

Categorically Perceiving Motor Actions

Chiara Brozzo

[forthcoming in F. Calzavarini & M. Viola (eds.), *New Challenges in Philosophy of Neuroscience*. Springer.]

ABSTRACT: In this chapter, I will present an empirical conjecture to the effect that some bodily actions are categorically perceived. These are bodily actions such as grasping or reaching for something, which I am going to call *motor actions*. My conjecture builds on one recently put forward about how the categorical perception of facial expressions of some emotions works. I shall motivate my own conjecture on the basis of both theoretical and empirical considerations, describe how it could be operationalised and what explanatory gain could be obtained from it.

1. Introduction

In this chapter, I am going to present an empirical conjecture about the way in which some bodily actions are perceptually processed. These are bodily actions such as grasping or reaching for something, which I am going to call *motor actions* (more on their characterisation in section 2). My conjecture has it that humans *categorically perceive* motor actions (the notion of categorical perception will be explained in section 3). This conjecture builds on and complements one recently put forward by Stephen Butterfill (2015), which is based on evidence to

the effect that humans categorically perceive facial expressions of some emotions (e.g., Ectoff and Magee 1992; Calder et al. 1996; see also Kotsoni et al. 2001). This evidence and the related conjecture will be presented in section 4.

In section 5, I shall present my own conjecture, to the effect that motor actions are categorically perceived. I will motivate this conjecture on the following grounds: a significant structural analogy exists across motor actions, the expressions of some emotions and the articulations of phonemes, insofar as all are actions directed to motorically represented outcomes (a notion that will be explained in due course), and may be categorically perceived as such.

Why should you be interested in a conjecture based on another conjecture? Because I believe these to be two sides of the same coin, and considering them jointly provides a useful unifying theoretical framework for interpretation and for further testing. In section 6, I will describe the explanatory gain that can be obtained from my proposed conjecture in terms of the interpretation of data about the neural mechanisms involved in the processing of motor actions.

2. What are motor actions?

I am now going to introduce the notion of motor action.¹ Consider a situation in

¹ The term *motor action* has been widely used in the study of action production, both in neuroscience (e.g., Gallese et al. 1996; Hamilton and Grafton 2007;

which I pick up a plum. What are the conditions under which it is correct to say that the plum has been *grasped* (as opposed to, e.g., *pushed away*)? In the most ordinary cases, these conditions will include that I end up with the plum held between my palm and my fingers. But it is also possible to imagine a situation in which my hands are occupied, or are somehow blocked, or do not exist at all, and I have to grasp the plum with my mouth or with my feet. Not *any* body part will do (I cannot grasp anything with my nose, due to anatomical constraints), but *some* body part has to be employed. The essential feature, then, for the characterisation of a motor action such as grasping something is that it involves one of a *circumscribed set of body parts*, which has to undergo a certain change—one, for instance, that brings it in a specific relation to a given object. In the light of the above, I shall call *motor actions* actions such as grasping, whose characterisation unavoidably involves mention of body parts and their configurations and an object in relation to which these configurations unfold. To clarify, compare a motor action—e.g., grasping—with the action of, e.g., designing an experiment. No specific body part is involved in the accomplishment

Jeannerod 1994, 2006; Rizzolatti et al. 1996) and philosophy (e.g., Butterfill and Sinigaglia 2014; Ferretti and Zipoli Caiani 2019; Mylopoulos and Pacherie 2017; Nanay 2013; Pavese 2015). I am, however, introducing this term with a specific meaning, to be illustrated shortly, that does not straightforwardly coincide with how this term has been employed in the aforementioned literatures, although it is likely consistent with it.

of the latter, nor is any circumscribed set of bodily movements.

Notice that motor actions cannot be specified exclusively in terms of the corresponding final bodily configurations. Consider as an example the action of catching something. One may believe it sufficient to characterise the action of catching a ball in terms of the ball being held between one's hands. But this is not sufficient, insofar as it cannot be said that a ball has been *caught* if the ball being held between one's hands is the result of someone carefully placing it there: without the specification that one has performed certain bodily movements, it is false that she has caught the ball (see Pacherie 2008). Because of this, a motor action should be specified by reference to a circumscribed set of *sequences* of bodily configurations, and not only the final one. So, motor actions are actions whose characterisation unavoidably involves mention of sequences of bodily configurations. This is a characterisation in purely behavioural terms. In the following sections, I shall propose a conjecture to the effect that humans categorically perceive motor actions. I will begin by characterising the notion of categorical perception.

3. The premise: humans categorically perceive facial expressions of some emotions

When human subjects are asked to discriminate between pairs of faces that express a certain emotion, their responses exhibit a very specific pattern of discrimination. Presented with several pairs of faces that differ by a fixed physical

amount, created by morphing a face expressing a certain emotion (happiness) into a face expressing another emotion (fear), subjects are better at discriminating pairs where each member expresses a different emotion, rather than pairs where both members express the same emotion (Etcoff and Magee 1992; see also Calder et al. 1996; Kotsoni et al. 2001). In a separate task, subjects are also asked to identify the stimuli that they are presented with as expressing one of two possible emotions (say, happiness or fear). Subjects are consistent in identifying stimuli that do not fall too close to the category boundary, whereas they are at chance (that is, they identify the face stimulus as happy or sad with equal frequency) when it comes to identifying stimuli that fall too close to the category boundary (Calder et al. 1996). When such patterns of discrimination are exhibited in relation to a certain domain—i.e., some pairs of stimuli are easier to discriminate than others, and, moreover, what explains this is that those pairs of stimuli fall in different categories recognised by the subjects—it is said that subjects have *categorical perception* for that domain (Repp 1984; Harnad 1987; McKone et al. 2001; Harnad 2003).² According to Harnad (2003), specifically, perceived differences

² Some important clarifications are in order. Hereafter I will discuss conjectures concerning the *recognition* of emotions on the part of an observer, but I shall not make any claims about the *nature* of emotions. As to the latter, there is a controversy on which I do not mean to adjudicate between whether the nature of emotions is *categorical* or *basic* rather than *dimensional*. According to the *basic* view of emotions (Ekman 1992; Izard 1971; Tomkins 1962), emotions fall into

between stimuli within a category are smaller than the actual physical differences between those stimuli, and/or perceived differences between stimuli across a category boundary are larger than the actual physical differences between those stimuli (Harnad 2003).³

discrete categories, which are reflected in the information provided by cues such as facial expressions and body postures. According to the *dimensional* view of emotions, by contrast, rather than falling into discrete categories, emotions arise from combinations of degrees of arousal and valence, two distinct dimensions whose values vary in a continuous way, without giving rise to clear-cut category boundaries (Russell 1980). It is crucial to notice that some authors have taken the aforementioned results supporting the view that the *recognition* of emotions takes place by means of categorical perception as support for a categorical view of the *nature* of emotions. Fugate (2013) points out that this is a mistake: she refers to evidence provided by Young and colleagues (1997) and Fujimura and colleagues (2011) suggesting that both categorical and dimensional information might be drawn on in categorical perception. I am grateful to the editors of this volume for pointing out this potential source of misunderstanding.

³ Harnad (2003) also offers a version of this definition to accommodate the case of learned categorical perception. In this version, the term of comparison is not actual physical differences between stimuli but rather perceived similarity between the stimuli within and across category boundaries before learning. I am grateful to a reviewer of this volume for inviting me to report Harnad's definition.

Evidence suggests that humans have categorical perception for a number of other domains in addition to facial expressions of emotion, including speech (Liberman et al. 1957; Eimas et al. 1971), colour (Bornstein and Korda 1984), orientation (Wolfe et al. 1992) and face identity (Beale and Keil 1995; Kikutani et al. 2008).

The colour case is illustrative of the sort of phenomena that the notion of categorical perception is supposed to explain. For example, I mentioned that pairs of stimuli falling in different categories are easier to discriminate. Bornstein and Korda (1984) showed that pairs of hues can be told apart comparably quickly, even though they may be more or less different in purely physical terms, so long as each belongs to a different colour category. Conversely, pairs of hues that belong each to a different colour category can be told apart more quickly than any pair of hues that belong to the same colour category. This is so in spite of the fact that the pair of hues belonging to the same colour category might be more different from each other in physical terms than the pair of hues that belong each to a different colour category. Thus, ease of discrimination is shown not to depend straightforwardly on physical differences, but, rather, is a function of the categories to which the stimuli belong.⁴

⁴ Studies that hinge on physical differences are subject to a potential objection: couldn't it be that same physical differences are treated differently by the retina and therefore end up being perceived differently by the subjects? A study by Witzel and Gegenfurtner (2014) counters this objection by using just-noticeable

Another phenomenon that the notion of categorical perception is supposed to explain is the occurrence of pop out effects. Daoutis and colleagues (2006) have shown that, given an array of coloured dots, all of the same colour except for one, the time it takes to find the odd one out does not increase as a function of the number of dots when the odd one out is of a different colour category with respect to the other dots. That is, hues falling into different colour categories pop out.

The evidence reported earlier in this section (Etcoff and Magee 1992; Calder et al. 1996) gives us reasons for thinking that humans have categorical perception of facial expressions of some emotions. The qualification *some* is justified by the fact that the stimuli employed in the earlier reported experiments typically involve happy and fearful faces. It might therefore be safer to claim that it is only for the expression of *some* emotions that humans have categorical perception.

There might be a principled reason behind this, namely that some emotions lend themselves to a more straightforward connection with their bodily expression than others (such as *Schadenfraude*).⁵ That is, it is plausible that some

differences instead of physical differences. A just-noticeable difference (JND) is the smallest difference between two stimuli that a subject can perceive.

⁵ This leaves it open that the connection in question could be mediated by factors such as conceptual knowledge (Brooks and Freeman 2018) or culture (see Caruana and Viola 2018).

emotions may have more easily identifiable characteristic expressions associated with them.

So far, I have introduced the notion of categorical perception and have reported evidence that the facial expressions of some emotions are categorically perceived. On the basis of this evidence, Butterfill (2015) puts forward the following conjecture: facial expressions of emotions could be categorically perceived insofar as they are actions directed to motorically represented outcomes (a notion that will be defined in the next section). I shall now present this conjecture, along with how it is supported by current evidence. This will provide the springboard for my own conjecture, which I will introduce in section 5.

4. What are facial expressions of emotions? The AMROs conjecture

In this section, I am going to present Butterfill's conjecture that facial expressions of emotions are actions directed to a motorically represented outcome (AMROs), and are processed as such within the context of categorical perception (Butterfill 2015). I shall explain the notion of AMRO first by reference to the case of speech, and will then show how this notion can be applied to the case of facial expressions of emotions.

4.1 Phoneme articulations as AMROs

Speech is one of the most extensively studied cases of categorical perception (e.g., Liberman et al. 1957; Eimas et al. 1971; Harnad 1987; Nygaard and Pisoni 1995;

Harnad 2003). Here is evidence that we categorically perceive speech. It is possible to create a series of test stimuli consisting in sounds that spread across the phonemes *ba* and *pa*. These are designed in such a way that each two neighbouring test sounds differ from one another by the same amount (in terms of frequency) as any other pair of neighbouring sounds (the test stimuli consisting in facial configurations described in section 3 were created on the basis of an analogous principle). Subjects find it hard to discriminate neighbouring pairs of test sounds, except when two neighbouring pairs fall on two different sides of a category boundary—i.e., one is perceived as *ba* and the other as *pa*. Within the same category, on the other hand, subjects will hear the same phoneme, e.g., *ba* (Liberman et al. 1957).

So, humans categorically perceive speech, and the categories consist in phonemes. But what are phonemes? An interpretation that has been put forward (e.g., by Liberman and Whalen 2000) is that a phoneme is an *outcome*, i.e. a state of affairs, to which an action is directed.⁶ What distinguishes outcomes (in the

⁶ This closely resembles the idea that phonemes are *intended gestures* of a speaker, which is the heart of the Motor Theory of Speech Perception (Liberman et al. 1957; Liberman and Mattingly 1985). The Motor Theory of Speech Perception has a complex history, and its evaluation is made difficult by the fact that it encompasses several different claims, whose fate has proved very different. Galantucci and colleagues (2006) helpfully break the Motor Theory of Speech Perception down into different claims: “(1) speech processing is special, (2)

case of speech, consisting in phonemes) from mere acoustic signals? The distinction is twofold. First, different acoustic signals could be employed to articulate the same phoneme. This is shown by the fact that we have categorical perception of speech: as mentioned earlier, a number of different acoustic signals will be treated as the same phoneme (e.g., *pa*) by a perceiver. In addition to this (and this goes beyond the idea that we categorically perceive speech), single acoustic signals by themselves may not be diagnostic of what phoneme is being articulated: the same single acoustic signal, depending on contextual factors such as speed of articulation or dialect, could result from the articulation of different phonemes (see, e.g., Repp and Liberman 1987).

So far, I have presented reasons in support of the idea that phonemes should be considered outcomes, and how this differs from considering them merely acoustic signals. The idea, in short, is that the same phoneme could be articulated through different acoustic signals, and the same acoustic signal could result from different phonemes being articulated. Building on this, Butterfill (2015) hypothesises that articulations of phonemes may be characterised as

perceiving speech is perceiving gestures, and (3) the motor system is recruited for perceiving speech.” Galantucci and colleagues argue that (1) is likely false, but that (2) and (3) still find support. Claim (3) has recently been vindicated by Whalen (2019). In this chapter, I am exploiting precisely claims (2) and (3) of the Motor Theory of Speech Perception, but not (1).

actions directed to motorically represented outcomes—henceforth, AMROs for short.

But what is a *motorically represented* outcome? It is an outcome represented by motor areas of the brain. The best evidence that an outcome is represented motorically is that a marker of motor processing, e.g. neuronal discharge (Butterfill and Sinigaglia 2014, p. 122) that is recorded in motor areas of the brain, or motor evoked potentials, can be found in correlation with an outcome being brought about.⁷ Butterfill suggests that phonemes could be motorically represented outcomes insofar as articulating a phoneme requires coordinated

⁷ These markers of motor processing are often discussed under the heading of *motor representations*. The idea that motor representations might represent outcomes rather than just fine-grained bodily movements has given rise to what Butterfill and Sinigaglia (2014) call the Interface Problem: how do the outcomes represented by intentions and the outcomes represented by motor representations non-accidentally match? Answers to this problem have been discussed by Butterfill and Sinigaglia themselves (2014), as well as by Mylopoulos and Pacherie (2017), Ferretti and Zipoli Caiani (2019) and Shepherd (2019). There are also motor representations representing an action in greater detail—for example, that represent grasping with a specific body part (e.g., one’s hand) and with a specific kind of grip (e.g., a precision grip—the one you would typically adopt to grasp a peanut; Rizzolatti et al. 1988). For a more extended discussion of what motor representations represent, see Ferretti (2016).

movements of the vocal organs—lips, tongue tip, tongue body, tongue root, velum and larynx (see, e.g., Goldstein and Fowler 2003).

To this, I would like to add that precisely this sort of rationale—that bringing about a certain outcome requires a series of movements coordinated around the outcome—has led neuroscientists to posit that actions should be represented in the brain, in a way that abstracts away from the details of the bodily movements but is also sensitive to a certain outcome being achieved (Jeannerod 1994, 2006). So, it seems plausible to suppose that not only are phonemes outcomes, but they are *motorically represented* outcomes.

On the basis of this idea, it is possible to hypothesise that the categories into which the categorical perception of speech sorts acoustic signals are motorically represented outcomes, and that articulations of phonemes are processed as AMROs in the context of categorical perception. This is an interpretation of what goes on in the categorical perception of speech that suggested by, e.g., Liberman and Whalen (2000).

4.2 Expressing emotion as an AMRO

How is this relevant to the categorical perception of facial expressions of emotions? Because Butterfill (2015) conjectures that facial expressions of emotions, as acoustic signals according to the interpretation reviewed in the previous subsection, could be categorically perceived as AMROs, rather than merely as facial configurations. This is based on the idea that an emotion being expressed is a motorically represented outcome. Let me show how this idea is

justified, before moving on to the conjecture in the next subsection.

A contrast between facial configurations and emotions being expressed can be set up, analogously with that between acoustic signals and phonemes. As with speech, the evidence to the effect that we categorically perceive facial expressions of emotions indicates that multiple facial configurations (e.g., more or less wide smiles) can be involved in the same emotion (e.g., happiness) being expressed. But, to support the idea that emotions being expressed are outcomes rather than just facial configurations, we also need evidence that single facial configurations are not necessarily diagnostic of emotions, i.e. may be taken to express different emotions depending on contextual factors.

Aviezer and colleagues (2008) provide just this sort of evidence. They show that the very same facial configuration can be taken to express different emotions depending on the context into which it is inserted—specifically, the overall bodily configuration of the individual exhibiting that facial configuration. For instance, the very same facial configuration on an individual’s face can be verbally classified by an observer as either disgusted or proud depending on the overall bodily configuration of the observed individual—i.e., whether the individual with this facial configuration is holding a disgusting object or is engaged in a power pose.

So, both conditions for emotions being expressed to be outcomes are fulfilled: the same emotion can be expressed through different facial configurations, and the same facial configuration can express different emotions depending on contextual factors.

But why think that emotions being expressed should be *motorically represented* outcomes? In response to this, Butterfill presents a line of reasoning analogous to the one provided in relation to the case of speech:

expressing an emotion by, say, smiling or frowning [...] involves making coordinated movements of multiple muscles [...]. That such an expression of emotion is a goal-directed action follows just from its involving motor expertise and being coordinated around an outcome [...]. (Butterfill 2015, p. 446)

That expressing emotions relies on motor expertise is further supported by evidence that motor programs seem to have a fundamental role in the production of emotions, so that tampering with motor programs imposes limits on one's own emotional experience (see Davis et al. 2010).⁸

4.3 Butterfill's conjecture: facial expressions of emotions are processed as AMROs in categorical perception

To sum up, so far I have presented reasons for thinking that emotions being expressed are outcomes, not reducible to facial configurations. Now on to Butterfill's conjecture. This has it that, when facial expressions of emotions are

⁸ I am grateful to the editors of this volume for bringing this evidence to my attention.

categorically perceived, these are processed as actions—specifically, AMROs—rather than merely as facial configurations. In other words, the stimuli consisting in facial configurations would trigger a hypothesis about which motorically represented outcome is being pursued—e.g., happiness being expressed—and, consequently, about which action is being performed in order to achieve that motorically represented outcome—e.g., expressing happiness.

Butterfill’s conjecture about the categorical perception of facial expressions of emotions, by his own admission, requires that “the things categorised in in categorical perception of expressions of emotions are events rather than configurations or anything static” (2015, p. 446). While the idea that acoustic signals are processed as actions may have seemed reasonable given that acoustic signals are dynamic stimuli, the idea that facial configurations are processed as actions might seem surprising. In response to this concern, Butterfill observes that his conjecture is not in principle incompatible with the fact that the categorical perception of expressions of emotions may be triggered by static stimuli, such as the facial configurations described in section 3. In support of this idea, he cites evidence to the effect that static stimuli are sufficient to trigger motor programs in an observer (Borghi et al. 2007).

In the light of this conjecture, the data about categorical perception of facial expressions of emotions reviewed in section 3 could be explained in the following way: pairs of stimuli that fall in the same category are treated in the same way because they can be interpreted as part of actions directed to the same motorically represented outcome: that happiness (or fear) is expressed.

Interpreting the data in this way makes room for the fact that, if the stimuli were made more complex so as to include wider bodily configurations, contextual factors affecting their categorisation could be taken into account, just as contextual factors may affect the categorical perception of speech.

Butterfill supports his conjecture on the basis of a few considerations. Among these, there is the idea that facial expressions of emotions and phonemes are analogous in a number of ways—e.g., facial configurations alone might not be diagnostic of emotions, in the same way in which isolated acoustic signals might not be diagnostic of phonemes, and both are open to the influence of contextual factors in determining which emotion or phoneme is detected by an observer. Moreover, Butterfill points out that when stimuli are chosen in order to test the categorical perception of facial expressions of emotions, the guiding principle is not which facial configuration is more likely to be associated with a given emotion, but rather which facial configuration is more likely to express a given emotion. Therefore, his conjecture is in line with how the stimuli are categorised in the first place, and makes sense for plausible analogies between facial expressions of emotions and phonemes. So, Butterfill's conjecture seems one worth exploring.

5. A complementary conjecture: humans categorically perceive motor actions

I would now like to go back to motor actions, introduced in section 2, and present a conjecture that builds on and complements Butterfill's one. According to my conjecture, humans would categorically perceive motor actions. This is based on the idea that motor actions are AMROs. Let me provide reasons in support of the latter idea first, and then explain why this motivates considering the possibility that motor actions could be categorically perceived. In the next section, I will show the explanatory gain to be obtained from this conjecture.

5.1 Motor actions are AMROs

In order to show why it is reasonable to consider motor actions AMROs, i.e. actions directed to motorically represented outcomes, let me start by showing why motor actions should be thought of as directed to outcomes.

This is easily done. Recall from section 2 that motor actions were defined in behavioural terms as actions whose characterisation unavoidably involves mention of sequences of bodily configurations. Grasping is a paradigm example of a motor action.

Now, something being grasped should be considered an outcome, as opposed to merely a bodily configuration (or series of bodily configurations), for the following reasons. First, multiple different bodily configurations may be employed to achieve the outcome of something being grasped. The latter could be

achieved by using thumb and index finger in different configurations (e.g., with a smaller or greater distance between the fingertips), or using all of the fingers on one's hand, or even using a different effector (e.g., the mouth as opposed to the hand). On the other hand, the same series of bodily configuration (e.g., one's fingers closing around the handles of a pair of pliers) may achieve different outcomes (e.g., something being grasped, or something being released) depending on contextual factors (in this case, the shape of the pliers).⁹ Therefore, motor actions such as grasping are directed to outcomes (something being grasped), which are interestingly different from bodily configurations: the same outcome can be achieved by different sequences of bodily configurations, and the same sequence of bodily configuration can lead to different outcomes.

Now, why think that these outcomes are motorically represented? As mentioned in section 4.1, the ideal evidence for an outcome being motorically represented is that a given marker of motor processing should be found in correlation with an outcome being brought about. For an outcome (as opposed to a mere sequence of bodily configurations) to be represented, two conditions need to be fulfilled (as suggested most recently by Butterfill and Sinigaglia 2014, and

⁹ This clever manipulation was used in an experiment by Umiltà and colleagues (2008): two different pairs of pliers were constructed, such that, with one pair of pliers, closing one's fingers around the handles would result in an object being grasped, and, with the other pair of pliers, the exact same sequence of bodily configurations would result in an object being released.

earlier, e.g., by Sinigaglia 2010). First, the same marker of motor processing (e.g., the same rate of neural discharge) should be found by holding the outcome fixed, but varying sequences of bodily configurations. Secondly, different markers of motor processes (e.g., markedly different rates of neural discharge) should be found by holding a sequence of bodily configurations fixed, but altering the outcome, e.g. by changing contextual factors.

In the case of motor actions such as grasping, we have precisely this sort of ideal evidence, under both conditions required to say that an outcome, as opposed to a sequence of bodily configurations, is represented motorically (as has been observed, e.g., by Rizzolatti and Sinigaglia 2008, as well as by Butterfill and Sinigaglia 2014). For example, there is evidence that in the premotor cortex of the macaque monkey brain—specifically, in the area F5—there are populations of neurons that activate in correlation with a grasping act regardless of whether grasping is executed with the hand as opposed to with the mouth (Rizzolatti et al. 1988),¹⁰ thus indicating that the same outcome is represented while varying sequences of bodily configurations.

But there is also evidence to the effect that there are neurons—also in the area F5—that, in response to the same sequence of bodily configurations—e.g.,

¹⁰ Here I am appealing to single-cell recordings in the macaque monkey brain based on the idea, supported by Rizzolatti and colleagues (2002), that there is a sufficient analogy between this particular region of the macaque monkey brain and the Brodmann area 44 of the human brain.

that involved in grasping an object—fire differentially depending on the context in which grasping is performed. The different contexts were constituted by the presence or absence of an object to be grasped (Umiltà et al. 2001; see also Villiger et al. 2011). Therefore, motor actions are AMROs, in virtue of their outcomes being represented motorically.

5.2 Motor actions could be categorically perceived

Let me take stock. In section 4, I reported Butterfill's (2015) observation that the articulating phonemes and expressing emotions are AMROs, as well as his conjecture that facial expressions of emotions could be processed as AMROs within the context of categorical perception, and, relatedly, sorted into categories consisting in motorically represented outcomes (e.g., happiness being expressed). In the previous subsection, I have pointed out that motor actions are AMROs, too. On the basis of this observation and of Butterfill's conjecture, it becomes plausible articulating phonemes, expressing emotions and motor actions should be species of the same genus—namely, AMROs. Given that both speech and facial expressions of emotions are categorically perceived, I put forward the conjecture that motor actions could be categorically perceived, too. By analogy with the case of speech and (according to Butterfill's conjecture) facial expressions of emotions, my conjecture has it that the categories into which categorical perception would subdivide motor actions is the motorically represented outcomes around which motor actions are coordinated—e.g., something being grasped.

Considering the possibility that motor actions could be categorically perceived might sound surprising, given that many instances of categorical perception that I have discussed in this chapter involve static stimuli, such as facial configuration or colour hues. Even though in section 4.3 I mentioned the possibility that static stimuli could trigger the perception of events in relation to Butterfill's conjecture, the fact remains that motor actions themselves *are* events. How could the idea that motor actions are categorically perceived be operationalised?¹¹

Let me now clarify that the notion of categorical perception is perfectly compatible with the idea that the stimuli to be categorised are events rather than objects. This is clearest if you think of the case of the categorical perception of speech. The stimuli employed to test this phenomenon, as said in section 3, are acoustic sounds that constitute phonemes. These are events. But how might this work in practice in the case of motor actions?

That motor actions are categorically perceived means that the following should in principle be possible. A pair of distinct motor actions should be identified—one could be grasping with the hand, since it is a widely studied case, and another could be pushing away an object with the back of one's fingers. The two different motor actions should be performed with the same hand. On the basis of these two different motor actions, a number of stimuli—either static, such as

¹¹ I am grateful to a reviewer of this volume for inviting me to discuss this important issue.

snapshots, or dynamic, such as short clips—should be obtained, such that pairs of neighbouring stimuli involve bodily configurations (or sequences of bodily configurations, if the stimuli are dynamic), that differ by the same amount in terms of their kinematic features (e.g., distance between finger tips). If it is true that humans categorically perceive motor actions, then pairs of neighbouring stimuli should be hard to tell apart when they fall within the same category (e.g., something being grasped), but easy to distinguish when each belongs to a different category (something being grasped vs. something being pushed away), despite the fact that, by design, all the neighbouring pairs of stimuli differ by the same amount.

As to what the categories could be beyond something being grasped and (maybe) something being pushed away, there is evidence that specific neural populations in the premotor cortex become active in correlation with different action types, such as grasping (Rizzolatti et al. 1988). As mentioned in section 5.1, the activation of these populations of neurons correlates with outcomes, such as something being grasped. The fact that the organising principle is outcomes rather than sequences of bodily configurations has been already expounded in section 5.1: the same neural activation can be observed in correlation with different sequences of bodily configurations bringing about the same outcome, and same sequences of bodily configurations are treated differently in terms of neural discharge depending on how the context shapes the overall outcome.

Taken together, these action types constitute what has been referred to as a *motor vocabulary*, or a *vocabulary of motor acts* (Rizzolatti et al. 1988; see

Jeannerod 2006). The outcomes to which the actions forming this motor vocabulary are directed are therefore plausible candidates for the categories in which humans subdivide motor actions, but it is again an empirical question whether they really provide the categories that humans are sensitive to.¹² More generally, of course, whether my conjecture holds is an empirical question, to be settled by means of experimental evidence.

6. How the conjecture would explain neural mechanisms involved in the processing of AMROs

In this last section, I am going to show that, if the conjecture I am proposing turned out to be true, this would provide a good explanation of data we currently have about the involvement of certain neural mechanisms in the processing of motor actions, as well as a unifying explanation for the involvement of certain neural mechanisms in the processing of other AMROs.

First of all, we need a bit more detail about how categorical perception occurs. A reasonable model of how this could occur, that has been put forward in

¹² Support for the aspect of the conjecture concerning the categories into which motor actions are sorted is given by evidence that the organisation of actions in the brain in terms of outcomes influences our processing of action-related language (e.g., Marino and colleagues 2017). I am grateful to a reviewer of this chapter for bringing this to my attention.

the case of the categorical perception of speech (Liberman and Mattingly 1985), is that, in the course of categorically perceiving a certain auditory stimulus, a hypothesis is made as to what phoneme is being articulated, and the hypothesis is checked against the available evidence. If the evidence is compatible with the hypothesis, the hypothesis is reinforced. If the evidence is incompatible with the hypothesis, the hypothesis is revised.

This model has been further supplemented in the following way: hypothesising which phoneme is being articulated would involve the activation of motor processes in the observer's brain that would normally be involved in the production of one's own speech. The Motor Theory of Speech Perception (Liberman and Mattingly 1985) makes precisely this suggestion. Thus could motor processes be involved in the categorical perception of speech.¹³

As part of his proposed conjecture, Butterfill (2015) suggests that the categorical perception of expressions of emotions could work in an analogous way: a hypothesis could be made about what emotion is being expressed, and the hypothesis would be checked against the available evidence. Specifically, Butterfill suggests that a hypothesis as to which emotion is being expressed could involve the activation of processes in an observer that would be activated were the observer to have that emotion herself (2015, p. 448). In particular, this would

¹³ As I mentioned in footnote 7, this aspect of the Motor Theory of Speech Perception still stands.

result in outcomes being motorically represented in an observer (he notes this has already been proposed by Adolphs 2001).

After putting forward this aspect of the conjecture, Butterfill notes that there is evidence suggesting that this is precisely what could occur in the case of the processing of expressions of emotions. He reports evidence that, on the one hand, processes that would occur when one is having a certain emotion also occur while observing other individuals' emotions (Bastiaansen et al. 2009; Gallese et al. 2004; Rizzolatti and Sinigaglia 2008; van der Gaag et al. 2007; Wicker et al. 2003). Moreover, there is evidence that *disrupting* the occurrence of these processes in an observer interferes with the recognition of others' emotions (Niedenthal et al. 2001; Oberman et al. 2007; Pitcher et al. 2008).

Let us therefore see whether an analogous model of the processing of motor actions is viable. Indeed, various sources have proposed a model according to which motor actions are recognised through a process of making hypotheses and checking them against the available evidence that involves the activation of motor processes in an observer (see, e.g., Kilner et al. 2007).¹⁴

¹⁴ The notion of *understanding from the inside* has been put forward to indicate cases in which an observer motorically represents an outcome that an observed individual is trying to fulfil (e.g., Rizzolatti and Sinigaglia 2010). An interesting topic of investigation, which is best left to another occasion, is the relationship between the motor processes hypothesized to be involved in the processing of motor actions and mindreading.

An especially pertinent source of evidence for motor process involvement in the categorical perception of motor actions is given by an experiment carried out by Cattaneo and colleagues (2010). In this experiment, a sensory-motor adaptation paradigm was employed: participants were trained to perform either a *push away* or a *pull towards* movement with their hand while blindfolded. This motor training was shown to have an impact on the subsequent visual recognition of analogous hand actions, and resulted in ambiguous stimuli being classified as *push away* movements following a training involving *pull towards* movements, and viceversa. The fact that activating a certain motor process by means of a repeatedly performed action impacts on the labelling of a subsequently observed action suggests that motor processes are involved in the identification of outcomes: if the recognition of a motor action was a purely visual phenomenon, it is unclear why motor training in the absence of visual stimuli should impact on it.

Moreover, transcranial magnetic stimulation (TMS) over the ventral premotor cortex suppressed the adaptation aftereffect in the recognition of an action. That is, a temporary disruption of areas involved in the production of motor actions impacted on the recognition of those actions (Cattaneo et al. 2010). This result dovetails nicely with that reported previously, to the effect that disrupting the occurrence of the processes involved in expressing an emotion in an observer interferes with the recognition of others' emotions.

Now, if motor actions turned out to be categorically perceived, along with speech and facial expressions of emotions, a unifying explanation could be given as to why areas involved in the production of speech, expressions of

emotions and motor actions seem to have a role in the perceptual processing of these stimuli. This would be explained in the following way. Articulations of phonemes, expressions of emotions and motor actions are all AMROs. Categorical perception, if Butterfill's conjecture and my proposed one turn out to be correct, sorts AMROs into categories corresponding to motorically represented outcomes. This categorisation relies on a process of hypothesis testing that draws on the very processes involved in the production of these stimuli. This is how the conjecture that motor actions are categorically perceived completes Butterfill's conjecture to that expressions of emotions are categorically perceived as AMROs, giving rise to a unifying theoretical framework. This same framework would accommodate the idea that categorically perceiving speech sorts stimuli into phonemes conceived as motorically represented outcomes.

7. Conclusion

In the foregoing, I have presented my conjecture that humans categorically perceive motor actions. I have done so by drawing an analogy with expressions of emotions and articulations of phonemes, which can be plausibly thought of as actions directed to motorically represented outcomes, as are motor actions. My conjecture builds on an interpretation of the categorical perception of speech whereby acoustic signals are processed as actions directed to a motorically represented outcome, and the categories consist in motorically represented outcomes. My conjecture also builds on the one recently put forward by Butterfill

(2015), which interprets the categorical perception of facial expressions of emotions in terms of facial expressions of emotions being processed as AMROs. The two conjectures and the interpretation of the categorical perception of speech in terms of the processing of actions naturally complement each other, and together give rise to a unifying theoretical framework, which could explain data showing the involvement of motor processes in an observer's brain when processing others' actions directed to motorically represented outcomes. While this does not show the conjecture to be true, it makes it worthy of consideration.

References

- Adolphs, R. (2001). The neurobiology of social cognition. *Current opinion in neurobiology*, *11*(2), 231-239.
- Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., Moscovitch, M., & Bentin, S. (2008). Angry, disgusted, or afraid? Studies on the malleability of emotion perception. *Psychological science*, *19*(7), 724-732.
- Bastiaansen, J. A., Thioux, M., & Keysers, C. (2009). Evidence for mirror systems in emotions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1528), 2391-2404.
- Beale, J. M., & Keil, F. C. (1995). Categorical effects in the perception of faces. *Cognition*, *57*(3), 217-239.

- Borghi, A. M., Bonfiglioli, C., Lugli, L., Ricciardelli, P., Rubichi, S., & Nicoletti, R. (2007). Are visual stimuli sufficient to evoke motor information?: Studies with hand primes. *Neuroscience Letters*, *411*(1), 17-21.
- Bornstein, M. H., & Korda, N. O. (1984). Discrimination and matching within and between hues measured by reaction times: Some implications for categorical perception and levels of information processing. *Psychological research*, *46*(3), 207-222.
- Brooks, J. A., & Freeman, J. B. (2018). Conceptual knowledge predicts the representational structure of facial emotion perception. *Nature Human Behaviour*, *2*(8), 581-591.
- Butterfill, S. A. (2015). Perceiving expressions of emotions: What evidence could bear on questions about perceptual experience of mental states? *Consciousness and Cognition*, *36*, 438-451.
- Butterfill, S. A., & Sinigaglia, C. (2014). Intention and motor representation in purposive action. *Philosophy and Phenomenological Research*, *88*(1), 119-145.
- Calder, A. J., Young, A. W., Perrett, D. I., Etcoff, N. L., & Rowland, D. (1996). Categorical perception of morphed facial expressions. *Visual Cognition*, *3*(2), 81-118.
- Caruana, F., & Viola, M. (2018). *Come funzionano le emozioni: da Darwin alle neuroscienze*. Il Mulino.
- Cattaneo, L., Barchiesi, G., Tabarelli, D., Arfeller, C., Sato, M., & Glenberg, A. M. (2010). One's motor performance predictably modulates the

- understanding of others' actions through adaptation of premotor visuo-motor neurons. *Social Cognitive and Affective Neuroscience*, nsq099.
- Davis, J. I., Senghas, A., Brandt, F., & Ochsner, K. N. (2010). The effects of BOTOX injections on emotional experience. *Emotion*, *10*(3), 433-440.
- Daoutis, C. A., Pilling, M., & Davies, I. R. (2006). Categorical effects in visual search for colour. *Visual cognition*, *14*(2), 217-240.
- Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, *171*(3968), 303-306.
- Ekman, P. (1992). Are There Basic Emotions? *Psychological Review* *99*, 550-553.
- Ectoff, N. L., & Magee, J. J. (1992). Categorical perception of facial expressions. *Cognition*, *44*(3), 227-240.
- Ferretti, G., & Caiani, S. Z. (2019). Solving the interface problem without translation: The same format thesis. *Pacific Philosophical Quarterly*, *100*(1), 301-333.
- Fujimura, T., Matsuda, Y. T., Katahira, K., Okada, M., & Okanoya, K. (2012). Categorical and dimensional perceptions in decoding emotional facial expressions. *Cognition & emotion*, *26*(4), 587-601.
- Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic bulletin & review*, *13*(3), 361-377.
- Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*(2), 593-609.
- Gallese, V., & Sinigaglia, C. (2011). What is so special about embodied

- simulation?. *Trends in cognitive sciences*, 15(11), 512-519.
- Goldstein, L., & Fowler, C. A. (2003). Articulatory phonology: A phonology for public language use. *Phonetics and phonology in language comprehension and production: Differences and similarities*, 159-207.
- Hamilton, A. F., & Grafton, S. T. (2007). Action outcomes are represented in human inferior frontoparietal cortex. *Cerebral Cortex*, 18(5), 1160-1168.
- Harnad, S. (1987). Psychophysical and cognitive aspects of categorical perception: A critical overview. In Harnad, S. (ed), *Categorical perception: The groundwork of cognition*, 1-52.
- Harnad, S. (2003) Categorical Perception. *Encyclopedia of Cognitive Science*. Nature Publishing Group/Macmillan.
- Izard, C.E. (1971). *The Face of Emotion*. New York, NY: Appleton-Century-Crofts.
- Jeannerod, M. (1994). The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain sciences*, 17(2), 187-202.
- Jeannerod, M. (2006). *Motor cognition: What actions tell the self*. Oxford University Press.
- Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). The mirror-neuron system: a Bayesian perspective. *Neuroreport*, 18(6), 619-623.
- Kikutani, M., Roberson, D., & Hanley, J. R. (2008). What's in the name? Categorical perception for unfamiliar faces can occur through labeling. *Psychonomic bulletin & review*, 15(4), 787-794.

- Kotsoni, E., de Haan, M., & Johnson, M. H. (2001). Categorical perception of facial expressions by 7-month-old infants. *Perception*, *30*(9), 1115-1125.
- Lieberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of experimental psychology*, *54*(5), 358-368.
- Lieberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, *21*(1), 1-36.
- Lieberman, A. M., & Whalen, D. H. (2000). On the relation of speech to language. *Trends in cognitive sciences*, *4*(5), 187-196.
- Marino, B. F., Borghi, A. M., Buccino, G., & Riggio, L. (2017). Chained activation of the motor system during language understanding. *Frontiers in psychology*, *8*, 199.
- McKone, E., Martini, P., & Nakayama, K. (2001). Categorical perception of face identity in noise isolates configural processing. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(3), 573-599.
- Mylopoulos, M., & Pacherie, E. (2017). Intentions and motor representations: The interface challenge. *Review of Philosophy and Psychology*, *8*(2), 317-336.
- Nanay, B. (2013). *Between perception and action*. Oxford University Press.
- Niedenthal, P. M., Brauer, M., Halberstadt, J. B., & Innes-Ker, Å. H. (2001). When did her smile drop? Facial mimicry and the influences of emotional state on the detection of change in emotional expression. *Cognition & Emotion*, *15*(6), 853-864.

- Nygaard, L. C., & Pisoni, D. B. (1995). Speech perception: New directions in research and theory. In J. Miller et al. (eds.), *Speech, Language and Communication* (London: Academic Press), 72-75.
- Oberman, L. M., Winkielman, P., & Ramachandran, V. S. (2007). Face to face: Blocking facial mimicry can selectively impair recognition of emotional expressions. *Social neuroscience*, 2(3-4), 167-178.
- Pacherie, E. (2000). The content of intentions. *Mind & Language*, 15(4), 400-432.
- Pacherie, E. (2008). The phenomenology of action: A conceptual framework. *Cognition*, 107(1), 179-217.
- Pavese, C. (2015). Practical Senses. *Philosophers' Imprint*, 15(29), 1-25.
- Pitcher, D., Garrido, L., Walsh, V., & Duchaine, B. C. (2008). Transcranial magnetic stimulation disrupts the perception and embodiment of facial expressions. *Journal of Neuroscience*, 28(36), 8929-8933.
- Repp, B. H. (1984). Categorical perception: Issues, methods, findings. *Speech and language: Advances in basic research and practice*, 10, 243-335.
- Repp, B. H., & Liberman, A. M. (1987). Phonetic category boundaries are flexible. In Harnad, S. (ed.), *Categorical perception: The groundwork of cognition*, 89-112.
- Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G., & Matelli, M. (1988). Functional organization of inferior area 6 in the macaque monkey: II. Area F5 and the control of distal movements. *Experimental Brain Research*, 71(3), 491-507.

- Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive brain research*, 3(2), 131-141.
- Rizzolatti, G., Fogassi, L., & Gallese, V. (2002). Motor and cognitive functions of the ventral premotor cortex. *Current opinion in neurobiology*, 12(2), 149-154.
- Rizzolatti, G., & Sinigaglia, C. (2008). *Mirrors in the brain: How our minds share actions and emotions*. Oxford University Press, USA.
- Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nature reviews neuroscience*, 11(4), 264-274.
- Rizzolatti, G., & Sinigaglia, C. (2016). The mirror mechanism: a basic principle of brain function. *Nature Reviews Neuroscience*, 17(12), 757.
- Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social psychology*, 39(6), 1161-1178.
- Shepherd, J. (2019). Skilled action and the double life of intention. *Philosophy and phenomenological research*, 98(2), 286-305.
- Sinigaglia, C. (2010). Mirroring and understanding action. In *EPSA philosophical issues in the sciences* (pp. 227-238). Springer, Dordrecht.
- Tomkins, S. S. (1962). *Affect, Imagery, Consciousness. Vol. 1: The Positive Affects*. New York, NY: Springer.

- Umiltà, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., & Rizzolatti, G. (2001). I know what you are doing: A neurophysiological study. *Neuron*, *31*(1), 155-165.
- Umiltà, M. A., Intskirveli, I., Grammont, F., Rochat, M., Caruana, F., Jezzini, A., Gallese, V. & Rizzolatti, G. (2008). When pliers become fingers in the monkey motor system. *Proceedings of the National Academy of Sciences*, *105*(6), 2209-2213.
- Van der Gaag, C., Minderaa, R. B., & Keysers, C. (2007). Facial expressions: what the mirror neuron system can and cannot tell us. *Social neuroscience*, *2*(3-4), 179-222.
- Villiger, M., Chandrasekharan, S., & Welsh, T. N. (2011). Activity of human motor system during action observation is modulated by object presence. *Experimental brain research*, *209*(1), 85-93.
- Whalen, D. H. (2019). The Motor Theory of Speech Perception. In *Oxford Research Encyclopedia of Linguistics*. (DOI: 10.1093/acrefore/9780199384655.013.404)
- Wicker, B., Keysers, C., Plailly, J., Royet, J. P., Gallese, V., & Rizzolatti, G. (2003). Both of us disgusted in My insula: the common neural basis of seeing and feeling disgust. *Neuron*, *40*(3), 655-664.
- Witzel, C., & Gegenfurtner, K. R. (2014). Category effects on colour discrimination. In Anderson, W., Biggam, C. P., Hough, C., & Kay, C. (Eds.) *Colour Studies: A broad spectrum*. John Benjamins Publishing Company, 200-11.

- Wolfe, J. M., Friedman-Hill, S. R., Stewart, M. I., & O'Connell, K. M. (1992).
The role of categorization in visual search for orientation. *Journal of
Experimental Psychology: Human Perception and Performance*, *18*(1),
34-49.
- Young, A. W., Rowland, D., Calder, A. J., Etcoff, N. L., Seth, A., & Perrett, D. I.
(1997). Facial expression megamix: Tests of dimensional and category
accounts of emotion recognition. *Cognition*, *63*(3), 271-313.