Interface 28*. Leiden: Brill, 1–58.

Sugamoto, N. (1989) 'Pronominality: A noun-pronoun continuum' in Corrigan R., Eckman F. and Noonan M. (eds.), *Linguistic categorisation*, 267-291. Amsterdam: John Benjamins.

Suzuki, T. (1978) *Japanese and the Japanese: Words in Culture*. Tokyo: Kodansha International.

Takeuchi, L. (1999) *The Structure and History of Japanese: From Yamatokotoba to Nihongo*. London: Longman.

Thommen, Tristan (2017) 'Are expressives presuppositional? The case of slurs'. Proceedings of ESSLLI 2017, 185-196. http://www2.sfs.uni-tuebingen.de/esslli-stus-2017/preproceedings_stus_2017.pdf

Toff, M. (2009) 'Through Different I's: Expressing the Self in Young Japanese Women's Life Writing in English' in Jensen, M. and Jordan. J. (eds.) *Life Writing: The Spirit of the Age and the State of the Art*. Newcastle Upon Tyne: Cambridge Scholars Publishing.

Tonhauser, Judith, David Bearer, Craige Roberts, and Mandy Simons (2013) 'Towards a taxonomy of projective content', Language, 89.1: 66–109.

Torre, Stephan (2016) 'De Se Thought and Communication: An Introduction'. In Manuel García-Carpintero and Stephan Torre (eds.) *About Oneself*. Oxford: Oxford University Press.

Travis, Charles (1975) *Saying and Understanding*. Oxford: Blackwell.

Travis, Charles (1981) *The True and the False: the Domain of Pragmatics*. Amsterdam: John Benjamins.

Travis, Charles (1985) 'On What Is Strictly Speaking True', Canadian Journal of Philosophy 15: 187–229.

Travis, Charles (1996) 'Meaning's Role in Truth', Mind 105: 451–466.

Travis, Charles (1997) 'Pragmatics', in B. Hale & C. Wright (eds.) *A Companion to the Philosophy of Language*: 87–107. Blackwell.

Tsohatzidis, Savas (1992): "Pronouns of address and truth conditions". Linguistics 30, 569–575.

Tsujimura, T. 1968. *Keigo no Shiteki Kenkyuu* [Historical Study of Honorifics], Tokyodo: Tokyo.

Uhlenbeck, E. M. (1968) *Personal pronouns and pronominal suffixes in old Javanese*. North-Holland Publishing Co.

van der Meulen, M. (2001). Self-references among children's first fifty words: Indications for an emerging sense of self in Dutch-speaking children. *Infant and Child Development, 10*(4), 161-171.

van der Sandt, Rob (1992) 'Presupposition projection as anaphora resolution'. Journal of Semantics 9(4). 333–377.

Vergis N. & Terkourafi M. (2015), The M-word: a Greek collocation between solidarity and insult. In: Terkourafi M. (Ed.) *Interdisciplinary Perspectives on Im/politeness.* AILA Applied Linguistics Series no. 14 Amsterdam: John Benjamins Publishing Company. 41-70.

von Fintel, Kai (2006) 'What is Presupposition Accommodation, Again?' Manuscript, MIT.

Vovin, A. (2003) (2003) *A reference grammar of Classical Japanese prose*. London: RoutledgeCurzon.

Wales, K. (1996) *Personal Pronouns in Present-Day English*. Cambridge: CUP.

Wang, Linton, Brian Reese & Elin McCready (2005) 'The projection problem of nominal appositives'. Snippets 10: 13–14.

Wechsler, Stephen (2010) What 'you' and 'I' mean to each other: Person indexicals, self-ascription, and theory of mind. Language, Volume 86, Number 2, June 2010, pp. 332-365.

Wettstein, H. (1986). Has Semantics Rested on a Mistake? *The Journal of Philosophy*. *83* (4), 185-209.

Wetzel, Patricia (1994) 'A movable self: the linguistic indexing of uchi and soto' in Jane Bachnik and Charles Quinn (eds) *Situated Meaning: Inside and Outside in Japanese Self, Society, and Language*. Princeton University Press Princeton: pp. 73-87.

Whitman, J. (1999) 'Personal Pronoun Shift in Japanese'. In Kamio, A. and Takami, Ken-Ichi (eds.) *Function and Structure*. Amsterdam: John Benjamins.

Williamson, Timothy (2009). 'Reference, inference and the semantics of pejoratives', in Joseph Almog and Paolo Leonardi (eds), *The Philosophy of David Kaplan*. Oxford: Oxford University Press, 137–58.

Yanovich, Igor (2010) 'On the nature and formal analysis of indexical presuppositions'. In Kumiyo Nakakoji, Yohei Murakami & Elin McCready (eds.), New Frontiers in Artificial Intelligence: JSAI-isAI Workshops, LENLS, JURISIN, KCSD, LILI, Tokyo, Japan, November 19-20, 2009, Revised Selected Papers, 272–291.

Zaenen, Annie, Lauri Karttunen & Richard Crouch (2005) 'Local textual inference: Can it be defined or circumscribed?' In ACL Workshop on Empirical Modelling of Semantic

Equivalence and Entailment. Ann Arbor, MI: Association for Computational Linguistics.

Zimmerman, Malte (2007) 'I like that damn paper – Three comments on Christopher Potts' The expressive dimension', Theoretical Linguistics 33(2): 237–245.

Zimmermann, Thomas Ede (1997): "The Addressing Puzzle". In: Künne, Wolfgang et al., eds.: Direct Reference, Indexicality, and Propositional Attitudes. Stanford: CSLI, 133–153.