

You, robot: on the linguistic construction of artificial others

Mark Coeckelbergh

Received: 20 March 2010 / Accepted: 26 July 2010 / Published online: 10 August 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract How can we make sense of the idea of ‘personal’ or ‘social’ relations with robots? Starting from a social and phenomenological approach to human–robot relations, this paper explores how we can better understand and evaluate these relations by attending to the ways our conscious experience of the robot and the human–robot relation is mediated by language. It is argued that our talk about and to robots is not a mere representation of an objective robotic or social-interactive reality, but rather interprets and co-shapes our relation to these artificial quasi-others. Our use of language also changes as a result of our experiences and practices. This happens when people start talking *to* robots. In addition, this paper responds to the ethical objection that talking to and with robots is both unreal and deceptive. It is concluded that in order to give meaning to human–robot relations, to arrive at a more balanced ethical judgment, and to reflect on our current form of life, we should complement existing objective-scientific methodologies of social robotics and interaction studies with interpretations of the words, conversations, and stories in and about human–robot relations.

Keywords Human–robot relations · Robot ethics · Language · Phenomenology · Hermeneutics

1 Introduction

I love you. Do you love me?

(sentence addressed to a robotic doll, reported in Turkle et al. 2006, p. 357).

The robots are coming. But if they enter your home, they may not kill you; there is a good chance that they simply want a hug. Robots are no longer confined to factories, laboratories, and—increasingly—battlefields. They gradually enter people’s daily lives, offering companionship, entertainment, sex, or health care. Some people prefer artificial friends or even artificial partners.

Although this scenario may remain science-fiction for many of us, robots are already used in these domains and we want to know what would happen if robots became not only more autonomous and intelligent but also increasingly ‘personal’ or ‘social’. The scenario invites us to reflect on what it would be like to live with ‘social robots’ (Breazeal 2003) and how we should evaluate what goes on between humans and robots. More importantly, however, it helps us to reflect about ourselves: what it is to be human, what we mean by social relations, and how we should live together.

How should we understand and evaluate human–robot relations? In response to new developments in social robotics, there is a growing literature on human–robot interaction and human–robot relations (Turkle et al. 2006; Breazeal 2003; Dautenhahn et al. 2005; Dautenhahn 2007; Levy 2007, etc.). Moreover, philosophy of information technology—in particular robot ethics—has also started to reflect on ‘artificial companions’ (e.g. Floridi 2008) and their use in domains such as elderly care and health care (e.g. Sparrow and Sparrow 2006). Some even consider the issue of robot rights (Brooks 2000; Levy 2009; Asaro 2006; Torrance 2008) or, more generally, raise the issue

M. Coeckelbergh (✉)
Department of Philosophy, University of Twente,
P.O. Box 217, 7500 AE Enschede, The Netherlands
e-mail: m.coeckelbergh@utwente.nl

concerning the protection of (some) robots from abuse (Whitby 2008).

However, while becoming increasingly interdisciplinary, most existing work in the area of personal and social robotics uses methods from social science or psychology. It remains close to the methodological naturalism of the robotics science it relates to. Furthermore, mainly aimed at description and explanation, it lacks a normative orientation. Can we view human relations from a different, more distinctly philosophical perspective that takes distance from scientific-objective approaches and contributes to the need of evaluation?

In previous work, I proposed a social–phenomenological approach to philosophy of robotics, which focuses on the philosophical relevance of robotic appearance and of human–robot relations as *social* relations (Coeckelbergh 2009, 2010a, b). In this paper, I wish to further develop this approach by exploring the potential benefits of a linguistic–hermeneutic turn in philosophy of robotics. I argue that the appearance of robots in human consciousness is mediated by language: how we use words interprets and co-shapes our relation to others—human others or artificial others. In addition, I discuss a common objection to such ‘robot talk’: are these relations unreal and deceptive, threatening authenticity? Then I draw conclusions for the study, design, and ethical evaluation of human–robot relations.

In course of my arguments, I engage with empirical research on human–robot relations, in particular the work by Sherry Turkle.

2 Human–robot relations as social relations

Human–robot relations can be defined as social relations. What does ‘social’ mean here? First, all robots are ‘social’ in the sense that they play a role in human society, in the same way as other artefacts are ‘part’ of society as instruments for human purposes. For instance, cars play a role in society. Second, some robots are also ‘social’ in a different sense: they seem to participate in ‘social’ interaction with humans: they are autonomous, interactive robots that follow social rules and interact with humans in a human-like way. However, the sense in which I shall use ‘social’ with regard to robots and human–robot relations concerns the consequences of these interactions for the way robots appear to us. There is a specific phenomenological sense in which some robots can be called ‘social’: some robots appear to us as more than instruments. They appear as ‘quasi-others’ (Ihde 1990) or artificial others. Interaction based on this appearance constitutes a (quasi-)social relation between us and the robot, regardless of the robot’s ontological status as defined by modern science and by traditional and modern metaphysics, which view the robot as a mere thing or machine.

For ethics, this approach implies that the moral status of robots does no longer depend on ‘objective’ features of the robot but on how the robot and the human–robot relation appear to human consciousness. What matters for robot designers who wish to design a ‘moral’ robot, then, is not the creation of a robot *mind* but the creation of an appearance-in-a-relation.

However, in order to further develop this view, we need a more precise account of how this phenomenological process unfolds. How is the ‘social’ appearance created? What are the conditions under which a robot appears to us as a quasi-other? How do we make sense of such human–robot relations? In the next sections, I will partly fill this gap by discussing the role of *language* in how robots and human–robot relations appear to us as social. By exploring the linguistic–hermeneutic dimension of human–robot social phenomenology, I hope to contribute to a better understanding and evaluation of what goes on or might go on between humans and robots.

3 Language and the social

Let me distinguish two opposing views on the relation between the social and language, which I shall name representationalism and constructivism. Both views differ from extreme idealism and naïve realism, which define the relation between language and world by absorbing the one into the other: extreme idealism (in its post-modern or structuralist version) ‘deletes’ the world outside language; naïve realism ‘abolishes’ the subject. Applied to the social (world), this would mean that the social is purely linguistic–conceptual or that we can know the social as an objective reality. Both views are misguided. The social exists also ‘outside’ language, although we have no unmediated access to it: we experience it through the lens of language, we talk about it. In this sense, language is intrinsically connected to the social. However, the two opposing views I have in mind differ fundamentally in the way they view this intrinsic relation.

According to the representationalist view, the social is prior to language. The relation between the social and language is one of representation: with words we represent social relations. Language is a ‘mirror’ of the social. Although we might not have *direct* access to the social, language brings the social to us without distortion. According to the constructivist view, by contrast, language is prior to the social. The social is constructed or even declared (Searle 1995): the speech act is prior to the social since it creates the social. According to this view, the social is subjective or inter-subjective in the sense that it is created by human intention and human agreement; however, *once created* it assumes an ‘objective’ character in the form of rules, laws and (other) institutions.

However, do we have to choose between these polar opposites? There is no real dilemma here. Rather than seeing either the social or language as ‘prior’, let us try to construe a synthesis of the best insights into the two positions. The representational view is right that the social is not *merely* a matter of linguistic construction or declaration; there is an extra-linguistic social reality. The social is not merely (inter-)subjective. However, we can only access that social reality through the medium of language (words and concepts are ‘glasses’ through which we see reality) and the constructivist view is right that our use of language co-shapes that reality. Language does not mirror the social; it also helps to create it. Moreover, even apart from this more ‘active’ dimension of language, the mediating role of language is not restricted to representation (or is not even adequately described as such): we also *interpret* the social. We can have different readings of social institutions and practices. In both cases, knowing the social requires an ‘act’ of interpretation or even construction.

4 Natural language and artificial companions

Applied to human–robot relations, this linguistic–hermeneutic perspective implies that it is consistent to hold the following two claims in conjunction: (1) our use of language constructs human–robot relations (understood as social relations) and (2) these relations have also an extra-linguistic dimension: the relations take on a quasi-objective reality but are not directly accessible to experience: they appear to us through the medium of language and are interpreted by us using the medium of language.

This argument about linguistic mediation reveals the possibility of at least two different linguistic-phenomenological ‘glasses’ or repertoires (or what Searle calls ‘status functions’) for relating to robots: we can see them as ‘things’ and declare them to be mere objects or machines, but we can also construct them as quasi-others. The first repertoire constitutes an ontology in which there is a strict division between (human) subjects and (robotic) objects (and between physical ontology and human-social ontology, as in Searle’s work on social ontology¹), whereas the second repertoire creates a social ontology of human–robot relations that is more ‘hybrid’ in nature. This ontological

¹ A full response to, and engagement with, Searle’s work on social ontology will require a longer work. For the purpose of this argument it suffices to say that Searle holds a ‘dualist’ view in the sense that he applies an objectivist approach to physical reality but a constructivist approach to social reality. The version of constructivism I explore in this paper, while remaining agnostic about the ultimate nature of reality, questions Searle’s dualist view. It suggests that we do not have unmediated access to any reality—‘objective’ reality or ‘facts’ have to be interpreted and constructed for them to *become* what we call knowledge.

hybridity ‘shows up’ in language (language as representation and interpretation) but at the same it is also constructed by language.

We cannot simply choose between these two possibilities or repertoires. In this sense Searle’s language of ‘declaration’ is misleading. We might declare whatever we want; experience can push us—that is, our language use and hence the social relation—in a different direction. In current human–robot relations, we can observe a shift from talking *about* robots and *about* human–robot relations to talking *to* robots. Let me explain this shift and bring out its linguistic dimension and philosophical significance by distinguishing between different ‘perspectives’.

4.1 Talking about human–robot relations

Consider the following example. The idea of ‘love’, ‘friendship’ or ‘marriage’ between humans and robots raises ethical questions (e.g. Levy 2007). Is love for robots *real* love? Is sex with robots acceptable? Some people feel it is even offensive to *talk* about it.

However, the question whether there can be real love between humans and robots is not the right question, since it assumes that such ‘love’ would be an objective reality that stands apart from our human–robot experiences and practices. Imagine that one day we *declare* a particular human–robot relation to be a ‘love’ relation. We might even declare a particular human and a particular robot ‘married’. In this sense, love with robots would be a social, linguistic construction. However, if this were to happen, it would only happen on the condition that the relation already *appears* to us and is *interpreted* by us as a ‘love’ relation in virtue of what really happens between the natural and the artificial ‘partners’. A different phenomenology and interpretation may not warrant the same declaration; it would render the ‘declaration’ empty, devoid of meaning (most current robots are *not* experienced by us as deserving our love and affection). But this is also true for human relations: we cannot ‘simply’ or ‘merely’ change the social by means of declaration; such a declaration has to be integrated with experience.

Note that a similar argument could be made for robot rights: a declaration of robot rights only makes sense if and only if the phenomenological–hermeneutic process, which is partly and importantly linguistic in nature (but not merely linguistic), supports the declaration. (For designers who aim at creating robots that warrant a declaration of rights, this means that the robot would have to have features that trigger this interpretative and declarative move. However, I will not further develop this point here.)

This argument suggests a novel way of looking at human–robot relations. Based on social-linguistic interpretations and constructions, one may distinguish the following perspectives on language and robots.

Impersonal third-person perspective (“it”) and *first-person perspective* (I, robot): The usual AI (artificial intelligence) perspective on language and robotics taken by robotics researchers focuses on programming (the ability of) natural language into a robot, that is, it is concerned with how the *robots* (“it”) talks, could talk, and should talk. If language is considered by designers and users of robots and other artificial systems, they mainly worry about what the robot says or could (not) say. The dream of traditional AI (and of contemporary complaints departments of large companies) was to build artificially intelligent systems that would be indistinguishable from a natural language user. Consider discussions about Turing’s game (Turing 1950) or Searle’s Chinese Room experiment (Searle 1980): the question is whether we can tell the difference between a human and a robot based on their use of natural language. The ultimate dream, then, is to build a conscious robot who says ‘I, robot’ (as in the title of a science-fiction film). Then there will be a first-person perspective: robots may declare that they are conscious and perhaps *demand* their rights.

However, the approach I propose here pays attention to how *humans* talk about robots. This shifts the focus from what the robot says to what humans say.

4.2 Talking about robots

Third-person perspective. As the development of robotics continues and robots become more ‘personal’, designers and users may shift from the first-person perspective (in the sense of trying to put a human-like mind into the robot) to the third-person perspective and from the impersonal pronoun to the personal pronoun. People no longer consider the robot as a machine and start to refer to robots in personal terms. “It” becomes “he” or “she”. We might say that “he” wants attention or food that “she” did something to me, and so on. (A similar process happens when we deal with intelligent animals. Some people also use this language when they talk to cars and other things.)

4.3 Talking to robots

Second-person perspective (You, robot): A next step (which also happens in some human–animal relations) is that we do not only talk *about* the robot but also talk *to* the robot. Here the robot is fully appearing as a quasi-other, being acknowledged as a quasi-other, and being constructed as a quasi-other. Speech is directed to the robot, not only to humans. We can easily imagine a human person addressing her companion or partner robot with “you”. Such a person can only do that given what the appearance of the robot ‘does’ to her. At the same time, by addressing the robot in this way, the robot and the relation are

constructed *as* companion and *as* partner relation (examples drawn from empirical research follow below).

If the robot appears as a quasi-subject, with its own consciousness, needs, desires, and thoughts, this might even develop into the appearance of (quasi-)inter-subjectivity. If the robot talks back, a ‘dialogue’ is created. Then we have the impression not only that we are talking *to* the robot but that we are talking *with* the robot—an experience that is similar to talking with human social others. Here we might declare our interaction to be a ‘discussion’ or ‘conversation’. Humans might refer to themselves and the robot as “we” (first-person plural) when talking to a human or artificial other. Someone may say: “we” have discussed this; “we” had a good time (however, I will not further develop this suggestion here and mainly focus on the second-person perspective).

This approach to human relations has several implications for researchers and designers in the field of social robotics.

5 Implications for the study and evaluation of human–robot relations

First, the approach pays more attention to how people speak about and to robots. Our ‘robot talk’ is not neutral but interprets and shapes our relations with robots; it has a hermeneutic and normative function.

This insight can throw new light on existing research in the field of social robotics. For example, research by Turkle and others reports elderly people and children talking to robots (Turkle 2005; Turkle et al. 2006). For example, one of the residents of a nursing home says to the robotic doll My Real Baby: “I love you. Do you love me?” (Turkle et al. 2006, p. 357). Turkle concludes from this and other observations that there is a great degree of attachment to robots as ‘relational artefacts’, as tools to build relationships. They can help to relate people. Turkle’s research also shows that people talk to robots as ‘evocative objects’ that remind them about humans who play or played a role in their lives.

However, next to these ways of relating to robots there is also the possibility that the robot comes to be regarded as a unique artificial *other* rather than a tool or an evocation of a human other. This ‘use’ is not well-researched yet but is indicated when Turkle suggests that people who develop such attachments want robots to talk back to them, e.g. say one’s name or say ‘I love you’ back to them: ‘Elders may come to love the robots that care for them, and it may be too frustrating if the robot does not say the words “I love you” back’ (Turkle et al. 2006, p. 360). (However, she recognises that this does not feel completely comfortable and that it raises the issue of authenticity. I return to a

similar ethical issue below.) It seems that an expectation arises within a social relation. The robot becomes a ‘you’—not as a stand-in for someone else (which we might call a ‘*delegated* second-person’) but a ‘you’ in its own right, an artificial second-person, which has a claim on me as a social being. Based on what I have said previously, we can understand this talking *to* robots as changing and shaping the human–robot relation as a social relation: people talk to robots not only because of the ‘personal’ or ‘social’ appearance of the robot (produced by the robot having certain features and capacities for certain behaviours); by talking to the robot in second-person terms, they also construct it as a quasi-other. *Partly by using language in this way* (e.g. using the words “you” and “love”), the robot does no longer appear to the elderly person as an object but as a quasi-other. This creates a social reality: it produces expectations that co-shape the quasi-social relation and are therefore no longer ‘merely subjective’.

This might not only throw new light on existing research, it can also suggest new research questions. For example, I suggest that we test the following hypothesis: the linguistic pre-construction of the human–robot relation (as a social relation) influences, co-constructs the actual relation. Usually researchers manipulate the (parameters of the) interaction and then see what happens in terms of language use, e.g. what people say to the robot. I propose to turn that around: manipulate the *linguistic* ‘environment’ and then see what happens to the interaction. For instance, what happens when the instructor already pre-defines the interaction as a ‘personal’ or ‘social’ relation by using words such as “she” or “he” or by giving the robot a name? Does the result differ from what happens in a control group (where the instructor uses “it” and the robot’s factory ‘name’)? Moreover, to explore the hermeneutic dimension one would need to conduct long-term studies of human–robot *relations*—with all their interpretations and narratives—as opposed to current short-term *interaction* studies.

Second, since languages and the cultures in which these languages are embedded differ, this approach naturally suggests paying attention to *cultural differences*. For example, many other languages have not one but two or more second-person forms. English has “you”, but in Dutch there is a difference between “you” (informal) and “u” (more formal and polite). Consider also the German “du” and “Sie”, the French “tu” and “vous”, or the Hindi “tum” and “aap”. Other languages may have a larger set of linguistic–social possibilities. Which forms are used and indeed *should* be used to address artificial others? These descriptive and normative questions are important since the precise form used expresses but also constitutes the social relations and structures. For example, with our words we might imply and constitute a master–slave relation or a

companion relation. There is also a gender dimension to this: when we address the robot from a third-person perspective, do we use “he” or “she”?

Language has a normative dimension, prescriptive dimension: words are related to ways of doing things. With Wittgenstein we might want to call these ways of doing things ‘forms of life’ (Wittgenstein 1953). Language and culture are closely connected. The meaning of words is created by language *use* and practices; at the same time language also shapes these practices. The way we do things, the *form* it takes (culture), depends on language. To imagine relations between humans and robots is to imagine a form of life: one that includes both human–human relations, human–robot relations, and the clouds of words and meaning that surround them. This form would probably partly mirror that of human–human relations, as suggested by the use of personal pronouns. That use is derived from how humans address one another. However, we also need to explore if and how human–robot relations may transform how we do things; they could contribute to a new form of life. Moreover, we may try to imagine what it would mean if robots had their *own* way of doing things, if they developed their own form of (artificial) life.

This may be hard to imagine for us. ‘Robot talk’ is very much connected to the worldviews/ontologies—including social ontologies—of the culture in which it vibrates. For example, in the West we tend to put a lot of emphasis on the subject–object distinction and most of us hold individualistic, non-relational ontologies. Robots are assigned to the ‘object’ category, humans to the ‘subject’ category. No crossovers or hybrids are allowed in this modern, dualist ontology; purity is preserved.² Languages already ‘contain’ such ontologies in their structures. Our linguistic grammar is also a moral, social and ontological ‘grammar’. Consider again the use of pronouns and the subject/object distinction in Western languages: the language we use co-shapes how we interpret and construct the social. This sets limits to imagining, shaping, exploring, and experiencing new forms of life in which robots assume the role of artificial others. At the same time, language is not fixed and our words and concepts can and do change as a result of changing social relations. For example, we can start using personal pronouns in our robot talk. How these relations change is partly but not entirely up to us. We cannot simply or entirely change our form of life. Living in a particular culture, being part of a particular generation, being the particular person one is, and so on, will influence the language one uses and how that language changes (I return to this point).

Third, therefore, this approach promotes the study of *changes* in human–robot relations. They change as a result

² See also the work of Latour (1993).

of many factors but are also and partly changing due to the ways we speak about robots and due to our declarations concerning these relations.

This approach brings in the time dimension and the personal identity dimension: it implies that there can be (*hi*)stories about human–robot relations (fictional and non-fictional) in which we interpret and shape human and robot identities and the relations in which these identities are embedded. We might use “they” when talking about a particular human–robot relation (with “they” referring to the human *and* the robot) or even the first-person “we”. For instance, people might refer to a man and his robot as “they” when there appears to be a personal, social relation between the two. The relation then assumes a kind of collective identity, in the same way as a couple of humans are collected by the word “they” (“I think they should get married”) (see also my point about inter-subjectivity in the previous section).

Robotics designers and scientists are advised to take these aspects into account when they consider use of robots: use does not only depend on the ‘object’ as such but on how that object appears to humans and how this appearance is mediated and shaped by human subjects and their use of language.

From a normative-ethical point of view, we may then ask: how *should* people talk about robots? An answer to this question can inform robot design. It might turn out, for instance, that there are good reasons why we should not want robots to appear as others or as persons, that we should not want that people *talk about them or to them in this way*. To explore this issue, let me end with a discussion of *one* ethical objection to personal robots. This will not only allow me to say more on human–robot relations; it will also give me an opportunity to show one way how the approach proposed here can contribute to the study and evaluation of *human* relations.

6 The deception objection to personal robots

One possible objection to personal robots (and the related linguistic–conceptual changes I discussed) concerns the dual charge that these robots, human–robot relations, human–robot ‘conversations’, etc. are not *really* persons, not *really* (social) relations or conversations (i.e. are *inauthentic*), and that giving them to people would be a matter of *deception*.

Although I am sympathetic to this objection, I doubt if it is tenable in its current form for the following reasons. This objection assumes at least the following: (1) that talking to things is always and necessarily morally problematic, (2) that only human relations are real, true, and authentic, (3) that there is an objective, external point of

view that allows us to judge the reality and truth of the human–robot relation, and (4) that to say that the robot is a thing is completely unproblematic. But these assumptions must be questioned.

First, many people already talk to things and other non-humans and this is not generally regarded as a moral problem. People talk to plants, animals, dolls, cars, navigation systems, DVD players, and—as Turkle already observed in the 1980s—computers (Turkle 1984).³ When people address current robots using a second-person perspective (“you”), the linguistic form is the same. Of course robots are embodied and relational objects invite more affective and social engagement (e.g. nurturing, see Turkle et al. 2006, p. 348), but it is not clear why this renders talking to robots itself *fundamentally* more morally problematic. In both cases, the object is experienced and treated as a quasi-other, there is only a gradual difference. Robots, to the extent that they are more interactive, invite strong(er) quasi-other experiences. In this sense, they are more ‘deceptive’ indeed. Further study of how we experience and treat different kinds of artefacts is needed, but at first sight the kind of ‘deception’ by robots as compared to the ‘deception’ involved when someone talks to a plant does not seem to be fundamentally different in kind.⁴

One might object that still the ‘perception’ of the robot is not real or not true. But what is more real or true: the (abstract) definition of a robot as an object or the actual experience of the robot as a quasi-other? From a constructive-phenomenological point of view, there is no ready and a priori answer to this question.

Paradoxically, those who are most outraged at the thought of talking to robots must have a strong quasi-other experience. If people *really* believed that the robots in question are ‘mere objects’, they would not have a moral problem with (others) talking to them. They would find it ‘silly’ or ‘childish’, or perhaps even ‘insane’, but they would not count it as *morally* problematic. The moral question arises only when the appearance as quasi-other

³ Note that some of these entities ‘talk to us’, for instance a navigation system in a car. Whether or not this is to be considered ‘talk’, it also shapes the particular human–technology relation that develops. There are many ways an artefact can ‘communicate’ to us and presumably each has its influence on the way we do things and perceive our environment. Note also that robot communication might partly happen by means of ‘body language’. If robot designers want to mimic human communication at all (they can make other design choices), this is an important aspect that should be taken into consideration.

⁴ The only exception I can think of is the possibility that sometimes people talk to robots or plants not because the robot, plant, or other object appears as a quasi-other but because they *use* it as an *instrument* to order their thoughts, to talk *with themselves*, to have an internal dialogue—indeed to think. Then the robot appears as a representation of an ‘internal’ dialogue partner, it is ‘internalised’ as an extension of the inner conversation.

has already taken place. The ethical issues arise when, phenomenologically speaking, the robot moves into the twilight zone between object and subject: it appears as ‘more’ than an object yet ‘less’ than a subject. For instance, sex robots can come to be seen as morally problematic only if they *already* appear as quasi-others. If they were interpreted or experienced as mere *tools* used in sexual activities, they would not receive the same outrage.⁵

Second, is reliance on appearance in social relations as such morally problematic? Appearance plays an important role in human–human relations. We do not always ‘really know’ the persons we socially engage with, yet appearance lubricates our inter-human social relations. Of course, we sometimes say that while someone *appeared* such and such, the person *really* is such and such. But how do we determine the real, the true or the authentic? Do we have unmediated access to pure reality, to the truth about a personal core, to an authentic self? The answer to this question is at least more complex than the objection suggests.

Third, one answer to this perennial philosophical question is the following. What we ‘have’ for sure are appearances of robots, of (ourselves as) humans, and of our relations to robots. Whether we can have access to anything else depends on whether we have an objective, external point of view outside the relations as they appear to us. However, both our view of human relations and our view of human–robot relations do not stand entirely apart from these relations themselves. Following from the arguments outlined in the previous sections, our response must be that our talk to humans and our talk to robots are not neutral *vis-à-vis* how we define the real and the truth and how and shape these relations. This does not imply that no meaning can be given to these terms, but rather that knowledge of the real or the true concerning humans and robot is not obvious, not given, and must be carefully constructed by taking into account concrete experience as appearances-in-relation.

Fourth, to say that robots “are” “mere” objects is not trivial and not unproblematic. As the previous discussion suggests, our resistance to talking *to* robots or talking *with* robots is shaped by (1) our experience of robots as mere machines (similar to the way Descartes could only experience animals as machines) and (2) our Western outlook on relations between humans and non-humans rooted in an ontology that makes strict subject/object distinctions and therefore excludes the possibility of ‘hybrids’ such as robots as quasi-others. However, while our current experiences and conceptual frameworks prevent us from seeing robots differently, this might change in the future. As new kinds of robots are being developed and used, both our

language and our experiences change and mutually influence one another. If we take this cultural and dynamic aspect into account in our evaluations, this might render these evaluations less stable than we would want them to be from a theoretical point of view, but they will be more relevant and better able to guide our decisions concerning the future.

Thus, although these experiential and conceptual changes are not entirely up to us since we cannot fully control the development of technology, social experience, cultural meaning, and personality, this position does not imply that we have to uncritically accept technological and linguistic–conceptual changes. The ethical-technological task we face is to try to guide, steer, and shape these changes in a desirable direction. However, our aim cannot be determined from an objective point of nowhere. Instead, from within our current situated frameworks, we must explore, imagine, and evaluate different human–robot possibilities: different kinds of robots and human–robot relations, perhaps different forms of life. Of course we might judge some possibilities to be ethically unacceptable. However, the intuitions and values we currently rely on to evaluate these possibilities are not fixed but can also change as possibilities and realities change. We might change our values or interpret them differently once other relational possibilities show up. We can only arrive at judgment within this dynamic, historical and interpretative moral-technological constellation. In the absence of the possibility to arrive at eternal, objective truths, our judgment must always remain provisional and vulnerable to criticism and challenge.

7 Conclusion

In this paper, I have explored a linguistic–hermeneutic turn to the study and evaluation of human–robot relations. In this way, I have opened up a perspective that differs from a scientific ‘objective’ approach but at the same time attends to concrete human–robot practices, experiences, conversations, and stories.

Some robots are revealed as artefacts that are co-constructed in at least the following ways: they are at the same time engineering constructs and social-linguistic constructs. Their appearance creates social relations that are linguistically mediated. The language we use allows us to take different perspectives on the robot: the language we use reveals the robot as an “it” but sometimes also as a “you”.

Which way of approaching robots has epistemic priority? Which one is truer: “it” or “you”? Although I am sympathetic to those who answer “it”, I suspect that this is the wrong question if we consider (knowledge construction

⁵ Levy puts sex robots within the history of sex aids (Levy 2007).

in) concrete human–robot practices. Quite similar to *Gestalt* switching,⁶ we are very well able to switch between these perspectives (but we cannot see them at the same time) and none of the *Gestalts* has epistemic priority. Depending on the culture and language we live in, one or the other perspective will be used more frequently, but there is no a priori order between the robot-as-thing and the robot-as-quasi-other.

Thus, to acknowledge the linguistic–cultural construction of robots and robot relations is limiting and freedom-enhancing at the same time. On the one hand, our practices and thinking are already embedded in a culture, in forms of life. On the other hand, once we are aware of the ‘grammar’ and what it does to us, we can try to stretch our language and thinking when exploring the implications of social robots and other technologies.

However, is there an *ethical* priority with regard to language use? Are there moral limits to linguistic–cultural imagination and practice? For instance, is it *morally* problematic to use the “you” perspective in relation to an artefact? From the previous discussion of the deception objection, we can conclude that using “you” is at least not *obviously* morally wrong. However, it may be that in relation to some technologies some uses of language and indeed some forms of life are morally unacceptable. It may be that some forms of life are to be preferred over others. In any case, further reflections on this issue should take into account how we talk to things (usually not seen as *morally* problematic), the role of appearance in human–human relations (again, usually not seen as very morally problematic), hard epistemological problems with notions such as the real, the true, and the authentic, and the dynamic nature of technological, societal, linguistic, and cultural possibilities.

Although this discussion raises many further questions, a social-phenomenological approach and the ‘linguistic turn’ proposed here seem promising conceptual tools to enhance our understanding of what goes on when we talk to robots. As such they can usefully inform, complement, and perhaps revise existing empirical approaches in interaction studies and social robotics. They may also assist philosophers of robotics to make sense of, and evaluate, the ‘personal’ and ‘social’ dimension of living with robots. They can make users aware of the variety of repertoires and perspectives that can be used in relating to robots. And even if the scenarios mentioned in the introduction remain science-fiction, we may transfer some insights from this discussion to understanding our relations to other technologies and artefacts.

Finally, any reconstruction of the object is likely to influence the construction of the subject. The emergence and adoption of more ‘personal’ or ‘social’ linguistic

repertoires in human–robot practices may actually change the subject’s understanding of itself. If a robot appears as a quasi-other, then the human subject can no longer understand itself as a ‘user’, as the one who controls the robot-as-instrument; instead, it is likely to re-constitute itself as a social actor living with (and living in the eyes of) the artificial social other. Talking to robots thus changes talking about humans—perhaps also *to* humans. The language we use reveals and shapes the social ontologies in which we live as much as it reveals and shapes ourselves.

Acknowledgments I would like to thank the anonymous reviewers for their suggestions, which helped me to fine-tune the presentation of my arguments and extend the range of my examples.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Asaro PM (2006) What should we want from a robot ethic? *Int Rev Inf Ethics* 6:9–16
- Breazeal C (2003) Toward sociable robots. *Robot Auton Syst* 42:167–175
- Brooks R (2000) Will robots rise up and demand their rights? *Time*, June 19 2000
- Coeckelbergh M (2009) Personal robots, appearance, and human good: a methodological reflection on roboethics. *Int J Soc Robot* 1(3):217–221
- Coeckelbergh M (2010a) Robot rights? Towards a social-relational justification of moral consideration. *Ethics Inf Technol*. doi:10.1007/s10676-010-9235-5
- Coeckelbergh M (2010b) Moral appearances: emotions, robots, and human morality. *Ethics Inf Technol*. doi:10.1007/s10676-010-9221-y
- Dautenhahn K et al (2005) What is a robot companion—friend, assistant, or butler? Intelligent robots and systems, IEEE/RSJ International Conference on intelligent robots and systems
- Dautenhahn K (2007) Methodology and themes of human–robot interaction: a growing research field. *Int J Adv Robot Syst* 4(1):103–108
- Floridi L (2008) Artificial intelligences’s new frontier: artificial companions and the fourth revolution. *Metaphilosophy* 39(4–5): 651–655
- Ihde D (1990) *Technology and the lifeworld*. Indiana University Press, Bloomington/Minneapolis
- Latour B (1993) *We have never been modern* (trans: Porter C). Harvard University Press, Cambridge, MA
- Levy D (2007) *Love and sex with robots: the evolution of human–robot relationships*. Harper, New York
- Levy D (2009) The ethical treatment of artificially conscious robots. *Int J Soc Robot* 1(3):209–216
- Searle JR (1980) *Minds, brains and programs*. *Behav Brain Sci* 3(3):417–457
- Searle JR (1995) *The construction of social reality*. The Penguin Press, London
- Sparrow R, Sparrow L (2006) In the hands of machines? The future of aged care. *Mind Mach* 16:141–161

⁶ See also Ihde’s concept of ‘multistability’.

- Torrance S (2008) Ethics and consciousness in artificial agents. *AI & Soc* 22:495–521
- Turing AM (1950) Computing machinery and intelligence. *Mind* 59:433–460
- Turkle S (1984) *The second self: computers and the human spirit*. Simon and Schuster, New York
- Turkle S (2005) Relational artifacts/children/elders: the complexities of cybercompanions. <http://www.androidscience.com/proceedings/2005/TurkleCogSci2005AS.pdf>. Accessed 19 March 2010
- Turkle S, Taggart W, Kidd CD, Dasté O (2006) Relational artifacts with children and elders: the complexities of cybercompanionship. *Connect Sci* 18(4):347–361
- Whitby B (2008) Sometimes it's hard to be a robot: a call for action on the ethics of abusing artificial agents. *Interact Comput* 20(3):326–333
- Wittgenstein L (1953) *Philosophical investigations* (eds and trans: Hacker PMS, Schulte J). Wiley-Blackwell, Oxford, 2009