

# THE DIRECT REFERENCE THEORY OF PEJORATIVES IN HATE SPEECH

Kanit (Mitinunwong) Sirichan  
Chulalongkorn University, Thailand

*The use of language in hate speech is understandably offensive. Though words do not kill, they convey an alarming message that can harm the victim. To understand how words can harm, it is necessary to understand the nature of the meaning of pejoratives or slurs that are used in hate speech. Pejoratives are undeniably offensive. However, they are puzzling as they can be used in two directions, namely, the offensive power preservation and the offensive power destruction. This paper proposes that the direct reference theory of pejoratives can solve the puzzle. A characterization of pejoratives is that it has the property of immediacy. They refer directly to the object of speech. Grounding on a shared context, any descriptions are unnecessary for understanding the offensive message of pejoratives. In this sense, pejoratives have indexical content as it is context-sensitive. The kind of indexical content that pertains to pejoratives is action-oriented. However, its object of reference is empty. In discussing pejoratives that are used in hate speech, some examples of Thai slurs are shown.*

## INTRODUCTION

The use of language in hate speech is understandably offensive. Though words do not kill, they convey an alarming message that can harm the victim. In order to understand how words can harm, it is necessary to understand the nature of the meaning of pejoratives<sup>1</sup> that are used in hate speech. The main question of this paper is what the reality of the meaning of pejoratives that are used in hate speech is. Answering this question requires an explanation for the offensive power of slurs in hate speech, though it might not straightforwardly lead to the practical solution of the problem of hate speech as a sociopolitical issue. The offensiveness of pejoratives is puzzling as pejoratives can be used in two directions, namely, the offensive power preservation and the offensive power destruction. In this paper, I propose that the direct reference theory of pejoratives can solve the puzzle. A characterization of pejoratives is such that it has the property of immediacy. They refer directly to the object of speech. Grounding on a shared context, any descriptions are unnecessary for understanding the offensive message of pejoratives. In this sense, pejoratives have indexical content as it

is context-sensitive. The kind of indexical content that pertains to pejoratives is action-oriented. However, its object of reference is empty. In the following, I first clear up the definition and scope of hate speech study relevant to this paper and show some examples of the use of Thai slurs in hate speech. The second part is a formulation of what is called the pejoratives-puzzle. Then the third part briefly shows some solutions to the puzzle and provides objections to such solutions. In the last part, I argue for my own solution: the direct reference theory but with a paradoxical claim that the reference of slurs in hate speech is empty; however, slurs have intentional content.

## WHAT IS HATE SPEECH? DEFINITION AND SCOPE WITH SOME EXAMPLES OF THAI SLURS

Literally, hate speech is the language of hate. According to the United Nations (United Nations 2019), there is no international legal definition of hate speech. However, in a legal context, such as in the US context, hate speech is characterized as linguistic or any form of expression (including other types of representation like pictures, films) that expresses hatred toward individuals or groups on the basis of race, ethnicity, gender, sexual orientation, religion, age, disability, or any other similar ground; in particular, it can instigate violence or harm target person or group.<sup>2</sup> According to Alexander Brown (2017, 424), the term ‘hate speech’ was initially employed in the US legal context concerning the case of racial discrimination. The issue in the legal debate is mainly on whether hate speech should be banned or censored in order to comply with the liberal democracy ideology where free speech is protected. In the legal context of Thailand, the applicability is not much different. (see e.g. Pirongrong Ramasutra 2015). According to Ramasutra, ‘hate speech’ is a new terminology for Thai society as the term has been formally translated by the Office of the Royal Society of Thailand as ‘speech that harms’ (*Pratussa Vaja*) around 2014. Moreover, hate speech is not regulated under the current constitution of the Kingdom of Thailand due to the protection of free speech.<sup>3</sup>

However, it is not necessary to think about the issue of hate speech within a legal context. An example is in the work of Jennifer Hornsby (2003, 1), who takes the study of hate speech as a study of a ‘principle of political morality,’ namely, a study of the moral dimension of hate speech in a political context. Some others, like Brown (2017, 422), have also remarked that the study of hate speech mainly grounding their account on the legal concept may not help to see what lies behind such concept. A study of an ordinary concept of hate speech is more promising than a legal one. It shows the myth behind speech that is used as a weapon against disapproval moral/political codes of conduct. One of the myths that Brown mentions is the myth of hatred as an emotional state.

Whatever approach is adopted, the debate seems to rest on the fluid nature of hate speech that fluctuates between speech that expresses hatred toward moral/political disapproved conducts and speech that harms or can provoke violence toward its target groups. However, it is possible to express hate speech without a feeling of hatred. So, the reason we care about hate speech seems to be concerned with the harm in hate speech rather than the hatred itself. If that is the case, it is

reasonable to reduce the scope of the talk about hate speech into the talk about harm in hate speech instead. Based on this understanding, my focus is on the component of hate speech that can be an instigator of the harm, namely, the offensiveness of pejoratives or slurring terms that are used in hate speech.

Concerning examples of hate speech in Thailand, they are expressed in various forms, including pictures and films<sup>4</sup>. However, this paper focuses on hate speech as being expressed in pejoratives or derogatory language. The hate speech phenomena in Thailand have been of much interest, especially during the years of political turmoil (around 2008), where hate speech was heavily employed by politicians and people from both opposite political views. The popular slurs in Thailand can be divided into three groups – political, gender, and racial slurs. Though religious slurs are also controversial, it seems they are less noticeable compared with other types of slurs and can be, in some cases, incorporated with racial slurs. However, within each group, there is also a difference between the non-pejorative use in hate speech and the pejorative use in hate speech. I focus on the type of pejoratives use in this paper. The obvious example of such a difference can be seen in political slurs. In the following, I show some examples of Thai slurs using in hate speech.

## Political slurs

### 1. The non-pejorative use in hate speech

The non-pejorative use can be in the form of poetic/metaphorical language.<sup>5</sup> A well-known example is *nàk pàen dīn* (“หนักแผ่นดิน”). This is the name of a controversial patriotic song that means useless/worst people of the land. The song was meant to blame the student movement in 1976 as being communists and has been believed to be the song that provoked brutal violence against the student movement on October 6, 1976, at Thammasat University. Once again, the song has been recited by the former Army Chief General Apirat Kongsompong to respond to the criticism of the Army budget spending by the Future Forward Party, a new progressive party.

Another famous example is *'kwayy-dang* (ควายแดง); its literal meaning is ‘red water buffalo’ which connotes the ignorant poor people who are led or bribed by political tycoons. The word *'kwayy* in Thai denotes buffalo but connotes the derogatory meaning of being stupid or an idiot. There are many pejoratives in the Thai language that use animal epithets or animal metaphors as an insult. The offensive sense of the word ‘buffalo’ might be equal to ‘rat’ in the English language.<sup>6</sup>

### 2. The pejoratives use in hate speech. Here is an example.

*'e-ngoh* (อีโง่) or ‘dumb bitch.’ This slur was used by the former Thai male prime minister to blame the female prime minister during the political turmoil in 2013 (see Bangprapa 2013). The prefix ‘e’ (อี) in the Thai language is an impolite prefix for female, and ‘ai’ (ไอ้) is for male. Though the term was uttered in a political context in order to discredit the female prime minister as a woman with low intelligence and undeserving of the position of the prime minister, it is criticized as the sexist slur being reproduced to subordinate women.

## Gender or sexist slurs

Here are some examples.

‘*e-ka-ree*’ (อีกะเรี) or ‘bitch/slut’: The word ‘*ka-ree*’ is believed to be from ‘curry’ which has a strong smell. Though there is no consensus on the origin of the term, the term is pejorative as it is an insult to a woman’s dignity.

‘*ai-ka-toey*’ (ไอกะเทย) or ‘faggot/fag’: The word denotes transgender people or homosexuality people, but connotes the abnormality of male/female sexuality, which is offensive.

## Racial slurs

Here are some examples.

‘*ai-jeak*’ (ไอเจ๊ก) or ‘chink’: The literal meaning denotes the Thai Chinese but connotes rude, inconsiderate manner. Though this term has been currently considered as less offensive, there is still the sense of an insult if it is used as name-calling.

‘*ai-lao*’ (ไอลาว): The literal meaning denotes the Laotian but connotes the humiliating sense of an inferior race to the Thai.

‘*ai-yun*’ (ไอยูน) or ‘Jap’: The literal meaning denotes the interference or nosy people, but the connotation is the Japanese with the satirical sense of being nosy on international affairs. Around the sixties or seventies in Thailand, it is one of the familiar racial slurs. It was the time when there was a sentiment of protesting against Japanese merchandise and probably the anti-Japan sentiment inherited from the Second World War when the Japanese army invaded Thailand.

## PEJORATIVES-PUZZLE: THE OFFENSIVE POWER

As shown above, there are two main components of slurs: the literal/denotative or descriptive component and the connotative or derogatory/offensive component.<sup>7</sup> The phenomenon of pejoratives-use in hate speech is puzzling. The puzzle concerns the latter component. To understand the pejoratives-puzzle, it is necessary to understand the background idea of the puzzle, that is, the difference between the use and the mention or the use-mention distinction thesis.

In the seminal work of Gottlob Frege (1892), there is a distinction between ‘sense’ and ‘reference.’ An expression does not merely have its reference as its meaning, but it also expresses meaning as sense. In grasping such a distinction, it is important to understand the difference between the use and the mention of a term, which can be called the use-mention distinction thesis. Whereas an expression is used *if and only if (iff)* it is about the reference directly, an expression is mentioned *iff* it is about the reference indirectly. In ordinary reference, a proper name in general like “Taro Mochizuki” is used in the sentence, say,

(1) Taro Mochizuki is a Descartes scholar.

In the case of an indirect reference, a proper name is mentioned, say,

- (2) "Taro Mochizuki" in the sentence (1) means Taro Mochizuki who is a Descartes scholar.  
 Or (3) Kanit said that Taro Mochizuki is a Descartes scholar.

It is obvious that general proper names operate under the use-mention distinction thesis. However, in the case of pejoratives or slurring terms, they strangely act as they violate the thesis, and this is what I call the pejoratives-puzzle. There are two ways of the breakdown of the use-mention distinction thesis. The first one is the offensive power preservation. The second one is the offensive power destruction.<sup>8</sup>

### The Offensive Power Preservation

There are at least three contexts of slurs use, showing that the offensiveness of pejoratives is preserved, even though they are used in contexts that seem to cancel out the offensiveness. A reason why slurs are offensive even in such contexts is that the nature of slurs does not conform to the use-mention thesis. Such contexts are the context of negative sentences, the context of indirect speech, and the co-referential context.

(1) Negative sentences: even though a slur is negated, the term is still offensive, e.g.

‘She is not a slut.’

‘He is an Afro-American, but he does not behave like a nigger.’

‘She is poor, but she is not a slut.’

(2) Indirect speech: even though a slur is not used directly, the term is still offensive, e.g.

‘He is a honkey.’ – ‘Kanit said that he is a honkey.’

(3) Co-referential context: even though a slur has the same reference as its descriptive counterpart, an inference of the co-referring term is still offensive.

1. He is Caucasian.

2. honkey = Caucasian

3. He is honkey.

### The Offensive Power Destruction

This is the melting situation of the offensive power where ‘mentioning’ the slur can be used as a destruction of the offensive power. However, it is not an ordinary ‘mention’ or the use of the term in an indirect reference context. There are two cases falling into this type, namely, the appropriation case (or the reclamation case) and the case of humor.<sup>9</sup> In the appropriation case – slurs are ‘reused’ or reclaimed by the victim’s group in order to fight back the offensive sense, e.g., the Boston Slut Walk in 2011. The offensiveness of slurs is destroyed only if slurs are used within the victim’s group. In the case of humor, slurs are often used in the context of making jokes among

trusted friends to express a sense of camaraderie.<sup>10</sup> In this case, the offensiveness is canceled out, though there is a thin line here between the sense of humor and the sense of humiliation or disrespect. However, this case is rather complicated and needs to be treated separately, so it will not be considered in this paper.

### **An Account of Pejoratives-Puzzle in Terms of Speech Acts**

It might be illuminating to understand the pejoratives-puzzle in terms of the illocutionary force of speech-acts. Slurs are considered to be illocutionary acts or a speech that performs an act that can induce effects on its hearer. (see e.g., Jennifer Hornsby and Rae Langton 1998; Rae Langton 2010). According to Austin (1962), there are three types of speech acts: locutionary acts, illocutionary acts, and perlocutionary acts.

1. The locutionary act is simply an act of “saying something” (John Langshaw Austin 1962, 94). An example by Austin is this. When someone utters a sentence like “Shoot her!”, she performs a locutionary act where the meaning of each term is just its reference.
2. The illocutionary act is an act of uttering an expression with a particular force; for example, in saying “Shoot her!” the speaker orders (or urged or advised) her hearer to shoot the woman. There is the illocutionary force of command in saying that sentence. The illocutionary force constitutes the meaning of the expression grounded on the speaker’s intention plus the external authority like social convention justifying the speaker’s intention. For the example, “Shoot her!”, if the speaker has the intention to command the hearer to shoot the woman but lacks the authority to command, her expression would fail to be performed, or in Austin’s terms, her expression is ‘abused.’
3. The perlocutionary act is an act that is performed by saying a locution, and it has real effects on her hearer, which can be either mental or physical effects. For example, the speaker who says, “Shoot her!” is succeeded in persuading her hearer so that the hearer picks up the gun and shoots the woman, or the speaker may shock her hearer. (see Rae Langton 2009, 28) (for the example “Shoot her!” see John Langshaw Austin 1962, 101).

Although Austin did not directly talk about pejorative language, he mentioned the case of insulting. In saying the sentence like “I insult you.”, the speaker is not performing a speech act of insulting because the term ‘insult’ in itself does not convey the force of insulting. But in saying the sentence like “You were cowardly.”, the speaker is performing an illocutionary act with the force of insulting (see John Langshaw Austin 1962, 30). Based on this understanding, pejoratives use in hate speech can be analyzed in the following way.

In uttering an expression with slurs, the speaker is performing an illocutionary act with a particular force, such as the force of insulting the whole group the victims belong to. However, it is puzzling. In the non-appropriation contexts, it is obvious that the illocutionary force of insulting is preserved. Even though one merely mentions a

slur, it is still offensive because the illocutionary force is partly constituted by an external factor that is independent of the speaker's intention. The external factor constituting the offensive component of slurs can be a social practice. However, in the appropriation contexts, the offensiveness is destroyed. The slurs are granted a new meaning or the second-order meaning by the victim's group to show that the victim's group is aware of the offensiveness of the first-order meaning. In this context, the illocutionary force of insulting is destroyed and is substituted by the illocutionary force of, probably, self-empowerment.

## PEJORATIVES-PUZZLE SOLUTIONS: SEMANTIC THEORY VS PRAGMATIC THEORY

Two types of theory suitable for solving the pejoratives-puzzle are the semantic theory and the pragmatic theory. The former treats meaning as truth-conditional bound, and the latter treats meaning as performatives. With regards to the semantic theory, there are two promising theories, i.e., the social externalist semantic theory and the indexical semantic theory. In the following, I briefly explain and criticize each theory and then propose an alternative for the solution.

### The Semantic Theory

#### 1. The Social Externalist Semantic Theory

According to this theory (see Christopher Hom 2008, 8), the meaning of a slurring term in a sentence is determined by its context of use. The kind of context in question is the social context playing a causal role in determining the content of the term. For example, for a racial slur to be offensive, it is the external social context like "institutions of racism" that causally determines the derogatory/offensive content of the slur.

Although this theory seems to be intuitive, a major criticism of this theory is that it is too broad to explain the non-derogatory use situation, e.g., the case of appropriation. The reason is that there is nothing like an institution of appropriation to guarantee the non-offensive sense of the appropriation. So, this theory cannot explain the case of offensive power destruction.<sup>11</sup>

#### 2. The Indexical Semantic Theory

The indexical semantic theory (see David Kaplan (2004); Michael Scott and Graham Stevens (2019)) is the most promising semantic theory, but it still requires some revisions in order to fully grasp the nature of slurs. According to this theory, the meaning of a slurring term in a sentence depends on the context of use in a similar way as the meaning of indexical terms, like, *I*, *this*, *here*, *now*. Such terms are called context-sensitive terms because the meaning is determined by the world or the context of reference. For example, the sentence like 'I am here.' refers directly to its object of reference. Understanding the meaning does not require a description of who 'I' am or where 'here' is. However, if someone else says the same sentence as what I said, 'I am here.', the truth condition of the sentence is varied to the reference, that is, the person

who said that sentence and the place she is in. So, the truth condition of the sentence or its meaning is determined by the world or the context of use. This theory is sometimes known as the direct reference theory. Within the indexical semantic theory, two preliminary approaches should be in focus.

The first one is an influential work of Kaplan (2004). The paper considers pejoratives as expressives. An expressive term expresses/displays/shows whether something is the case. This is similar to how indexicals/demonstratives, pejoratives work as well. It is to be contrasted with a descriptive term that describes/says what the case is. However, the main limitation of Kaplan's view is that – in saying that pejoratives are expressives, it entails that only the case of pejoratives that express contempt is in consideration. However, pejoratives can be used either with contempt or without contempt. So, Kaplan's approach does not seem to be an appropriate solution to the pejoratives-puzzle.

The second approach is the (racial) pejoratives as indexicals written by Scott and Stevens (2019).<sup>12</sup> According to this approach, racial pejoratives function like indexical terms; that is, they are context-dependent. A context (c) is defined by an agent (a), world (w), time (t), location (l) and target (u), and formally written as “A context  $c = \langle ac, wc, tc, lc, uc \rangle$ .”

The meaning of pejoratives in a sentence is determined by truth conditions, not by social facts surrounding the use of the term. The meaning of pejoratives is truth-conditional bound in the sense that (racial) pejoratives pick out or refer to a racial group. Racial pejoratives negatively represent the target agent just because of the target being a member of that group. For a racial pejorative like ‘*Ai-Jek*,’ the reference of the term is the Thai-Chinese group, and the offensive sense of the term is caused by the negative representation of the group. As Scott and Stevens (2019, 396-7) say, “..racial pejoratives always encode their derogatory meanings ...racist meanings.” In other words, the fact that racial pejoratives are offensive or express racist judgments “..is a straightforward semantic (truth-conditional) fact” (Scott and Stevens 2019, 396-7). The indexical semantic theory is thus able to account for the offensive preservation case.

As for the case of the offensiveness destruction like appropriation, this theory explains that in the appropriation context, the racial pejoratives do not have standard truth-conditional content. It is the context in which a positive attitude of camaraderie or empowerment is expressed (Scott and Stevens 2019, 395). That is the reason why the offensiveness is destroyed. This is quite similar to what Katherine Ritchie (2017, 158) said that the appropriation expresses the social identity of the plural first-person indexicals “we.” Although the content of racial pejoratives in this context is non-standard truth-conditional, it is still truth-bound as it primarily refers to a racial group, but the identity of the speaker melts its offensiveness.

It seems the indexical theory provides a sufficient solution to the pejoratives-puzzle. Despite this, the theory overlooks a crucial aspect of indexical terms: indexicals are action-oriented terms. This aspect can be called the ‘performative’ aspect of indexicals. A well-known example for elaborating this point is the trolley shopper's thought experiment from John Perry (1979, 33). The story is this. Suppose you are shopping in a supermarket with a trolley cart. Suddenly you see a trail of sugar on the floor, so you think someone has a leaky sugar bag, and you would like to tell that



person not to make a further mess; but then after you follow the trail for a while, you found that the trail is coming from you, and you finally realized that it is 'you' who made the mess. The moment you realize that it is you or 'I' who makes the mess, there is an immediate consequence, for instance, an act of trying to clear up the floor. The sense of 'I' is different from 'the person who makes the mess.' There is no action-oriented aspect of the latter.

According to the indexical theory, pejoratives like racial or sexist slurs have a similar nature as indexicals. If it is correct that indexicals are action-oriented terms, then we cannot overlook this action-oriented component in such pejoratives. There is an immediate effect in expressing a slur, i.e., the offensive effect, which can harm or instigate a violent response. Such an effect can be taken as the performative acts of hate speech. As for the question of how the performative acts work requires an explanation basing on the pragmatic theory.

### **The Pragmatic Theory of Pejoratives**

The pragmatic theory in consideration here is the view proposed by Jennifer Hornsby (2001, 2003), Jennifer Hornsby and Rae Langton (1998), Rae Langton (2009, 2012). According to the pragmatic theory basing on Austin's theory of speech acts, as explained in the above section, language or speech is viewed as communicative where both speaker and hearer play an essential role. Pejoratives that are used in hate speech are illocutionary acts, namely, the performative of an utterance with a particular force that constitutes the meaning of the utterance. As regards the pejoratives use in hate speech, in saying 'e-ka-ree' or 'slut', you are performing an act of insulting women.

According to Rae Langton (2009), in uttering racial pejoratives, one is performing an act of racial discrimination. In uttering sexist slurs, one is performing an act of subordination. The offensive component of slurs, e.g., racial discrimination, the subordination of others, is their illocutionary forces. In saying, "She's a slut!" the speaker is performing an illocutionary act with the force of subordinating women. Even though it is possible that one has no intention to insult women as a whole, the offensive component of slurs just does the work. The external factor, say a social practice of suppressing women, constitutes the offensive component of slurs. For Langton (2018), an external authority supporting the offensive component is actually the bystanders who keep silent in front of the use of hate speech. Not protesting or complaining in such a situation is like granting authority to the speakers of hate speech to use and reuse slurs with offensive force.

There is a salient point that Hornsby and Langton emphasize. It is what is called the silencing thesis or the illocutionary silencing; namely, the performative of an illocutionary act is not possible. Originally, Hornsby and Langton's example concerns the effect of pornography on the subordination of women. As pornography creates a false belief among men that women saying 'no' when being asked for sex would mean 'yes,' the illocutionary act of the word 'no' uttered by a woman in that situation is silenced or cannot be performed.<sup>13</sup> In the context of hate speech, the silencing thesis can be explained on the assumption that there is an asymmetry of power relation among speakers so that a speech of the less privileged would be silenced by the more

privileged. The more privileged have words that can silence the speech of the less privileged. This explains why there are no slurs of an equal offensive power like 'slut'. This phenomenon is also true in Thai hate speech.<sup>14</sup>

However, a negative implication of the silencing thesis is that it silences the possibility of the offensiveness destruction as well. That is because the presupposition of the thesis is the unavailability of the counter speech of equal offensiveness. In comparison, the case like appropriation suggests the availability of the counter speech, though with the same word but different sense. It is the situation where slurs can be used without being offensive. The slurs are granted new meaning or the second-order meaning by the victim's group in order to show that the victim's group is aware of the offensiveness. In this context, the first-order meaning or the offensive meaning is rejected by the victim's group. If that is the case, the pragmatic theory does not seem to offer the solution to the pejoratives-puzzle. Hence, we need an alternative theory that is able to solve both puzzling directions, that is, the direct reference theory.

## THE DIRECT REFERENCE THEORY

The direct reference theory is a combination of the indexical semantic theory and the pragmatic theory but encounters no problems as both theories. It proposes two paradoxical claims: first, the indexical component of pejoratives, and second, the empty reference of hate speech.

The indexical component of pejoratives that are used in hate speech means that the meaning of pejoratives is context-dependent and truth-bound. For example, the truth-condition of the sentence containing a sexist slur like '*e-ka-ree*' is determined by the context in the Thai language; that is, it refers 'immediately' (or without descriptions) to a group of women who are prostitutes or sexually promiscuous. The offensive sense of the term is already inhered in its content because it is an expressive term or the term which expresses offensive attitudes. The offensive attitudes are expressed or performed in uttering the slur, in this case, the act of insulting, humiliating a woman, which can be considered as an "attack on human dignity" (see Jeremy Waldron 2012, 5). This is the action-oriented aspect that can explain why the term harms the target.

To support my point here, according to Christopher Potts (2007, 166), one of the six characteristics of pejoratives or expressive terms is immediacy. Potts (2007, 167) says, "Immediacy: Like performatives, expressives achieve their intended act simply by being uttered; they do not offer content so much as inflict it." This feature shows that expressives or pejoratives can be treated like performatives. Once performed, there are consequences. This is the idea that expressives behave like indexicals and demonstratives. It is in line with the realization of 'I' who makes the mess in Perry's example of the trolley shopper.

The second claim is paradoxical. Though pejoratives are used in hate speech has a direct reference, the reference is empty. Primarily, this is supposed to be the solution to the pejoratives- puzzle concerning the appropriation context. The main reason is that the use of slurs in hate speech is a kind of generic use. It is a weak generalization of the property of a group, which does not exist. This can be said as what Hom and May

(2018, 299) suggest that the truth-conditions of a sentence containing slurs is a kind of fictional truth. For example, in ascribing the property of being a slut to a group of women, it refers to the property like casual or promiscuous sexual behavior which is deemed by those who are morally conservative as the negative or disapproval property. But such behavior is deemed as a negative property in the eyes of one group against another. Being a 'slut' is created to exist as the reference of the term; similarly, 'unicorn' is created to exist as the reference of the term 'unicorn'. In this sense, the truth-condition of slurs is simply a fictional truth.

Donald Davidson's idea of metaphors may help illuminate the empty reference of pejoratives in hate speech. For Davidson (1978), metaphors do not have propositional content. They convey a sense of perspectival content or a way of 'seeing as,' not 'seeing that.' The concept of 'seeing as' can be portrayed in gestalt pictures, e.g., the duck-rabbit picture, where one can see the picture as a duck or as a rabbit. The content of what one sees is merely perspectival. The use of slurs or pejoratives in hate speech can be taken as metaphorical language in Davidson's sense. The content of a sentence containing a slur like 'She is a slut!' is mere perspectival content. It is a judgment from a perspective of a group toward another. There is no propositional content as the reference is empty.

However, even though the reference of slurs is empty, this does not mean that such terms are meaningless. Under the condition of sharing the same language, the language users know what they mean when they express a slur. At least, there is what can be called 'intentional content.' (see Sirichan and Lepore (2015)) The intentional content is the object of thought in the Meinongian sense; that is, being in an intentional state involves what the state is about. In thinking about a unicorn, the intentional content of your thought is the unicorn. Whether the unicorn exists outside your thought is not relevant because, in thinking, there is, necessarily, an object of thinking. Regards the semantics of pejoratives or slurs in hate speech, slurs can be expressed without hatred as an intentional state. However, they are intentionally created, slurs hence have at least intentional content, though without existed object of reference. An upshot of the second claim is that it can be an explanation for the case of appropriation. In that situation, the offensive sense of the appropriated slur (or the second-order slur) is destroyed, but still, the speaker who appropriates the slur is aware of the first-order meaning, which is offensive. The empty reference of slurs allows the possibility of the use without the offensive sense to be fixed. Therefore, the direct reference theory is the most plausible solution to the pejoratives puzzle in both directions.

There might be an objection that some pejoratives are not clearly defined by their context of use, e.g., 'pussy', which may refer to cowards or to prostitutes. And there are also some pejoratives that do not automatically have pejorative connotations, e.g., 'pig,' which may mean the pig as an animal or connote the sense of being lazy or dirty.<sup>15</sup> Such pejoratives are the case of words with double meaning. However, the objection cannot be applied here as I have stated at the outset that pejoratives used in hate speech are indeed grasped on the condition that speakers and hearers must share a common background of language and social norms surrounding the use of the language. In particular, the hate speech speakers usually intend to make explicit that the pejoratives they use have pejorative connotations. They need to make sure that the pejoratives they use convey the force of insult. Otherwise, hate speech would be powerless.

## CONCLUSION

As the main aim of this paper is to answer the question about the reality of the meaning of pejoratives or slurs in hate speech, the direct reference theory provides the analysis that can cope with the puzzle of pejoratives. The pejoratives-puzzle is puzzling in both directions, that is, the offensive power preservation and the offensive power destruction. Two main theories are suitable to solve the puzzle, namely, the semantic and the pragmatic theory. On the one hand, among the semantic theories on offer, the social externalist semantic theory can account for the former direction but not the latter; the indexical semantic theory, in particular, the one which treats pejoratives as indexicals, is more promising, but it is insensitive to the performative aspect of indexicals. On the other hand, the pragmatic theory is realistic in treating speech as performatives – in expressing a slur; one is doing harm to the target by attacking the human dignity of others. Another strong thesis of the pragmatic theory is the silencing thesis, which can account for different degrees of the offensiveness of slurs, such that the degree of offensiveness is relative to the power status of the speaker. However, the silencing thesis rules out the appropriation case where an expression of a slur can be non-offensive. Accordingly, I propose the direct reference theory to solve the puzzle in both directions, though its claim sounds paradoxical.

## NOTES

1. There is a complication concerning the use of the term ‘pejorative’ and ‘derogatory.’ Pejoratives target individuals, not groups, e.g., ‘stupid’; but derogatory terms target groups, not individuals, e.g., slut, nigger. (See Bianchi 2014). Some literatures on hate speech focus on ‘derogatory’ terms in hate speech (e.g., Bianchi 2014, Langton 2012, Hornsby and Langton 1998), but some literatures do not make a clear distinction. For example, Scott & Stevens 2019 talks about racial pejoratives expressing derogatory content (See also Corredor 2014). In this paper, I follow Scott and Stevens’s use of pejoratives as I understand that derogatory remarks or epithets are a kind of pejorative.

2. See, e.g., “American Library Association,” 2017; “USLEGAL.COM,” 2020; see also Jennifer Hornsby (2003) for a sufficient condition of hate speech.

3. See, e.g., Constitution Drafting Commission (2017). Please note that the limitation of free speech in Thailand is somewhat controversial. Instead of being used to protect the lesser privileged groups, the term ‘hate speech’ has been frequently used in the opposite. (See Smithi Skunnawat (2020).

4. A good example of hate speech against women through the medium like pictures and films is pornography. (See, e.g., Rae Langton 2000) for a philosophical study on pornography.

5. Another example is the phrase “Khae Pha Chet Toa” (“แค่ผ้าเช็ดเท้า”) or ‘just a foot rag.’ This phrase, though is not a well-known phrase; it is an excellent example of non-pejorative use or metaphorical use in hate speech. I found these terms from the right-leaning women’s magazine Sakulthai. The full sentence is actually “The red shirt

people are ‘just a foot rag.’” The phrase ‘just a foot rag’ is meant to humiliate the red shirt movement as being a worthless lower class. It is offensive as Thai consider the foot as the dirtiest or demeaning part of a body.

6. A similar example in the English language is an offensive calling of the French as ‘frog.’

7. Noted here that the similarity of derogation and offensiveness is taking for granted. According to Christopher Hom and Robert May (2018, 118), there is a distinction between derogation and offensiveness – while the former is semantic, the latter is psychological, which can be subjective. However, this is not a substantial issue here, for though the offensiveness seems to be subjective, there is a sense that it is context-dependent. If we share the same context (same language), we share the sense of offensiveness in the use of slurring term.

8. In order to communicate with international hearers, the following examples of slurs are mainly slurs in English. (See also Kanit Sirichan and Ernest Lepore 2015).

9. Bianchi (2014)’s analysis of non-offensive use is quite close to mine. She puts the non-offensive use in two kinds of contexts, i.e., friendship contexts and appropriation contexts, and calls both as community uses. However, Bianchi’s aim is to defend the semantic and pragmatic account of slurs or the content view against the deflationary view (Anderson and Lepore 2013). As my paper does not discuss the deflationary view, I will not engage with Bianchi’s proposal.

10. Also, in a game of playing with words or a pun.

11. See some other criticism of Hom’s theory, e.g., Michael Scott and Graham Stevens (2019), Christina Corredor (2014).

12. Though Scott and Stevens’s main concern is racial pejoratives; it is applicable to other types of pejoratives as well.

13. See the debate on the illocutionary silencing in Alexander Bird (2002), Mary Kate McGowan, Alexandra Adelman, Sara Helmers, and Jacqueline Stolzenberg (2011).

14. In English, the terms like ‘honkey’ or ‘womanizer’ do not have a comparable offensive sense as ‘slut’. An obvious example of the asymmetry in Thai hate speech is the political slur like ‘kwai dang’ is not equally offensive to the opposite slur ‘Sa-lim’, which connotes the yellow shirt people or the royalist.

15. I agree with the anonymous referee for this objection.

## REFERENCES

- American Library Association. 2017. *Hate speech and hate crime*. Available at <http://www.ala.org/advocacy/intfreedom/hate>. Accessed: September 11, 2017.
- Austin, John Langshaw. 1962. *How to do things with words*. Edited by J.O. Urmson and M. Sbisà. Oxford: Oxford University Press, 2<sup>nd</sup> edition 1975.
- Bangprapa, Mongkol. 2013. Abhisit under fire for ‘dumb bitch’ jibe. *Bangkok Post* [online]. 13 September, 01.07. Available at <https://www.bangkokpost.com/thailand/politics/369440/dumb-bitch-protests-dog-abhisit>. Accessed: September 13, 2019.

- Bianchi, Claudia. 2014. The speech acts account of derogatory epithets: Some critical notes. In *Liber Amicorum Pascal Engel*. Edited by J. Dutant, D.Fassio, & A. Meylan. Geneva, Switzerland: University of Geneva. 465-480.
- Bird, Alexander. 2002. Illocutionary silencing. *Pacific Philosophical Quarterly* 83: 1-15.
- Brown, Alexander. 2017. What is hate speech? Part 1: The myth of hate. *Law and Philosophy* 36: 419-468. Available at <https://link.springer.com/article/10.1007/s10982-017-9297-1>. Accessed: September 22, 2019.
- Constitution Drafting Commission. 2017. *Constitution of the Kingdom of Thailand (B.E.2560)*. Available at [https://cdc.parliament.go.th/draftconstitution2/ewt\\_dl\\_link.php?nid=1460&filena me=index\\_](https://cdc.parliament.go.th/draftconstitution2/ewt_dl_link.php?nid=1460&filena me=index_). Accessed: October 1, 2020.
- Corredor, Cristina. 2014. Pejoratives and social interaction. In *Issues in Philosophy of Language and Linguistics*. Edited by P. Stalmaszczyk. Lodz: Lodz University Press. 39-55.
- Davidson, Donald. 1978. What metaphors mean. *Critical Inquiry* 5 (1), Special Issue on Metaphor, 31-47.
- Frege, Gottlob. 1892. On Sinn and Bedeutung. In *Frege Reader*. Edited by Michael Beaney. 1997. Oxford: Wiley-Blackwell, 151-171.
- Hom, Christopher and Robert May. 2018. Pejoratives as fiction. In *Bad Words: Philosophical Perspectives on Slurs*. Edited by David Sosa. Oxford: Oxford University Press, 108-131.
- Hom, Christopher. 2008. The semantics of racial epithets. *The Journal of Philosophy* 105: 416-440.
- Hornsby, Jennifer and Rae Langton. 1998. Free speech and illocution. *Legal Theory* 4: 21-37.
- Hornsby, Jennifer. 2001. Meaning and uselessness: how to think about derogatory words. *Midwest Studies in Philosophy* 25: 128-141.
- Hornsby, Jennifer. 2003. Free speech and hate speech: Language and rights. In *Normativita Fatti Valori*. Edited by Rosaria Egidi and M. dell'Utri and M. de Caro. Macerata, Italy: Quodlibet. 297-310. Available at <https://eprints.bbk.ac.uk/158/1/hornsby7.pdf>. Accessed: September 3, 2019.
- Kaplan, David. 2004. The meaning of ouch and oops. In *Unpublished Transcription of the Howison Lecture in Philosophy at UC Berkeley, transcribed by Elizabeth Coppock*. Available at <http://eecoppock.info/PragmaticsSoSe2012/kaplan.pdf>. Accessed: June 12, 2020.
- Langton, Rae. 2012. Beyond belief: Pragmatics in hate speech and pornography. In *Speech and Harm: Controversies over Free Speech*. Edited by McGowan and Maitra. Oxford: Oxford University Press, 72-93.
- Langton, R. 2009. *Sexual solipsism: Philosophical essays on pornography and objectification*. Oxford: Oxford University Press.
- Langton, Rae. 2018. The authority of hate speech. In *Oxford Studies in Philosophy of Laws Volume 3*, Oxford University Press, 123-152.
- McGowan, Mary Kate, Alexandra Adelman, Sara Helmers and Jacqueline Stolzenberg. 2011. A partial defense of illocutionary silencing. *Hypatia* 26 (1): 132-149.

- Perry, John. 1979. *The problem of the essential indexical and other essays*. Oxford: Oxford University Press.
- Potts, Christopher. 2007. The expressive dimension. *Theoretical Linguistics* 33 (2):165-198.
- Ramasutra, Pirongrong. 2015. *Pratussa vaja (hate speech) buengton samrap sangkom Thai* (An Elementary Introduction of Hate Speech for Thai Society). Available at <https://www.tcijthai.com/news/2015/11/archived/5385>. Accessed: August 5, 2020.
- Ritchie, Katherine. 2017. Social identity, indexicality, and the appropriation of slurs. *Croatian Journal of Philosophy* 17 (50): 155-180.
- Scott, Michael and Graham Stevens. 2019. An indexical theory of racial pejoratives. *Analytic Philosophy* 60 (4): 385-404.
- Sirichan, Kanit and Ernest Lepore. 2015. "In Hate Hated": Hate speech and its content. In *Proceedings of the International Symposium "Philosophy in Dialogue: Bridging the Great Philosophical Divides, 26-28 March 2015 Department of Philosophy, Chulalongkorn University, Thailand*. Edited by Suwanna Satha-Anand, Kanit Sirichan and L. Skar. Bangkok: Chulalongkorn University. 55-70.
- Skunnawat, Smithi. 2020. *Opinion: Free speech vs. hate speech in Thailand and why it is important*. Thai Enquirer. September 9, 2020. Available at <https://www.thaienquirer.com/18010/opinion-free-speech-vs-hate-speech-in-thailand-and-why-it-is-important/>. Accessed: October 12 2020.
- United Nations. 2019. *United Nations strategy and plan of action on hate speech*. United Nations. Available at <https://www.un.org/en/genocideprevention/documents/UN%20Strategy%20and%20Plan%20of%20Action%20on%20Hate%20Speech%2018%20June%20SYNOPSIS.pdf>. Accessed: October 3, 2020.
- USLEGAL.COM. 2020. *Hate speech law and legal definition*. Available at <https://definitions.uslegal.com/h/hate-speech/>. Accessed: October 3, 2020.
- Waldron, Jeremy. 2012. *The harm in hate speech*. Cambridge: Harvard University Press.

## ACKNOWLEDGEMENTS

This paper is a revised version of the original paper "The Direct Reference of Hate Speech and Its Emptiness" presented for "The International Conference on Hate Speech in Asia" organized by Asia Centre, Bangkok, Thailand, 7-9 October 2020. I would like to thank Professor Taro Mochizuki, Department of Philosophy, Osaka University, for his collaboration and for making the panel on "Philosophical Perspectives of Hate Speech in/for Asia" possible. As for the improvement of the paper, I would like to thank the editor, Jove Jim S. Aguas, for the help in editing and proofreading this paper, and especially the anonymous referees of the journal for the constructive comments that have helped improve the paper.