Tesis doctoral

HIDING IN PLAIN SIGHT:

A Non-propositional Presuppositional Account of Slurs

TESIS PRESENTADA POR

Alba Moreno Zurita

PARA OPTAR AL GRADO DE DOCTOR CON MENCIÓN INTERNACIONAL
EN EL PROGRAMA DE DOCTORADO EN FILOSOFÍA (BO2.56.1)

DIRECTOR: Neftalí Villanueva Fernández

**UNIVERSIDAD
DE GRANADA**

Facultad de Filosofía y Letras
Departamento de Filosofía I

WITCHGRASS

Something

comes into the world unwelcome

calling disorder, disorder—

If you hate me so much

don't bother to give me

a name: do you need

one more slur

in your language, another

way to blame

one tribe for everything—

Louise Glück

Table of contents

Agradecimientos

Haber podido realizar la tesis doctoral como parte del grupo de investigación Filosofía y Análisis ha sido un auténtico privilegio. Es una inmensa suerte que este proceso haya ocurrido bajo la mirada de María José Frápolli, Manuel de Pinedo y Neftalí Villanueva. Solo puedo expresar una gratitud y cariño infinito. Me siento en deuda. Gracias a mi director Neftalí Villanueva por toda la paciencia, comprensión y generosidad. Verlo pensar, trabajar y escribir es algo tan guay para mí como escuchar una de mis canciones favoritas.

Mi agradecimiento a Manolo Heras, David Bordonaba y Víctor Fernández. Una mención especial a este último por sus intervenciones cuando la situación lo ha requerido y porque las experiencias filósoficas más valiosas han sido gracias a su insistencia en enseñarme las cosas buenas de este trabajo.

Gracias a Llanos Navarro, Amalia Haro, Manuel Almagro, Ana Redondo, Xavi Osorio, Eduardo Pérez y Trunkel por hacer que los años 2018 y 2019 fueran y vayan a ser, sin ninguna duda, de los mejores años vividos y disfrutados. Tener unas compañeras como Llanos y Amalia en este trabajo es algo impagable.

Muchas gracias a Saray Ayala, Bianca Cepollaro, Marta Jorba, Claudia Picazo, Andrés Soria, Isidora Stojanovic y Dan Zeman por su generosidad y disponibilidad. Sus contribuciones son un sitio al que volver una y otra vez.

No one could have imagined the strange circumstances in which my stay at the University of Tartu in the spring of 2020 would take place, but Alex Davies made it all so much easier. Thank you very much, Alex, for your kindness and advice.

Gracias a mi compañero Eduardo por su paciencia impertérrita y por enseñarme a tenerla para considerar otras ideas. Si he podido ponerle punto y final a este trabajo ha sido gracias a tu apoyo y compañía. El alivio y la alegría de poder ver terminada la tesis es tan tuyo como mío. Puedo imaginarme cómo serán mis días cuando ya no tenga que revisar más este documento pero no puedo, ni quiero, imaginármelos sin ti.

Muchas gracias a mis padres y mis hermanos por estar siempre tan pendientes y hacerme sentir tan querida. Mi agradecimiento eterno a Eulalia y Eduardo, que han hecho posible que pueda terminar la tesis en las mejores circunstancias. No podría haber estado mejor cuidada.

Publications

Almagro-Holgado, A., & Moreno, A. (2022). Affective polarization and testimonial and discursive injustice. In D. Bordonaba Plou, V. Fernández Castro, & J. R. Torices (Eds.), *The Political Turn in Analytic Philosophy* (pp. 259–278). De Gruyter.

Moreno, A. & Pérez-Navarro, E. (2021). Moreno Zurita, A. & Pérez-Navarro, E. (2021). The resistant effect of slurs: A nonpropositional, presuppositional account. *Daimon*, 84, 31–46.

Moreno, A. & Pérez-Navarro, E. (2021). Beyond the conversation: The pervasive danger of slurs. *Organon F*, 28(3), 708–725.

Summary

Fear of not being able to speak because we are always going to offend someone has smilingly become a fashionable topic in recent years. This work is a contribution to this debate, but not for the reasons those who have made this fear part of their identity. There is a difference between believing that, in the right circumstances, almost anything we say can be offensive to someone and maintaining that our linguistic activity carries risks on many occasions. In my view, the difference is that, according to the first perspective, it is the audience's sensitivity that is at fault, whereas in the second case the focus is on the harms of our linguistic activity. The ideas discussed in this thesis are motivated by this second perspective. The main arguments that are developed in it are not intended to specify in detail what harms language entails, but to show that there are harms even in situations where we have reasons to believe that there should not be any.

The object of study of this thesis is slurs and their behaviour. To a first approximation, when an insult is targeted at a person just by virtue of her belonging to a given group, it is called a slur. But my perspective on what counts as a target group requires them to be systematically discriminated against on the basis of their identity traits. The identity traits that have historically been considered susceptible to discrimination have been those related at least to race, gender, religious beliefs and sexual orientation. To be systematically discriminated against means, for the purposes of this work, that one of the elements that make up the identity of the person is an impediment to enjoying a quality of life that seems reasonable to us (Fricker 2007: 27). This is a criterion that is not met by all target groups that are derogated. In my view it makes sense, then, to distinguish between derogatory terms and slurs, for in order for a term to be considered a slur it must not only derogate but also exert oppression (Popa-Wyatt 2016; McGowan 2019). This affects the definition of slur given above. From the position I want to take, it is not enough to say that slurs are characterised as insults

that derogate by virtue of belonging to a given group, but that this group must be systematically discriminated against on the basis of some identity trait.

There is a methodological decision in this work that is partly, but not only, motivated by this conception of slurs. As in previous works (Moreno & Pérez-Navarro 2021a; Moreno & Pérez-Navarro 2021b), there is a conscious decision not to mention any slur. One of the hypotheses I offer evidence for is in this thesis that slurs derogate in every possible context. As we shall see, there are semantic reasons I rely on to explain this idea (section 2.4). However, there are also reasons related to the practical features of slurs (section 2.3) that help me justify this decision. In short, practical features are related to the effects that occur in the conversation and are experienced by the target group due to the exposure to utterances of slurs.

The evidence I offer to consider slurs derogatory in all contexts is based on their normalising potential. The normalising potential of slurs refers to the ease of subsequent appearances in the course of a conversation of a particular slur that has appeared earlier, without having to contemplate the possibility that our listener might reasonably attribute a certain discriminatory attitude to us. For me, this normalising component is not automatically deactivated by the fact that we find ourselves in an academic environment, such as the context of a lecture or a paper. The use or mention of slurs is never free; it always entails a moral cost. In some cases, this moral cost could be justified, but this does not seems to be the case in contexts of discussion among experts in the philosophy of language.

The proposal to be developed is called Non-Propositional Presuppositionalism (from now on, NPP). NPP falls within the family of theories that place derogatory content at the level of meaning that corresponds to what is presupposed (Cepollaro & Stojanovic 2016; Marques & García-Carpintero 2020). In most presuppositional theories of slurs, a propositional conception of the notion of presupposition is maintained, since it is understood that derogatory content is also propositional by nature. However, one of the main stakes of this thesis is to show that the best way to account for the

derogatory effect of slurs is to conceive of derogatory content as a type of *non-propositional* pre-suppositional content. NPP is not the only non-propositional presuppositional theory on the market. As we shall see, the proposal developed by Marques and García-Carpintero (2020) explains derogatory content by means of what they call *expressive presuppositions*.

It has been assumed historically in the philosophy of language that all the information we share for the purposes of conversation is propositional in nature (García-Carpintero 2015). The majority tendency has been to maintain that the component of the common ground that could determine the felicity of our utterances is a set of propositions. Accepting a proposition means eliminating from the common ground the possible worlds that are incompatible with its truth. It is natural that the common ground was initially understood only as the logical space because the most obvious factor determining which utterances are acceptable is what possibilities are open. To treat the accommodation of utterances of slurs as pertaining to the possible worlds available is to treat derogatory content as carrying *locational information* (Lewis 1979; Charlow 2014: 639). This means that by accepting this proposition we can place ourselves in logical space by reducing the set of possible worlds we might be in. From my perspective, however, the derogatory effect of slurs does not result from their helping us know what possible world we are in, so treating its accommodation as the accommodation of a proposition is insufficient. The information carried by derogatory content is of an *orientational* (see again Charlow 2014: 639), not locational, nature. That is, what makes utterances featuring slurs derogatory does not have to do with their telling us which world we are in, but with their telling us which world to turn ours into.

The main reason why I believe this information is orientational is because the information carried by derogatory content has a practical nature. Specifically, it tells us to disparage a target group because of some aspect of their identity. As we will see shortly, exposure to the discriminatory attitude recommended whenever a slur is uttered has certain negative consequences for the target

groups (section 2.3). The world utterances featuring slurs tell us to turn ours into is thus a morally worse world than the world as it currently is.

Chapter 2 sets out the desiderata an account of the meaning of slurs like the one presented here should meet. In particular, the proposal should account for the descriptive, practical and semantic features of slurs, which the chapter is devoted to discussing. Special emphasis is placed on descriptive and semantic features, although one practical feature—the normalising potential of slurs—will eventually play an important role in chapter 6. Descriptive features are understood as those that allow us to distinguish between a slur and a mere insult and explain why the possibility of factual error arises with the former but not with the latter. Once identified, descriptive features allow us to introduce the neutral counterparts of slurs, which in turn allow us to refer to the target group in a non-offensive way, but are also problematic both for ideological reasons and in light of McDowell's (1981) anti-disentangling argument. However, for the purposes of this thesis I assume that slurs do have neutral counterparts (even if I do not take them to be neutral in any deep sense) for two reasons. First of all, we need them to specify what is said when we utter a sentence in which a slur appears. Secondly, we may need them to refer to the target group and point out the oppression it suffers.

Among the semantic features of slurs, the focus of chapter 2 is on non-displaceability, which is the feature that makes slurs be derogatory in every context. The analysis of this feature allows us to draw several lessons. The first is the distinction between non-displaceability and the projection profile of the slurs, which makes their occurrences be derogatory no matter the linguistic environment in which they are found. The projection profile of slurs follows from their non-displaceability, but they are two distinct things. The second lesson is that non-derogatory contexts do not exist and slurs are therefore non-displaceable. This is presented as an hypothesis in support of which some empirical studies are provided.

In chapter 3, the different theories on the meaning of slurs found in the literature are classified into three families. These three families are those formed by content theories, force theories and deflationary theories. The proposal that has been defended in this thesis is located within the theories of content and, specifically, within the theories that place the derogatory content at the presuppositional level of meaning. However, NPP goes in a way beyond the classification presented in this chapter, as the non-propositional character it attributes to derogatory content makes it also a heir of expressivism, which was deemed a force theory.

Chapter 4 is devoted to developing NPP. The notion of presupposition is reviewed and some of the proposals that use propositional presuppositions to explain the derogatory content are exposed. Presuppositional theories have an advantage in not having to identify a slur's contribution to the truth-conditions of the sentences in which it appears with those of its neutral counterpart. This is related to the problem of neutral counterparts, dealt with in chapter 2.

Two non-propositional presuppositional proposals are presented: Marques and García-Carpintero's (2020), which relies on reactive attitudes, and NPP, which relies on world-orderings. The difference between these proposals is in the way in which they explain the impact of non-propositional presuppositions on the common ground. An explanation based on ordering of possible worlds is deemed preferable because this explanation is more parsimonious in at least two ways: ontologically and with respect to intervention. The first is because talking about reactive attitudes means introducing a new element into the common ground, where we already have possible worlds. The second is because an explanation based on possible worlds is neutral between structural and individual perspectives on intervention.

If chapter 4 presents the account of the meaning of slurs favoured in this thesis, chapter 5 states the reasons for favouring it having to do with the semantic features of these terms. NPP is, I argue in that chapter, specially well suited to account for slurs' resistance to cancellation, rejection and retraction. Slurs' resistance to cancellation, which follows straightforwardly from the projection

profile first referred to in chapter 2, is the ability of the derogatory effect of slurs to resist in every single linguistic environment. Slurs are also resistant to rejection because no reply to an utterance involving them seems able to target their derogatory effect, and to retraction because saying "I take that back" after uttering a slur never seems to accomplish its purpose.

I argue that the derogatory effect of slurs projects because it is due to their presupposed content, and if it seems to do so even in cases in which a presupposition would not project is because we have construed the examples wrong—we are looking at cases in which any presupposition would project, and the non-propositional character of derogatory content makes it impossible to build the example the opponent of presuppositional theories has in mind. The presuppositional character of derogatory content, for its part, is what makes very difficult for it to become the target of rejection and retraction.

Chapter 6, finally, faces the task of offering a philosophical argument for the hypothesis laid out in chapter 2 that slurs are non-displaceable. To do so, I first offer a workable construal of the opposite hypothesis that slurs are in fact displaceable. I interpret those who have thought that there are non-derogatory occurrences of slurs, such as Hom (2010), as relying on a distinction between controlled and uncontrolled contexts and assuming that utterances of slurs taking place in controlled contexts are not derogatory.

The argument appeals to the normalising potential of slurs, which was mentioned in chapter 2 as one of the practical features of slurs. In particular, the argument is that occurrences of slurs are imbued with normalising potential even in controlled contexts, where one might expect them not to be derogatory. If this is so, then utterances of slurs are derogatory even if they take place in a controlled context, and thus always derogatory. Slurs are hence non-displaceable.

Resumen

Podría decirse que el miedo a no poder hablar porque siempre vamos a ofender a alguien se ha convertido en un tema de moda en los últimos años. Este trabajo es una contribución a ese debate, pero no por las razones que creen quienes han hecho de ese miedo parte de su identidad. Para mí hay una diferencia entre creer que, en las circunstancias adecuadas, casi cualquier cosa que digamos puede resultar ofensiva para alguien y sostener que en muchas ocasiones nuestra actividad lingüística entraña riesgos. En mi opinión, la diferencia estriba en que, según la primera perspectiva, es la sensibilidad de la audiencia la que tiene la culpa, mientras que en el segundo caso la atención se centra en los perjuicios de nuestra actividad lingüística. Las ideas que se discuten en esta tesis están motivadas por esta segunda perspectiva. Los principales argumentos que se desarrollan en ella no pretenden especificar con detalle qué daños conlleva el lenguaje, sino mostrar que existen daños incluso en situaciones en las que tenemos razones para creer que no debería haberlos.

El objeto de estudio de esta tesis son los *slurs* y su comportamiento. En una primera aproximación, cuando un insulto se dirige a una persona por el mero hecho de pertenecer a un grupo determinado, se denomina *"slur"*. Pero mi perspectiva de lo que cuenta como grupo objetivo exige que se le discrimine sistemáticamente en función de sus rasgos de identidad. Los rasgos de identidad que históricamente se han considerado susceptibles de discriminación han sido los relacionados al menos con la raza, el género, las creencias religiosas y la orientación sexual. Sufrir discriminación sistemática significa, a efectos de este trabajo, que alguno de los elementos que conforman la identidad de una persona le supone un impedimento para disfrutar de una calidad de vida que nos parezca razonable (Fricker 2007: 27). Se trata de un criterio que no cumplen todos los grupos objeto de derogación. Tiene por tanto sentido distinguir entre términos derogatorios y *slurs*, ya que para que un término se considere un *slur* no solo debe ser derogatorio, sino también ejercer opresión (Popa-Wyatt 2016; McGowan 2019). Esto afecta a la definición de *"slur"* dada anteriormente. Des-

de la posición que quiero adoptar, no basta con decir que los *slurs* se caracterizan por ser insultos que derogan por el hecho de pertenecer a un grupo determinado, sino que este grupo debe ser sistemáticamente discriminado por algún rasgo identitario.

Hay una decisión metodológica en este trabajo que está motivada en parte, pero no solo, por esta concepción de los *slurs*. Como en trabajos anteriores (Moreno & Pérez-Navarro 2021a; Moreno & Pérez-Navarro 2021b), hay una decisión consciente de no mencionar ningún *slurs*. Una de las hipótesis en favor de las que ofrezco evidencia en esta tesis es que los *slurs* derogan en todos los contextos posibles. Como veremos, hay razones semánticas en las que me baso para explicar esta idea (sección 2.4). Sin embargo, también hay razones relacionadas con las características prácticas de los *slurs* (sección 2.3) que me ayudan a justificar esta decisión. En una palabra, las características prácticas están relacionados con los efectos que se producen en la conversación y que experimenta el grupo objetivo debido a la exposición a proferencias de *slurs*.

La evidencia que ofrezco para considerar los *slurs* derogatorios en todos los contextos se basa en su potencial normalizador. El potencial normalizador de los *slurs* refiere a la facilidad con la que, en el transcurso de una conversación, puede volver a aparecer un *slur* que ha aparecido con anterioridad, sin tener que contemplar la posibilidad de que nuestro oyente nos atribuya razonablemente una determinada actitud discriminatoria. Para mí, este componente normalizador no se desactiva automáticamente por el hecho de encontrarnos en un entorno académico, como puede ser el contexto de una conferencia o una ponencia. El uso o mención de *slurs* nunca es gratuito; siempre conlleva un coste moral. En algunos casos, este coste moral podría estar justificado, pero no parece ser el caso en contextos de discusión entre expertos en filosofía del lenguaje.

La propuesta a desarrollar se denomina Presuposicionalismo No Proposicional (en adelante, NPP). NPP se enmarca dentro de la familia de teorías que sitúan el contenido derogatorio en el nivel de significado que se corresponde con lo que se presupone (Cepollaro & Stojanovic 2016; Marques & García-Carpintero 2020). En la mayoría de las teorías presuposicionales de los *slurs* se mantiene

una concepción proposicional de la noción de presuposición, ya que se entiende que el contenido derogatorio es también proposicional por naturaleza. Sin embargo, una de las principales apuestas de esta tesis es mostrar que la mejor manera de dar cuenta del efecto derogatorio de los *slurs* es concebir el contenido derogatorio como un tipo de contenido presuposicional no proposicional. NPP no es la única teoría presuposicional no proposicional que existe en el mercado. Como veremos, la propuesta desarrollada por Marques y García-Carpintero (2020) explica el contenido derogatorio mediante lo que ellos denominan "presuposiciones expresivas".

En la filosofía del lenguaje se ha asumido históricamente que toda la información que compartimos a efectos de la conversación es de naturaleza proposicional (García-Carpintero 2015). La tendencia mayoritaria ha sido mantener que el componente del *common ground* que puede determinar la felicidad de nuestros enunciados es un conjunto de proposiciones. Aceptar una proposición significa eliminar del *common ground* los mundos posibles incompatibles con su verdad. Es natural que el *common ground* se entendiera inicialmente solo como el espacio lógico, porque el factor más obvio que determina qué enunciados son aceptables es qué posibilidades están abiertas. Tratar la acomodación de las proferencias de *slurs* como concerniente a los mundos posibles disponibles es tratar el contenido derogatorio como portador de *información locacional* (Lewis 1979; Charlow 2014: 639). Esto significa que aceptando esta proposición podemos situarnos en el espacio lógico reduciendo el conjunto de mundos posibles en los que podríamos estar. Desde mi perspectiva, sin embargo, el efecto derogatorio de los *slurs* no resulta de que nos ayuden a saber en qué mundo posible estamos, por lo que tratar su acomodación como la acomodación de una proposición es insuficiente. La información que lleva consigo el contenido derogatorio es de naturaleza *orientacional* (véase de nuevo Lewis 1979; Charlow 2014: 639), no locacional. Es decir, lo que hace despectivas a las proferencias que contienen *slurs* no tiene que ver con que nos digan en qué mundo estamos, sino con que nos digan en qué mundo debemos convertir el nuestro.

La principal razón por la que creo que esta información es orientacional es porque la información que lleva consigo el contenido derogatorio tiene un carácter práctico. Concretamente, nos dice que menospreciemos a un grupo objetivo por algún aspecto de su identidad. Como veremos en breve, la exposición a la actitud discriminatoria recomendada cada vez que se pronuncia un *slurs* tiene ciertas consecuencias negativas para los grupos objetivo (sección 2.3). El mundo en el que las proferencias que incluyen *slurs* nos dicen que convirtamos el nuestro es, por tanto, un mundo moralmente peor que el mundo actual.

El capítulo 2 expone los *desiderata* que debe cumplir una teoría del significado de los *slurs* como la que aquí se presenta. En concreto, la propuesta debe dar cuenta de las características descriptivas, prácticas y semánticas de los *slurs*, que el capítulo está dedicado a discutir. Se hace especial hincapié en los rasgos descriptivos y semánticos, aunque un rasgo práctico —el potencial normalizador de los *slurs*— acabará desempeñando un papel importante en el capítulo 6. Por rasgos descriptivos se entienden aquellos que nos permiten distinguir entre un *slur* y un mero insulto y explicar por qué la posibilidad de error de hecho surge en el caso del primero y no en el del segundo. Una vez identificados, los rasgos descriptivos nos permiten introducir las contrapartidas neutrales de los *slurs*, que a su vez nos permiten referirnos al grupo objetivo de una manera no ofensiva, pero también son problemáticos tanto por razones ideológicas como a la luz del argumento *anti-disentanglement* de McDowell (1981). Sin embargo, para los fines de esta tesis asumo que los *slurs* tienen contrapartidas neutrales (aunque no las considere neutrales en un sentido profundo) por dos razones. En primer lugar, las necesitamos para especificar lo que se dice cuando proferimos una oración en la que aparece un *slur*. En segundo lugar, podemos necesitarlas para referirnos al grupo objetivo y señalar la opresión que sufre.

Entre las características semánticas de los *slurs*, el capítulo 2 se centra en la *non-displaceability*, que es la característica que hace que los *slurs* sean derogatorios en todos los contextos. El análisis de este rasgo nos permite extraer varias lecciones. La primera es la distinción entre la *non-dis-*

placeability y el perfil de proyección de los *slurs*, que hace que sus apariciones sean derogatorias independientemente del entorno lingüístico en el que se encuentren. El perfil de proyección de los *slurs* se deriva de su *non-displaceability*, pero son dos cosas distintas. La segunda lección es que los contextos no derogatorios no existen y que, por tanto, los *slurs* no son *displaceable*. Esto se presenta como una hipótesis en apoyo de la cual se aportan algunos estudios empíricos.

En el capítulo 3 se clasifican en tres familias las distintas teorías sobre el significado de los *slurs* que se encuentran en la bibliografía. Estas tres familias son las formadas por las teorías del contenido, las teorías de la fuerza y las teorías deflacionarias. La propuesta que se ha defendido en esta tesis se sitúa dentro de las teorías del contenido y, en concreto, dentro de las teorías que sitúan el contenido derogatorio en el nivel presuposicional del significado. Sin embargo, NPP va en cierto modo más allá de la clasificación presentada en este capítulo, ya que el carácter no proposicional que atribuye al contenido derogatorio la convierte también en heredera del expresivismo, que se ha considerado una teoría de la fuerza.

El capítulo 4 está dedicado al desarrollo de NPP. Se revisa la noción de presuposición y se exponen algunas de las propuestas que utilizan presuposiciones proposicionales para explicar el contenido derogatorio. Las teorías presuposicionales tienen la ventaja de no tener que identificar la contribución de un *slur* a las condiciones de verdad de las oraciones en las que aparece con las de su contrapartida neutral. Esto está relacionado con el problema de las contrapartidas neutrales, tratado en el capítulo 2.

Se presentan dos propuestas presuposicionales no proposicionales: la de Marques y García-Carpintero (2020), que se basa en actitudes reactivas, y NPP, que se basa en ordenamientos de mundos. La diferencia entre estas propuestas radica en la forma en que explican el impacto de las presuposiciones no proposicionales en el *common ground*. Se defiende como preferible una explicación basada en los ordenamientos de mundos posibles porque esta explicación es más parsimoniosa al menos en dos sentidos: ontológicamente y con respecto a la intervención. Lo primero es porque ha-

blar de actitudes reactivas supone introducir un nuevo elemento en el *common ground*, donde ya tenemos mundos posibles. Lo segundo, porque una explicación basada en mundos posibles es neutral entre las perspectivas estructural e individual de la intervención.

Si el capítulo 4 presenta la teoría del significado de los *slurs* favorecida en esta tesis, el capítulo 5 expone las razones para favorecerla, que tienen que ver con las características semánticas de estos términos. En ese capítulo, sostengo que NPP es especialmente adecuada para explicar la resistencia de los *slurs* a la cancelación, el rechazo y la retractación. La resistencia de los *slurs* a la cancelación, que se deriva directamente del perfil de proyección mencionado por primera vez en el capítulo 2, es la capacidad del efecto derogatorio de los *slurs* para resistir en todos los entornos lingüísticos. Los *slurs* son también resistentes al rechazo porque ninguna respuesta a una proferencia que los incluya parece capaz de dirigirse a su efecto derogatorio, y a la retractación porque decir "retiro lo dicho" después de pronunciar un *slur* nunca parece cumplir su propósito.

Sostengo que el efecto derogatorio de los *slurs* proyecta porque se debe a su contenido presuposicional, y si parece hacerlo incluso en casos en los que una presuposición no proyectaría es porque hemos interpretado mal los ejemplos: estamos ante casos en los que cualquier presuposición proyectaría, y el carácter no proposicional del contenido derogatorio hace imposible construir el ejemplo que el oponente de las teorías presuposicionales tiene en mente. El carácter presuposicional del contenido derogatorio, por su parte, es lo que hace muy difícil que se convierta en objeto de rechazo y retractación.

El capítulo 6, por último, se enfrenta a la tarea de ofrecer un argumento filosófico a favor de la hipótesis expuesta en el capítulo 2 de que los *slurs* no son *displaceable*. Para ello, primero ofrezco una interpretación factible de la hipótesis contraria, según la cual los *slurs* son de hecho *displaceble*. Interpreto que quienes han pensado que hay apariciones no derogatorias de *slurs*, como Hom (2010), se basan en una distinción entre contextos controlados y no controlados y asumen que los *slurs* que se pronuncian en contextos controlados no son derogatorios.

El argumento apela al potencial normalizador de los *slurs*, que se mencionó en el capítulo 2 como una de sus características prácticas. En concreto, el argumento es que los insultos están imbuidos de potencial normalizador incluso en contextos controlados, en los que cabría esperar que no fueran derogatorios. Si esto es así, los insultos son derogatorios aunque se produzcan en un contexto controlado y, por lo tanto, siempre son derogatorios. Por tanto, los *slurs* son *non-displaceable*.

Chapter 1

Introduction

1.1. Introduction

Fear of not being able to speak because we are always going to offend someone has smilingly become a fashionable topic in recent years. This work is a contribution to this debate, but not for the reasons those who have made this fear part of their identity. There is a difference between believing that, in the right circumstances, almost anything we say can be offensive to someone and maintaining that our linguistic activity carries risks on many occasions. In my view, the difference is that, according to the first perspective, it is the audience's sensitivity that is at fault, whereas in the second case the focus is on the harms of our linguistic activity. The ideas discussed in this thesis are motivated by this second perspective. The main arguments that are developed in it are not intended to specify in detail what harms language entails, but to show that there are harms even in situations where we have reasons to believe that there should not be any.

The object of study of this thesis is slurs and their behaviour. To a first approximation, when an insult is targeted at a person just by virtue of her belonging to a given group, it is called a slur. In this first chapter, I make explicit my criterion for distinguishing slurs from other kinds of words (section 1.2), justify my decision not to mention any slurs from here on (section 1.3), describe briefly the framework in which this thesis is inscribed (section 1.4) and present its structure and main arguments (section 1.5).

1.2. What is a slur?

In the definition advanced above, it is group membership that makes a fundamental distinction between slurs and other insults. However, there are different ways of understanding what makes a group a target group. Hom and May (2018: 113) put forward three options that it would make sense

to follow to demarcate what can count as a target group, which they call "*G*". The first one is there are no restrictions on *G*; it can be instantiated by any group whatsoever. The second one is based on the restriction on *G* supplied by a theory of natural groups.[1] This theory would isolate racial, religious, gender or sexual orientation groups as targets of derogatory statements in virtue of their being natural groups. Finally, there is a restriction on *G* provided by ideologies that are active in sociocultural contexts. A group could be a value of *G* only insofar as there is a discriminatory cultural norm that supports it.

Cepollaro (2020: 6) chooses the first option for determining what counts as a target group. According to her, firstly, it is not necessary for a slur to exist that the assessment of the target group is wrong or unfair. There is nothing morally reprehensible with slurs in themselves. Secondly, it is also not necessary that the target group is the result of a discriminatory ideology. It is only sufficient that a group of speakers have reasons to use a certain term to systematically express an evaluation of a certain property. Thirdly, it is not an essential characteristic to be considered a target group that the evaluation of the property at issue is negative, because, according to Cepollaro, there are also slurs that carry positive evaluations.

My perspective on what counts as a target group is determined by Hom and May's third option instead. I will only consider groups that are systematically discriminated against on the basis of their identity traits as a target group for slurs. The identity traits that have historically been considered susceptible to discrimination have been those related at least to race, gender, religious beliefs and sexual orientation. To be systematically discriminated against means, for the purposes of this

---

[1] According to Hom and May, this option has the least to offer largely because of the problematic assumption that there are biological criteria to individuate the categories distinguished therein. Hom & May, while rejecting this option, consider the possibility that there are ways to defend it. "While there may be ways to make sense of this option, we leave it aside for others to make the case" (Hom & May 2018: 113). Since they are not more explicit on this point, my view here is that any attempt to defend this option is morally unjustifiable.

work, that one of the elements that make up the identity of the person is an impediment to enjoying a quality of life that seems reasonable to us (Fricker 2007: 27). This is a criterion that is not met by all target groups that are derogated. In my view it makes sense, then, to distinguish between derogatory terms and slurs, for in order for a term to be considered a slur it must not only derogate but also exert oppression (Popa-Wyatt 2016; McGowan 2019). This affects the definition of slur given above. From the position I want to adopt it is not enough to say that slurs are characterised as insults that derogate by virtue of belonging to a given group, but that this group must be systematically discriminated against on the basis of some identity trait.

The reasons offered by Cepollaro to explain her choice of the first option described by Hom and May would only serve, from my perspective, to determine the target groups of derogatory terms, but not of slurs. Example of derogatory terms would be "posh" or "facha" (a Spanish derogatory term that refers to right-wing people). Indeed, these terms derogate, that is, we express negative attitudes by uttering them, but the identity traits that pick out these groups are not properties that are used in our social structure to oppress. This is not to say that the use of derogatory terms does not serve to harass, point fingers or humiliate, and to the extent that this is the case, the thesis maintained in this work regarding the derogatory effect of slurs is most likely applicable. However, it is not within the scope of this thesis to specify whether derogatory terms behave in the same way as slurs.

It could be objected that being systematically discriminated against is not an acceptable criterion for distinguishing between derogatory terms and slurs. Someone might say that it is not possible to assess whether a group is systematically discriminated against. Someone might ask *who* decides that a group is or is not discriminated against, or say that it is ultimately a matter of ideology and any target group can be said to be discriminated against. It is true that the derogatory force of slurs is subject to historical and social-ideological changes. It may happen, as we will see in section 2.2, that a slur becomes, over time, a derogatory term because the identity traits that used to characterise that group are no longer considered to ground systematic discrimination. However, there are

identity traits that were historically considered susceptible to discrimination and continue to be so today. In this sense, it is possible that the distinction fits with positions on the political spectrum that are committed to the idea that it is the traits derived from sexual orientation, gender, geographic origin or race, and not others, that contribute to the creation and use of pejorative and discriminatory terms. However, it seems to me that this coincidence is a consequence of making the distinction, rather than its primary motivation.

1.3. On the decision not to mention any slur

There is a methodological decision in this work that is partly, but not only, motivated by this conception of slurs. As in previous works (Moreno & Pérez-Navarro 2021a; Moreno & Pérez-Navarro 2021b), there is a conscious decision not to mention any slur. One of the hypotheses I offer evidence for in this thesis is that slurs derogate in every possible context. As we shall see, there are semantic reasons I rely on to explain this idea (section 2.4). However, there are also reasons related to the practical features of slurs (section 2.3) that help me justify this decision. In short, practical features are related to the effects that occur in the conversation and are experienced by the target group due to the exposure to utterances of slurs.

The evidence I offer to consider slurs to be derogatory in all contexts is based on slurs' normalising potential. The normalising potential of slurs refers to the ease of subsequent appearances in the course of a conversation of a particular slur that has appeared earlier, without having to contemplate the possibility that our listener might reasonably attribute a certain discriminatory attitude to us. For me, this normalising component is not automatically deactivated by the fact that we find ourselves in an academic environment, such as the context of a lecture or a paper. The use or mention of slurs is never free; it always entails a moral cost. In some cases, this moral cost could be justified, but this does not seems to be the case in contexts of discussion among experts in the philosophy of language.

This position has consequences for philosophical practice as we know it. My point is that mentions of slurs in academic papers are potentially harmful too: even if, not being used, they are not as dangerous as uses of these terms, they facilitate ulterior occurrences of the slur in question too.[2] Of course, not all philosophers are comfortable with mentioning slurs even if they take it to be necessary. Quill Kukla, writting as Rebecca Kukla, for instance, says:

> By flagging that I will be mentioning slurs and reminding the reader that even the mention of slurs can harm, I hope to frame these mentions in a way that allows readers to be conscious of such effects and to try to minimise it. I also use scare quotes around the slurs throughout, to help avoid normalising them as part of everyday speech, and in the hope of marking them at the visual level as problematic term that I am not uttering in my own voice and that are not to be taken for granted as readable. (Kukla 2018: 24)

I do not think the use of scare quotes blocks normalisation, but this does not mean that Kukla's mentions of slurs are necessarily unjustified. The moral cost might be worth assuming in this case. I keep out doubts, however, that it is worth assuming in cases in which more examples than the strictly necessary are given.

There are reasons as to why we can distinguish a cultural product from an academic work. However, while a significant number of cultural products have changed their relationship to the use and mention of slurs, I do not find it obvious that there has been a drastic change in the use and mention of slurs in academic spaces, as we can always take refuge in the idea that we do it for the sake of advancing knowledge. If the differences between one case and the other seem clear enough, the reasons why changes have taken place in the former but not in the latter are not so clear. That said, the following two objections might be raised. Firstly the decision not to mention any slur can

---

[2] Herbert (ms) points out in the same direction when she argues that we should be careful even when merely talking about slurs. Her argument is that, just by mentioning these words, we already trigger harmful implicit associations.

make it difficult for the examples in this thesis to trigger the desired intuitions especially if the theory to be favoured is a presuppositional positions. Not specifying the slur may result in not knowing what the presupposition is. Notwithstanding, I believe that the debate on slurs is at such an advanced stage that linguistic intuitions are still available, even if not in the most accessible way.

In second place, I believe that taking this decision may lead some to consider that my reasons for doing so are due to a desire to adopt a position of moral superiority. It seems to me that the best I can do in the face of this criticism is to accept it. I think it is indeed better not to use or mention slurs in academic spaces than to do so. This follows from the decision not to mention slurs in this thesis, so not to accept this criticism or the problems arising from this decision would be dishonest on my part. I also believe that the reasons for my decision are justified. Firstly, there is a manifest discomfort with talking about slurs in the philosophy of language that has led scholars to do such things as introducing disclaimers before starting a talk where slurs would be used or mentioned, mentioning slurs without writing the whole word, or mentioning just a couple of examples at the beginning. One might think that the decision not to mention any slur is to push these strategies to the limit in the study of these expressions, but my motivations are no different from those who adopt such measures, which are, perhaps, less problematic. I believe that any alternative we seek to the use or mention of slurs in the academic environment involves a decision that has ethical motivations and purposes. It seems natural to me that such strategies have been developed given the kind of terms that slurs are. Secondly, in addition to ethical reasons, there are theoretical reasons for this decision. The hypothesis I will introduce in chapter 2 and develop in chapter 6 is that there are no neutral contexts in which slurs do not derogate, so I have no reason to consider that the context of this thesis is not one of them.

1.4. The framework

The purpose of this section is to present the theoretical framework within the philosophy of language in which the proposal that I am going to develop is situated. This proposal is called Non-Propositional Presuppositionalism (from now on, NPP). NPP falls within the family of theories that place derogatory content at the level of meaning that corresponds to what is presupposed (Cepollaro & Stojanovic 2016; Marques & García-Carpintero 2020). In most presuppositional theories of slurs, a propositional conception of the notion of presupposition is maintained, since it is understood that derogatory content is also propositional by nature. However, one of the main stakes of this thesis is to show that the best way to account for the derogatory effect of slurs is to conceive of derogatory content as a type of *non-propositional* presuppositional content. NPP is not the only non-propositional presuppositional theory on the market. As we shall see, the proposal developed by Marques and García-Carpintero (2020) explains derogatory content by means of what they call *expressive presuppositions*.

In chapters 4 and 5 I will explain why the non-propositional character of derogatory content, understood as presuppositional content, is an advantage in allowing us to explain certain features of the derogatory effect of slurs, such as its resistance to cancellation, rejection and retraction. But before going into the specific features associated with the derogatory effect, I want to start by explaining the configuration of the common ground when we try to accommodate it. The common ground can be understood as a theoretical construct that accounts for the current state of our conversations. We can imagine the common ground as a kind of canvas that contains a record of the movements and changes that take place in the conversation (McGowan 2019: 39–40). Suppose I say to my partner "I am going to pick up my brother at the train station this afternoon". It is likely that we are both starting from the same point, since we are likely to share the information that I have at least one brother, so it goes without saying. In fact, it will be because we share this information that my utter-

ance is felicitous. This shared information that we take for granted with our interlocutors is what is commonly known as a conversation's presupposition.

It is not necessary for us to know our interlocutors personally in order for us to take certain information to be shared for the purposes of the conversation. There are two reasons for this. Firstly, the set of information we presuppose can be extremely varied and extensive, ranging from information that we only share with the people closest to us to social conventions that we can take our interlocutors to be aware of even if we hardly know them. Secondly, due to the mechanism of accommodation (Lewis 1979), the participants in the conversation can adjust the information in the common ground to be at the same point as the speaker who already had that information. For example, someone who has just met me but hears me tell my partner that I am going to pick up my brother at the train station will, through a process of accommodation, be able to incorporate into the set of information we share that I have a brother.

It has been assumed historically in the philosophy of language that all the information we share for the purposes of conversation is propositional in nature (García-Carpintero 2015). The majority tendency has been to maintain that the component of the common ground that could determine the felicity of our utterances is a set of propositions. Thus, for the utterance "I am going to pick up my brother at the train station this afternoon" to be felicitous, the proposition *Alba has a brother* must be part of the common ground. Accepting that proposition means eliminating from the common ground the possible worlds that are incompatible with its truth. It is natural that the common ground was initially understood only as the logical space because the most obvious factor determining which utterances are acceptable is what possibilities are open. If we take this model to correctly represent the presuppositions triggered by utterances featuring slurs, we will accommodate the impact of an utterance such as "There are many Ss in this neighbourhood", where "*S*" is a slur, by eliminating certain possible worlds from the common ground—for instance, those where the speaker is not intolerant towards the *N*s, where "*N*" would function as the neutral counterpart of

"*S*". To treat the accommodation of this utterance as pertaining to the possible worlds available is to treat derogatory content as carrying *locational information* (Lewis 1979; Charlow 2014: 639). This means that by accepting this proposition we can place ourselves in logical space by reducing the set of possible worlds we might be in. From my perspective, however, the derogatory effect of slurs does not result from their helping us know what possible world we are in, so treating its accommodation as the accommodation of a proposition is insufficient. The information carried by derogatory content is of an *orientational* (see again Lewis 1979; Charlow 2014: 639), not locational, nature. That is, what makes utterances featuring slurs derogatory does not have to do with their telling us which world we are in, but with their telling us which world to turn ours into.

The main reason why I believe this information is orientational is because the information carried by derogatory content has a practical nature. Specifically, it tells us to disparage a target group because of some aspect of their identity. As we will see shortly, exposure to the discriminatory attitude recommended whenever a slur is uttered has certain negative consequences for the target groups (section 2.3). The world utterances featuring slurs tell us to turn ours into is thus a morally worse world than the world as it currently is. We will also see that the expression of this discriminatory attitude has the capacity to break something in the context of the conversation that it is not possible to repair (chapter 6).

We can model the discriminatory attitudes expressed by utterances featuring slurs by saying that these utterances commit the speaker to a particular ordering of the worlds in the common ground, such that, roughly, those worlds where the participants in the conversation have the less possible contact with the *N*s are preferred to those in which there is more contact. Given the speaker's commitment to this way of ordering possibilities, it makes sense to attribute certain courses of action to her. For example, the speaker who utters "There are many Ss in this neighbourhood" would be expected to avoid that neighbourhood or to recommend that visitors to the city do not pass through it. As already advanced and we will see in detail in chapter 4, the world-ordering approach

to accommodate the impact of utterances featuring slurs on the common ground is not the only non-propositional option for doing so. But my point here is that derogatory content does not have to do with *what is the case* but with *what to do*, whether we implement that point through world orderings or in some other way is another line in this debate.

The idea that through derogatory content we obtain information of an orientational nature commits NPP to an *expressivist* conception of meaning. I think there are different ways of approaching the relationship between NPP and expressivism, but it seems to me that the most obvious is the following. Expressivism individuates the meaning of evaluative language according to the use speakers makes of it (see Geach 1965; Horwich 2005). If we look at the use of evaluative language, we will see that it has a special connection to action. If someone says "It is wrong to start eating if not everyone is seated at the table", what we can reasonably assume is that this person is committed to this rule and will behave in this way when the situation arises. To say that evaluative uses of language have a special connection with action implies assuming that there are other uses of language that do not, that is, that there are other uses of language that are not evaluative. This makes the type of expressivism that NPP engages with a local expressivism (Chrisman 2007; Gibbard 1990, 2003; Frápolli & Villanueva 2012) rather than a global expressivism (Price 2011; Price et al. 2013). In fact, NPP assumes that, apart from the derogatory content, which will be treated as non-propositional, slurs also have a propositional content that accounts for the descriptive component of their meaning (section 2.2).

The proposal of NPP is meant to be *ideal* in a certain sense. Keiser (2022: 4) distinguishes two senses in which theories can be ideal or non-ideal. In the first sense, theories are ideal insofar as they are *idealised*—insofar as they distort some features of the phenomenon they seek to explain. They can do so be because the features at issue are not considered essential to the phenomenon or because it is convenient to ignore them at an early stage. The fact that at least most theories are idealistic in this sense is not necessarily a bad thing, as idealisations may be justified given the re-

searcher's purposes. In the second sense, ideal theories are those that manage to live up to some normative standard. In this sense, theories that are ideal are necessarily good, since what we value are the normative and social consequences that we can attribute to the theory.

Expressivism, understood as a family of theories to which the proposal I develop in this thesis is heir, was motivated by an attempt at avoiding what it considered an inappropriate idealisation of the dominant philosophy of language, namely the assumption that the sole or at least the main function of language was to describe the world. *Descriptivist* positions did not in this sense serve their purpose well, at least if we take their aim to be the characterisation of language. The original motivation to abandon a descriptivist stance on language, and explore other forms of expression, is, for the expressivist, that descriptivists ignore some essential features of language—what we do when we speak exceeds describing how things are in our surroundings.

But to avoid leaving these essential features aside, expressivism has to deal with the Frege-Geach problem, which can be considered as the main objection to expressivism. This objection can be understood as taking expressivism to also commit an idealisation by neglecting in the characterisation of the evaluative domain of language its striking similarities with the descriptive domain. A large body of literature has addressed the Frege-Geach problem and many solutions have been offered (see Geach 1965; Horwich 2005; Chrisman 2007; Frápolli & Villanueva 2012; Hom & Schwartz 2013; Charlow 2014). Since I believe that expressivism has succeeded in answering this objection, I consider that the idealisations that expressivism made with respect to the evaluative uses of language were justified from the point of view of explaining the phenomenon. However, this does not mean that delivering an ideal theory in the second sense was among expressivists' motivations, nor was it among the ones of those who criticised expressivism. But it is among mine. I want the proposal advanced here to avoid the pernicious consequences of descriptivism.

1.5. Argument and structure

This thesis is divided into seven chapters. After this introduction, chapter 2 discusses a number of features of slurs that a theory that aims at accounting for their behaviour should accommodate. These features are grouped into three families: *descriptive*, *practical* and *semantic* features. Descriptive features are those that distinguish slurs from mere insults; practical features pertain to the consequences of uttering a slur, and semantic features encompass the interaction with context of the derogatory effect of slurs. What I will take to be essential for the view I will advance is for it to be able to accommodate descriptive features and some semantic ones. Within the latter category, special emphasis will be put on *non-displaceability* and some manifestations of *ineffability*. I will also get back to practical features, and in particular to the *normalising potential* of slurs, later on the thesis.

In chapter 3, I will survey a number of theories that have aimed at accounting for the derogatory effect of slurs. I will group these views according to whether they attribute some kind of content that is specifically derogatory to slurs, as in *content theories*, they explain the derogatory effect in terms of the force of speech acts, as *force theories*, or they try to explain their behaviour in some other way, as in *deflationary theories*. Most of the views I will consider belong to the former family. I will not test each theory's ability to account for each and every feature discussed in chapter 2. Instead, I will use the features that are most obviously accommodated by each theory to motivate it, and point to those features that seem hard to account for in each view.

One of the families of content theories I will introduce in chapter 3 is the one encompassing *presuppositional* views. Most of what I have to say about these theories, however, will have to wait until chapter 4, where I present my own view, which belongs to this family—*non-propositional presuppositionalism* (NPP) about the meaning of slurs. I will survey the history of presuppositional accounts of the meaning of slurs, starting with views that consider the presuppositions triggered by

slurs propositional, complicating them to accommodate the special character of these presuppositions, and finally arriving at my own proposal, which relies on non-propositional presuppositions. The reasons for going non-propositional pertain to the accommodation of both descriptive and semantic features. I will discuss the advantages of non-propositional views along the first dimension as I introduce them.

Chapter 5 is devoted to NPP's advantages in accounting for some semantic features of slurs; in particular, their *resistance* to *cancellation*, *rejection* and *retraction*. I will characterise cancellation, rejection and retraction and contrast the difficulty of making any of these moves in response to an utterance featuring a slur with the lack of problems we have when it is an ordinary assertion we are facing. After characterising each feature, I will show how NPP can account for it.

Resistance to cancellation has sometimes been identified with non-displaceability, the other semantic feature of slurs discussed in chapter 2 that I said would play a central role in the thesis. After distinguishing these two features in chapter 5, chapter 6 is devoted to exploring, in a way that is somewhat independent from the theory advanced in chapter 4, but closely connected to the arguments on non-displaceability made in chapter 2, the hypothesis that there are no non-derogatory contexts. I try to disentangle the intuition behind Hom's (2010) distinction between *orthodox* and *non-orthodox* occurrences of slurs, the second of which he takes to be non-derogatory, in terms of the context in which they take place, and argue that even in *controlled* contexts, where we might think slurs fail to derogate, there is derogation after all. This finding, which as I will show is supported by the some empirical evidence, I will explain in terms of the normalising potential all occurrences of slurs have—a practical feature that, as promised in chapter 2, I will rescue in this chapter. Finally, chapter 7 concludes the thesis by summarising its main findings.

Chapter 2

The features of slurs

2.1. Introduction

The main task of this chapter is to explain a number of traits that have been attributed to slurs (see Potts 2007; Hom 2010: 164–170; Croom 2011: 345–347; DiFranco 2014: sections 1.b, 1.c and 1.d). The classification I develop in this chapter, inspired by the way DiFranco (2014), sets out the different desiderata that should be met by a theory of slurs. Although all the features he discusses are features of pejoratives in general, just like the ones Hom (2010) discusses, the way he organises these features allows for a better understanding of the behaviour of slurs in particular. There is nothing wrong in considering slurs as a kind of pejorative term, because they are. However, it is a mistake to equate the characteristics of pejoratives and slurs without due specification because the behaviour of slurs differs from that of some other pejoratives—mere insults—too. We can explain this difference between slurs and mere insults by appealing to the fact that the former are mostly considered thick terms (Williams 1985), as we will see shortly, and the latter are not. This will go a long way to differentiating between slurs and mere insults, but it is not all that can be said. For example, the effects of exposure to slurs and hate speech in general that have been studied do not apply to mere insults. These differences in behaviour between slurs and mere insults make me maintain that the characteristics I will discuss are attributable to slurs, even if some of them might also be applicable to insults in general.

The characteristics I am going to talk about are divided into three groups, namely descriptive features, practical features and semantic features. This division corresponds to the order of the sections in this chapter. In dealing with the first type of characteristics (section 2.2), I discuss why slurs are considered thick terms and how this conception forces us to account for the possibility of factual error in the use of slurs. Because of this possibility, I anticipate here the problem of neutral counter-

parts, to which I will return to in subsection 4.3.2. The second type of characteristics encompasses those having to do with the effects that utterances of slurs have (section 2.3). Among these, I will put special emphasis on the normalising potential of slurs. Finally, in the section on semantic features, I address the particularities of slurs that restrain the way we can characterise derogatory content (section 2.4). In particular, I address the notion of non-displaceability, which will play a fundamental role in the main argument of this work.

2.2. Descriptive features

As I indicated in chapter 1, slurs are not mere insults. The most salient difference between a slur like "boche" and a mere insult like "jerk" is that there seems to be a descriptive component in the meaning of the former that is missing in the meaning of the latter—namely, *being of German origin*[1]. To see this, note that someone would be making a moral error by calling a person either a jerk or a boche. But, if someone used the word "boche"[2] to refer to a person who is not of German origin, along with the moral error she would also be making a factual one. Thinking of slurs as terms that combine description and evaluation is the reason they are commonly grouped together with *thick terms* (Williams 1985). If this is accurate, one could say that the evaluative component of the meaning of a slur is the part that has the capacity to derogate, while the descriptive component would consist in the objective conditions under which one does not make a factual mistake by using the slur, i.e., its conditions of application.

---

[1] We can also think of examples of mere insults that are thick terms, for example, "pig". I will leave these aside here.

[2] If we take into account the distinction I have made between slurs and derogatory terms (section 1.2), "boche" would be an instance of a term that used to be a slur and has become a derogatory term. This is why I considerer that mentioning it will not pose the problems that mentioning slurs in academic contexts does. In the next section, I will comment on the reasons why I think it makes sense to think that the derogatory force of this term has changed.

The fact that the meaning of slurs includes a descriptive component might be thought to imply that, for every slur, there will be a *neutral counterpart* that will share the slur's descriptive meaning but lack its evaluative one. However, slurs can have descriptive meaning without having neutral counterparts. In the so-called "anti-disentangling argument" (see Roberts 2011), McDowell points out that, although thick terms have by definition a descriptive component, this component cannot be isolated in such a way that we can establish the term's application conditions without taking into account the evaluative component of its meaning (McDowell 1981: 143–144). McDowell's argument can be construed as follows:

1. If we could separate the descriptive component and the evaluative component of a thick term, we could introduce a term (its neutral counterpart) that we could competently use just by knowing what objects fall under its extension and what objects do not.

2. However, we cannot master the extension of a thick term without taking an evaluative perspective into account.

3. Hence, we cannot separate the descriptive component and the evaluative component of a thick term so that we can introduce a neutral counterpart.

This argument marries well with the intuitions we have about certain slurs, especially racial slurs, in which there seems to be something descriptive that makes it possible to talk about factual error, but we do not quite know what it is—it is difficult to specify what the slur's neutral counterpart is. For instance, with some racial epithets in Spanish, it is not clear which target group the slur is used to derogate; it may refer to people from certain North African countries, who speak a certain language, or who profess a certain religion.

If the conclusion of McDowell's argument was to be applied to slurs, the descriptive and the evaluative components of slurs would be inseparable. I take it to be an open question whether this means that there cannot actually be neutral counterparts of slurs, in the sense that I will adopt in this

thesis. For its purposes, I will assume the conclusion of the anti-disentanglement argument as applied to slurs, as it gives theoretical support to the intuition that the act of classification that we perform whenever we use a slur itself involves the adoption of an evaluative, bigoted perspective. I thus do not take neutral counterparts to be truly neutral, as they can only acquire their meaning in a community that is at least partly structured in a hierarchical way. But we need them to point to this structure in, at least, a non-offensive way. I do not think that slurs and their *sui generis* neutral counterparts are semantically alike. Using a slur and using a neutral counterpart involves significantly different consequences that will be developed during the course of this dissertation.[3]

If we take slurs to have neutral counterparts, it might be natural to think that, since the slur and its neutral counterpart share their conditions of application, the latter will make the same contribution to the truth-conditions of the sentences in which it occurs as the slur does. However, the fact that we keep talking about neutral counterparts does not mean that we take them to share with the corresponding slurs their contribution to the truth-conditions of the sentences in which they occur. If we did, we would be subject to Richard's (2008) criticism that, since a sentence like "*N*s are *N*s", where "*N*" is the neutral counterpart of a slur "*S*", is trivially true, a sentence like "*N*s are *S*s" will also be true, for "being an *N*" and "being an *S*" are extensionally equivalent, that is, substitutable *salva veritate*. If the slur is a discriminative slur in the sense in which we defined them in chapter 1, though, it would be desirable that the proposal we develop about the meaning of slurs does not depend on a sentence like "*N*s are *S*s" being true. After all, it is impossible to accept that it is without committing to, e.g., a racist attitude, if "*S*" a racial epithet.

The proposal I will advance in chapter 4 does not have this problem, as it assumes the existence of neutral counterparts, but does not take them to share their contribution to the truth-conditions of the sentences in which they occur with that of the corresponding slur. We will get to this in

---

[3] I will say more about this in subsection 4.3.2.

due time, but the reasoning will be, in a nutshell, as follows. A sentence of the form "A is an *N*" and one of the form "A is an *S*" will share their truth-conditions *provided that a certain condition is met*. If it is not, "A is an *N*" might be true or false, but "A is an *S*" will lack truth-value. The same will happen with "*N*s are *N*s" and "*N*s are *S*s": if the condition is not met, the first will be trivially true, but the second will lack truth-value.

For the rest of this dissertation, then, I will accept that the meaning of slurs has a descriptive component, and assume that there is such a thing as a slur's neutral counterpart. However, I will not take neutral counterparts to contribute to the truth-conditions of the sentences in which they occur in the same way in which the corresponding slur does. I will return to the issue of neutral counterparts in subsection 4.3.2.

2.3. Practical features

Features such as the normalising potential of slurs (DiFranco 2014: section 1.a), the discomfort that their utterance can produce in the course of a conversation (DiFranco 2014: section 1.a and b) or the change in derogatory capacity due to historical variability (Hom 2010: 165) have been considered practical features of slurs. I started this dissertation with the perception that, through the use of slurs, speakers express a discriminatory attitude towards a target group. I think, however, that the practical characteristics of slurs also pertain to effects that go beyond the mere expression of a discriminatory attitude, and that they can be useful if we want to specify in some way what it means to suffer oppression. I am not just talking about how easy it is for slurs to spread prejudice; I am talking about how they can affect the personal conditions of people who are part of the target group. Therefore, along with explaining the common practical characteristics that have been attributed to slurs in the literature, I will also mention some empirical studies that have been carried out that are closely related to these types of characteristics.

First, when we talk about the normalising potential of slurs, we mean the ease with which, in the course of conversation, a slur that has already been used before can be used. It is very easy in the context of a conversation for participants to be willing to use a slur that has already been used because there is a perception that it has been implicitly agreed that in that context the slur has become neural or has lost its derogatory effect. Some of the reasons for this perception could be that, since it has already been uttered, we should not contemplate the possibility that our listeners might reasonably attribute a discriminatory attitude to us if we continue to use it. It could be said that a free way is established for their use, as in a way the responsibility for the derogatory effect has been diluted as they are part of the topic of conversation and no particular speaker has to bear the moral and social cost of uttering slurs. This can also generate a sense of complicity among the participants in the conversation; we know that we are doing something that would be reprehensible in any other context, but nothing is going to happen in this one (Cepollaro 2020: 47). We could relate the normalising potential of slurs to studies (Soral et al. 2018; Winiewski et al. 2018) that have concluded that continued exposure to spaces dedicated to the transmission of hate speech significantly increases the desensitisation of people who do not belong to a derogated group, as well as further legitimising violence and discrimination. In section 6.4, we will have the opportunity to discuss the normalising potential of slurs and why this gives support to the hypothesis that there are no non-derogatory contexts, which will be introduced in the next section.

Second, against the ease with which the slur can be repeated naturally due to the normalising potential of slurs, it can also occur that a sense of discomfort can arise among the participants in a conversation when a slur is uttered. These two elements may seem to be in tension. However, it is not incompatible to feel discomfort or uneasiness when a slur is first uttered and feel legitimised to use slur once its use has been normalised in conversation. It seems to me that the relationship between these two perceptions in the face of the utterance of a slur makes sense precisely because of what has been said just above, namely, we may believe that the context has changed in a way that

has made it possible for the use of the normalised slur not to derogate anymore. In addition to the audience's feeling of discomfort, it has been found that target groups may experience emotional responses too, which include discomfort but also anger (Swim et at. 2001, 2003). Other long-term effects that have also been reported include a deterioration in the cognitive, emotional and psychological health of people who are part of a target group that has been exposed to the use of slurs (Garnets, Herek & Levy 1990; D'Augelli 1992; Herek, Gillis & Cogan 1999; Cowan & Mettrick 2002).

Third, as discussed in section 1.2, the derogatory force of slurs is subject to historical and sociological changes. The case of "boche" is a representative example that shows that target groups are determined according to prevailing ideologies. The social and political situation in Europe in the inter-war period made "boche" a slur, as people with the characteristic of being of German origin were systematically discriminated against. However, the social and political situation may change in such a way that people bearing this identity feature are not systematically discriminated against. This does not mean that this term has become a term empty of derogation. If we take into account the distinction I made earlier, we can consider it a derogatory term, and a number of consequences follow from its having a derogatory effect. The fact that the social context has changed and this trait is no longer considered a reason to be systematically discriminated against makes it reasonable to think that the derogatory force of the slur is not the same, although it has not entirely lost it. Conversely, the variability of derogatory force may not only affect terms that were used as slurs in the past but may also create new ones.

2.4. Semantic features

So far we have discussed the descriptive and practical characteristics of slurs. This section is devoted to their semantic features, which I understand as follows. The projection profile of slurs has been considered the main reason to explain derogatory content in presuppositional terms because of

its similarity with the projection profile of presuppositions (Lasersohn 2007; Marques & García-Carpintero 2020; see also subsection 4.3.2). The projection profile of slurs encompasses the behaviour slurs exhibit when embedded in complex linguistic environments. This behaviour can be explained by a property that slurs possess, namely *non-displaceability*. I will thus consider the feature of non-displaceability as one of the essential properties of slurs precisely because of what it allows us to explain and the role it has played in the debate around presuppositional accounts of derogatory content. Other features that are among the semantic properties of slurs and that are related—in some cases more, in some cases less—to the property of non-displaceability are *expressive autonomy*, *immediacy* and ineffability.

Let us then start with non-displaceability. We can find a characterisation of non-displaceability in different works that are representative of the literature on slurs, such as (Potts 2007: 13), (Cepollaro 2020: 140), (Croom 2011: 346) and (Hom 2010: 171). I consider these three references a fundamental support for understanding and characterising the feature of non-displaceability. However, it seems to me that it is worthwhile to stop and think about this notion in the light of the original definition we find in Cruse (1986):

> Another characteristic distinguishing expressive meaning from propositional meaning is that it is valid only for the utterer, at the time and place of utterance. This limitation it shares with, for instance, a smile, a frown, a gesture of impatience, or a dog's bark (all of which, as it happens, are also continuously gradable). The capacity of language to transcend the immediate context of utterance (sometimes referred to as the capacity for displacement), which enables me to talk about the pain I felt yesterday, or the pain Arthur will feel tomorrow in Australia, depends entirely on propositional meaning. (Cruse 1986: 272)

Two things deserve comment in this passage. First of all, the utterer seems to be important; at least, this is something we seem entitled to extract from Cruse's saying "that it is valid only for the utterer". I will get to this later. Secondly, there is also the comparison that follows between the limitations shared by expressive content and actions such as smiling or frowning. But all Cruse does is to state a difference between expressive content and smiles and frowns, on the one hand, and propositional content, on the other. He does not explain what feature it is that expressive content and smiles and frowns have but propositional content fails to have and that accounts for this difference. It seems to me, then, that it makes sense to ask what expressive content has in common with smiles and frowns that makes it as non-displaceable as they are.

One of the ways of understanding what it means to be non-displaceable is to appeal to the difficulty of embedding that content in certain complex linguistic environments. This might remind us of the Frege-Geach problem (section 1.4) faced by emotivism (Ayer 1936/2001), but Cruse is pointing elsewhere. Non-displaceability is not a problem about how expressive content challenges the truth conditions of truth-functional contexts and, hence, the validity of arguments involving expressive terms. Instead, non-displaceability is conceived as a distinctive feature of expressive content that prevents it from behaving as it should under the scope of a truth-functional operator such as the conditional or negation. Suppose someone says:

(1) That fucking Lorelai is always late!

This sentence has expressive content. But if in reply to the utterance of this sentence someone just uttered its negation, the resulting sentence would have expressive content as well:

(2) That fucking Lorelai isn't always late!

Cruse's point is that the expressive content in (1) is not displaceable to a context like (2), it's "only valid for the utterer". Of course, there is also expressive content in (2), but if Cruse is right, this expressive content is different from the one conveyed by an utterance of (1). This nevertheless gives

47

slurs a very peculiar projection profile. What happens in this case is that the force of the insult is not modified in a linguistic environment that usually blocks that force. If we take this behaviour into account, the comparison with smiling or frowning is naturally understood because it seems to make sense to think that if we could put this emoji "🙂" in the antecedent of a conditional, it would still be a smile, or have the effect of a smile, even if it were placed in that space. We could test how this works:

(3) If 🙂 you have studied, then the test will go well.

In a nutshell, inserting expressive content into complex linguistic environments should be one way of displacing this content. What we see, however, is that this content does not seem to be subject to certain truth-functional operators and, therefore, cannot be displaced to other contexts that include another time or another possible world, such as the context to which the conditional normally displaces content. As I interpret Cruse's words, this is what we are to understand when he states that "(t)he capacity of language to transcend the immediate context of utterance (sometimes referred to as the capacity for displacement), which enables me to talk about the pain I felt yesterday, or the pain Arthur will feel tomorrow in Australia, depends entirely on propositional meaning".

So far, I have explored the second aspect of Cruse's excerpt: his comparison of expressive content with other actions that express emotions. However, it remains to be seen what role Cruse's emphasis on the person who delivers the expressive content plays in all this. Arguably, the main function of expressive content is to convey our emotions, attitudes or feelings, and there is a sense in which we should be able to do this as directly as possible. It thus makes sense to think that expressive content cannot be modified or blocked in a multitude of linguistic environments; that we can always extract what is being expressed and that this is not transferable.

Having said that, it seems to me that to attach more importance here to the role of the speaker is too far removed from what Cruse is saying. I therefore choose to understand that what Cruse is

doing in pointing to the role of the speaker is simply to refer to one of the components of the con- text of utterance (Kaplan 1977/1989). Arguably, to speak of context of utterance at this point is con- troversial since Cruse makes no explicit reference to this notion as it was already being used in the philosophy of language debate. By context of utterance we should understand the situations in which speakers find themselves when they utter a sentence. Each situation will be characterised by the place, the time, the speaker's social standing, by whom our audience is formed or who the inter- locutors of the conversation are, etc. If we stick to a minimal characterisation, the context of utte- rance is the situation which is determined by the time and place from which the speaker makes her utterance, and these elements all appear in the fragment. If, as it seems reasonable to understand, Cruse is talking about the context of utterance, what we could conclude is that the expressive con- tent cannot be moved from one context of utterance to another.

Now, let us move away from the original characterisation of the feature of non-displaceability and turn to how this notion has been treated in the analysis of slurs through some of the works I mentioned at the beginning of this section. Specifically, let us now see how Croom and Hom ap- proach the characterisation of non-displaceability. Croom claims that some of the linguistic tests proposed by McCready (2010) to detect non-descriptive terms (Croom 2011: 345) can be used to show that expressive content has the property of non-displaceability. He claims this, in particular, about the test by which we place the term in question in the antecedent of a conditional to check its projection profile and the test by which we look at the behaviour of expressive terms when they are negated. Let us now look at the case of the first test and see what happens when we insert a slur in the antecedent of a conditional and what happens when we insert its neutral counterpart.

(4) If I didn't like lesbian women, then I'd probably be homophobic.

(5) If I didn't like *S*s, then I'd probably be homophobic.

In the first case, where we put the neutral counterpart in the antecedent, by the force of the conditional, the speaker does not commit herself to not liking lesbian women. We cannot say that an attitude of contempt projects in (4). As Croom points out, the attitude "does not project out to generate inferences about the speaker's communicative intent" (Croom 2011: 345). However, in (5), where the neutral counterpart "lesbian women" has been replaced by a slur associated with it, the derogatory effect of the slur is capable of extending even if the slur is placed in the antecedent of a conditional. The fact that the derogatory effect cannot be restricted by the conditional points to the same idea as when Cruse states that expressive terms cannot be embedded, i.e., we cannot embed the attitude.

If the attitudes we express through slurs (their derogatory effect) could be embedded like the attitudes we report, they would be displaceable and therefore would not project because they would be locked under the scope of the conditional. As for negation, let's see what happens when we try to negate a slur:[4]

A: Maya came to the party last night.

B: That's not true/That's false.

= Maya didn't come to the party.

A: Maya is an *S*.

B: That's not true/That's false.

≠ Lesbian people are not bad.

According to McCready, the negation test allows us to discriminate expressive terms from non-expressive terms because when they are under the scope of negation it is not their expressive content that we negate. In the first case, what we accept or deny is the fact that Maya did or did not go to

---

[4] We will return to the negation test in chapter 5, where we will discuss slurs' resistance to rejection. Moreover, this test, together with the biding test, is used as evidence to support the idea that derogatory content is not part of what is said (Croom 2011: 345–346).

the party. In the second case, what we want to deny is that lesbian people are bad. However, all we can do is deny that Maya is lesbian. We have no way, through negation, to reject what is wrong with using *S*.

Hom (2010, 2012) arrives to a similar conclusion as Croom taking a different path. His starting point is to understand non-displaceability as a feature of particular occurrences of slurs. In this sense, an occurrence of a slur will be non-displaceable if and only if every utterance of the slur in that same context is derogatory. This includes cases in which the slur is embedded in a complex linguistic environment. For instance, most occurrences of the slur "*S*" will be derogatory. Now, if an occurrence of "*S*" is derogatory, it will be so whether the slur appears under the scope of negation, within the antecedent of a conditional, etc.[5]

Unlike Hom, thus, I will understand non-displaceability as a feature of every possible occurrence of slurs, not of particular occurrences of them; as implying that there are no non-derogatory uses of slurs. Non-displaceability, as I understand the term here, amounts to saying that every possible use of a slur is a derogatory one. In this sense, a slur will be non-displaceable if all of its occurrences are derogatory, i.e., if it is derogatory in every context. Once we understand non-displaceability in this way, the feature that Hom calls "non-displaceability" becomes just the ability of slurs' derogatory effect to project under different linguistic environments. To be able to shift context is to be able to capture the attitude in such a way that it does not "slip" into the actual context. The most

---

[5] In my view, the characterisations of non-displaceability offered by both Potts (2007) and Cepollaro (2020) remain faithful to the original one, as witnessed by the following quotes:

> Expressives cannot (outside of direct quotation) be used to report on past events, attitudes, or emotions, nor can they express mere possibilities, conjectures, or suppositions. They always tell us something about the utterance situation itself. (Potts 2007: 5)

> Expressive content is always referred to the situation of utterance: even when it appears to be *syntactically* embedded, the expressive content of slurs is *semantically* unembedded (which is not surprising, given the fact that it cannot interact with at-issue content. (Cepollaro 2020: 141)

obvious mechanism for doing this is linguistic, that is, inserting expressive content into a complex linguistic environment, such as a conditional or negation. However, directly identifying non-displaceability with the behaviour exhibited by slurs under the scope of truth-functional operators, as Croom and Hom seem to do, has the consequence that we tend to confuse the property of non-displaceability with the hyper-projective profile (Camp 2018a: 39) displayed by slurs (DiFranco 2014: section 1.c). The hyper-projective profile of slurs follows from their non-displaceability, but the relationship between these two characteristics is one of consequence, not identity.

Cruse's characterisation points to a broader thesis, not limited to specific linguistic mechanisms. What strictly follows from Cruse' characterisation of non-displaceability is that expressive content does not have the capacity to be captured and moved to another context —recall how it is not possible to block the effect of the emoji even when it is embedded in the antecedent of the conditional. When I talk about slurs being non-displaceable (along the lines of the characterisations by Potts (2007) and Cepollaro (2020)), what I mean, then, is that the derogatory effect of slurs will always "jump out" of whatever context we want to lock it in. We can imagine mechanisms other than the insertion in a specific linguistic environment that we might try to use to capture the derogatory effect. For example, let us think of a case of reported speech, which is another mechanism by which we can shift content to another context. In particular, let us imagine a situation in which this sentence is uttered:

(6) B says: A said that M is an *S*.

According to an empirical study carried out by Cepollaro et al. (2019), occurrences of slurs such as that in (6) turn out to be less offensive than predicative occurrences of slurs, such as "M is an *S*". The explanation they offer for the lesser offensiveness of (6) is that the audience perceives that, with the act of reporting, B makes explicit her distance from the attitude expressed by A (Cepollaro et al. 2019: 36-39).

What a case like (6) tells us is that variations in the strength of the derogatory effect of slurs can occur. The observations about the variation in the strength of the derogatory effect of slurs can be interpreted as reasons against slurs' being non-displaceable in the sense of it being impossible to capture the derogatory effect in such a way that it does not creep into the actual context. After all, if the derogatory effect can be modulated, it is because we have been able to capture the derogatory content. However, I think the point of the study is instead that, although derogation can be modulated, there is *always* derogation when a slur is uttered, no matter the context.

The data we obtain from this study allow me to make the following points. On the one hand, they allow us to test the feature of non-displaceability outside a strictly linguistic environment, which is a counterexample for the identification of the feature of non-displaceability and the projection profile of slurs. On the other hand, it allows me to put forward one of the main hypotheses that will guide this work, namely that there are no contexts in which slurs do not derogate. I will take a slur will be non-displaceable if it is derogatory in every context. Since I think this is the case, I will take a slur to be non-displaceable.

In chapter 6 we will have the opportunity to see why even in situations where we have very good reasons to consider that the use of slurs is not derogatory, it is not clear that it is. This can generate some resistance because makes us have to confront our conviction that our purest motivations are not enough to block derogation. The fact that I will consider that there are no neutral contexts or that there are no contexts in which slurs do not derogate does not mean that derogation will always be equally dangerous. Understanding that the consequences of slurs and their dangerousness can change seems to me to be in line with the empirical evidence we have on the repeal of slurs (Cepollaro et al. 2019).

The other three semantic characteristics that are also related to non-displaceability are the features of immediacy, expressive autonomy and ineffability. Let us start with immediacy because of its strong connection to non-displaceability. *Immediacy* (Potts 2007: 13; Croom 2011: 346) high-

lights the performative character of expressive content, namely, when we utter a slur what we do is expressing an attitude. In the case of slurs, what we carry out is the expression of a discriminatory attitude towards a target group. If we remember, in previous lines I pointed out that expressive content serves to transmit our emotions, attitudes or feelings. To the extent that non-native content belongs to us, it is not transferable. Consider, for example, the case of pain. The feature of immediacy of expressive content points to the fact that I cannot express the pain that my friend feels at a given moment but this does not mean that I cannot experience and express the same pain. In the same way, I cannot express the attitude that my friend expressed at a given moment, although this does not mean that I cannot express the same attitude. As noted above, expressive content cannot transcend the context of utterance. The reason why I cannot express their attitude is that we have no way of picking up the expression of an attitude that was carried out by the utterance of a slur in a particular context of utterance. Having said that, I think we can now understand why the feature of non-displaceability and the feature of immediacy are so closely connected. Ultimately, immediacy reinforces the idea that expressive content is not transferable from one context of utterance to another.

The *expressive autonomy* of slurs (DiFranco 2014: section 1.d) refers to the ability to derogate independently of the speaker's attitudes. Imagine a situation where someone utters a slur with the intention of saying something good about the target group. For example, "There *S*s are very good at mathematics". What is problematic here is not only the generalisation of a quality related to the geographical origin of a population group but also that, even if we want to say something good about the target group, we use the slur to do so. Another example in which expressive autonomy is put to the test is the following. Let's imagine that someone says "Institutions that treat *N*s as *S*s are morally depraved" (Hom 2008: 429), where "*N*" is the neutral counterpart for the slur "*S*". In this case, the speaker's intention to point out the discriminatory attitude of institutions towards a certain

target group does not block the derogatory effect of the slur.[6] Finally, a case of reported speech such as the one discussed above, (6) "A said that M is an *S*", is also a case of expressive autonomy. The distance from the original utterance believed to be taken by the speaker who makes the reported speech, probably motivated by the intention of pointing out the discriminatory attitude towards a target group expressed by another person, does not make the use of slur neutral either. Such cases are of particular interest because they tend to be perceived as less offensive than the original speech. Cepollaro et al. (2019) discuss whether what the reporter does is a use or mention of the slur, and Tenchini and Frigerio (2020) discuss what kind of relation is established between the reporter and the original utterance. In any case, what is at issue here is whether the intention of the speaker is relevant to the blocking of the derogatory effect of the slur, and the conclusion of both studies seems to be that it is not.[7]

Finally, *ineffability* (Potts 2007: 10) refers to the impossibility of paraphrasing expressive content. Specifically, in the case of slurs, it would be the impossibility of using other words to convey the evaluative component. How can we express a discriminatory attitude if not with another term that also expresses this kind of attitude? As for pejoratives in general, we have already seen that there are insults such as "jerk" which, unlike slurs, we would not say are made up of a descriptive part and an evaluative part, but only of the evaluative part. This would bring them closer to behaving like *thin therms* (Williams 1985). The impossibility of paraphrasing, in this case, would be equivalent to the impossibility of giving a definition for the concept "good" (Moore 1903/1993: section 11; see also Gibbard 2003: 23–29).

---

[6] According to Hom, this is an example of unorthodox occurrence of the slur, which means that in this case there is no derogatoion. I will analyse this case in section 6.1.

[7] It is quite likely that nearly all the semantic features I attribute to slurs, however, are such only from certain semantic positions, expressive autonomy being the most obvious case. This feature of slurs can only be recognised to the extent that we believe, as I do, that the intention of the speakers does not determine the meaning of the terms and expressions.

2.5. Conclusion

At the beginning of this chapter I indicated that the different features of slurs that I was going to distinguish would guide the development of NPP as a theory that sets out to explain the meaning of slurs. First, the descriptive features tell us the main difference there can be between a slur and an insult. The factual error that the descriptive features mark in relation to the use of slurs led us to introduce the neutral counterparts of slurs and the anti-disentangling argument (McDowell 1981; Roberts 2011). According to this argument, it is not possible to separate descriptive content from evaluative content. This fits with my intuition that there is not a clear neutral counterpart for some slurs, especially for certain racial epithets. Eliminating the descriptive features that set neutral counterparts would imply that slurs are purely evaluative. From my perspective this would be good for two reasons. The first is that we would have to give up the idea that by using slurs we are doing some kind of disguised description. The second is that we would avoid setting descriptive characteristics on the basis of people's identity traits, which I think is problematic in itself.

In spite of these reasons, however, I am assuming that there are neutral counterparts and, therefore, that slurs have descriptive features. The reasons are the following. First, a theory of a presuppositional nature, such as NPP, needs to specify a content in case the presupposition is given. Secondly, recognising certain descriptive characteristics can be useful in identifying situations of injustice that are generated by people's identity traits. To the extent that this is useful in addressing oppression, neutral counterparts are worth keeping. One possible criticism of this position is to point out that since it needs this content for the theory to work, I make this need into something morally good. But the moral reasons I have to favor the theory are independent from the theoretical ones.

A little further on, the practical characteristics allow us to specify the harm we associate with the act of derogating. These characteristics not only provide us with information about the conse-

quences of the use of slurs but are also explanatory in accounting for the dynamics that are established when we use them. For example, the normalisation of slur use is one reason that supports the hypothesis that there are no non-derogatory contexts. Moreover, the sense of complicity that is generated when a slur has been normalised in a conversation plays an important role in the idea that derogation is something sticky and that we cannot get rid of once we have committed it. This is related to the explanation we will see in chapter 5 about the resistance of slurs to retraction.

Finally, within the semantic features, we have discussed the feature of non-displaceability. This discussion allowed us to put forward one of the main hypotheses of this work, which will guide chapter 6, namely the absence of non-derogatory contexts. We also saw that the feature of non-displaceability must be distinguished from the ability of slur's derogatory effect to project under different linguistic environments. In chapter 5, I will argue that the derogatory effect of slurs indeed projects under every possible linguistic environment. So slurs exhibit both non-displaceability and hyper-projectivity: they are derogatory in every context and every linguistic environment. As introduced in this chapter, it would be conceptually possible to imagine contexts in which slurs do not derogate, even if once we have a derogatory use of a slur, this cannot be moved to a different linguistic environment without the derogatory component. In this sense, we could have hyper-projectivity without non-displaceability. Non-displaceability, as I am using the term, seems, on the other hand, to imply hyper-projectivity. If every possible use of a slur is a derogatory one, then it is a fortiori impossible to move a slur into a linguistic context where the derogatory content is absent. So, even if these two characteristics are not related by identity, they are in a relationship —one of consequence.

<center>**Chapter 3**</center>

<center>**Theoretical landscape**</center>

3.1. Introduction

The main objective of this chapter is to present some of the most relevant positions that can be found in the literature on slurs. Since this literature on slurs has been very prolific, the reconstruction of a map that accounts for the different positions can be done in different ways. In my opinion, there are two proposals that stand out for the organisation they propose. On the one hand, DiFranco (2014) divides the positions into two broad groups. Some authors, DiFranco says, maintain that we can explain the derogatory effect by the fact that slurs have a specifically derogatory content. DiFranco calls these *content theories*, among which we can find most proposals. The second group of theories, those that are not based on derogatory content, are known as *deflationary theories*. On the other hand, Cepollaro (2020) classifies the theories following a different criterion: she groups, on the one hand, theories that place the derogatory content at the level of the asserted content —semantic positions—, and, on the other hand, positions that do not —pragmatic positions—, whether they understand the derogatory content as implied or presupposed or they do not rely on derogatory content at all. Cepollaro's criterion is more restrictive because semantic content is narrower than content in general, as it includes only what is said. This should not be understood as a criticism of Cepollaro's (2020) approach, since her classification is only instrumental to her work's aim of explaining thick terms and slurs as hybrid evaluatives.[1]

It seems to me that the organisation I propose in this chapter benefits from both distinctions. I am going to classify theories into three groups, namely, content theories, *force theories* and deflationary theories. The group of content theories is characterised, as stated, by its explaining derogatory effect in terms of content. Furthermore, I will specify at which level of meaning these theories

---

[1] I will present Cepollaro and Stojanovic's account in subsection 4.3.4.

place derogatory content. The second group of theories, force theories, is characterised by its explaining derogatory effect in terms of the force of the speech acts we perform when we utter sentences containing slurs. The third group, deflationary theories, is characterised by explaining the derogatory effect through social rules or conventions.

As I survey the different salient positions on slurs I try to explain some of the objections that have been made to these positions in the literature and that Hom (2010) and DiFranco (2014) take up in their analyses. However, I do not do this exercise with all the positions because, although I think Hom and DiFranco offer a good overview of the different theories, it seems to me that we cannot put them all on the same level and, therefore, we cannot demand the same from them. In a space devoted to commenting on different positions on the meaning of slurs, it makes sense to test certain ideas that have been defended in other contexts to see whether they can be applied to slurs. However, if these intuitions have not developed into theoretical positions that have contributed in a significant way to debate on slurs, we should not occupy that space with possible objections to those yet-underdeveloped positions. As we shall see, the Fregean-based position (subsection 3.3.1) and the gesture-based position (subsection 3.3.4) are underdeveloped positions in relation to derogatory content —we can only imagine what could be said along those lines.

There are different ways of ordering the groups and proposals within this chapter. The order I have chosen is as follows. Firstly, we will find content theories. Within this category, I will first discuss proposals based on implicatures (subsection 3.2.1) and presuppositions (subsection 3.2.2) and, second, theories of content that place content at the level of what is said (subsection 3.2.3). It may seem counterintuitive not to start with theories that locate derogatory content as part of what is said, but it seems to me that this order respects the chronological priority that implicature theories have had in the history of theories of expressive content, at least if we take into account the work of Stenner (1981). Secondly, we will find force theories (section 3.3) and finally, a deflationary proposal (section 3.4).

3.2. Content theories

3.2.1. Implicature theories

The first group of proposals on the meaning of slurs is the implicature theories. Broadly speaking, implicatures refer to information that is not asserted, but merely suggested. Grice (1967/1989) distinguishes between two types of implicatures, namely conversational implicatures and conventional implicatures.Within conversational implicatures, Grice draws a further distinction between *particularised* and *generalised* conversational implicatures. The examples of conversational implicatures I will offer in what follows will be examples of particularised conversational implicatures, but the arguments against views based on them I will consider apply to views relying on generalised conversational implicatures too. Väyrynen's (2013) account of thick terms resembles one based on generalised conversational implicatures, but the kind of "generalised pragmatic implication" he relies on is kind of *sui generis* and, in fact, he considers and rejects using generalised conversational implicatures to account for the behaviour of thick terms (Väyrynen 2013: 97–107). Although, as pointed out in section 2.1, I conceive of slurs as a kind of thick term, getting into the particularities of Väyrynen's proposal would require introducing mechanisms that go beyond implicature and demand more space than I can devote to this kind of account here.

Consider the following example. A friend who lives abroad and has come to visit me asks me if it is compulsory to wear a mask on public transport. I answer that I wear it because when we go to catch the bus it's rush hour and you can't keep a proper distance from the other people on the bus. In this case, I would have flouted the Gricean maxim of quantity that tell us not to make our contributions more informative than required, since the question whether is it compulsory to wear a face mask on public transport or not does not require that much information as a response. (There is another maxim of quantity that tell us to make our contributions as informative as required.) We

need to follow conversational maxims such as this in order to be cooperative. If my friend wants to consider me cooperative, she must assume that I have skipped the maxim in order to convey more than what I have literally said. Therefore, she must infer that I have meant to say that it is not compulsory to wear a mask on public transport.

Maxims of quantity constitute only one among four categories of maxims. The other three are quality, relation and modality. The maxims of quality refer to the truthfulness of what the speaker says; the maxim of relation refers to the relevance of what is being said, and the maxim of modality refers to how we convey the information. When a conversational implicature is triggered, our interlocutors must reconstruct what we must have meant—in the example above, that it is not mandatory to wear a mask. This is because the alternative is to think that we have only meant to convey the literal meaning of our words (in this case, that it is advisable to wear the mask), and if so we would have flouted some maxim.

Among the features of conversational implicatures pointed out by Grice, three have received special attention in the literature. First, conversational implicatures are cancellable: it is admissible for me to add "and on top of that it's compulsory to wear it" after enumerating the reasons why one should wear a mask (Grice 1967/1989: 44). Second, they are calculable: it must be possible for my friend to infer, taking the conversational maxims into account, what I must have meant by saying what said (Grice 1967/1989: 31). Third, they are (unless they depend on the maxim of modality) non-detachable: since what allowed the hearer to extract the implicature was the content of my words and not the particular words I chose to convey that content, it is not possible to find another way of saying what I said that lacks the implicature in the same context (Grice 1967/1989: 39).

The first problem we encounter in explaining the derogatory content of slurs in terms of conversational implicatures is that they are based on the intuition that what we literally convey is not what we really mean. However, when we utter a slur it does not seem that what we are doing is resorting to a mechanism to avoid literally conveying the proposition that Ns are despicable, since, in

fact, that is what we are saying. Secondly, conversational implicatures are, as we have seen, cancellable. But derogatory content does not seem to be cancellable. Let us compare the following examples:

(1) It's rush hour and you can't keep a proper distance from the other people on the bus, and on top of that it's compulsory to wear it.

(2) The Zaidín neighbourhood is full of *S*s, but I don't mean to say anything bad about the *N*s.

What differentiates one case from the other is that in the former we can say that p is suggested, so by saying that ¬p we would not be contradicting ourselves. However, in the second case, by using the slur we do not seem to be suggesting, but rather committing to that p, so that by denying ¬p in the second half we would be contradicting ourselves. This is what makes it so that in the first case we can say that the implicature is cancellable and in the second case it is not. The derogatory content behaves in such a way that it cannot be concealed and leaves a trace that cannot be erased.

Having ruled out conversational implicatures as grounding an explanation of the derogatory character of slurs,[2] let us try with conventional implicatures. Conventional implicatures are contents that are systematically suggested each time we use a certain word, because of the conventional meaning of that word. For instance, by using "but" we always suggest that there is a contrast between what comes before and what comes after this connective. Conventional implicatures are characterised by the following. First, unlike conversational implicatures, they are not cancellable without triggering a feeling of oddity, as we have seen is also the case with derogatory content. Second, we cannot infer conventionally implied content by means of the maxims that have been flou-

---

[2] Remember that there are both particularised and generalised conversational implicatures. Although the arguments above have been stated as concerning particularised conversational implicatures, they apply to generalised ones too, since relying on them would equally have the result that the speaker has not said that *N*s are despicable and they are cancellable as well.

ted, as was the case with conversational implicatures. This is information that, although not being part of the asserted content, is communicated via conventional meaning. Third, they are detachable from the asserted content, i.e., the same thing can be said in an alternative way without triggering the implicature. Fourthly, conventional implicatures are triggered in the vast majority of contexts, i.e., they share with derogatory content at least part of being non-displaceable. As we can see, just like it happened with presuppositions but unlike what we have seen that happens with conversational implicatures, conventionally implicated content and derogatory content show similar behaviour.

According to a standard view of derogatory content based on conventional implicatures (Stenner 1981; Whiting 2007, 2013; Williamson 2009), two sentences like "A is an *S*" and "A is an *N*" would be extensionally equivalent, or equivalent at the level of what is said. However, the first of these triggers an implicature whose content is the proposition that A is despicable because of being an *N*, where "*N*" is the neutral counterpart of "*S*".

To explain one of the objections to the implicature approach, we must make reference to Bach's (1999) argument that it makes no sense to talk about conventional implicature because the content at issue is part of what is said. Consider the following example:

(3) Alice says: Sean is the landlord, but he's decent.

(4) Saoirse says: Alice said Sean was the landlord and he was decent.

It seems that what Saoirse says in (4) does not faithfully reflect what Alice says in (3); if "and" and "but" made the same contribution to the truth-conditions of the sentences in which they occur, however, Saoirse's paraphrase should be legitimate. Thus, Bach concludes, "and" and "but" are not truth-conditionally equivalent. This example is the result of applying what has been called the indirect quotation (IQ) test (Bach 1999). Analogously, Hom applies the IQ test to slurs to prove that the slur and its neutral counterpart do not make the same contribution to the truth conditions and, there-

fore, we cannot maintain the explanation of the derogatory content in terms of conventional implicatures.

Consider now this example:

(5) Alice says: Clara is an *S*.

(6) Saoirse says: Alice said Clara was an *N*.

Again, it seems that in (6) Saoirse is leaving out important information about what Alice said in (5), since (6) fails to convey the derogatory effect of the slur. According to Hom (2008: 8), this proves that the derogatory content is not conventionally implicated, but asserted.

DiFranco, for his part, points to the difficulty that proposals that rely on conventional implicatures to explain derogatory content meet when it comes to account for slurs' practical characteristics (DiFranco 2014: section 2.f). If we recall, the practical features of slurs were broadly shaped by the risk of normalising attitudes and subsequent uses of slurs. If I say "Sean is the landlord but he's decent", I may be conventionally implicating that landlords are not usually decent, but there seems to be no invitation to share this belief that would eventually result in the belief's spread. We do not feel complicit when someone uses the word "but". But we do feel complicit when someone uses a slur. There is thus something missing in the conventional implicature account.

One way to overcome this difficulty could be to replace the standard conception of conventional implicatures on which Stenner (1981), Whiting (2007, 2013) or Williamson (2009) rely with Potts' (2007) *expressive* conventional implicatures. Potts' proposal is characterised by the understanding that the content of the implicature triggered by slurs is not a proposition but an attitude. The implicated content takes the form of a function that models the attitudes of the speakers. This function or expressive index is represented as "<*a* I *b*>", where "a" and "b" are subjects and "I" is the interval that is modified according to the attitudes expressed in the conversation. For instance, the expressive index of a conversation could be <Alice [-1, 1] Clara>, where the interval [-1, 1] repre-

sents Alice's baseline state of indifference to Clara. If Alice utters "Clara is an *S*", however, the index would turn into <Alice [-0.5, 0] Clara> where the interval [-0.5, 0] represents Alice's negative attitude towards Clara.

Before discussing the virtues and defects of Pott's proposal, note that, as we will see in the next chapter, what we mean by "expressive" or what characterises a proposal as expressive can be understood in different ways. It may be that among Potts' motivations for developing an expressive approach to implicatures is to be able to explain the practical features of slurs. However, these are not the only reasons we can find behind the expressive version of the conventional implicatures proposal. We can also find in this proposal an expressive conception of derogatory content, since the implicature that triggers the slur takes the form of the expression of a certain attitude. Therefore, we could say that Potts' proposal is expressive insofar as he understands that the derogatory content cannot be collected in a propositional way and, also, insofar as the expression of an attitude facilitates the explanation of the practical features of slurs.

A proposal such as Potts', which characterises derogatory content as non-propositional, seems to be in a good position to explain the normalisation of certain attitudes and the use of slurs in the course of conversation. However, the expressive index has the problem of not being able to distinguish slurs from other pejoratives, such as insults (DiFranco 2014: section 2.f). If, instead of saying "Clara is an *S*", Alice says "Clara is a jerk", the expressive index would be represented in the same way as for the slur, namely as <Alice [-0,5,0] Clara>. In a nutshell, we cannot within this framework characterise slurs as terms that derogate by virtue of membership in a certain group.

3.2.2. Presuppositional theories

The next family of theories to be discussed comprises presuppositional theories. As we saw in section 1.4., the position I will develop, namely NPP, is a presuppositional theory. Following the expository nature of this chapter this subsection will be only a sample of the main points of the discus-

sion that this group of positions has given rise to. Details of the arguments for and against such theories will be given in chapters 4 and 5. We can say that a presupposition is a relation between two sentences such that for one of them to have truth-value the other sentence must be true. For example, if Víctor says "I had a bad night because my baby is teething", for this sentence to have truth-value, the speaker, Víctor, must have a baby. We can also consider as presuppositions the propositions that are commonly assumed to be true by the participants in a conversation. In this case, it is assumed in the context of the conversation that Víctor has a baby.

Both Hom (2010: 175) and DiFranco (2014: section 2.g) focus on the latter characterisation of presuppositions, which has a pragmatic slant, to present the framework on which presuppositional proposals about derogatory content rely. As we will see in the next chapter, however, whether we consider presuppositions from a semantic or pragmatic perspective is, in a sense, irrelevant for the task of accounting for derogatory content in presuppositional terms.

Presuppositions have been considered a suitable semantic vehicle to explain the derogatory content of slurs because they allow us to communicate it in a devious way. In the next chapter, we will talk more specifically about how derogatory content, understood as presuppositional content, becomes part of the common ground without the participants in the conversation having the opportunity to reject it. Some of the most representative presuppositional theories are expressive presuppositions proposals such as Schlenker's (2007) or Marques and García-Carpintero's (2020) and Cepollaro and Stojanovic's (2016) evaluative presuppositions proposal.

Most of the literature has maintained that slurs and presuppositions behave very similarly when they have were in complex linguistic environments (Lasershon 2007; Cepollaro & Stojanovic 2016; Marques & García-Carpintero 2020). This is what we mean when we say that they have a very similar projection profile. However, as we will see below, slurs manage to project in complex linguistic environments where presuppositions are blocked. So, the main objection directed at presuppositional theories has been this: why should we consider it an advantage to explain derogatory

content in presuppositional terms if they do not have the same projection profile, which was the main motivation for doing so?

Consider the following sentences:

(7) Amy has stopped wearing her mask on the street.

(8) If Amy has stopped wearing her mask on the street, then she will be able to run better in the park.

(9) Amy has not stopped wearing a mask on the street.

(10) Has Amy stopped wearing a mask on the street?

In all these cases we can say that the original presupposition, namely that *Amy wore a mask outdoors*, is projected. It does not matter whether the trigger of the presupposition appears in the antecedent of a conditional, negated or in the context of a question in all of these cases we get the presupposition. Let us now compare these cases with the following appearances of slurs:

(11) A is an *S*.

(12) If A is an *S*, she shouldn't go to that restaurant.

(13) A is not an *S*.

(14) Is A an *S*?

As in the case of presuppositions, the derogatory content survives even if it appears in a conditional, negated or questioned. However, if we apply one of the linguistic tests that allow us to know whether we are dealing with a presupposition or not, such as the *binding test* (van der Sandt 1992), what we observe is that we obtain different results. Consider the following cases:

(15) If Amy wore a mask on the street, then she has stopped wearing it.

(16) If *N*s are bad because they are *N*s, A is a *S*.

The binding test is applied by placing in the antecedent of the conditional the presupposition triggered in the original sentence. So to get (9), we would place the presupposition we get in (1), namely *that Amy wore the mask in the street*, and to get (10), we would place the presupposition we get in (5), arguably *that Ns are bad because they are Ns*. In the first case, the presupposition gets blocked, but this does not happen in the case where we have a slur. As noted, the fact that the derogatory content manages to pass the binding test when the presuppositional content does not, then, has been taken to show that the derogatory content cannot be explained by presuppositions. In chapter 5 we will look at different explanations for the outcome of applying this test to slurs. NPP will maintain an explanation based on the non-propositional nature of the derogatory content.

3.2.3. What is said

3.2.3.1. Combinatorial externalism

We now turn to discuss a position which locates the derogatory content of slurs at the level of what is said. This is Hom's (2008, 2012) *combinatorial externalism.* According to this proposal a sentence featuring a slur such as "A is an S" means "A ought to be subject to $p^*_1 + \ldots + p^*_n$ because of being $d^*_1 + \ldots + d^*_n$ all because being N^*", where $p^*_1 + \ldots + p^*_n$ are discriminatory prescriptions licensed by a bigoted ideology, $d^*_1 + \ldots + d^*_n$ are the negative traits that ideology attributes to members of the target group, and "N^*" is S's alleged neutral counterpart (Hom 2008: 431). When we say that A is an S, then, we say that A ought to be treated in a discriminatory way because of the properties A has in virtue of her belonging to the group she belongs to. Since slurs are derogatory in virtue of this being the content we convey by using them, combinatorial externalism is a content-based account of slurs. This content is located at the level of what is said, since by saying that A is an S we *say that* such and such is the case.

Combinatorial externalism is a species of externalism because the meaning of slurs does not depend on the speaker's attitudes, but on the social practices at work (see Putnam 1975). This

means that combinatorial externalism can accommodate slur's expressive autonomy. It also accommodates the natural intuition that slurs have empty extensions: since no one ought to be treated in a discriminatory way because of having the negative traits that bigoted ideologies attribute to certain groups, no one can be said to be an *S* (Hom 2012: 394).

Things do not go that smoothly for combinatorial externalism when it comes to accounting for the projection profile of slurs (cf. Richard 2008: 17; DiFranco 2014: section 2.i). If the translation of "A is an *S*" offered above is accurate, "A is not an *S*" should mean "A ought not to be subject to $p^*_1 + \ldots + p^*_n$ because of being $d^*_1 + \ldots + d^*_n$ all because being N^*". But, as noted above, this is of course true: A ought not to be treated in a discriminatory way because of having such and such properties, because no one ought to be treated in a discriminatory way because of having such and such properties. It seems, then, that by saying "A is not an *S*" we would be evincing our disagreement with the attribution of certain negative traits to the target group. But this does not necessarily seem to be what we do when we say that A is not an *S*. We can say this because we disagree that A is an N^* even if we agree that, if she were so, she would have the negative traits that the ideology at issue attributes to this group. In fact, a lot of context must be built to avoid this being the natural interpretation of such an utterance.

3.2.3.2. Inferentialism

Inferentialism is a semantic theory in which inference has explanatory priority, i.e., conceptual content can only be understood in the light of inferential reasoning. This makes it possible to deal with semantic content without a representational vocabulary: words do not have meaning because they represent entities other than themselves. Instead, they acquire their meaning because they are within an inferential network, that is, they can be part of premises or conclusions. We can say that we know the meaning of an expression because we know the conditions under which speakers are justified in using it and the consequences of having used it.

This is how Williamson (2009) characterises the inferentialist account of the slurs. This scheme, in turn, is based on the inferentialist characterisation of slurs proposed by Dummett (1981) by means of an introduction rule and an elimination rule. We use the slur "boche" because by the introduction rule we know that A is German, therefore A is a boche and, by the elimination rule, we know that it follows that A is cruel.

Introduction Rule

A is German.

Therefore, A is a boche.

Elimination Rule

A is a boche.

Therefore, A is cruel.

Brandom (2000) maintains that the only way for the non-bigoted to avoid the inference that A is a boche, and therefore A is cruel, is not to use the slur. We cannot reject what follows from using the slur "boche" if we use the word. Inferentialism is in this way remarkable for the ease with which it can account for the independence of speakers' attitudes towards the use of slurs. No matter what the speaker's intentions were, the fact remains that from the use of a slur such as "boche" it follows that A is cruel. The importance given by inferentialism to the set of inferences that follow from the terms we use allows us to specify what is wrong with using them. We could thus say that inferentialism allows us to account for the derogatory effect in a much clearer way than other proposals.

However, inferentialism has problems too. One pointed out by Williamson (2009) is that, according to inferentialism, we need to know the introduction rule for a slur to count as knowing what the slur means. This does not mean that we must be able to make the introductory rule explicit, namely, "If A is German, then A is a boche", but we must be willing to say of A that he is a boche if we think that she is German. Now, suppose we have been raised in a community in which there is

no neutral counterpart for "boche". This is the only word available to refer to Germans, and we use it competently. Since we do not know the word "German", though, we might wonder whether it makes sense to say that we know the rule. If we do not, this would, according to inferentialism, mean that we do not know what "boche" means, which seems counterintuitive.

Another problem that Hornsby (2001: 136) sees with inferential explanation is that she believes it is unfeasible to specify for each slur the consequences to which we commit ourselves once we have used it. As I have commented above, inferentialism can specify the derogatory effect of slurs given the relevance, from this point of view, of the consequences that follow from what we say. However, it does not follow from this ability to do so that the inferentialist account must have, for each slur, a characteristic that captures its derogatory effect. This does not make an inferential explanation of the derogatory content any worse.

3.2.3.3. A perspectival theory

The perspectival theory of slurs is represented by Camp (2013). According to Camp, when a speaker utters a slur she is making explicit her commitment to a certain perspective on the Ns. This perspective is based on the idea that speakers using a slur identify and reduce a certain target group to a characteristic g it has, such as its geographical origin or sexual orientation. The characteristic g determines the conditions of application of the slur, which we can understand as its descriptive features (section 2.1). This reduction to some kind of characteristic g by the bigots lays the foundation for generating evaluative traits explained by group membership. For example, we can associate an evaluative trait or characteristic such as promiscuity, indecisiveness or perversity with a characteristic g that makes reference to sexual orientation. This combination of, on the one hand, the reduction of a given target group to a characteristic g and, on the other hand, the association of certain evaluative features from that feature g makes this proposal able to account for both the descriptive features and the derogatory capacity of the slurs.

Camp contrasts her approach with Anderson and Lepore's (2013a, 2013b) proposal, which, as we will see a little later in section 3.4, is that derogatory effect is not to be explained through derogatory content but because speakers break rules by using slurs, which are considered taboo terms. On the one hand, against the position that accounts for derogatory effect on the basis of elements that are external to meaning, Camp maintains that her perspectival theory is a content-based proposal insofar as when we use a slur we show adherence to a certain perspective on the basis of the meaning of that slur. On the other hand, in relation to the level of meaning at which we situate derogatory content, Camp maintains, in contrast to positions that understand derogatory content in terms of implicatures, that since derogatory content is not cancellable it must be situated at the level of what is said. In relation to the latter position, it can be objected that, although derogatory content is shown to be resistant to cancellation, this is only a reason not to explain this content in terms of conversational implicatures, but not to reject the option of understanding the content in terms of conventional implicatures. Therefore, if Camp wants to reject this group of theories she should, firstly, take into account the distinction between conversational and conventional implicatures and, secondly, give additional reasons why the difficulty of cancellation counts against understanding derogatory content by means of conventional implicatures.

According to DiFranco (2014: section 3), one of the weaknesses of this proposal is that talking about allegiance to a perspective is not sufficient to explain what was referred to as the practical features of slurs in section 2.3. In particular, DiFranco focuses on and maintains that pointing out some allegiance does not necessarily generate any complicity in the people observing it. DiFranco's example to illustrate his criticism is a case where someone wears a sticker on their car in support of a political candidate. According to him, this does not make observers feel any kind of complicity with the perspective supported by the person who has decided to put the sticker somewhere visible.

Let us stop and look at this example in more detail. We could say that two important ideas are at work in it. Firstly, we could appeal to the capacity of certain symbolism to manifest allegiance to a certain perspective. Secondly, we could think about how the case proposed by DiFranco resembles what happens when a slur is uttered in a conversation, that is, what attitudes are mobilised in the context of the conversation. I understand that DiFranco's aim is to propose an example that leads us to think about the latter, and what I want to discuss next is whether his example works or not.

I think what makes it difficult for DiFranco's case to generate the intuition of complicity is the kind of space in which we find ourselves in this example. We are not sharing a space with the person who wears the sticker, but we share a space with the person who utters a slur. To see this let's now imagine that we get into a taxi and we do not see a sticker, but we realise that the station on the radio is one that is characterised by its conservative tone. In this case, we can understand that putting this station on, on the part of the taxi driver, is a way of showing allegiance to a certain ideology. If this kind of thing matters to us, we will feel discomfort. But, unlike in the sticker case, we will feel it because we will feel complicit. We have chosen to share our space with someone who exhibits certain attitudes in her choice of radio station. We could opt out; by not doing so, we are legitimising those attitudes and thus becoming accomplices. We could consider this case analogous to the utterance of a slur in conversation. We could opt out of the conversation; by not doing so, we legitimise the utterance of the slur. We feel complicit in the taxi case and in the case of slurs because there is a shared space we could abandon. In the sticker case, by contrast, there is no shared space we could leave.

Another example that could be considered analogous to the use of slurs is the following. Following the events that took place in Catalonia on 1 October 2017, there was a proliferation in many shops across the country of placing a Spanish flag on the outside of shops in order to show their support for a certain idea of the unity of the Spanish State and, especially, for the police action in

74

the name of that unity. On many occasions, deciding to enter a shop with a flag of this type was to accept allegiance to a certain perspective. If for some reason one had to enter any such shop while not sharing the owners' allegiance, one would certainly feel uncomfortable, and it seems reasonable to attribute this feeling to the feeling of complicity.

3.3. Force theories

3.3.1. A Fregean account

The first position I will discuss within this group of theories is an application of the notion of nuance to derogatory effect. In "On sense and meaning", Frege, in addition to distinguishing between the sense and reference of a word, also points out the positive or negative associations that are attached to that word by what he calls the representation or, later, the colouring of a word (Frege 1892). According to Frege, the colouring of a word is not part of its meaning, but rather the mental image we associate with it. The classic example that has been used to explain colouring has been the case of "dog" and "cur". Although "dog" and "cur" have the same meaning and reference, they do not have the same colouring, as the word "cur" triggers certain negative associations such as *dirty* or *unkempt*. The first position I will discuss is characterised by an attempt to explain the derogatory content of slurs through their colouring, i.e., in virtue of the negative associations that slurs trigger.

The main problem with this position is that it does not seem to be able to account for the expressive autonomy of slurs (see Hom 2008; DiFranco 2014: section 2.a). As mentioned above, slurs' expressive autonomy is the feature whereby they have their derogatory effect independently of the speaker's attitudes. However, the Fregean approach to slurs seems to imply that the derogatory effect only manifests itself when either the speaker or the hearer has a negative association of the target group. Therefore, within this proposal we cannot explain cases where, despite not having a negative association, we still derogate by using the slur. One possible response that could save the Fregean approach from this objection would be to maintain that the utterances of slurs generate a

certain predisposition in speakers to have certain negative associations (DiFranco 2014: section 2.a). In this way, we would relax the commitments of the original Fregean approach by maintaining that it is not necessary to have the negative association in order for the slur to derogate; it is enough to be predisposed to have it. However, this is not a position that has as of today been developed, and so, as DiFranco says, it would still require further development in order to account for the derogatory capacity of slurs.

3.3.2. Expressivism

Classical expressivism is characterised by the claim that sentences containing a moral term are not susceptible of being declared true or false. What one would expect from an expressivist proposal about the meaning of slurs would therefore be that sentences containing slurs are not susceptible of being declared true or false (DiFranco 2014: section 2.b). This interpretation of the expressivist thesis is historically outdated, as expressivist positions do admit truth-values at least since the work of Blackburn (1984, 1993) and Gibbard (1990, 2003). However, since what I intend to do in this chapter is to set out how the expressivist position has been interpreted in relation to slurs, I will assume, only here and for this purpose, a version of expressivism that denies that the sentences at issue have a content capable of being declared true or false. As indicated in section 1.4, NPP is a proposal that inherits the expressivist conception of content and attributes truth-values to sentences containing slurs.

The main problem with applying the thesis of classical expressivism and equating slurs with moral vocabulary to understand them as purely evaluative terms is that from this perspective we would not be able to explain the descriptive features of slurs and, therefore, we would not be able to distinguish between a mere insult and a slur (section 2.1). A possible solution to this would be to explain the meaning of slurs through a dual approach, as proposed by Saka (2007). According to this proposal, by uttering a sentence such as "A is an *S*" we would express, on the one hand, the be-

76

lief that *N*s are *S*s, and, on the other hand, an attitude of contempt towards *N*s. Such a dual approach would solve the lack of descriptive features, but this is not enough. According to Hom (2010: 170) and DiFranco (2014: section 2.b), an expressivist approach should also be able to cope with the Frege-Geach problem (section 1.4).

Hom identifies the Frege-Geach problem as stemming from the fact that expressivism individuates the meaning of moral terms by the *force* of the speech acts containing them. According to Geach, theories that defend that the force of a proposition determines its *content*, such as expressivism is supposed to be, cannot give a satisfactory explanation of how certain inferential processes such as *Modus Ponens* works when they include certain expressions (Geach 1965: 462–463). *Modus Ponens* is a clear example of an argument where the force associated to an expression varies, but not its content. According to Geach (1965: 449), all expressivism can say about a sentence like "It is wrong to increase the minimum wage" is that it expresses an attitude of disapproval towards the act of increasing the minimum wage. However, "It is wrong to increase the minimum wage" fails to express this attitude then it appears as part of the conditional "If it is wrong to increase the minimum wage, then the government has made a fool of themselves". According to Geach's critique, if content depend on force, we would be forced to say that the content also varies, and, so, we would end up with arguments like the following:

It it is wrong to increase the minimum wage, then the government has made a fool of themselves (p → q)

It is wrong to increase the minimum wage (r)

Thus, the government has made a fool of themselves (q)

The kind of argument, construed in this manner, is not a *Modus Ponens*. In order for it to be a *Modus Ponens*, the propositional content must be the same for each case, despite the difference in force. Otherwise, expressivism is to blame for a fallacy of equivocations (Geach 1965: 462–464). Ho-

wever, Hom says a little later that the Frege-Geach problem is related to the feature of non-displaceability (Hom 2010: 171). To relate this objection to non-displaceability, however, is to admit that expressivism does not have a problem in attributing stable meaning to slurs, since non-displaceability precisely means that they will always derogate. Slurs do not seem subject to the Frege-Geach problem in the way moral terms are, since slurs and moral vocabulary behave differently: the former are always derogatory, while the latter do not always carry an evaluation.

If A is an *S*, then she should not go to that restaurant (p → q)

A is an *S* (p)

Thus, she should not go to that restaurant (q)

In this case, unlike the previous example which contained moral terms, we will not be able to say that this is an invalid argument because the premises and the conclusion have different contents. If content depends on force, "A is an *S*" should have the same content in both premises, as it is derogatory in both. We would thus have constructed a valid *Modus Ponens* argument.The particular projection profile of slurs, which, as discussed in the previous chapter, derives from the feature of non-displaceability, means that in all cases the slur projects, and therefore, derogates. This feature is what makes slurs behave in this particular way and, for most of the phenomena we want to explain in relation to slurs, this feature poses an additional difficulty. However, for the conception of expressivism that has been maintained in this section, non-displaceability is a convincing answer to the objection attributed to expressivism.

3.3.3. Slurring as misrepresenting

The third force-based approach to the meaning of slurs I will consider is represented by Richard (2008). This proposal is characterised by the thesis that, when speakers use a slur, they misrepresent the target group. According to Richard, by using a slur we perform an act different from assertion because misrepresentations are not susceptible to being declared true or false, but this does not

mean that we should think of them in terms of expressions of attitudes. The criticism to the intro-duction of neutral counterparts (section 2.1) could be said to have its point of origin in Richard's approach (2008: 12–13), since it is him who points out the problem that, if slurs and their neutral counterparts make the same contribution to truth conditions, we are obligated to accept the truth of a sentence like "*N*s are *S*s". According to Hom (2010: 172) and DiFranco (2014: section 2.c), the weaknesses of this proposal are, on the one hand, that the only thing we do when we utter a slur is to misrepresent the target group, and on the other hand, that we do not find in the proposal a deve-loped characterisation of what the act of misrepresenting consists of.

In relation to the first point, Hom points out that, since all we do in uttering a slur is to per-form a misrepresentation, we cannot explain the cases where in uttering a slur we allegedly do not make a derogatory use of it. Consider the following example:

(17) Institutions that treat *N*s as *S*s are morally depraved.

According to Hom (2010: 165), this is a case where it would make no sense to negatively evaluate the speaker uttering the slur. One could object that the speaker has chosen to use the word "*S*" whe-re there was no need to, and thus can be said to be behaving in a bigoted way; but Hom might reply that we cannot construct a case like (17) by means of a neutral counterpart, since what we would obtain would be a nonsensical sentence:

(18) Institutions that treat *N*s as *N*s are morally depraved.

Since Richard offers no clue as to how he would cover these cases, his proposal can be said to be incomplete.

Let us look at Hom's objection in a little more detail. The objection is based on the distinction between orthodox appearances —appearances where slurs derogate— and non-orthodox appearan-ces—appearances where slurs do not derogate. Hom considers that a theory that is presented as a

proposal about the meaning of slurs should be able to account for both cases, where (17) is a non-orthodox occurrence.

Although the distinction between orthodox and non-orthodox appearances of slurs will be discussed in detail in chapter 6, at this point we could reply to this by referring to slurs' expressive autonomy. According to this feature, the slur will derogate independently of the attitudes expressed by the speaker. (17) is a case where the speaker intends to make a complaint about how institutions treat a certain population group, but it takes more than this to automatically override the repeal of the slur. And, even if we were to consider that (17) is not a case where regardless of the speaker's attitudes the slur derogates because only a non-bigot could make such use of the slur, we can reply that there are other alternative formulations that allow us to *do*, if not to say[3], exactly the same thing. For instance,

(19) Institutions that treat *N*s in a racist manner are morally depraved.

If Hom's objection was focused on pointing out that the Richard's proposal cannot cover allegedly non-derogatory uses of slurs, DiFranco, for his part, focuses on pointing out that it does not provide enough information about what exactly it means to carry out misrepresentation of a certain group (DiFranco 2014: section 2.c). Firstly, one way of understanding misrepresentations could be that the speaker uttering the slur refers to certain negative ideas or associations that are attributed to a certain target group with the aim that the audience shares or feels complicity towards these associations. In any case, this would only be speculation about what the acts of misrepresentation we perform through slurs consist of, since Richard's proposal does not seem sufficiently clear in this respect. Moreover, according to DiFranco (2014: section 2.c), even if this were the case, it would still have to be explained in what way listeners feel challenged by the associations that are supposed to trigger the misrepresentations. Secondly, DiFranco maintains that misrepresentations would only

---

[3] The feature of ineffability prevents us from being able to say the same thing without using another expressive term (section 2.4).

work for certain problematic slurs, such as discriminatory slurs, but we would not want to be able to say that when faced with a slur like "facha" what we are doing is misrepresenting this group, or people with a conservative ideology. The distinction between slur and derogatory term (section 1.2) makes me not have to worry about this objection, as I do not consider this term to be a slur.

3.3.4. A gestural theory

The next proposal can be seen suggested but not explicitly maintained in Hornsby's (2001: 140) work. According to this author, the utterance of a slur is equivalent to an utterance of the slur's neutral counterpart accompanied by a gesture of contempt. This gesture of contempt cannot be captured propositionally, but it makes explicit some of the speaker's attitudes. The criticisms that will be levelled at this proposal relate to the lack of clarity about the idea surrounding the gesture that, according to Hornsby, explains the derogatory effect of slurs. According to DiFranco (2014: section 2.d), we can interpret the gestural theory in two ways. First of all, we could understand that it is not a gesture that we can physically appreciate, such as giving the finger, but something that somehow remains hidden even though it has the same effects as if we had made a physical gesture. Secondly, we could understand the derogatory effect of slurs to be analogous to making a derogatory gesture. According to DiFranco, for the first of the interpretations, Hornsby would have to specify what exactly it means to make a derogatory gesture; for the second interpretation, Hornsby would have to specify in what sense the derogatory effect is analogous to a derogatory gesture. In any case, although we cannot say that Hornsby's proposal is a clear-cut one, the idea of linking the derogatory effect to the representation or expression of derogatory gestures is intuitive and can be a good starting point to theoretically argue for the idea that the use of slurs dehumanises (Jeshion 2013a, 2018; Tirrell 1999, 2012).

3.4. A deflationary theory

Not all views have aimed at accounting for the derogatory character of slurs by attributing derogatory content to them. The most discussed non-content-based account of slurs is due to Anderson and Lepore (2013a, 2013b). According to Anderson and Lepore, slurs are not derogatory because they have derogatory content, but because they are taboo words: since their utterances are prohibited, using or mentioning them means to break a social convention and is thus perceived as offensive (Anderson and Lepore 2013a: 38). Anderson and Lepore's account can be called "deflationary" inasmuch as it does not rely on the notion of derogatory content to explain what makes slurs derogatory words.

A salient advantage of Anderson and Lepore's account is that it applies to occurrences of slurs across the board, whether they are uses or mentions of these terms (Anderson and Lepore 2013b: 353; see also Saka 1998: 122). This serves to nicely account for the uneasiness that many of us feel when faced with a slur, even if it is only mentioned. Although content-based views might be enriched so as to accommodate this fact, it seems dubious that they can do so in as natural a way as Anderson and Lepore's view does. However, Anderson and Lepore are forced to lose other of the theoretical virtues that an account of slurs might have in this trade-off.

Three disadvantages of this view are these. First, it seems to reverse the order of explanation: slurs are taboo because they are derogatory, not derogatory because they are taboo (Croom 2011; see also Cepollaro 2020: 114). Second, slurs are not necessarily taboo words: a bigoted enough society may use words that we might recognise as slurs, and thus as derogatory, even if no social convention forbids utterances of these words (Whiting 2013). Third, there is no way of drawing the distinction between slurs and derogatory terms (see section 1.2) if we characterise the former merely in terms of prohibition, as this applies to the latter too (DiFranco 2014: section 3). As smoothly as Anderson and Lepore may explain why mentions of slurs strike us as offensive just like uses of them

do, therefore, it would be reasonable to prefer explaining this fact in a less natural way but avoiding these three objections.

3.5. Conclusion

The theoretical options that have tried to explain the derogatory effect of slurs in the literature are varied. Some of them have tried to account for this effect by pointing out a specifically derogatory content, while some of them have tried to do so by taking into account the force of the speech acts we perform when we utter sentences containing slurs, and yet some others have directly attempted to address the derogatory effect by looking at social norms. Within the first family of theories, some have identified derogatory content have located at the level of presuppositions or implicatures while others have located as part of what is said. The proposal I will make in this thesis is precisely a presuppositional one. The whole of next chapter will focus on presenting different presuppositional views to arrive at my own.

Theory			Advantages	Disadvantages
Content theories	Implicature theories		• Explain the elusive character of the derogatory effect.	• Have difficulty accommodating the normalising potential of slurs. • Wrongly predict that slurs pass the IQ test.
	Presuppositional theories		• Explain the elusive character of the derogatory effect.	• Wrongly predict that slurs pass the binding test.
	What is said	Combinatorial externalism	• Explains the expressive autonomy of slurs. • Accommodates the intuition that slurs have empty extensions.	• Has difficulty accounting for the projection profile of slurs.
		Inferentialism	• Explains the expressive autonomy of slurs. • Provides an intuitive explanation of the consequences of uttering a slur.	• We can know the meaning of a slur without knowing the corresponding inferential rule.
		A perspectival theory	• Provides an intuitive explanation of the ideological commitments involved in the utterance of a slur.	• Has difficulty explaining the feeling of complicity.

Theory		Advantages	Disadvantages
Force theories	A Fregean account	• Distinguishes between slurs and neutral counterparts.	• Underdeveloped.
	Expressivism	• Explains the expressive autonomy of slurs.	• Does not explain the descriptive features of slurs.
	Slurring as misrepresenting	• Identifies the problem for neutral counterparts following from the anti-disentangling argument.	• Underdeveloped.
	A gestural theory	• Explains the elusive character of the derogatory effect.	• Underdeveloped.
A deflationary theory		• Explains why mentions of slurs are problematic too.	• Reverses the order of explanation. • Cannot distinguish between slurs and derogatory terms.

Chapter 4

A non-propositional, presuppositional account of slurs

4.1. Introduction

This chapter is fully dedicated to the development of NPP. As indicated in section 1.4, NPP is a semantic proposal that stands at an intersection between an expressivist conception of derogatory content and the location of this content at the presuppositional level of meaning. The specific theses assumed and characterising this proposal are the following:

(i)   Slurs have derogatory content, but this content is not propositional.

(ii)   Derogatory content is located at the presuppositional level.

(iii) The impact on the common ground corresponding to the non-propositional presuppositions that slurs trigger should be modelled in terms of the ordering of possible worlds.

This chapter and the next one are intertwined insofar as the purpose of the latter is to test these theses. In this chapter, by developing NPP, we will give an account of the descriptive features of slurs (section 2.1). In the next chapter, by showing how NPP works with respect to certain linguistic phenomena, we will deal with the semantic features of slurs (section 2.3).

The order you will find in this chapter is as follows. Section 4.2 is devoted to introducing the notion of presupposition. In section 4.3 I comment on a few presuppositional proposals that ascribe propositional content to the presuppositions triggered by slurs, such as those by Macià (2002) (subsection 4.3.1), Schlenker (2007) (subsection 4.3.3), and Cepollaro and Stojanovic (2016) (subsection 4.3.4). I will also offer in subsection 4.3.2 some preliminary considerations in favor of presuppositional accounts in general. Section 4.4 is devoted to discussing the reasons why it makes sense to consider an extension of the notion of presupposition so that they can have non-propositional content. In section 4.5, finally, I present non-propositional presuppositional proposals, among which

I include the one developed by Marques and García-Carpintero (2020) and my own NPP. I also establish the main differences between the two accounts, which I think speak in favour of the latter.

4.2. Slurs and presuppositions

The main debates about presupposition took place in the second half of the 20th century. The debate about presuppositions has been motivated by the study of their characteristics, such as the possibility of accommodating them and their projection profile. All these basic characteristics of presuppositions will be discussed in this and the next chapter. The relationship between presuppositions and expressive terms, for its part, is relatively novel if we consider Macià's (2002) proposal as the first one that applies a presuppositional treatment to explain the meaning of expressive terms.

Presuppositions are a very broad phenomenon. We speakers continually presuppose information in our linguistic exchanges: that our interlocutor speaks the same language as us, that face masks prevent the transmission of certain viruses, that we have to get up early on Mondays, etc. Presuppositions can be either *semantic* or *pragmatic*. On the one hand, semantic presuppositions are relations between sentences that hold in virtue of the truth-values of the sentences involved. In particular, a sentence presupposes another whenever the latter has to be true in order for the former to have a truth-value. For instance, consider the following pair of sentences:

(1) The shop around the corner specialises in piano sheet music, you might find what you are looking for.

(2) There is a shop around the corner.

For us to say that (1) has a truth-value, (2) must be true. After all, it makes no sense to ask whether it is true or false that the shop on the corner is devoted to sheet music if there is no shop on the corner to start with. Since (1)'s truth-value depends on (2)'s in this way, we say that (1) presupposes (2).

On the other hand, we have pragmatic presuppositions, which are things that speakers do through their utterances, rather than relations between sentences. Pragmatic presuppositions capture the intuition that speakers share a significant amount of information that allows their communication to be successful. Speakers assume in making their utterances certain information in the absence of which such utterances would not be felicitous. For instance, if I utter (1) I will normally be assuming (2), as my utterance of (1) would be infelicitous otherwise. What we presuppose in this sense is part of the context in which utterances takes place. In particular, the presuppositions that are active at a certain point in a conversation are modelled through the *common ground* (Stalnaker 1978). The common ground can be understood as a theoretical construct that accounts for the current state of our conversations. At first, it was assumed that it is sufficient to model the propositions whose truth is taken for granted, as it seemed that this was the only thing that determines what counts as an acceptable or unacceptable contribution. As we will see in section 4.4, however, this does not seem to be enough and the common ground must be enriched.

One might think that this semantic-pragmatic distinction is, in effect, a distinction between two phenomena: some of the presuppositions we find are semantic, while others are pragmatic. However, this distinction should only serve to orient us theoretically about the phenomenon of presupposition. The explanation of semantic and pragmatic presuppositions becomes meaningless if they are understood as two completely unrelated notions.[1] On the one hand, whenever we speak of pragmatic presuppositions we are also trivially speaking of semantic presuppositions. For instance, I assume (2) as I utter (1) because I would not making a truth-apt assertion if (2) were not true. The pragmatic presupposition here, then, exploits a semantic presuppositions. Semantic presuppositions are associated with the conventional meanings of terms, so a semantic approach will be useful if it is the origin of the presupposition that we are interested in.

---

[1] In (Moreno & Pérez-Navarro 2021a), we took slurs to trigger semantic presuppositions as opposed to pragmatic ones. However, I no longer think it is productive to see things in this way.

But, on the other hand, it makes no sense either to deny that semantic presuppositions are many times mediated, in one way or another, by a context of utterance. For example, imagine we are with friends and acquaintances and at some point Víctor says "I slept badly tonight because my baby is teething". The presupposition triggered by this utterance is *that Víctor has a baby*, because it is an utterance of a sentence whose truth-aptness depends on whether the referent of "my" indeed has a baby. But, of course, it is the speaker of the context that is assigned as the referent of "my". So we must assume a context of utterance to assign a content to the semantic presupposition.

Adopting a pragmatic perspective also allows us to incorporate a diachronic dimension that makes it possible to adscribe how we update the information in the common ground via *accommodation* (Lewis 1979). Accommodation is a mechanism by which speakers update the information in the common ground so as to make certain utterances felicitous. Imagine that there are people in the situation described above who did not know, prior to Víctor's utterance, that he has a baby. It seems unlikely that they will take Víctor to have talked nonsense. Instead, they will just learn that Víctor has a baby and proceed in accordance with that information. That is, the presupposition is accommodated in such a way that Víctor's utterance is made felicitous. It is the diachronic dimension associated with a pragmatic perspective that allows us to describe how the information in the common ground is updated.

4.3. Propositional presuppositional accounts of slurs

As noted in section 4.1, I will distinguish two groups within presuppositional theories. In the first of these, to which this section is devoted, are those positions that have been built using the existing notion of presupposition; these views are thus propositional presuppositional accounts of slurs. The proposals that fall into this group are the following. First of all, I will present Macià's (2002) account, which, I have already pointed out, can be considered the first in which a presuppositional treatment is applied to expressive content. Second, we have Schlenker (2007), who develops an ex-

pressive presuppositional proposal that is intended to compete with proposals that explain the meaning of expressive terms by means of conventional implicatures. Finally, I will discuss Cepollaro and Stojanovic's (2016) proposal, which is a hybrid account of the meaning of slurs and thick terms inasmuch as it distinguishes a descriptive and an evaluative component in the meaning of these kinds of terms, the latter of them located at the presuppositional level.

4.3.1. Presuppositions and partial functions

(Kaplan 1999) can be considered the first works that explicitly addresses the need for a semantic account of the meaning of expressive terms. In his paper, Kaplan draws attention to the status of couples of arguments such as these:

(A)

    (3) Toni Morrison is a Nobel laureate in spite of being African American.

    (4) Toni Morrison is a Nobel laurate and African American.

(B)

    (4) Toni Morrison is a Nobel laurate and African American.

    (3) Toni Morrison is a Nobel laureate in spite of being African American.

While the conclusion of (A) follows from its premise, this does not occur in (B). Kaplan explains this by saying that, in addition to having semantic content, the expression "in spite of" has another kind of content, which is expressive. (4) thus shares all its content with (3), but not the other way around. This is why (A) is a valid argument but (B) is not.

Macià, however, rejects splitting the content of "in spite of" and proposes to explain the contrast between (A) and (B) by means of presuppositions, which, in his account, are to be modelled through partial functions. While the function associated with a proposition is one that assigns a truth-value to each possible world, partial functions assign a truth-value only to those possible worlds where a certain presupposition is true. For instance, consider the following sentence:

(5) Dani went to the pool to watch his brother swim.

We can point out the worlds for which the function determined by (3) will give the value true and the worlds for which it will give the value false, but only within the worlds where the function is defined, which are those in which the presupposition *that Dani has a brother* is true. Within those worlds, the partial function will assign truth to those worlds in which Dani indeed went to the pool to watch his brother swim and falsity to those in which he did not.

Let us now consider (3), where we find the expression "in spite of". This sentence can be taken to determine a partial function that assigns a truth-value only to worlds in which being African American makes it less likely that one can win a Nobel prize. Within those worlds in which a truth-value can be assigned, those in which Toni Morrison is African American and a Nobel prize winner will be assigned truth.

Once we render the meaning of "in spite of" in this way, it is straightforward to explain the contrast between (A) and (B). If we say that an argument is valid whenever there is no context in which the premise is true while the conclusion is not, (A) will be valid, because this is the case, while (B) will be invalid, because there are context in which (4) is true but (3) is not, viz., those in which it is not presupposed that you are unlikely to win a Nobel prize if you are African American. Of course, (3) is not false in these contexts, but it is not true either.

Making (3) presuppose *that African Americans are less likely to win a Nobel prize* is what "in spite of" does in that sentence. Although Macià himself does not address slurs in his 2002 paper, it does not require a lot of effort to offer a parallel account of slurs. To keep complexities to a minimum, let us assume that utterances of sentences of the form "A is an *S*", where A is a person and "*S*" is a slur, trigger the (admittedly too simple, as we will see) presupposition *that Ns are bad*, where "*N*" is the neutral counterpart of "*S*". If this is so, "A is an *S*" will determine a partial function

90

that assigns a truth-value only to those worlds in which *N*s are bad. Within these worlds, the truth-value assigned will be truth if A is indeed an *N* and falsity if she is not.

4.3.2. Advantages of a presuppositional account

Macià's view is enough to start discussing some of the advantages of presuppositional accounts. The main reason to choose a presuppositional account of the meaning of slurs over its alternatives is that it allows us to account nicely for slurs' resistance to cancellation, rejection and retraction, to which the next chapter will be devoted. However, I would like to set out now some preliminary intuitions regarding the decision to situate derogatory content at the presuppositional level of meaning.

Choosing presuppositions as the mechanism through which to explain derogatory content is explained, first, by a basic intuition about how presuppositions behave with respect to asserted content. That is to say, unlike asserted content, with presuppositional content the participants in a conversation do not have the option of considering whether or not it becomes part of the common ground, because presuppositional content slips right in to become part of it. The most direct consequence of this behaviour is that, since it becomes part of the common ground without prior consideration on the part of the speakers, this makes it impossible for us to make the presupposed content the subject of the conversation in a natural way. In the same way, we can neither reject nor retract the presupposed content (sections 5.4 and 5.5). These are arguably among the most explored reasons in the literature for choosing a presuppositional explanation, along with the similarity in projection profile between derogatory content and presuppositional content that we discussed in subsection 3.2.2.

However, we can adduce another reason that would provide a new horizon to look at when trying to account for derogatory content. The aim is to think of presuppositional views as a group of theories that, while assuming the existence of neutral counterparts, can explain why slurs and their

neutral counterparts do not make the same contribution to truth conditions. At the beginning of chapter 2, it was noted that, as a presuppositional account, NPP would be able to account for the descriptive characteristics of slurs without falling into problematic compromises. Recall that the descriptive characteristics of slurs account for the fact that we can make, in addition to the moral error that we always make, factual errors in applying the slur. So there seems to be something descriptive in their meaning. But let us also remember that this does not mean that they must have neutral counterparts and, even if they did, this would not mean that the counterparts make the same contribution to the truth conditions as the corresponding slur.

Presuppositional theories assign a descriptive meaning to slurs, thereby accounting for the possibility of factual error, and do so by accepting the existence of neutral counterparts even if they do not share truth conditions with slurs. The development of this explanation consists of two steps. First, I will discuss how we can have neutral counterparts without them making the same contribution to the truth conditions. Second, I will explain why I accept the existence of neutral counterparts despite reasons to the contrary.

First, the way in which a presuppositional theory can prevent slurs and their neutral counterparts from making the same contribution to truth conditions is made evident by simply pointing out the basic mechanism by which presuppositions work. As I will point out when introducing every presuppositional proposal, for an utterance of a sentence like "A is an S" to be felicitous the common ground must contain a certain component (either a proposition, a reactive attitude or an ordering of possible worlds). Now, slurs will make the same contribution to truth conditions as their neutral counterparts as long as this component can be found, but otherwise, they do not share truth conditions. This makes it the case that, if we do not share the presupposition against Ns, "All Ns are Ns" and "All Ns are Ss" differ in truth-value, since the first would be true but the second lacks truth-value.

Secondly, as we saw in section 2.2 after introducing the "anti-disentangling argument" (McDowell 1981; Roberts 2011), accepting neutral counterparts means that we can do the corresponding classification from a non-evaluative perspective. This is, so we argued, compatible with the conclusion of the disentangling argument, namely that the descriptive and the evaluative cannot be separated in thick terms. The existence of neutral counterparts, in this sense, does not imply that the slur and the neutral counterpart behave in exactly the same way. When cancelled, the evaluative component involved in every use of a slur renders the whole proposition meaningless, rather than false. I will accept the existence of neutral counterparts for the following reasons. On the one hand, we lack theoretical tools that allow us to otherwise assign an asserted content to a sentence such as "A is an *S*". If we did not assign such a content, we would only have the presupposition, in whatever form we want to adopt it, but only the presupposition after all. I remarked at the beginning of this chapter that part of the appeal of presuppositional explanation was that it allowed us to explain how derogatory content hides behind the asserted content and becomes part of the common ground directly. If we want to derogatory content to hide, we need an asserted content behind which it can do so. On the other hand, neutral counterparts are rendered meaningful by a classification that, although not neutral and probably problematic, must be used if we want to identify oppressed groups. To the extent that this identification is successful, neutral counterparts can become terms that are accepted by oppressed groups and, in this way, allow people who do not belong to these groups to identify the victims of derogation in a non-offensive way. This can be seen as a necessary step towards ending oppression.

4.3.3. Expressive presuppositions

Schlenker (2007) begins by recognising that his proposal is in line with the work previously done by Macià (2002) and Sauerland (2007) and his own (Schlenker 2003), although, as we will see, he will characterise his proposal as an *expressive* one. Schlenker aims to develop a presuppositional

treatment that can have as much success as Potts' (2007) proposal, which is based on conventional implicatures (see subsection 3.2.3), in accounting for the features of expressive content.

Schlenker's proposal can be characterised as follows. First, Schlenker maintains that the expressive presuppositions that expressive terms trigger are a type of semantic presupposition. Second, these presuppositions have an indexical character insofar as they are tied to the context of utterance. Third, the presuppositions at issue have an attitudinal component, since they refer to an agent's mental state. Fourth, this agent does not have to be the speaker; it can be any other person. So, if someone utters "A is an *S*" and it is part of the common ground that the agent (whoever that is) believes that *N*s are bad, where "*N*" is again the neutral counterpart of "*S*", the sentence "A is an *S*" will be true if A is an *S* and false otherwise. But, if it is not part of the common ground that the agent believes that *N*s are bad, the sentence will lack truth-value, unless the presupposition is made part of the common ground as the sentence is uttered through a process of accommodation.

Williamson (2009) and Nunberg (2018) find Schlenker's proposal defective for the following reason. If, previously to the utterance of "A is an *S*", it is not presupposed in the context that the agent believes that *N*s are bad, the presupposition will be accommodated and become part of the common ground. However, while it is problematic to accept an utterance of "A is an *S*", it doesn't seem to be that much of a problem accepting that a certain agent has a certain belief about *N*s being bad (Williamson 2009: 151–152; Nunberg 2018: 284; see also Marques and García-Carpintero 2020: 142–143). The criticism is that accommodating that an agent believes that *N*s are bad is far from what it means to accommodate the derogatory effect of a slur. It is as if there were no difference between accommodating the derogatory effect and accommodating an agent's belief that spaghetti must cook for 30 minutes. Despite this criticism, Schlenker's proposal is the first proposal that tries to explain the derogatory effect by taking into account the attitudes of the people involved in the conversation. This is a difference with previous theories, even if the expression of attitudes is

understood from a propositional framework. I will shortly analyse the way in which we can understand this proposal as an expressive proposal.

4.3.4. Evaluative presuppositions

The third propositional presuppositional account I will consider is the one defended by Cepollaro and Stojanovic (2016). The aim of these authors is to account for the behaviour of hybrid evaluatives, among which slurs and thick terms are found. Hybrid evaluatives are characterised by their expressing two types of content, namely descriptive and expressive or evaluative content. A sentence featuring a slur such as "A is an *S*" would therefore communicate two types of content. On the one hand, we would have a descriptive content whose truth-conditions coincide with those of the sentence "A is an *N*", where "*N*" is again the neutral counterpart of "*S*". On the other hand, we would have an expressive content that Cepollaro and Stojanovic render as *Ns are bad just because they are Ns*. Utterances of "A is an *S*" trigger a presupposition with this content. In this view, an utterance of a sentence of the form "A is an *S*" will be felicitous if and only if the proposition *Ns are bad just because they are Ns* is in the common ground. If this happens, the sentence "A is an *S*" will be true if A is in fact an *N* and false otherwise. If the presupposition *Ns are bad because they are Ns* is not on the common ground, however, the sentence "A is an *S*" will have no truth-value.

In a more extensive work on the subject, her book *Slurs and Thick Terms* (2020), Cepollaro continues the line of analysis opened in the 2016 paper. According to Cepollaro, one of the reasons in favour of adopting a presuppositional treatment rather than an implicature theory to explain derogatory content is that, unlike the latter, a presuppositional treatment can explain certain conversational dynamics that occur when we use slurs. In particular, Cepollaro refers to the sense of complicity that as discussed in section 2.2 and subsection 3.2.3.3. If something is presupposed content we can assume that other participants in the conversation will share that information, as presupposed content is, as Sbisà says, "what ought to be shared" (1999: 501). When we talk about slurs, we

could say that this security becomes complicity, which gives rise to the use of slurs without qualms and, furthermore, normalises their subsequent uses in the course of conversation. We feel comfortable using slurs because we assume common information with our interlocutors. What we can explain so naturally from a presuppositional perspective, according to Cepollaro, cannot be explained from an implicature approach without *ad hoc* arguments (Cepollaro 2020: 145). A proposal such as the one based on expressive conventional implicatures (Potts 2007) gives us a lot of information about speakers' mental states. However, from this theory we cannot make the leap to assume that the mental state expressed by the speaker is shared by the other participants in the conversation. Presuppositions enjoy a normative character that we cannot attribute to expressive conventional implicatures. According to Cepollaro, this makes a presuppositional explanation preferable to one in terms of implicatures.

I agree with this motivation for Cepollaro's account. However, the way they characterise the presuppositions at issue, the only thing that distinguishes them from ordinary presuppositions is the introduction of an evaluative term in the sentence that is used to identify its content—"*N*s are bad just because they are *N*s". It could be that we could find within this proposal a semantic position that individuated the meaning of "bad" in such a way that we could somehow point out what is the difference between ordinary presuppositions and evaluative ones. However, this is not something we find in the proposal as it stands. In the next section, I will develop this criticism and offer an alternative to Cepollaro and Stojanovic's view.

4.4. Going non-propositional

In section 4.2, I characterised the common ground as including only the propositions whose truth is taken for granted in the course of a conversation. We need to enrich the common ground, however, to point out what makes expressive or evaluative presuppositions special. They are special because they are *non-propositional*.

Before introducing what I understand to be a suitable notion of expressive presupposition, I will qualify the different ways in which a proposition can be presented as expressive. The first sense of "expressive" that I identify is related to the expression of the speakers' attitudes. The use of expressive terms, such as pejoratives or slurs, or moral vocabulary in its evaluative use, provides us with information about the speakers, about the attitudes they show when using them. The second sense of "expressive" is complementary to the first and concerns the non-propositional character of evaluative uses of language. The two senses need not necessarily go hand in hand. Distinguishing two senses of "expressive" is intended to clarify that talking about expressive content does not always imply talking about non-propositional content. Potts' (2007) and Schlenker's (2007) characterisations of expressive content only lead them to engage in the first sense because it is only the attitudes expressed by speakers that are relevant to individuate the content of expressive terms.

The two senses of "expressive" are primarily motivated by the distinction seen in section 1.4 between locational and orientational information (Lewis 1979; Charlow 2014: 639). Locational information tells us what world we are in, while orientational information tells us what world we should turn our world into; it motivates us to make modifications to the world. For example, if we take "I am hungry" to be expressive in the first sense, we should understand that the person is expressing some kind of locational information by describing the particular mental state she is in, and telling us that we inhabit a world where she is hungry. Understanding this utterance in this way is compatible with its not being connected to any particular action. That is, it would be compatible with the speaker going to get food or not. Meanwhile, if we take "I am hungry" to be expressive in the second sense, we should understand that it is necessarily connected to action, and the kind of information it expresses is thus orientational. The person who utters that she is hungry would thus in a sense be rationally required to do something to stop being hungry. For the specific purposes of this and the following chapter, where we aim to explain certain behavioural traits of slurs, an ex-

pressive proposal that is only expressive in terms of the speakers' attitudes is insufficient, so the proposal we seek to develop should also be non-propositional.

Drawing attention to the non-propositional nature of some of our speech acts is nothing new at this point of the debate in philosophy of language. In particular, a parallelism can be drawn here with classical expressivim's distinction between saying and expressing (Ayer 1936/2001). This distinction stems from the idea that not all of our speech acts are descriptive, as some of them express attitudes. But it is also important, from an expressivist point of view, to distinguish between expressing an attitude and saying that one is in it. If we did not enrich the common ground so that it not only contains propositions, those speech acts that express an attitude could only be modelled as conveying a proposition—the proposition that the speaker has the attitude. But then the speech act would not express the attitude, but rather report it. So enriching the common ground is a way of respecting the expressivist idea that expressing an attitude is not to say that one is in it.

The standard way of understanding what the common ground contains has been to maintain that it is constituted by the propositions that are presupposed in the conversation, that is, the set of propositions that are taken for granted by the participants in the conversation and that make their utterances felicitous. The paradigmatic example of how to implement the effect of an utterance on the common ground has been that of assertion. Assertions have traditionally been modelled as a proposal to eliminate from the common ground those worlds that are incompatible with the asserted content.

However, we may think that to maintain that the common ground is only made up of propositions and that for these propositions to be taken for granted is the only requirement for our utterances to be felicitous is a limiting approach, which García-Carpintero (2015) dubs "info-centrism".[2] Positions that reject this idea have developed alternative explanations that do not rely on the incor-

_____

[2] As we saw in chapter 1, Keiser (2022) takes what we are calling "info-centrism" here to be an idealisation, but also a dangerous one in that it leads to theories that are less than ideal in their normative consequences.

poration of a proposition into the common ground to model, for example, the impact of commands or questions. The following are examples of each of these. First, let us imagine four possible worlds in which the following is true:

w_1: I bought a bike and I ate a bag of Gublins.

w_2: I bought a bike and I did not eat a bag of Gublins.

w_3: I did not buy a bike and I did eat a bag of Gublins.

w_4: I did not buy a bike and I did not eat a bag of Gublins.

If Amalia asks me if I bought a bike, the impact this question has on the common ground is to make a partition that distinguishes the worlds in which this occurred, namely w_1 and w_2, and the worlds in which it did not, namely w_3 and w_4. If to this question I answer that I ate a bag of Gublins, my utterance is not felicitous because it does not discard worlds in such a way that all the worlds that remain in the common ground belong to the same side of the partition. If, instead, I said that I did buy a new bike, all the worlds that could be the real world remain on one side of the partition. What is relevant about this type of explanation is that it is a way of modelling the impact of a question on the common ground that does not rely on the incorporation of a proposition of the type *The speaker wants to know if I bought a bike* (Davidson 1979; Lewis 1970).

Imperatives and directives are arguably, together with questions, the most representative example of kinds of utterances whose effect has been implemented in a non-propositional way. Different proposals have been made as to how to model their impact on the common ground. The first proposal we are going to explain is Portner's (2007), which relies on To-Do Lists. To understand the proposal, consider the following example. Víctor has a young child, Pablo. In the common ground shared by Víctor and Pablo, we find shared propositional information such as, for example, *Pablo is safe at home*, *When we put on our pyjamas it is time to sleep*, *León is Pablo's little brother*, etc. In addition to the set of shared propositions, though, the To-Do Lists approach has it that we can also

find one list of tasks corresponding to each participant. On Víctor's side, we have bathing the children, making dinner, etc. On Pablo's side, we have picking up the toys after playing, etc.

Now, Pablo likes to eat a lot, but Víctor is afraid that he will snack too much in the afternoon and not eat a proper dinner. At a certain point in the afternoon, Víctor says "Pablo, don't eat any more cookies". According to the To-Do List approach, this imperative modifies the common ground in such a way that a new item is added to the list of existing tasks that Pablo has in that common ground—stop eating cookies.

Another way of modelling the impact of imperatives is the one proposed by Starr (2020), which is to rank the possible worlds on the common ground according to their preferability. According to this proposal, Víctor's command would impact on the common ground by ordering the possible worlds in it in such a way that those worlds where Pablo stops eating cookies are preferable to those worlds where he does not comply with the command. That is, in this case, the criterion for ordering the worlds is based on whether Pablo fulfills the command.

As noted, info-centrism is a tendency to model the impact of all speech acts as if they were assertions. Marques and García-Carpintero (2020: 142) call "flattening proposals" the positions that try to model the impact of slurs on the common ground as the addition of a proposition. That is, if the effect of a question would be the incorporation into the common ground of a proposition like *The speaker wants to know if I bought a bike*, the effect of a command would be the incorporation into the common ground of a proposition like *The hearer has to pick up the toys*, and the effect of an utterance featuring a slur would be the incorporation into the common ground of a proposition of the form (according to Schlenker, to take an example) *A has the belief that Ns are bad*.

We can attribute to Cepollaro and Stojanovic, whose account I discussed in subsection 4.3.4, the motivation for developing the proposal of evaluative presuppositions that is expressive in the first sense —as having to do with someone's attitudes—, but, in the end, this motivation does not result in a substantive proposal. Firstly, it is not clear what relationship should be established bet-

ween the addition of a moral term in the sentence that is meant to capture the presupposition and the expression of a certain attitude on the part of the speakers. Secondly, the account is developed within a propositional framework, which in the end makes the presuppositions at issue not truly expressive. In the sense favoured here already mentioned in the introduction of this thesis (section 1.4), something has to be non-propositional to be expressive, because otherwise we would be talking about a report of the attitudes at issue rather than the expression of them.

4.5. Non-propositional presuppositional theories

It is the second sense of "expressive"—expressive as non-propositional—that I think it makes sense to exploit from a presuppositional perspective that aims to explain the behaviour of slurs. In this section, I will discuss a couple of theories that are developed around a non-propositional notion of presupposition. First of all, I will present the proposal based on expressive presuppositions developed by Marques and García-Carpintero (2020), who maintain that when we use slurs we presuppose some reactive attitude that we take to be appropriate. Secondly, we have the proposal based on non-propositional presuppositions that holds that when we use slurs we presuppose a certain ordering of possible worlds on the common ground. Both theories are developed in accordance with the second sense of "expressive" that I have distinguished; both treat derogatory content as non-propositional content. Now, I will situate the differences between the two proposals in the way they explain how non-propositional presuppositions impact the common ground. While non-propositional presuppositions modify it by adding reactive attitudes according to the first proposal, they modify it by adding world-orderings in the second one. As we will see, this gives the second proposal an edge over the first insofar as it makes the common ground simpler.

4.5.1. Expressive presuppositions and reactive attitudes

I consider the proposal defended by Marques and García-Carpintero as the first one that develops a presuppositional approach with a fully expressive notion of presupposition in the two senses mentioned above. These authors' particular thesis is that through the utterance of slurs we presuppose a set of reactive attitudes as appropriate. The best way to contextualise this proposal is to refer first to the debate on the extension or enrichment of the common ground. Referring to the enrichment of the common ground will also allow us to show that there is a debate on which to rely in order to develop non-propositional presuppositions. The novel point here is not that the non-propositional presuppositional proposals will make a contribution to the debate on the enrichment of the common ground itself. Therefore, it could be said that the enrichment of the common ground is a transversal discussion to that of non-propositional presuppositions, but it provides a whole theoretical framework from which to state the proposal of expressive presuppositions and NPP.

What Marques and García-Carpintero propose to add to the common ground for modelling the impact of slur utterances is a set of reactive attitudes (Strawson 1962) that the participants in the conversation consider appropriate. Reactive attitudes are emotions and, as such, they consist of two aspects, namely, an intentional or particular object and a formal object. The intentional object is the target of the attitude and the formal object is the attitude that is directed towards that object (Kenny 1963). Reactive attitudes can be seen as a reaction on the part of agents to a certain moral behaviour. A classic example proposed by Strawson is indignation. We react to a certain moral wrong by being indignant. We can also express positive reactive attitudes, such as pride. We can show pride as a reaction to a certain moral behaviour that we consider right. According to Marques and García-Carpintero we have in the common ground a set that includes the attitudes accepted as appropriate by the participants in the conversation, among which may be the attitude of contempt toward a given group (Marques & García-Carpintero 2020: 144).

Let us dwell a little on the decision to specify the reactive attitude that we presuppose in relation to the utterance of slurs as the emotion of contempt. There is a debate about whether contempt can be considered a reactive attitude or not (Mason 2018; Manela 2019: section 4.2). One of the reasons why we might think that considering contempt as a reactive attitude is not quite in line with what a reactive attitude is is that these are emotional reactions that we have in response to the expectations we have regarding the moral behaviour of others. Therefore, it seems necessary that reactive attitudes have a moral character associated with them. It is easy to understand indignation and pride as reactive attitudes because it is difficult for us to imagine cases that are not marked by a moral component, but we can imagine cases of contempt where it does not make sense to attribute a moral character. For example, someone may feel contempt for people who listen to trap music because they consider that they have no taste, but they do not consider that there is anything morally reprehensible about them; it is not a question of these people violating a moral requirement that we consider to be basic. So, if there are reasons why we might refuse to consider contempt as a reactive attitude, we should indicate in what sense Marques and García-Carpintero treat contempt as one of them. Marques and García-Carpintero rely on the position held by Roseman (2018) and Manne (2017), which is characterised by understanding contempt as a morally justifiable emotion, so if we accept these reasons, we could also speak of contempt as a reactive attitude.

Although Marques and García-Carpintero's decision to specify the reactive attitude as contempt is argued, I think that, since we are talking about slurs, talking about indignation would make the position of expressive presuppositions more intuitive. Moreover, the emotion of indignation

helps to explain the motivation behind many cases of slur appropriation.[3] Many cases of appropriation are explained by the fact that the target group has taken the outrage that motivates bigots to use the slur against them and made it their own. But it is not even necessary to go to cases of appropriation to make the reactive attitude concrete in the emotion of indignation rather than contempt. For example, as Manne (2017) has argued, the use of some slurs referring to women is not motivated by contempt for them, but by indignation that their behaviour is not subject to gender norms.

In Marques and García-Carpintero's proposal, for an utterance of a sentence such as "A is an *S*" to be felicitous it must be the case that in the set of reactive attitudes that we find in the common ground the attitude of indignation towards the *N*s is included. If this happens, the sentence will be true if *A* is an *N* and false otherwise. If this is not the case, i.e. if the attitude of indignation towards *N*s is not in the set of reactive attitudes that are considered appropriate, the sentence will have no truth-value. This is not unlike what we say about ordinary presuppositions, namely, the propositions that we assume as shared by the rest of the participants in the conversation.

---

[3] It has been discussed in the literature whether it would not be better to use the terms "reclaimed" or "reappropriated" instead of "appropriated" to refer to theses uses, as "appropriation" can be interpreted as if something has been taken ilegitimately (Zeman 2022: 944, n.3). I will follow Zeman in continuing to use "appropriation" given the lack of consensus on this.

Within appropriated uses, we can distinguish two varieties (Bianchi 2014; Anderson 2018; Jeshion 2020). On the one hand we have what is known as "pride reclamation", and on the other hand we have "insular reclamation". The first of these is the most common way in which appropriated uses take place. It is about positively empashising the belonging to a certain target group. The main objective of this practice is to bring about political and social changes that will result in rights for the discriminated target groups. The process of appropriation of slur, understood in this way, can result in a slur being the politically correct way of referring to the target group, as was the case with the term "queer". The second is motivated by a sense of solidarity and camaraderie among the members of a target group. In contrast to the first use, it is understood that this use of appropriation is only allowed for members belonging to a target group. The main difference between the two senses is based on the objectives to be achieved.

4.5.2. Non-propositional presuppositions and world-orderings

I will do two things in this section. The first is to implement a formal account of the effect of slurs that relies on orderings of the possible worlds in the common ground. The second is to argue that this account provides a better understanding of what is involved in the use of slurs than a treatment based on reactive attitudes.

One way of modelling the impact of slurs is to do so through the ordering of possible worlds. In doing so, I draw on Stanly's (2015) use of Starr (2020) for his own approach to political propaganda. Stanley's proposal is based on the distinction between *at-issue content* and *not-at-issue content* (Potts 2005). At-issue content is that understood as proposed to be added to the common ground. By contrast, not-at-issue content is directly added to the common ground, without any need for the audience to accept any proposal (Stanley 2015: 135). Speakers do not have the opportunity to decide whether to accept not-at-issue content or not, as it becomes part of the common ground without any mediation. One way to see that at-issue content is proposed to be incorporated into the common ground while not-at-issue content directly becomes part of it is through negation (Strawson 1950; Shanon 1976; Karttunen & Peters 1979; Chierchia & McConnell-Ginet 2000). Not-at-issue content is either already part of the common ground or treated as if it was, i.e., it is not what is being *talked about* in this exchange. We will have the opportunity to look at this type of case in more detail in the next chapter. For now, suffice it to say that not-at-issue content is the genus of which presupposed content is a species.

According to Stanley, we can consider political propaganda as relying on not-at-issue content. In particular, Stanley models the impact of propaganda on the common ground on the approach to imperatives developed by Starr (2020). If we recall from the section 4.4, this position models the impact of imperatives by ranking the possible worlds in the common ground according to their preferability. In the example we used, the preferability criterion for ordering the worlds was that, in this

case, Pablo would stop eating cookies. In the case of political propaganda, Stanley maintains that the preferability criterion for the ordering of worlds is related to socialisation: worlds in which the speaker has less contact with the target group are privileged over worlds in which she has more contact.

What I propose, then, is to apply this model to slurs. NPP is thus characterised by its maintaining that when we utter a sentence like "A is an S", what we are doing is presupposing a certain ranking of possible worlds where the worlds in which the speaker has less relation with the Ns, the Ns being the neutral counterpart of S, are above the worlds where she socialises more with the derogated group. According to this proposal, an utterance of a sentence like "A is an S" will be felicitous if the common ground includes an ordering of possible worlds where the worlds with which the speaker has less contact with the Ns are placed above those with which she has more contact. Given this ranking, the sentence "A is an S" will be true if A is an Ns and false otherwise. If there is no such ranking in the common ground, the worlds will be rearranged in such a way to accommodate the slur so that the utterance of "A is an S" is felicitous. Thus, when we accommodate an utterance featuring a slur, the common ground is modified in such a way that those worlds where the speaker socialises less with the derogated group are placed above the worlds where socialisation is greater.

The main reason for understanding the effect of slurs in terms of world-orderings rather than reactive attitudes is that it results in a more parsimonious ontology. World-orderings also serve to model the effect of directives, while we know of no speech act, other than utterances of sentences featuring slurs, whose effect we can model in terms of reactive attitudes. By modelling the effect of slurs in terms of world-orderings, we subsume these speech acts under the more general class of directives, even if the associated command is made in a back-handed way. Moreover, enriching the set of operations that utterances can perform on the common ground instead of enriching the common ground itself may be seen as a generally preferable move.

Additionally, NPP has the virtue, in comparison with the reactive attitudes proposal, of being compatible with a conception of the intervention in response to utterances of slurs that combines both structural (e.g., Banks & Ford 2009; Anderson 2010; Haslanger 2015; Ayala 2016, 2018; Mallon ms) and psychological elements (Brownstein et al. 2019). This makes NPP a more parsimonious proposal also with regard to the kind of intervention it recommends, and not just how to model the impact of derogatory effect. By intervention I mean what kind of treatment, in a practical sense, we can give to deal with the consequences of slurs. One way to understand how reactive attitudes work is to consider that speakers share them, at least for the purposes of the conversation. That is, accommodation would only be possible if the reactive attitude (in the case of slurs, an attitude of indignation) is shared by the participants in the conversation. This can be interpreted as if the derogatory effect of slurs depends on the particular implicit attitudes of the speakers. It therefore seems that this proposal is committed to a psychologistic perspective, and that all we can hope to do is to modify the speaker's attitudes. But intervention does not necessarily involve specific measures for each situation; it can have to do with how we can intervene in the context in general so that, ideally, fewer and fewer slurs are uttered or, if the damage is already done, how best to understand the consequences of slur use. Having said that, it is not within the scope of this work to propose specific intervention measures in response to utterances of slurs. However, I consider it an advantage of NPP its not being committed to one type of intervention from the outset and, therefore, to be able to integrate other elements.

Most of the theoretical debate on slurs has focused on studying the intricacies of their semantic behaviour; we have only a few exceptions that reflect on what the consequences of utterances of slurs are are and what can be done about them (see e.g. Corredor 2014; Kukla 2018; Herbert ms). One possible explanation as to why we do not have a discussion regarding intervention in response to utterances of slurs may be the perception we have that the tools that philosophy of language provides us with are not adequate when it comes to this. They may be insufficient for us to say anyt-

hing relevant about what to do about the harm caused by the utterance of a slur, or we may feel that there is nothing to say because it is not within the discipline's competence. This is not the point of my complaint. It seems to me that our contributions to the theoretical debate on slurs, which is clearly marked by the kind of dangerous behaviour they exhibit, owe it to the more general debate on slurs what can be done about the consequences of their use. Therefore, let this serve as an invitation to think of ways in which to understand the consequences of derogation and their possible redress. As stated (section 1.3), I conceive slurs as a linguistic phenomenon that is linked to our social and political practices. In particular, we understand situations where slurs are uttered as situations where harm is caused by the derogatory nature of slurs. If we believe that derogating someone is tantamount to generating aggression at the linguistic level, we also believe that there should be mechanisms that allow us to intervene in these situations.

4.6. Conclusion

The aim of this chapter has ultimately been to present NPP, the proposal around which this thesis revolves. To do so, I have first introduced the notion of presupposition in more detail than I did in the previous chapter and surveyed a few accounts that rely on presuppositions as they have traditionally been conceived, i.e., as propositional. While doing so, I have discussed some reasons why we might go for a presuppositional account, leaving the most substantial reasons for the next chapter. In particular, I have talked about the advantage that presuppositional theories have in not having to identify a slur's contribution to the truth-conditions of the sentences in which it appears with that of its neutral counterpart. I have then discussed a limitation of propositional versions of the presuppositional view: the fact that they do not flesh out what makes expressive presuppositions different from ordinary ones. To overcome this limitation, two versions of a non-propositional presuppositional account have been offered: Marques and García-Carpintero's, which relies on reactive attitudes, and NPP, which relies on world-orderings. The chapter has ended up by discussing some advanta-

ges that I think NPP has over Marques and García-Carpintero's view. In particular, the explanation in terms of ordering of possible worlds was more ontologically parsimonious and remained neutral with respect to the kind of intervention required by this kind of case. These advantages I believe are what brings NPP closer to an ideal proposal in the second sense distinguished by Keiser (2022: 4) and to which I made reference in section 1.4. In the first sense, theories are ideal insofar as they are idealised —insofar as they distort some features of the phenomenon they seek to explain. In the second sense, ideal theories are those that manage to live up some normative standard. In this sense, theories that are ideal are necessarily good, since what we value are the normative and social consequences that we can attribute to the theory.

Chapter 5

Slurs' resistance to cancellation, rejection and retraction

5.1. Introduction

In chapter 2 I discussed the relationship between non-displaceability and the projection profile of slurs. The latter, which is derived from the former, tells us how slurs behave in certain linguistic environments. We have seen that slurs tend to project into linguistic environments where other terms do not (subsection 3.2.2). What we are going to explore in this chapter is why this happens. To do so, we will test NPP in relation to three specific linguistic phenomena, namely resistance to cancellation, rejection and retraction. Let us recall the concrete theses that made up NPP:

(i)   Slurs have derogatory content, but this content is not propositional.

(ii)   Derogatory content is located at the presuppositional level.

(iii) The impact on the common ground corresponding to the non-propositional presuppositions slurs trigger should be modelled in terms of the ordering of possible worlds.

The way in which I will test these theses is by showing that a non-propositional presuppositional explanation is the best way to account for slurs' resistance to cancellation, rejection and retraction. The first thing we can do in order to explain the particularities of slurs with respect to cancellation, rejection and retraction is to compare what happens when we block the effect of any speech act and when we try to block the derogatory effect of slurs. What we can observe is that the strategies we use in the first case do not work in the second. Consider the following examples to see this intuitively.

(1) The cat is on top of the piano.

(2) A is an *S*.

(3) If the cat is on top of the piano, we should force it to come down before practising.

(4) If A is an *S*, she shouldn't go to that restaurant.

(5) No, it's not.

(6) No, she's not.

(7) I take that back. (Targeting one's previous utterance of (1).)

(8) I take that back. (Targeting one's previous utterance of (2).)

Cases (1) and (2) are cases of assertion. In (3) and (4) we have the sentences in (1) and (2) embedded in a complex context, viz. a conditional. In sentence (3) we should not say that the speaker is committed to the proposition that the cat is on the top of the piano; however, it seems that in (4), despite being a conditional as well, the derogatory effect of the slur persists. The difficulty of blocking the derogatory effect of slurs in a case like (4) is what I mean when I talk about slurs' resistance to cancellation. Cases (5) and (6) are responses to (1) and (2), respectively, which aim to reject the proposition expressed by each of them. But, while (5) serves to reject the proposition that cat is on the piano, one cannot successfully reject the derogatory effect of (2) by uttering (6). This is what I mean when I talk about the resistance of slurs to be rejected. Finally, with (7) and (8) what we are trying to do is to take back the effect of (1) and (2). Again, this seems feasible with respect to (1) but not with respect to (2), and this is what I mean when I say that slurs present resistance to retraction.

The structure of this chapter is quite simple. In section 5.2, I give a first hint at what it means for the effect of slurs to be resistant by depicting the way in which the effect of ordinary assertions is not. After this, sections 5.3–5.5 are devoted to three varieties of resistance exhibited by slurs: resistance to cancellation (section 5.3), resistance to rejection (section 5.4) and resistance to retraction (section 5.5). In each of these sections, I first present the phenomenon at issue and then explain how NPP manages to account for it. As we will be able to observe, my position will become independent

from the explanation based on expressive presuppositions offered by Marques and García-Carpinte-ro (2020) as we approach slurs' resistance to retraction.

5.2. The resistant effect of slurs

In this section, I describe the peculiar behaviour of slurs I've made reference to. Let's take a closer look at how we might block the effect of an assertion through cancellation, rejection and retraction. I take as a paradigmatic example an assertion such as (1):

(1) The cat is on top of the piano.

As indicated in the previous chapter, the effect of an assertion is to add the proposition expressed to the common ground shared by the participants in the conversation.This move can be divided into two different phases. First, (a) the speaker utters sentence (1), and as this happens, the proposition that the cat is on the top of the piano is proposed to be added to the common ground. Secondly, (b) the proposition will be added to the common ground if the speaker's proposal is accepted by her audience. There are a number of places where communication can go awry. Correspondingly, there are three ways in which I could utter the words in (1) without them having their default effect. First, something could go wrong before even the first step takes place. If the words are uttered in a linguistic environment that cancels their effect, the proposal to add to the common ground the proposition that the cat is on the mat will not even be made. This is what happens, for instance, when the words appear in the antecedent of a conditional, as they do in

(3) If the cat is on top of the piano, we should force it to come down before practising.

We summarise this by saying that the effect of an assertion is cancelable. Second, something could go wrong between the first and the second step. In particular, it could be that my audience dos not accept my proposal. This is what happens when they reply to my assertion with something like (5) "No, it's not". We summarise this by saying that the effect of an assertion is rejectable. Lastly,

something could go wrong after the second step. This is what happens when, after the proposition that the cat is on the mat has gone into the common ground, I say something like (7) "I take that back". If this later assertion is accepted by my audience, the proposition will be taken out of the common ground. We summarise this by saying that the effect of an assertion is retractable. The backbone of this chapter, then, is to explain why ordinary assertions are susceptible to these moves but when it comes to the case of slurs, by contrast, we cannot block the derogatory effect that they have.

It must be observed that although I speak of these three phenomena together in order to talk about the resistance of the derogatory effect of slurs, they are different linguistic acts. Cancellation and retraction have in common that they are carried out by the speaker. In contrast, rejection is something that takes place in the linguistic exchange; it is the listener who does it. Specifically, for resistance to cancellation, we find in the literature a line of debate that has been guided by the works of Jeshion (2013a), Sennet and Copp (2015), Cepollaro and Stojanovic (2016), Camp (2018a) and Marques and García-Carpintero (2020). Furthermore, this linguistic act has arguably received increased attention because it has been used to motivate presuppositional explanations (Lasershon 2007; Cepollaro & Stojanovic 2016; Marques & García-Carpintero 2020; see subsection 4.3.2). The account of slurs' resistance to rejection I will elaborate will be based to some extent on the proposals defended by Cepollaro and Stojanovic (2016) and Marques and García-Carpintero (2020), although the authors of the latter proposal do not deal specifically with this issue. Other notable works in this respect are the proposals defended by Camp (2013) and Jeshion (2013a, 2013b). Finally, regarding retraction, it could be said that the explanation I develop building from Marques and García-Carpintero (2020)'s proposal adds a new perspective to address this feature of slurs, alt-

hough we can find references in the literature that have addressed the issue, such as McGowan (2009) and Simpson (2013).[1]

5.3. Resistance to cancellation

5.3.1. The phenomenon

As has just been noted, it has been considered that derogatory content behaves with respect to projection in much the same way as presuppositions do (Lasershon 2007; Cepollaro & Stojanovic 2016; Marques & García-Carpintero 2020). The fact that this similarity has been considered more or less relevant to the development of a presuppositional proposition is only important here for the explanation it may offer to us of slurs' resistance of cancellation.

As we know, the phenomenon of projection occurs when the sentence we obtain by embedding a sentence in a complex context, such as (3), triggers the same presuppositions as the original sentence. Consider the following examples:

(9)   Marga has stopped eating meat.

(10) If Marga has stopped eating meat, then we have to adapt the original lasagne recipe.

(11) Marga has not stopped eating meat.

(12) Has Marga stopped eating meat?

As we can see, (9) triggers the presupposition *that Marga used to eat meat*, and this also occurs in (10)–(12). The presupposition triggered by (9) survives when the sentence is in the antecedent of a conditional as in (10), negated as in (11) or put in the form of a question as in (12). Now, let's compare how derogatory content behaves in this series of sentences containing a slur:

---

[1] Caponetto (2020) advances a different taxonomy of ways of blocking the effect of a speech act. She discusses retraction like I do (Caponetto 2020: 2407–2410) and her *amendment* (Caponetto 2020: 2410-2412) bears a (passing) resemblance with the cancellation, but instead of rejection she considers annulment (Caponetto 2020: 2404-2406), which seems a wholly different phenomenon.

(1)   The cat is on top of the piano.

(4)   If A is an *S,* she shouldn't go to that restaurant.

(6)   No, she's not.

(13) Is A an *S*?

We could say that (2) presupposes that the worlds in the common ground are ranked so that they are more preferable the less contact the speaker has with the *N*s in them, and this presupposition is also projected in (4), (6) and (13), i.e. in the same complex context in which the presupposition triggered by (1) does. But there are more linguistic tests that we can use to find out whether a given content actually behaves as a presupposition. In particular, the binding test (van der Sandt 1992), introduced in subsection 3.2.2, allows us to check whether a given content behaves as a presupposition by analysing its projection profile. Let's look at the following examples:

(14) If Marga used to eat meat, then she has stopped eating meat.

(15) If *N*s are bad because they are *N*s, A is a *S*.

These cases are composed as follows. We place in the antecedent of the conditional a sentence whose content is the presupposition triggered by the original sentence; in (14), the antecedent would be a sentence expressing the content presupposed by (9), and in (15), the antecedent would be a sentence that expresses the content of the presupposition triggered by (2), at least as Cepollaro and Stojanovic (2016) understand it. And in the consequent of the conditional, we place the original sentence, i.e., (2) and (9) respectively. This is how the binding test is applied.

However, although (14) and (15) have the same structure, they behave differently. An ordinary presupposition, as is the case in (14), should not project and, in fact, does not. Still, the derogatory effect of (2) remains in (15). Therefore, what this test tells us is that the derogatory effect of slurs does not behave like a presupposition, since this effect still projects in cases where presuppositions do not. This particular behaviour of the derogatory effect is known as the hyper-projectivity of

116

slurs (Jeshion 2013a; Sennet & Copp 2015; Cepollaro & Stojanovic 2016; Camp 2018a; Marques & García-Carpintero 2020). This objection, i.e., the problem of hyper-projectivity, and resistance to cancellation are connected. Resistance to cancellation is an instance of the hyper-projectivity of slurs. The expressive content of slurs scopes out negation, as it does with other operators. Resistance to cancellation must thus be distinguished from the feature of non-displaceability (section 2.4). Resistance to cancellation means that, if we are in a context where a slur derogates, usual linguistic devices to express distance with respect to the content will not be successful to separate ourselves from the derogatory content, because it will always derogate. This is nevertheless compatible with there being non-derogatory uses of slurs. If we take slurs to be non-displaceable, by contrast, we take them to derogate in all possible contexts. Since resistance to cancellation is implied by the hyper-protective profile of slurs, resistance to cancellation is also derived from non-displaceability.

5.3.2. The account

As mentioned above, there have been presuppositional proposals, such as the theory put forward by Cepollaro and Stojanovic (2016) or the theory held by Marques and García-Carpintero (2020), which do adopt the similarity of behaviour between derogatory content and presuppositional content as the main reason for the development of a presuppositional explanation of slurs. So, what I will do next is present the answers regarding hyper-projectivity offered by these two proposals and argue that Marques and García-Carpintero's answer is the most complete explanation of slurs' resistance to cancellation.

What Cepollaro and Stojanovic (2016) propose is to make a comparison between the behaviour of the derogatory effect of slurs in a case like (15) and a case involving a gendered personal pronoun (Heim 2008). Consider the following examples:

(16) If *N*s are bad because they are *N*s, A is an *S*.

(17) If Julia is a woman, then she can't do jiu-jitsu.

(17) shares a relevant features with (14) and (15), since we have placed in the antecedent of the conditional the presupposition that its consequent expresses, which is *that Julia is a woman*. This presupposition is triggered by the pronoun "she" appearing in the consequent. What we observe is that (15) and (17) have behaviour similarity with respect to projection, since the fact that we place sentences expressing the presuppositions in the antecedent of the conditional does not prevent these presuppositions from projecting. What these authors argue, therefore, is that if we have no difficulty understanding gendered personal pronouns as triggering presuppositions, neither should we have any difficulty in understanding the derogatory effect of slurs in presuppositional terms. This explanation revolves around the idea that slurs are not the only expressions that exhibit hyper-projectivity. To reinforce this response to the objection, Cepollaro and Stojanovic rely on the distinction between soft and hard triggers made by Abusch (2002) and Abbott (2006). Soft triggers give rise to cancellable presuppositions, while hard triggers are triggers like the ones occurring in (15) and (16), which means that the presuppositions they give rise to are non-cancellable. According to Cepollaro and Stojanovic (2016: 480), this occurs because soft triggers give rise to pragmatic presuppositions while hard triggers give rise to semantic presuppositions.

However, Cepollaro and Stojanovic do not pursue an explanation for why slurs and gendered pronouns are hard rather than soft triggers. Besides, their explanation is based on the distinction between two types of presuppositions and, as it said in the previous chapter (section 4.2), I see no reason to maintain such a sharp distinction between semantic presuppositions and pragmatic presuppositions. I find Marques and García-Carpintero (2020: 149–150) closer to offering a properly explanatory account. Their point is that, when we place (2) in a context such as (17), we are not really applying the binding test to it. For the binding test to be correctly applied, the antecedent of (17) should have as its content the presupposition in (2) (see McCready 2010: 9). But, as noted in

the previous chapter (subsection 4.5.1), this presupposition is not propositional (it is a reactive attitude that is presupposed), while the antecedent of (17) will necessarily have propositional form[2]—otherwise, (17) will be ungrammatical. Thus, no matter what we place in the antecedent of (17), it will not have the presupposition in (2) as its content. It should not worry us, therefore, if no sentence with the form of (17) is free of derogatory character. No such sentence will be the result of applying the binding test. Thus, we cannot say that slurs fail the binding test just because the presupposition that *N*s are bad because of being *N*s projects in (17).

Despite appearances, though, Cepollaro and Stojanovic's and Marques and García-Carpintero's explanations can be said to be complementary rather than incompatible. We could say that slurs are hard triggers because the presuppositions they trigger are non-propositional. Moreover, Marques and García-Carpintero's account can also give us a hint as to why gendered pronouns are hard triggers too. Rather than propositional presuppositions, gendered pronouns can be taken to trigger presuppositions concerning the assignment of values to variables—they pose certain restrictions on this assignment for utterances of sentences containing such pronouns to be felicitous. Since we can only place sentences with propositional form in the antecedent of a conditional like (17), no conditional will allow us to apply the binding test to a sentence featuring a gendered pronoun. As I pointed out in the introduction to this chapter, NPP will distance itself from the proposal based on expressive presuppositions as we approach resistance to retraction. But, as far as resistance to cancellation is concerned, NPP is fully in line with the explanation offered by Marques and García-Carpintero (2020).

Theory	Explanation	Advantage	Disadvantage
Evaluative presuppositions	Comparison of slurs with gender pronouns as hard triggers.	Predictive power.	No explanatory power.

---

2 This is related to Potts' (2007, 176–9) discussion of the *ineffability* of the meaning of expressives, a feature that was discussed in section 2.4.

Theory	Explanation	Advantage	Disadvantage
Expressive presuppositions	We cannot pick up a reactive attitude in a propositional way.	Both predictive and explanatory power.	Less ontologically parsimonious, commitment to individualist intervention measures.
NPP	We cannot pick up a world-ordering in a propositional way.	Both predictive and explanatory power, as well as ontological parsimony and compatibility to different intervention measures.	-

5.4. Resistance to rejection

5.4.1. The phenomenon

The derogatory effect of slurs is not only hard to cancel. It is also hard to reject. In this section, I flesh out what resistance to rejection amounts to and show how a non-propositional, presuppositional account of slurs like NPP can make sense of this resistance.

Remember (2):

(2) A is an *S*.

An utterance of (2) could be replied by saying

(6) No, she is not.

However, we cannot be sure that, by doing so, it is the derogatory effect of the utterance that we are rejecting. A natural interpretation of (6) is just that A is not an *N*, but this is compatible with our sharing whatever negative and discriminatory attitude toward *N*s the speaker of (2) holds (see Stanley 2015: 135-136). Arguably, by uttering (6) we only point out the factual error (section 2.2) that it seems we can make in uttering a slur when we "get the target group wrong". Something else is needed if we want to make sure that we are understood as rejecting this attitude; if we want to make sure that we are understood as rejecting the obvious moral error that is also committed in uttering the slur.

This fact is smoothly accommodated by a presuppositional account like NPP. When we say something like (6), we are targeting the content asserted by the speaker of (2). If the derogatory effect of this sentence is explained through its presuppositions, it is then natural that replies like (6) do not manage to block that effect, as it is not through replies like this that we target presuppositions. This is why (6) does not block the derogatory effect of (2) and, moreover, why the presuppositional account helps to detect why a response like (6) does not work.

Presuppositions are usually assumed to be identifiable through the "Hey, wait a minute!" test (von Fintel 2004), which provides us with a way of rejecting them. Let us recall the following case:

(9) Marga has stopped eating meat.

If the interlocutor, for whatever reason, does not share the assumption *that Marga used to eat meat*, she can respond to (9) in this way:

(18) Hey, wait a minute! I didn't know Marga used to eat meat.

One could say that the interlocutor has thus rejected the presupposition of (9). But let us now try applying the same test to (2) using Cepollaro and Stojanovic's (2016) version of the presupposition triggered by this utterance:

(2) A is an *S*.

(19) ?? Hey, wait a minute! I didn't know *N*s are bad because of being *N*s.

The result of applying the "Hey, wait a minute! I didn't know that…" test, in this case, does not seem as natural as in the previous case. Again, what I will do below is to present both the response offered by Cepollaro and Stojanovic (2016) and the response that we can develop from the theory of expressive presuppositions held by Marques and García-Carpintero (2020).

Cepollaro and Stojanovic (2016: 467) point out that the "Hey, wait a minute! I didn't know that…" move, which as we saw in the previous subsection is used to reject the presupposed content, only makes sense for descriptive presuppositions, such as the one triggered by (9), but for evaluative presuppositions, such as the one triggered by (2), we do not get good results. It seems to me, however, that, by saying this, Cepollaro and Stojanovic are again describing the behaviour of evaluative presuppositions rather than explaining why they behave in this way. I think that we can give an explanation along the same lines as the one Marques and García-Carpintero (2020) gave for the projective behaviour of slurs. Note, first, that, since "to know" is a propositional attitude verb, "I didn't know that" requires a propositional complement. Thus, no matter what we place after "I didn't know that", it will not have the presupposition in (2), which is not propositional, as its content. Replies like (19) will therefore always fail to target this presupposition.

But also, propositional presuppositions are meant to model the information shared by participants in a conversation as to how the world is. They serve to describe the stage at which participants are in a process of inquiry, which is aimed at knowing which of the candidate worlds we inhabit. In this sense, propositional presuppositions carry locational information (Lewis 1979; Charlow 2014: 639). Non-propositional presuppositions are nothing like this. Suppose we take these presuppositions to concern the participants' reactive attitudes, as Marques and García-Carpintero do. In what kind of process do they play a role? Coordination is a natural candidate: we could say that part of the aim of a conversation is for speakers to align their reactive attitudes (cf. Gibbard 1990: 110), and expressive presuppositions allow us to characterise the point they are at. However, it is far from settled that this is the role that reactive attitudes play in conversations (Pérez Carballo & Santorio 2016). At any rate, we can be quite certain that, whatever process we should be talking about, it is not a process of inquiry. Speakers who engage in non-propositional communication do not aim at

discovering what the appropriate reactive attitudes are. In other words, the information carried by non-propositional presuppositions is not locational, but *orientational*. This explains that utterances of sentences like (2) cannot be challenged with replies like (19). If instead of reactive attitudes, we use world-orderings to model the effect of slurs, we will reach a similar conclusion: these presuppositions do not play a role in a process of finding out how worlds are actually ordered.

Cepollaro and Stojanovic maintain that speakers do not opt for this type of rejection for evaluative presuppositions and make use of metalinguistic responses such as the following (Cepollaro & Stojanovic 2016: 466):

(2) A is an *S*.

(20) Hey, wait a minute! You shouldn't talk about *N*s like that.

According to these authors, (20) would be a more natural response to the utterance of a slur, although they point out that speakers are motivated to give this kind of response in pedagogical contexts mostly. I agree that replying to (2) with (20) is better than replying to it with (19), but it also comes at a cost—it interrupts the flow of the conversation (Cepollaro & Stojanovic 2016: 468). An answer like (20) diverts the topic of conversation, as we are no longer talking about whether A is an *S* or not but about what words we should use. This is what makes (20) an answer of a metalinguistic nature. Now, the fact that this answer is the only one that makes sense to contemplate in the face of (2) suggests that we cannot explicitly point out the presupposition triggered by (2), and a plausible explanation for this is this presupposition's non-propositionality.

Now, in section 4.2, we looked at the following example in order to explain what the semantic presuppositions consisted of:

(21) The shop around the corner specialises in piano sheet music, you might find what you are looking for.

(22) There is a shop on the corner.

As noted in that section, if the second sentence were false, the first sentence would have no truth-value. Strawson (1952) and Yablo (2006), however, maintain that the fact that the second sentence is false does not render the first sentence meaningless or necessarily lose its truth-value. Yablo takes these failures to give rise to *non-catastrophic* presuppositions (Yablo 2006: 164). Using Strawson's (1952: 225) example, let us imagine a linguistic exchange between A and B. A, who wants to put pressure on B to buy something, says "The neighbour who lives on the corner has bought it too". B, who realises what A intends to do, says "That's a lie, nobody lives in the house on the corner". This would be a paradigm case of non-catastrophic presupposition because rejecting the presupposition triggered by what A says, namely *that there is a neighbour in the house on the corner*, does not make her utterance infelicitous: it still conveys a truth-apt content, albeit a false one. The existence of non-catastrophic presuppositions, besides posing a problem for a strictly semantic view of presuppositions, indicate that it is possible to identify and reject a presupposition. Still, no parallel to this example seems at hand in the case of slurs. What is it then that makes it possible to identify and reject the presupposition in Strawson's example while, in (6), the most we can do is to deny the asserted content? The answer, again, seems to me to be due to the non-propositional nature of derogatory content. We cannot construct such an example by means of a slur.

I take these considerations to explain why it is so hard to reject the derogatory effect of a slur. Direct negation, as in (6), does not target the presupposition, but standard ways of rejecting presuppositions, as in (19), do not seem to work either. The fact that the derogatory effect of slurs is explained through presuppositions and that these presuppositions are non-propositional is what is needed to account for slurs' resistance to rejection.

Theory	Explanation	Advantage	Disadvantage
Evaluative presuppositions	Distinction between descriptive presuppositions and evaluative presuppositions.	Predictive power.	No explanatory power.

Theory	Explanation	Advantage	Disadvantage
Expressive presuppositions	We cannot pick up a reactive attitude in a propositional way.	Both predictive and explanatory power.	Less ontologically parsimonious, commitment to individualist intervention measures.
NPP	We cannot pick up a world-ordering in a propositional way.	Both predictive and explanatory power, as well as ontological parsimony and compatibility to different intervention measures.	-

5.5. Resistance to retraction

We have finally arrived at the last and least discussed variety of resistance exhibited by slurs, which we should also be able to account for from a proposal such as NPP. The phenomenon of retraction is a widely discussed phenomenon in the debate about evaluative language (von Fintel & Gillies 2008; Ross & Schroeder 2013; MacFarlane 2014; Marques 2014, 2018). Retraction is the mechanism whereby a speaker says something like "I retract that" or "I take that back" (MacFarlane 2014: 108). By doing so, according to MacFarlane, the speaker cancels the normative effect of the speech act to which "that" refers, which is the target of the speech act of retraction. For example, by retracting a proposal we withdraw a commitment we had previously acquired to contribute to carrying out that proposal. Let's imagine that we are gathered with our friends. At a certain point, I say "Shall we watch the last episode of *Atlanta*?" and one of them replies "It's not out yet". Given this information, I then say "I take that back" in order to remove the normative effects of my proposal, which exempts me from doing my best to watch the episode instead of, for instance, actually boycotting the plan.

In the case of utterances featuring slurs, the normative effects associated with such words seem hard to cancel in this way. In other words, you cannot "unring a bell" once it has been rung (see McGowan 2009: 403; Simpson 2013: 570). Let us look at an example:

(2)   A is an *S*.

(23) You shouldn't talk about *N*s like that.

(24) Oh, I'm sorry, I take that back.[3]

Has the derogatory effect produced by (2) disappeared? It does not seem like it has. A presuppositional account like the one I have advanced in this chapter can explain this fact along the lines drawn by the explanation of slurs' resistance to rejection. When I say "I take that back", what I am taking back is the asserted content, not the presupposed content.[4] The expression "I take that back" is naturally read as elliptical for "I take what I said back", not for "I take what I presupposed back", which would be a philosophical construct. If we say "I take that back" trying to retract a presupposition, all we will do is retract the assertion that triggered that proposition. Now, if the derogatory content of slurs is presupposed, this would explain why we cannot undo the damage associated with the utterance of a slur.

Note, however, that a sufficiently rich context might allow an utterance of "I take that back" to target the presupposition triggered by a previous utterance instead of the content asserted through it. Consider the following dialogue:

(25) Julia has stopped going to jiu-jitsu classes.

(26) In fact, she signed up for them but didn't get to attend a single class.

(27) Oh, is that so? I take that back then.

Given (26), it is natural to think that (27) does not target the content asserted by (25), but the presupposition that Julia used to go to jiu-jitsu classes. By uttering (27), the speaker is not introducing

---

[3] I think this is a representative case where not mentioning slur makes it difficult (but not impossible) for us to have the right linguistic intuitions to understand the phenomenon but in section 1.4 I explained the reasons for not mentioning any slur.

[4] We could also say that the retraction targets the speech act itself. In this case, however, the speech act would be the assertion, and not the presupposition. It is in fact awkward to talk about presuppositions as acts that the speaker carries out. Thanks to Claudia Picazo for suggesting this alternative.

back into the common ground worlds in which Julia has kept going to jiu-jitsu classes, but worlds in which she has never even gone to them. These worlds had been removed from the common ground by the speaker's utterance of (25) followed by a process of accommodation.

It is not clear, though, that this mechanism can be used to retract a non-propositional presupposition like the ones I am interested in. So, again, when we move to the field of expressive content we get the following. Consider this dialogue, built to be structurally parallel to the one presented above:

(2)   A is an *S*.

(28) In fact, *N*s are valuable members of our society.

(29) ?? Oh, is that so? I take that back then.

(29) feels weird if intended to target the content asserted by (2)—why would the fact that *N*s are valuable members of our society speak against the claim that A is an *N*? But, if (29) is intended to target the presupposition triggered by (2), it is hard to see how it could achieve its purpose. *N*s will have been derogated regardless of one's efforts to "go back in time".

Up to this point, I could be accused of what, in the previous two sections, I accused Cepollaro and Stojanovic of. I have described *how* slurs behave with respect to retraction, but I have not explained *why* they do so. As with rejection, however, I think that the fact that the presuppositions triggered by slurs are non-propositional allows us to account for the way in which they differ from other presuppositions when it comes to retraction. What makes the context of the first dialogue allow (27) to target the presupposition of (25) is that, in (26), the interlocutor provides the speaker of (25) with certain information. Cooperative principles then obligate the speaker to reintroduce into the common ground the worlds excluded through presuppositions accommodation. There is no information that the interlocutor of the second dialogue could offer, though, the automatically forced the speaker of (2) to cancel the non-propositional presupposition triggered by her utterance.

Anything the interlocutor says is strictly compatible with the speaker of (2)'s negative evaluation of *N*s. Of course, knowing certain facts may drive the speaker to reconsider her position. But, unlike in the first dialogue, refusing to do so does not make her uncooperative.

What if, instead of (29), the interlocutor replies denying precisely the presupposition triggered by (2)? This would require this presupposition to be made explicit in propositional form. But, as we saw when discussing slurs' resistance to cancellation, this is not possible, as the presupposition triggered by (2) is not propositional. Thus, presupposition and non-propositionality together allow us to account for the difficulty of retracting a slur.

So far, I have dealt with slurs' resistance to retraction, but I have not discussed why, as noted at the beginning of this section, we should aim at accounting for this kind of resistance. Let us pause, then, to think a little more about the relationship between slurs and retraction. I believe that there are two directions that can be explored in relation to the possibility of retracting a slur, and that both stem from the same motivation: since we have committed a grievance by uttering a slur, we could use retraction to undo the harm we have done. Now, on the one hand, we can ask ourselves whether we have the ability, whether we have the linguistic tools, to retract the utterance of a slur. On the other hand, we can think about whether we are morally obliged to do so.

Regarding the first question, the answer given here is no: we cannot retract the utterance of a slur. In relation to the second question, although I do not have a clear position on this issue and the answer is beyond the scope of this chapter, I am inclined to think that, in the context of public communication, we are morally obliged to try. In other words, we know that we will not be able to undo the damage we have done but we have an obligation to try so that, even if it does not undo the expression of the attitude, we at least apologise for it.

It might seem that there is a tension between maintaining that we *must* try to retract an utterance that has been offensive and maintaining that we *cannot* do so. If all we are going to get by attempting a retraction is an apology, it might seem better to make the apology directly. But I think

there is a subtle difference between trying to retract and apologising. By attempting to retract, there is an acknowledgement of exactly what has been wrong with our utterance. In any case, I do not think that this tension between trying to do something and not being able to do it is a bad thing in this case. I believe that retraction in public contexts in relation to hate speech in general, and slurs in particular, is based precisely on the recognition that we are expressing a certain discriminatory attitude that we will not be able to undo. I think there is indeed a tragic point in holding this position, but it need not necessarily be terrible.

There is a kinder reading about having to retract for being offensive and not being able to do so, and that is that it allows us, if we are willing, to get to know ourselves a little better. For example, it could be the case that we did not know that we were willing to use a certain slur. However, this gentler reading runs the risk that behaving in a morally reprehensible way could give us the opportunity to achieve something good. This interpretation is not only problematic because of the goal we can supposedly achieve, but also because we may find the internal moral exercise that the bigots must go through irrelevant. So, having put all these cards on the table, I conclude the following. In the context of public communication, we have a duty to try to retract even if we are unable to do so and, moreover, we most likely do not have the right to make explicit the lesson we may have learned from it.

5.6. Conclusions

Following Cepollaro and Stojanovic (2016) and Marques and García-Carpintero (2020), I have offered an account of the meaning of slurs that, by relying on presuppositions, explains much of their behaviour, and explains the behaviour that ordinary presuppositions leave unexplained by understanding the presuppositions triggered by slurs as non-propositional. As noted in the previous chapter, I believe that a non-propositional, presuppositional account of slurs, such as NPP, allows us to explain the resistance to these three linguistic mechanisms. Being able to do this is proof that NPP

works by assuming the theses I listed at the beginning of this chapter. Maintaining that NPP works is the best way to account for the behaviour of the derogatory content and thus also the derogatory effect. Its elusive character allows it to hide, it resists being pointed out, a previous move that is needed whenever we want to block it.

Chapter 6

Beyond the conversation: The pervasive danger of slurs

6.1. Introduction

In the previous chapter we discussed the projection profile of slurs, which is the feature whereby the derogatory effect of slurs can be found in every linguistic environment. However, derogation occurs no matter what words surround the slur, but also no matter who the speaker or their audience are or what they know about each other. It could be said, then, that in the previous chapter we focused on why slurs derogate in all linguistic environments and in this chapter we will focus on why slurs derogate in all possible linguistic contexts. All this, as has been pointed out, has its origin in the feature of non-displaceability. The analysis of this feature in chapter 2 led me to hypothesis that there are no non-derogatory contexts for slurs.

In (Moreno & Pérez-Navarro 2021b) we argued that factors such as the speaker or their audience might make an occurrence of a slur non-derogatory. However, my intuitions have changed regarding the possibility of finding contexts in which this happens. The change of position that has led me to maintain in this thesis that all occurrences of slurs are derogatory is motivated by the following reasons. First of all, a natural consequence of understanding non-displaceability as I defined it in chapter 2 is that we cannot find non-derogatory occurrences of slurs, and there is empirical evidence (Cepollaro et al. 2019) that occurrences of slurs are derogatory even if their derogatory force can vary depending on whether they appear as part of reported speech or in an isolated utterance. What can change from one context to another is the derogatory force of slurs, given that there is always some. In this chapter, I will explore the differences between contexts that lie behind this variability, but ultimately conclude that they are not enough to make any occurrence of a slur completely non-derogatory. Secondly, it seems to me that a relationship can be established between the existence of neutral counterparts and the claim that occurrences of slurs are always derogatory. As

noted, I understand neutral counterparts as terms that are not truly neutral because they are the result of a classification that responds to a certain unjust social structure. But, even if neutral counterparts are not innocent, as long as they exist there will always be an alternative to uttering the slur, so its occurrences will always be derogatory. Only in the hypothetical case that we had no neutral counterparts and could only refer to a certain feature through the slur could we perhaps admit that there are neutral uses of slurs. Of course, this is a lot to assume, because to make sense of this case we have to assume that we used the slur because we felt it completely necessary to point out an identifying feature of a person or group. At any rate, since we have neutral counterparts, we always have the option of doing so in a non-offensive way.

The hypothesis that there are no non-derogatory contexts runs counter to what has generally been assumed in the literature. Cases of mention, in which we talk about the word rather than applying it to anybody, are usually taken to be non-derogatory (see e.g. Hornsby 2001: 129–130). Moreover, some full-fledged uses of slurs are standardly taken to be non-derogatory too. Among these, two kinds of uses have been most discussed. On the one hand, members of the target group can appropriate a slur in order to demarcate the group or foster solidarity or feelings of belonging, thus being able to use it in a non-derogatory way (see Bianchi 2014; Cepollaro 2017). But, on the other hand, we can also find non-derogatory uses of slurs that are not instances of appropriation—what have been called non-derogatory, non-appropriated (NDNA) uses of slurs (Hom 2008; see also Croom 2011 and section 6.2 for examples). Other non-derogatory uses of slurs that have been discussed in the literature are referential (Anderson 2018) and identificatory (Zeman 2022) uses. All these occurrences that in (Moreno & Pérez-Navarro 2021b) we assumed to be non-derogatory but potentially dangerous occurrences now take to be derogatory occurrences of slurs with diminished derogatory force. We will see this in more detail in section 6.3. That being said, I will again rely on the distinction between controlled and uncontrolled contexts that we made in (Moreno & Pérez-Na-

varro 2021b). But, instead of resulting in the existence of non-derogatory occurrences of slurs, the distinction will now serve to account for the variability in derogatory force and its dangerousness.

The fact that I take all occurrences of slurs to be derogatory does not mean that this proposal must be considered a kind of prohibitionist position (see Anderson & Lepore 2013a, 2013b; see also Cepollaro et al. 2019: 33). This is, I do not think that it should be forbidden to utter a slur even when it is mentioned, for instance, for pedagogical purposes. But I do think that the utterance of a slur always comes at a moral cost, which in cases like this may be worth paying.

6.2. Non-orthodox occurrences of slurs

In this section, I survey the different cases in which occurrences of slurs have been said to be non-derogatory. These will be the cases on which I will focus on subsequent sections to discuss whether they are in fact non-derogatory. A the end of the section, I will argue that the cases under discussion include uses of slurs, despite attempts to reduce all allegedly non-derogatory occurrences of slurs to cases of mention (Hornby 2001).

According to Hom (2010: 168–169), some occurrences of slurs, which Hom calls "orthodox occurrences", are non-displaceable, while others, which Hom calls "non-orthodox occurrences", are displaceable. Orthodox occurrences, Hom says, are non-displaceable because they are derogatory even when embedded, while non-orthodox occurrences are displaceable in the sense that they are not always derogatory. It is important to recall that, as noted in section 2.4, the sense in which Hom talks about displaceability and non-displaceability is different from the sense in which I do so. Hom's non-displaceability is closer to what I called "hyper-projectivity", which is the feature of slurs that distinguishes them from the presuppositional triggers we are familiar with and which makes their occurrences be derogatory, if the context makes them so, independently of the linguistic environment in which the slur is found. What is important about non-orthodox occurrences of slurs, then, is that they are not always derogatory.

Let "*S*" be a slur, and let us substitute it for the word that Hom uses in his examples of orthodox occurrences of slurs:

(1) If there are *S*s in the building, then A will be relieved.

(2) There are no *S*s in the building.

(3) Are there *S*s in the building?

(4) John said that there are *S*s in the building.

(5) John said: "There are *S*s in the building."

(6) In the novel, there are *S*s in the building.

"*S*" is derogatory in (1)–(6). It is so even if it appears in the antecedent of a conditional in (1), embedded under negation in (2), as part of a question in (3), reported in indirect style in (4) and indirect style in (5), and embedded under an "in the fiction" operator in (6) (see also Hornsby 2001: 129–130; Potts 2007: 166; Croom 2011: 347; Hom 2012: 384–385; Anderson & Lepore 2013a: 30; Croom 2014: 228). However, Hom does not consider "*S*" to be derogatory in his example of a non-orthodox occurrence of a slur, which he takes from (Hom 2008: 429) and we reproduce here substituting "*N*" for the neutral counterpart of "*S*":

(7) Institutions that treat *N*s are *S*s are morally depraved.

Hom takes (7) to contain a non-derogatory occurrence of "*S*". [1]Note that, in this case, "*S*" is not embedded in any of the ways depicted in (1)–(6). Hom thus takes (7) to include a non-orthodox occurrence of a slur.

A couple of categories can be distinguished within allegedly non-orthodox occurrences of slurs. One of the cases that most readily come to mind is that of appropriated uses of slurs. Appropriated uses of slurs are those that take place when speakers belonging to the target group use the

---

[1] Note that my commitment that there are no non-derogatory occurrences of slurs requires me to take (7) to contain a derogatory occurrence, *pace* Hom.

corresponding slur to demarcate the group or foster solidarity of feelings or belonging (see Bianchi 2014: Cepollaro 2017; Anderson 2018). Speakers who have appropriated a slur are supposed to be able to use it to refer to themselves or other members of the group without derogating anyone.

But we can also find occurrences of slurs that are supposed to be non-orthodox without being instances of appropriation. These have been aptly labelled "non-derogatory, non-appropriated" (NDNA) uses of slurs, an umbrella term that is meant to cover all non-orthodox uses of slurs that do not involve appropriation (Hom 2008; see also Croom 2011). An example of an alleged NDNA use of a slur, given by Hom (2008: 429), is the following:

(8) There are lots of *N*s at Y, but not *S*s.

If truly non-orthodox, (7) Institutions that treat *N*s are *S*s are morally depraved, would be another example of an NDNA use of a slur.

Appropriated and NDNA uses are both *uses* of slurs, but paradigmatic non-orthodox occurrences of slurs are *mentions* of them. In fact, Hornsby apparently endorses the idea that all non-orthodox occurrences of slurs are at the end of the day cases of mention:

> Certainly there are occurrences of derogatory words that are utterly inoffensive: "He is not [an *S*]" can be said in order to reject the derogatory "[*S*]"; one can convey that "[*S*]" is not something one calls anyone by saying "There aren't any [*S*s]." But these examples do not count against their uselessness as I mean this, because they are examples in which it is part of the speaker's message that she has no use for the word "[*S*]". We might gloss the two sentences so that the word is mentioned rather than used: ""[*S*]" is not what he ought to be called"; ""[*S*]" has no application". (Hornsby 2001: 129)

Hornsby seems to reduce all cases of allegedly non-orthodox occurrences of slurs to cases of mention rather than use. A plausible paraphrase of (7) in which "*S*" is mentioned rather than used would be this:

(9) Institutions that treat Ns as deserving to be called "Ss" are morally depraved.

However, I do not think that Hom's distinction between orthodox and non-orthodox occurrences of slurs should coincide with the distinction between use and mention, so that every sentence featuring an allegedly non-orthodox occurrence of a slur can be paraphrased as a case of mention. In fact, in the sense that a criterion to distinguish the latter from the former can be given just in terms of the form of the sentence used. If we take Hom's examples seriously, a sentence in which a slur is used can be derogatory or non-derogatory independently of whether the slur appears in the antecedent of a conditional, embedded under negation, or as part of a question, and the same happens when the slur is merely mentioned.

The way Hom draws the line, it seems that the relevant factor when distinguishing between orthodox and non-orthodox occurrences of slurs is the context. (1) and (7) are both cases of use, but Hom takes (1) to be derogatory and (7) not to be. (5) and (9) are both cases of mention, but Hom takes (5) to be derogatory and (9) not to be. The difference must thus lie in the kind of context that we most plausibly associate with each sentence: (1) is most easily imagined as uttered in a context in which an act of derogation is taking place, while for (7) we tend to imagine contexts in which which the speaker is in fact denouncing derogatory practices, and something parallel to this can be said about (5) and (9).

Thus, the difference between orthodox and non-orthodox occurrences of slurs does not lie in the form of the sentence used, but in the context in which they take place. In the next section, I flesh out what exactly should distinguish contexts in which occurrences of slurs are derogatory from context in which they are not. However, as we shall see, the difference has a limited impact, since no context matches the profile that would make the occurrences of slurs taking place in it lack derogatory character.

6.3. Controlled and uncontrolled contexts

The upshot of the previous section was that Hom's distinction between orthodox and non-orthodox occurrences of slurs draws on the context at which they take place. In this section, I take a closer look at the kinds of context that could make an occurrence of a slur orthodox or non-orthodox. In particular, I identify Hom's non-orthodox occurrences of slurs with those that take place in what I call "controlled contexts" and orthodox occurrences of slurs with those that take place in "uncontrolled contexts" (cf. Camp 2018b: 43). However, we will see in the next section that, even if the distinction between controlled and uncontrolled contexts can help us rank occurrences of slurs according to their derogatory force, no occurrence exhibits no derogatory force at all.

Communication is a risky business. There are a number of factors that can have an impact on the kind of effect that a given utterance will have, and most of them escape our control. When communicating, we often have to manage without knowing what our audience knows or what their expectations are. Still, even if rare, contexts can be found in which we can predict with reasonable accuracy what the consequences of given utterances will be. I call these "controlled contexts". When I am talking about utterances including a slur, an example of a controlled context would ideally be a pedagogic one, and another, more contentious one would be that in which a slur is successfully used in an ironic way.

Let us begin, then, by examining a case of controlled context. Imagine the following situation. Our daughter comes home from school one day saying that a classmate of hers has said that the parents of another classmate are *S*s. When our daughter tells us that this has happened, she does so as part of her telling us all that has happened today at school. It is just one anecdote among many. So, once our daughter has told us this, we feel it is necessary to explain to her why she should not utter that word. At this point, it seems to me that there are basically two things that can happen. One is that, as soon as our daughter utters the slur, our reaction is something like "We don't say that word

at home". This seems to me to be the ideal reaction, as it allows the adult not to utter the slur again. However, I think it makes sense to consider that, in this kind of situation, adults sometimes stop to think about the best way to deal with the issue, and this prevents us from reacting immediately. In this case, later in the day we tell her that she should not say that word again, to which our daughter might reply, "What word?". Her reaction of not knowing what we are referring to does not mean that she has forgotten the word, but she honestly does not know which of the words she has uttered we have forbidden her to say again. Since she does not know which specific word we are referring to, we end up uttering "*S*". Specifically, we say something like "It's not nice to say "*S*"" or "It's not nice to call anyone "*S*", that's an ugly thing to say". We have uttered a slur, even if we have only mentioned it.[2]

Here is another kind of case in which we can say that the utterance of a slur has taken place in a controlled context. This time, I am not talking about a mere mention of a slur, but about a full-blown use—an ironic use. Let us imagine that we are a group of friends who perceive ourselves as progressive, who would never mention a slur in front of strangers, let alone use it to insult a person because they belong to a certain group. However, we find fun in imitating bigots' mannerisms and enjoy inner jokes that include ironic uses of "*S*". We are completely sure that all our friends in the group share our sensibility and that none of them will take us to aim at insulting anyone. We think it

---

[2] It could also be said, especially if we have in mind the second case in which we rather want to explain to our daughter what a slur is, that pedagogical contexts do not only allow the mention of slurs, but also their use. For example, we could also have said "There are no *S*s, only *N*s". However, if, as we have said, our daughter does not know what a slur is, the fact that we have seen fit to categorise her classmate's parents as belonging to a target group, albeit in a supposedly neutral way, makes the sentence problematic (see Mühlebach 2021). Among other things, through categorisation we have confirmed membership of a particular derogated group and thus revealed characteristics that were previously irrelevant to our daughter. Discussions with Marta Jorba, Eduardo Pérez, Manuel de Pinedo and Neftalí Villanueva on pedagogical contexts have been very useful for the development of this part.

is intuitive to take occurrences of slurs such as these to take place in a controlled context, whatever the form of the sentences in which they appear.

Other uses of slurs that have recently been described as non-derogatory are referential (Anderson 2018) and identificatory (Zeman 2022) uses. Referential uses take place when members of the target group use a slur to address other members without any intention to appropriate the term, while identificatory uses take place when they simply take the word, to be the one that refers to the group they themselves belong to. I see these uses as taking place in controlled contexts too, as the speaker's group membership is salient enough for her to be confident that the audience will understand that she did not mean to insult, just as it happened in our daughter's case.

If controlled contexts are those in which we can be sure about the other participants' knowledge and expectations, almost all contexts in which we can find ourselves are uncontrolled ones. It is difficult to know anyone as well as we know our own children or our close friends, and in many cases, we hardly have any relevant information about our audience. Consider our daily interactions with strangers and the limiting case of the completely uncontrolled context in which public communication takes place. When we utter a slur in an uncontrolled context, our audience has every reason to attribute to us a discriminatory attitude toward a given group, and we cannot reasonably expect not to be attributed such an attitude, which is what, in an intuitive sense, means to derogate. Thus, in uncontrolled contexts, which are most of the contexts, occurrences of slurs have full derogatory force.

When in (Moreno & Pérez-Navarro 2021b) we made the distinction between controlled contexts and uncontrolled contexts in order to determine whether the occurrence of a slur was derogatory or not, a plausible objection was that slurs can occur in a derogatory way in controlled context too. Our response was that it may indeed be the case that the occurrence of a slur in a controlled context is derogatory. For example, I may know exactly what the reaction of the audience to my utterance of a slur will be, and know this reaction to be one that will precisely result in derogation. In

this case, the occurrence of the slur will be derogatory even if it takes place in a controlled context. This objection does not affect the position defended here, though, since slurs derogate in every context and the distinction between controlled and uncontrolled contexts only tell us how strong the derogatory force of the slur is.

Another plausible objection to the proposal here, mirroring the one above, was the following. In uncontrolled contexts, we cannot be sure that the uptake of our utterance will fail to derogate, but this does not mean that it will derogate. It may happen that, just by chance, every single member of the audience understands the occurrence of the slur as non-derogatory, even if we are not able to predict that this will be the case. For instance, all bystanders who hear our daughter utter a slur could assume that she does not know what the word means. Note, however, that we have characterised uncontrolled contexts as those in which hearers have every reason to attribute to us a discriminatory attitude toward the target group, which we cannot in turn reasonably expect not to be attributed to us. However, inasmuch as they would be warranted in so doing, our utterance can be taken to be derogatory.

The idea that, in uncontrolled contexts, it is reasonable to attribute the utterer a discriminatory attitude toward the target group no matter what her actual attitudes are is similar to one that has been defended by Lasersohn (2007). According to Lasersohn, slurs are emotionally charged terms, so uttering them entails a social risk. Lasersohn believes that speakers are aware of the social burden of slurs, and this is the main reason why most speakers avoid uttering slurs—because, whatever their particular attitudes, they are aware that they can reasonably be attributed to bigotry. Precisely because of this, when a speaker does utter a slur, it makes sense to think that she is comfortable with being identified as a bigot, and this is how the derogatory content of the slur projects where most presupposed content does not (Lasersohn 2007: 228). Like Lasersohn, I think that, if a speaker utters a slur in an uncontrolled context in which it is even merely possible that someone understands

the occurrence of the slur as derogatory, it makes sense to take the speaker to be comfortable with this possibility, and thus to take the occurrence to be actually derogatory.

6.4. The normalising potential of slurs

We have seen that the derogatory force of slurs is greater in uncontrolled contexts than it is in controlled context such as the ones involving pedagogic and ironic occurrences. Still, no matter how carefully we arrange the current context to make sure that the utterance of a slur does not have the kind of effect we want to avoid, it will facilitate ulterior occurrences of the term. In particular, it will make the slur more likely to appear in uncontrolled contexts in which the utterance of the slur has full derogatory force. In this section, I explore how this could be the case with the two kinds of occurrences of slurs that I presented in the previous section—pedagogic mentions and ironic uses of slurs. If even the apparently most "innocent" occurrences of slurs, such as those that take place in pedagogic contexts, are at the end of the day derogatory, it is natural to conclude that all occurrences of slurs are.

Let us start with the irony case. Remember that, in this case, we are a progressive group of friends who enjoy using slurs in an ironic way to make fun of bigots. In (Moreno & Pérez-Navarro 2021b) we considered it quite intuitive to think of this case as involving a non-derogatory occurrence of the slur. However, it seems a mistake to me to take this intuition for granted with respect to this case. As I noted, the main motivation for uttering slurs in these cases is to have fun with our friends. We have ironically used "*S*" in our friend group so many times that it no longer makes us uncomfortable to hear the word, which makes its utterance in uncontrolled contexts more likely now[3]. This less-derogatory use of a slur thus normalises more-derogatory occurrences of the word, and is potentially harmful in this sense. Of course, our friend group could be careful enough not to let uses of slurs slip out of the controlled context. This is why ironic uses of slurs such as these are

---

[3] Moreover, it is these practices that foster the desensitisasion of the target groups (section 2.2).

not harmful *tout court*, but only potentially harmful. But, since potential harm implies actual danger, these uses are dangerous *tout court*. In other words, ironic uses of slurs may have less derogatory force than full-blown uses of these terms, but they are still derogatory.

Now, take the example involving a pedagogic mention of a slur, also described in the previous section. Remember that, in this case, we ended up uttering the word "*S*" in order to make our daughter aware that she should not call anyone an *S*. Again, in (Moreno & Pérez-Navarro 2021b) we maintained that in such cases we can be sure that by doing this we are not insulting anyone. Now, I think that in such a case the most we can be sure of is that our utterance is not going to be as offensive as if we uttered the slur in an uncontrolled context, but that it will still derogate. The reason why even in controlled contexts slurs derogate is that even in a context where we know our daughter well enough to ensure that she has understood why we have uttered the slur, we cannot remove the *normalising* component of the slur. No matter how much we want to enrich the context in a case like this, the normalising component of the slur will not disappear. This is probably due to the fact that in pedagogical cases communication is not established horizontally between the participants of the conversation. Therefore, if the person with the most authority utters the slur, there will always be a risk of normalising its use. We have taught our daughter what "*S*" means, thus giving her the tools to use the word to insult if at some point she wants to do so. Note that, at least in this case, the risk that our daughter grows up to use "*S*" as an insult may be worth it: preventing an actual risk may be preferable to preventing a virtual one. In this sense, this case might strike us as clearer than the previous one. The normalising potential is similar in both cases, though, so this occurrence of the slur is also derogatory. The difference is that, in the pedagogic case, it is clearer how the benefits could outweigh the potential harm. In the irony case, all we have on the positive side is the fun we have with our friends.

It can still generate some resistance just to consider that pedagogical occurrences of slurs are also non-derogatory. Thinking of these kinds of situations as derogatory contexts makes us have to

confront our conviction that our purest motivations, which are trying to prevent slurs from being used and normalised, are not enough to block derogation. However, the harm that can be caused by uttering a slur in isolation is not comparable to the harm that can be caused by doing so in a pedagogical context. It is necessary to distinguish between the two situations because not doing so could risk distorting what derogation entails. However, there is not much we can do in a pedagogical context to make derogation non-existent. In pedagogical contexts, a non-horizontal relationship is established between the person who teaches something and the person who receives the learning. Insofar as, by definition, pedagogical contexts are constituted in this way, the simple utterance of the slur by the person teaching can have a very concrete effect on her audience.

Referential and identificatory uses of slurs are closer to pedagogic mentions than ironic uses in this respect. As with pedagogic mentions, however, there is still the risk that these uses facilitate ulterior occurrences of the term in contexts in which the group membership of the speaker, although salient, does not make the audience understand such occurrences as non-derogatory because the speaker does not belong to the target group.[4]

Hence, the moral here is that we may have full control over the present context, but we do not have control over all possible future contexts. Thus, slurs always have normalising potential. The slur might not be problematic in the context in which it is uttered, but it may reveal itself to be so as we look beyond the original conversation and consider other exchanges that might be facilitated by the original utterance. We take something to be dangerous whenever it may cause harm, even if it

---

[4] The only reason I situate referential uses (Anderson 2018) and identificatory uses (Zeman 2022) here is as a matter of theoretical coherence. Although I have no developed position on this, and only a mere intuition, I do not feel entirely comfortable talking about the harmful potential of appropriated, referential and identificatory uses. I believe that, in these cases, the normalising potential of slurs is irrelevant and that the moral cost assumed, if any for these cases, will always be justified. I just don't think I have the right to say, even if I belonged to one, whether the derogated groups commit any kind of infringement, however minor and justified.

does not actually do so. Insofar as occurrences are always potentially harmful, therefore, I take them to be always dangerous, and thus always derogatory.

As noted in section 6.1, the idea that there are no non-derogatory occurrences of slur should not suffice to classify me as a prohibitionist (see Anderson & Lepore 2013a, 2013b; see also Cepollaro et al. 2019: 33)—I do not think utterances of a slur should be forbidden *tout court*. I think there are some practical consequences to the categorisation of some uses of slurs as appropriated or NDNA, and that the distinction between use and mention has practical consequences when it comes to slurs too. It may be permissible to mention a slur in certain contexts, just like it may be permissible to make an appropriated or an NDNA use of a slur. This marks a difference between these occurrences of slurs and those where slurs have their full derogatory force. But we should be aware that these practices come with a moral cost too. The price may be worth paying, of course. It just misrepresents our moral life to assume that it consists in choosing the only permissible thing to do in each case; rather, we assess the moral costs and benefits of each course of action, decide what weight to give to each, and act in consequence. The moral benefits of performing a certain utterance of a slur might outweigh the pervasive moral cost I have described, and so it might be worth it to utter the slur.

6.5. Conclusion

In this chapter, I have argued that we cannot consider any isolated sentence, in the absence of a context of utterance, as an example of a non-derogatory occurrence of a slur. We have to take the context into account; but the context will always make the occurrence derogatory, as the evidence surveyed shows. I have argued that this is because the mere utterance of a slur has a certain kind of impact—it normalises further occurrences of the word. This is so even in controlled contexts. Nothing prevents the ulterior occurrences of the slur that are normalised by those taking place in con-

trolled contexts to take place in uncontrolled contexts, so even this less-derogatory occurrences are derogatory after all. In the terms introduced in section 2.4, slurs are non-displaceable.

If derogation is not the criterion that distinguishes acceptable from unacceptable occurrences of slurs, what is? I do not think that a criterion can be found by which certain occurrences of a slur should be allowed, but by the same token I do not think that there is a criterion that forbids the rest of occurrences *tout court*. The point is that uttering a slur always comes at a moral cost, and it is the responsibility of the speaker to assess such a cost and decide whether it is worth mentioning a word to explain to a child that it should never be used.

Chapter 7

Conclusion

Communication is a risky business. This idea has run through all the chapters of this thesis. I began this work by pointing out that one of the issues that has had the greatest social impact in recent years is the idea that it is no longer possible to speak without offending anyone. The realisation of how offensive we can be has given rise to a sensitivity that makes us think about what words to use, and this does not only apply in the case of public communication. This increased sensitivity in speech has also highlighted a certain dimension along which we can evaluate what we say. This evaluation can become a tricky subject and it is not within the scope of this thesis to comment on it more than is strictly necessary. However, to decide not to mention any slurs (section 1.3) or to talk about the linguistic act of retraction in relation to the utterance of slurs (section 5.4) is to assume that this work takes part in some such evaluation.

The way in which I have proposed to specify whether what we say is offensive or not has been to asses, in a given context, the moral cost of uttering a slur. The decision not to mention any slur indicates that I have not been willing to assume this cost in this thesis. This decision is not only motivated by my conviction that it is not necessary to mention any slur in an academic context in order to reach the correct linguistic intuition, but also because this is a way of bringing NPP closer to being an ideal theory in Keiser's (2022: 4) second sense. In this sense, a theory can be said to be ideal if it lives up to some normative standard. The normative standard to which this thesis adheres is the non-normalisation of the use of slurs by mentioning them, so not to do so has been the way to comply with it. This is where the methodological motivations for this thesis intertwine with its theoretical motivations.

All of this determines the perspective from which the main objective of this work, which has been to account for the derogatory effect of slurs, has been pursued. This task has required, first, to

specify who is affected by the derogatory effect of slurs. This answer depended on what we consider as the target group (section 1.2). The perspective adopted is that only consider groups that are systematically discriminated against on the basis of their identity traits as a target group for slurs. The identity traits that have historically been considered susceptible to discrimination have been those related at least to race, gender, religious beliefs and sexual orientation. Target groups that are not formed by virtue of these traits are not necessarily exempt from derogation, but they will not be objects of derogation by slurs. The distinction between slurs and derogatory terms was intended to point out that the derogatory effect of slurs, unlike that of derogatory terms, is always accompanied by a discriminatory attitude.

More often than not, when we use a slur, and I would argue that this is true of hate speech in general, we express some kind of unease about a particular aspect of the world. It has been a central task in this thesis to explain what happens in the context of the conversation when such unease is expressed. The ability to explain this depended on the ability to explain that the derogatory effect of slurs is connected with action and that it is somewhat elusive. The first is because by uttering a slur we commit to an idea of what the world should be like, and this licenses treating people in a certain way. The second is because the derogatory content is impossible to capture because it is non-propositional and presuppositional in nature. In order to explain the connection with action, NPP has been framed within an expressivist theory of meaning on which derogatory content is not propositional. But, in order to explain the elusiveness of this content, not only is use made of a non-propositional position of the meaning, but also the content is placed outside of what is said. This combination is what makes derogatory content so difficult to capture and confront.

Chapter 2 set out the desiderata an account of the meaning of slurs like the one presented here should meet. In particular, the proposal should account for the descriptive, practical and semantic features of slurs, which the chapter was devoted to discussing. Special emphasis was put on descriptive and semantic features, although one practical feature—the normalising potential of slurs—

would eventually play an important role in chapter 6. Descriptive features were understood as those that allow us to distinguish between a slur and a mere insult and explain why the possibility of factual error arises with the former but not with the latter. Once identified, descriptive features allow us to introduce the neutral counterparts of slurs, which in turn allow us to refer to the target group in a non-offensive way, but are also problematic both for ideological reasons and in light of McDowell's (1981) anti-disentangling argument. However, for the purposes of this thesis I have assumed that slurs do have neutral counterparts (even if I do not take them to be neutral in any deep sense) for two reasons. First of all, we need them to specify what is said when we utter a sentence in which a slur appears. Secondly, we may need them to refer to the target group and point out the oppression it suffers.

Among the semantic features of slurs, the focus of chapter 2 was on non-displaceability, which is the feature that makes slurs be derogatory in every context. The analysis of this feature allowed us to draw several lessons. The first was the distinction between non-displaceability and the projection profile of the slurs, which makes their occurrences be derogatory no matter the linguistic environment in which they are found. The projection profile of slurs follows from their non-displaceability, but they are two distinct things. The second lesson is that non-derogatory contexts do not exist and slurs are therefore non-displaceable. This was presented as an hypothesis in support of which some empirical studies were provided.

Three groups have been distinguished in chapter 3 to classify some of the different theories on the meaning of slurs found in the literature These three families were those formed by content theories, force theories and deflationary theories. The proposal that has been defended in this thesis is located within the theories of content and, specifically, within the theories that place the derogatory content at the presuppositional level of meaning. However, NPP goes in a way beyond the classification presented in this chapter, as the non-propositional character it attributes to derogatory content makes it also a heir of expressivism, which was deemed a force theory.

149

Chapter 4 was devoted to developing NPP. The notion of presupposition was reviewed and some of the proposals that use propositional presuppositions to explain the derogatory content were exposed. As we saw, presuppositional theories have the following advantage: the truth-conditions assigned to utterances of sentences containing slurs are not identical to the truth-conditions assigned to the utterances of sentences containing their neutral counterparts. However, it was also pointed out that proposals based on propositional presuppositions (Schlenker 2007; Cepollaro & Stojanovic 2016) do not flesh out what makes expressive presuppositions different from ordinary ones. This makes their explanation of the derogatory effect insufficient. Non-propositional presuppositional accounts are thus vindicated.

Two non-propositional presuppositional proposals were introduced: Marques and García-Carpintero's (2020), which relies on reactive attitudes, and NPP, which relies on world-orderings. The difference between these proposals lies in the way in which they explain the impact of non-propositional presuppositions on the common ground. An explanation based on ordering of possible worlds was deemed preferable because this explanation is more parsimonious in at least two ways: ontologically and with respect to intervention. Firstly, talking about reactive attitudes implies introducing a new element into the common ground, where we already have possible worlds. Secondly, an explanation based on operations performed on possible worlds is neutral between structural and individual perspectives on intervention.

If chapter 4 presented the account of the meaning of slurs favoured in this thesis, chapter 5 stated the reasons for favouring it having to do with the semantic features of these terms. NPP is, I argued in that chapter, specially well suited to account for slurs' resistance to cancellation, rejection and retraction. Slurs' resistance to cancellation, which follows straightforwardly from the projection profile first referred to in chapter 2, is the ability of the derogatory effect of slurs to resist in every single linguistic environment. Slurs are also resistant to rejection because no reply to an utterance

involving them seems able to target their derogatory effect, and to retraction because saying "I take that back" after uttering a slur never seems to accomplish its purpose.

I argued that the derogatory effect of slurs projects because it is due to their presupposed content, and if it seems to do so even in cases in which a presupposition would not project is because we have construed the examples wrong—we are looking at cases in which any presupposition would project, and the non-propositional character of derogatory content makes it impossible to build the example the opponent of presuppositional theories has in mind. The presuppositional character of derogatory content, for its part, is what makes very difficult for it to become the target of rejection and retraction.

Chapter 6, finally, faced the task of offering a philosophical argument for the hypothesis laid out in chapter 2 that slurs are non-displaceable. To do so, I first offered a workable construal of the opposite hypothesis that slurs are in fact displaceable. I interpreted those who have thought that there are non-derogatory occurrences of slurs, such as Hom (2010), as relying on a distinction between controlled and uncontrolled contexts and assuming that utterances of slurs taking place in controlled contexts are not derogatory.

The argument appealed to the normalising potential of slurs, which was mentioned in chapter 2 as one of the practical features of slurs. In particular, the argument was that occurrences of slurs are imbued with normalising potential even in controlled contexts, where one might expect them not to be derogatory. If this is so, then utterances of slurs are derogatory even if they take place in a controlled context, and thus always derogatory. Slurs are hence non-displaceable.

Conclusiones

La comunicación es un negocio arriesgado. Esta idea ha recorrido todos los capítulos de esta tesis. Comencé este trabajo señalando que uno de los temas que más impacto social han tenido en los últimos años es la idea de que ya no es posible hablar sin ofender a nadie. La constatación de lo ofensivos que podemos llegar a ser ha dado lugar a una sensibilidad que nos hace reflexionar sobre qué palabras utilizar, y esto no solo se aplica en el caso de la comunicación pública. Esta mayor sensibilidad a la hora de hablar también ha puesto de relieve una cierta dimensión con respecto a la que podemos evaluar lo que decimos. Esta evaluación puede convertirse en un tema peliagudo y no entra en el ámbito de esta tesis comentarlo más de lo estrictamente necesario. Sin embargo, decidir no mencionar ningún *slur* (sección 1.3) o hablar del acto lingüístico de la retractación en relación con la proferencia de *slurs* (apartado 5.4) es suponer que este trabajo participa en alguna evaluación de este tipo.

En el capítulo 2 se expusieron los *desiderata* que debe cumplir una teoría del significado de los *slurs* como la que aquí se presenta. En concreto, la propuesta debe dar cuenta de las características descriptivas, prácticas y semánticas de los *slurs*, que el capítulo estuvo dedicado a discutir. Se hizo especial hincapié en los rasgos descriptivos y semánticos, aunque un rasgo práctico —el potencial normalizador de los *slurs*— acabó desempeñando un papel importante en el capítulo 6. Por rasgos descriptivos se entienden aquellos que nos permiten distinguir entre un *slur* y un mero insulto y explicar por qué la posibilidad de error de hecho surge en el caso del primero y no en el del segundo. Una vez identificados, los rasgos descriptivos nos permiten introducir las contrapartidas neutrales de los *slurs*, que a su vez nos permiten referirnos al grupo objetivo de una manera no ofensiva, pero también son problemáticos tanto por razones ideológicas como a la luz del argumento *anti-disentanglement* de McDowell (1981). Sin embargo, para los fines de esta tesis he asumido que los *slurs* tienen contrapartidas neutrales (aunque no las considere neutrales en un sentido pro-

fundo) por dos razones. En primer lugar, las necesitamos para especificar lo que se dice cuando proferimos una oración en la que aparece un *slur*. En segundo lugar, podemos necesitarlas para referirnos al grupo objetivo y señalar la opresión que sufre.

Entre las características semánticas de los slurs, el capítulo 2 se centró en la *non-displaceability*, que es la característica que hace que los slurs sean derogatorios en todos los contextos. El análisis de este rasgo nos permitió extraer varias lecciones. La primera fue la distinción entre la *non-displaceability* y el perfil de proyección de los *slurs*, que hace que sus apariciones sean derogatorias independientemente del entorno lingüístico en el que se encuentren. El perfil de proyección de los *slurs* se deriva de la *non-displaceability*, pero son dos cosas distintas. La segunda lección es que los contextos no derogatorios no existen y que, por tanto, los *slurs* son *non-displaceable*. Esto se presentó como una hipótesis en apoyo de la cual se aportaron algunos estudios empíricos.

En el capítulo 3 se clasificaron en tres familias las distintas teorías sobre el significado de los *slurs* que se encuentran en la bibliografía. Estas tres familias son las formadas por las teorías del contenido, las teorías de la fuerza y las teorías deflacionarias. La propuesta que se ha defendido en esta tesis se sitúa dentro de las teorías del contenido y, en concreto, dentro de las teorías que sitúan el contenido derogatorio en el nivel presuposicional del significado. Sin embargo, NPP va en cierto modo más allá de la clasificación presentada en este capítulo, ya que el carácter no proposicional que atribuye al contenido derogatorio la convierte también en heredera del expresivismo, que se ha considerado una teoría de la fuerza.

El capítulo 4 estuvo dedicado al desarrollo de NPP. Se analizó la noción de presuposición y se expusieron algunas de las propuestas que utilizan presuposiciones proposicionales para explicar el contenido derogatorio. Se presentaron dos propuestas presuposicionales no proposicionales: la de Marques y García-Carpintero (2020), que se basa en actitudes reactivas, y NPP, que se basa en ordenamientos de mundos. La diferencia entre estas propuestas radica en la forma en que explican el impacto de las presuposiciones no proposicionales en el *common ground*. Se defendió la posición

basada en los ordenamientos de mundos posibles porque esta explicación es más parsimoniosa al menos en dos sentidos: ontológicamente y con respecto a la intervención.

Si en el capítulo 4 se presentó la teoría del significado de los *slurs* favorecida en esta tesis, en el capítulo 5 se expusieron las razones para favorecerla, que tienen que ver con las características semánticas de estos términos. En ese capítulo argumenté que NPP es especialmente adecuada para explicar la resistencia de los *slurs* a la cancelación, el rechazo y la retractación. El capítulo 6, por último, se enfrentó a la tarea de ofrecer un argumento filosófico a favor de la hipótesis expuesta en el capítulo 2 de que los *slurs* no son *displaceable*. Para ello, primero se ofreció una interpretación factible de la hipótesis contraria, según la cual los *slurs* son de hecho *displaceble*. Interpreté que quienes han pensado que hay apariciones no derogatorias de *slurs*, como Hom (2010), se basan en una distinción entre contextos controlados y no controlados y asumen que los *slurs* que se pronuncian en contextos controlados no son derogatorios. El argumento hacía referencia al potencial normalizador de los *slurs*, que se mencionó en el capítulo 2 como una de sus características prácticas. En concreto, el argumento fue que los *slurs* están imbuidos de potencial normalizador incluso en contextos controlados, en los que cabría esperar que no fueran derogatorios. Si esto es así, los *slurs* son derogatorios aunque se produzcan en un contexto controlado y, por lo tanto, siempre son derogatorios. Por tanto, los *slurs* son *non-displaceable*.

References

Abbott, B. (2006). Where have some of the presuppositions gone. In B. Birner & G. Ward (Eds.), *Drawing the Boundaries of Meaning: Neo-Gricean Studies in Pragmatics and Semantics in Honor of Laurence R. Horn* (pp. 1–20). John Benjamins.

Abusch, D. (2002). Lexical alternatives as a source of pragmatic presuppositions. In B. Jackson (Ed.), *Proceedings of SALT XII* (pp. 1–20). CLC Publications.

Anderson, E. (2010). *The Imperative of Integration*. Princeton: Princeton University Press.

Anderson, L. (2018). Calling, addressing, and appropriation. In D. Sosa (Ed.), *Bad Words: Philosophical Perspectives on Slurs* (pp. 6–28). Oxford University Press.

Anderson, L., & Lepore, E. (2013a). Slurring Words. *Noûs*, 47(1), 25–48.

— (2013b). What did you call me? Slurs as prohibited words. *Analytic Philosophy*, 54(3), 350–363.

Ayala, S. (2016). Speech affordances: A structural take on how much we can do with our words. *European Journal of Philosophy*, 24(4), 879–891.

— (2018). A structural explanation of injustice in conversations: It's about norms. *Pacific Philosophical Quarterly*, 99(4), 726–748.

Ayer, A. J. (1936/2001). *Language, Truth and Logic*. London: Penguin.

Bach, K. (1999). The myth of conventional implicature. *Linguistics and Philosophy*, 22(4), 327–366.

Banks, R. R., & Ford, R. T. (2008). (How) does unconscious bias matter: Law, politics, and racial inequality. *Emory Law Journal*, 58(5), 1053–1152.

Bianchi, C. (2014). Slurs and appropriation: An echoic account. *Journal of Pragmatics*, 66, 35–44.

Blackburn, S. (1984). *Spreading the Word: Groundings in the Philosophy of Language*. Clarendon.

— (1993). *Essays in Quasi-Realism*. Oxford University Press.

Brandom, R. (2000). *Articulating Reasons: An Introduction to Inferentialism*. Harvard University Press.

Brownstein, M., Madva, A., & Gawronski, B. (2019). What do implicit measures measure? *WIREs Cognitive Science*, e1501.

Camp, E. (2013). Slurring perspectives. *Analytic Philosophy*, 54(3), 330–349.

Camp, E. (2018a). A dual act analysis of slurs. In D. Sosa (Ed.), *Bad Words: Philosophical Perspectives on Slurs*. Oxford University Press.

Camp, E. (2018b). Insinuation, common ground, and the conversational record. In D. Fogal, D. W. Harris, & M. Moss (Eds.), *New Work on Speech Acts*. Oxford University Press.

Caponetto, L. (2020). Undoing things with words. *Synthese*, 197(6), 2399–2414.

Cepollaro, B. (2017). Let's not worry about the reclamation worry. *Croatian Journal of Philosophy*, 17(2), 181–193.

— (2020). *Slurs and Thick Terms: When Language Encodes Values*. Lexington Books.

Cepollaro, B., & Stojanovic, I. (2016). Hybrid evaluatives: In defense of a presuppositional account. *Grazer Philosophische Studien*, 93(3), 458–488.

Cepollaro, B., Sulpizio, S., & Bianchi, C. (2019). How bad is it to report a slur? An empirical investigation. *Journal of Pragmatics*, 146, 32–42.

Charlow, N. (2014). The problem with the Frege-Geach problem. *Philosophical Studies*, 167(3), 635–665.

Chierchia, G., & McConnell-Ginet, S. (2000). *Meaning and Grammar: An Introduction to Semantics*. MIT Press.

Chrisman, M. (2007). From epistemic contextualism to epistemic expressivism. P*hilosophical Studies*, 135(2), 225–254.

Cowan, G., & Mettrick, J. (2002). The effects of target variables and setting on perceptions of hate speech. *Journal of Applied Social Psychology*, 32(2), 277–299.

Croom, A. M. (2011). Slurs. *Language Sciences*, 33, 343–358.

— (2014). The semantics of slurs: A refutation of pure expressivism. *Language Sciences*, 38, 26–35.

Cruse, D. A. (1986). *Lexical Semantics*. Cambridge University Press.

Davidson, D. (1979). Moods and performances. In A. Margalit (Ed.), *Meaning and Use* (pp. 9–20). Dordrecht: D. Reidel.

D'Augelli, A. R. (1992). Lesbian and gay male undergraduates' experiences of harassment and fear on campus. *Journal of Interpersonal Violence*, 7(3), 383–395.

DiFranco, R. (2014). Pejorative language. *Internet Encyclopedia of Philosophy*. Retrieved from . Last accessed 23 April 2023.

Dummett, M. (1981). *Frege: Philosophy of Language*. Harvard University Press.

Frápolli, M. J., & Villanueva, N. (2012). Minimal expressivism. *Dialectica*, 66(4), 471–487.

Frege, G. (1892). Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, 100(1), 25–50.

Fricker, M. (2007). *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford University Press.

García-Carpintero, M. (2015). Contexts as shared commitments. *Frontiers in Psychology*, 6, 1932.

Garnets, L., Herek, G. M., & Levy, B. (1990). Violence and victimization of lesbians and gay men: Mental health consequences. *Journal of Interpersonal Violence*, 5(3), 366–383.

Geach, P. (1965). Assertion. *The Philosophical Review*, 74(4), 449–465.

Gibbard, A. (1990). *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Harvard University Press.

— (2003). *Thinking How to Live*. Harvard University Press.

Grice, P. (1975/1989). Logic and conversation. In *Studies in the Way of Words*. Harvard University Press.

Haslanger, S. (2015). Social structure, narrative and explanation. *Canadian Journal of Philosophy*, 45(1), 1–15.

Herbert, C. (ms). How to talk about slurs.

Herek, G. M., Gillis, J. R., & Cogan, J. C. (1999). Psychological sequelae of hate-crime victimization among lesbian, gay, and bisexual adults. *Journal of Consulting and Clinical Psychology*, 67(6), 945–951.

Hom, C. (2008). The semantics of racial epithets. *The Journal of Philosophy*, 105(8), 416–440.

— (2010). Pejoratives. *Philosophy Compass*, 5(2), 164-185.

— (2012). A puzzle about pejoratives. *Philosophical Studies*, 159(3), 383–405.

Hom, C., & May, R. (2018). Pejoratives as fiction. In D. Sosa (Ed.), *Bad Words: Philosophical Perspectives on Slurs* (pp. 138–158). Oxford University Press.

Hom, C., & Schwartz, J. (2013). Unity and the Frege–Geach problem. *Philosophical Studies*, 163(1), 15–24.

Hornsby, J. (2001). Meaning and uselessness: How to think about derogatory words. *Midwest Studies in Philosophy*, 25(1), 128–141.

Horwich, P. (2005). The Frege-Geach point. *Philosophical Issues*, 15(1), 78–93.

Jeshion, R. (2013a). Slurs and stereotypes. *Analytic Philosophy*, 54(3), 314–329.

— (2013b). Expressivism and the offensiveness of slurs. *Philosophical Perspectives*, 27(1), 231–259.

— (2018a). Slurs, dehumanization, and the expression of Contempt. In D. Sosa (Ed.), *Bad Words: Philosophical Perspectives on Slurs* (pp. 77–107). Oxford University Press.

— (2020). Pride and prejudiced. *Grazer Philosophische Studien*, 97(1), 106–137.

Kaplan, D. (1977/1989). Demonstratives: An essay on the semantics, logic, metaphysics and epistemology of demonstratives and other indexicals. In J. Almog, J. Perry, & H. Wettstein (Eds.), *Themes From Kaplan* (pp. 481–563). Oxford University Press.

160

— (1999). The meaning of "ouch" and "oops". Retrieved from

. Last accessed 23 April 2023.

Karttunen, L., & Peters, S. (1979). Conventional implicature. *Syntax and Semantics,* 11, 1–56.

Keiser, J. (2022). *Non-Ideal Foundations of Language*. Routledge.

Kenny, A. (1963). Action, Emotion and Will. Humanities Press.

Kukla, R. (2018). Slurs, interpellation, and ideology. *Southern Journal of Philosophy*, 56(S1), 7–32.

Lasersohn, P. (2007). Expressives, perspective and presupposition. *Theoretical Linguistics*, 33(2), 223–230.

Lewis, D. (1970). General semantics. *Synthese*, 22(1/2), 18–67.

— (1979). Scorekeeping in a language game. *Journal of Philosophical Logic*, 8(1), 339–359.

MacFarlane, J. (2014). Assessment Sensitivity: Relative Truth and its Applications. Oxford University Press.

Macià, J. (2002). Presuposicion y significado expresivo. *Theoria: Revista de Teoría, Historia y Fundamentos de la Ciencia*, 17(3), 499–513.

Mallon, R. (ms). Psychology, accumulation mechanisms, and race.

Manela, T. (2021). Gratitude. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Retrieved from   . Last accessed 23 April 2023.

Manne, K. (2017). *Down Girl: The Logic of Misogyny*. Oxford University Press.

Marques, T. (2014). Doxastic disagreement. *Erkenntnis*, 79(S1), 121–142.

— (2018). Retractions. *Synthese*, 195(8), 3335–3359.

Marques, T., & García-Carpintero, M. (2020). Really expressive presuppositions and how to block them. *Grazer Philosophische Studien*, 97(1), 138–158.

Mason, M. (2018). Contempt: At the limits of reactivity. In M. Mason (Ed.), *The Moral Psychology of Contempt* (pp. 173–192). Lanham, MD: Rowman and Littlefield.

McCready, E. (2010). Varieties of conventional implicature. *Semantics and Pragmatics*, 3(8), 1–57.

McDowell, J. (1981). Non-cognitivism and rule-following. In S. Holtzman & C. M. Leich (Eds.), *Wittgenstein: To Follow A Rule* (pp. 141–162). Routledge.

McGowan, M. K. (2009). Oppressive speech. *Australasian Journal of Philosophy*, 87(3), 389–407.

— (2019). *Just Words: On Speech and Hidden Harm*. Oxford University Press.

Moore, G. E. (1903/1993). *Principia Ethica*. Cambridge University Press.

Moreno, A., & Pérez-Navarro, E. (2021a). The resistant effect of slurs: A nonpropositional, presuppositional account. *Daimon*, 84, 31–46.

— (2021b). Beyond the conversation: The pervasive danger of slurs. *Organon F*, 28(3), 708–725.

Mühlebach, D. (2021). Semantic contestations and the meaning of politically significant terms. *Inquiry*, 64(8), 788–817.

Nunberg, G. (2018). The social life of slurs. In D. Fogal, D. Harris, & M. Moss (Eds.), *New Work on Speech Acts* (pp. 237–267). Oxford University Press.

Pérez Carballo, A., & Santorio, P. (2016). Communication for expressivists. *Ethics*, 126(3), 607–635.

Popa-Wyatt, M. (2016). Not all slurs are equal. *Phenomenology and Mind*, 11, 150–156.

Portner, P. (2007). Imperatives and modals. *Natural Language Semantics*, 15(4), 351–383.

Potts, C. (2005). *The Logic of Conventional Implicatures*. Oxford University Press.

— (2007). The expressive dimension. *Theoretical Linguistics*, 33(2), 165–198.

Price, H. (2011). *Naturalism Without Mirrors*. Oxford University Press.

Price, H., Blackburn, S., Brandom, R., Horwich, P., & Williams, M. (2013). *Expressivism, Pragmatism and Representationalism*. Cambridge University Press.

Putnam, H. (1975). The meaning of 'meaning'. *Minnesota Studies in the Philosophy of Science*, 7, 131–193.

Richard, M. (2008). *When Truth Gives Out*. Oxford University Press.

Roberts, D. (2011). Shapelessness and the thick. *Ethics*, 121(3), 489–520.

Roseman, I. J. (2018). Rejecting the unworthy: The moral psychology of contempt. In M. Mason (Ed.), *The Moral Psychology of Contempt* (pp. 107–130). Rowman and Littlefield.

Ross, J., & Schroeder, M. (2013). Reversibility or disagreement. *Mind*, 122(485), 43–84.

Saka, P. (1998). *Reference and Definite Descriptions*. MIT Press.

— (2007). How to do things with slurs: Studies in the way of derogatory words. *Language and Communication*, 27(3), 223–242.

Sbisà, M. (1999). Ideology and the persuasive use of presupposition. In J. Verschueren (Ed.), *Language and Ideology: Selected Papers from the 6th International Pragmatics Conference* (pp. 492–509). International Pragmatics Association.

Schlenker, P. (2003). A plea for monsters. *Linguistics and Philosophy*, 26(1), 29–120.

— (2007). Expressive presuppositions. *Theoretical Linguistics*, 33(2), 237–245.

Sennet, A., & Copp, D. (2015). What kind of a mistake is it to use a slur? *Philosophical Studies*, 172(4), 1079–1104.

Shanon, B. (1976). On the two kinds of presuppositions in natural language. *Foundations of Language*, 14(2), 247–249.

Simpson, R. M. (2013). Un-ringing the bell: McGowan on oppressive speech and the asymmetric pliability of conversations. *Australasian Journal of Philosophy*, 91(3), 555–575.

Soral, W., Bilewicz, M., & Winiewski, M. (2018). Exposure to hate speech increases prejudice through desensitization. *Aggressive Behavior*, 44(2), 136–146.

Stenner, A. J. (1981). A note on logical truth and non-sexist semantics. In M. Vetterling-Braggin (Ed.), *Sexist Language: A Modern Philosophical Analysis* (pp. 299–306). Littlefield, Adams & Co.

Stanley, J. (2015). *How Propaganda Works*. Princeton University Press.

Starr, W. (2020). A preference semantics for imperatives. *Semantics and Pragmatics*, 20, 6.

Strawson, P. F. (1950). On referring. *Mind*, 59(235), 320–344.

— (1954). A reply to Mr. Sellars. *The Philosophical Review*, 63(2), 216–231.

— (1962). Freedom and resentment. *Proceedings of the British Academy*, 48, 187–211.

Swim, J. K., Hyers, L. L., Cohen, L. L., & Ferguson, M. J. (2001). Everyday sexism: Evidence for its incidence, nature, and psychological impact from three daily diary studies. *Journal of Social Issues*, 57, 31–53.

Swim, J. K., Hyers, L. L., Cohen, L. L., Fitzgerald, D. C., & Bylsma, W. H. (2003). African American college students' experiences with everyday racism: Characteristics of and responses to these incidents. *Journal of Black Psychology*, 29(1), 38–67.

Tenchini, M. P., & Frigerio, A. (2020). The impoliteness of slurs and other pejoratives in reported speech. *Corpus Pragmatics*, 4(3), 273–291.

Tirrell, L. (1999). Derogatory terms: Racism, sexism and the inferential role theory of meaning. In K. Oliver & C. Hendricks (Eds.), *Language and Liberation: Feminism, Philosophy and Language* (pp. 215–232). SUNY Press.

— (2012). Genocidal language games. In I. Maitra & M. K. McGowan (Eds.), *Speech and Harm: Controversies over Free Speech* (pp. 174–221). Oxford University Press.

van der Sandt, R. (1992). Presupposition projection as anaphora resolution. *Journal of Semantics*, 9(4), 333–377.

von Fintel, K. (2004). Would you believe it? The king of France is back! Presuppositions and truth-value intuitions. In M. Reimer & A. Bezuidenhout (Eds.), *Descriptions and Beyond* (pp. 315–341). Oxford University Press.

von Fintel, K., & Gillies, A. S. (2008). CIA leaks. *The Philosophical Review*, 117(1), 77–98.

Väyrynen, P. (2013). *The Lewd, the Rude and the Nasty: A Study of Thick Concepts in Ethics*. New York, NY: Oxford University Press.

Whiting, D. (2007). Inferentialism, representationalism and derogatory words. *International Journal of Philosophical Studies*, 15(2), 191–205.

— (2013). It's not what you said, it's the way you said it: Slurs and conventional implicatures. *Analytic Philosophy*, 54(3), 364–377.

Williams, B. (1985). *Ethics and the Limits of Philosophy*. Harvard University Press.

Williamson, T. (2009). Reference, inference and the semantics of pejoratives. In J. Almog & P. Leonardi (Eds.), *The Philosophy of David Kaplan* (pp. 137–159). Oxford University Press.

Winiewski, M., Hansen, K., Bilewicz, M., Soral, W., Świderska, A., & Bulska, D. (2017). Contempt speech, hate speech: Report from research on verbal violence against minority groups. Stefan Batory Foundation.

Yablo, S. (2006). Non-catastrophic presupposition failure. In J. J. Thomson & A. Byrne (Eds.), *Content and Modality: Themes From the Philosophy of Robert Stalnaker* (pp. 411–436). Oxford University Press.

Zeman, D. (2022). A rich-lexicon theory of slurs and their uses. *Inquiry*, 65(7), 942–966.