

Behaviour reading and the role of motor cognition in understanding action and speech

s.butterfill@warwick.ac.uk; revision 1

[***must read (Biro and Leslie 2007). re-read Sommerville and Decety 2006. Go over (Browman and Goldstein 1992) carefully.]

Abstract

What is the role of mirror neurons, and of mirroring motor cognition more generally, in understanding intentional action? This paper argues that behaviour reading is the mechanism by which such motor cognition enables us to understand action and ascribe propositional attitudes. Behaviour reading involves segmenting the flow of movements into behavioural chunks and discerning structure from patterns in the chunks. Despite considerable research in multiple disciplines, the nature of these chunks and how they are represented has not been explained. This paper develops and defends the conjecture that the role of mirroring motor cognition is to segment streams of bodily motion into chunks, which are object-directed motor actions.

I. Introduction

This paper identifies connections between behaviour reading and mirroring motor cognition. Behaviour reading involves segmenting the flow of movements into behavioural chunks and discerning structure from patterns in the chunks. Mirroring motor cognition is the activation of motor commands in the observer of an action which mirror the motor commands of the agent executing the action. These two phenomena have been investigated in isolation from each other.

Connecting them enables us to fill gaps in the research on both sides.

A range of theorists have argued that our ability to predict and interpret others' actions and utterances in terms of their beliefs, desires and intentions rest on behaviour reading, which is more basic than intentional explanation (D. Baldwin and Baird 2001; Byrne 1999; Newtson, Engquist and Bois 1977; Povinelli and Giambrone 2001). I assume that they are correct at least insofar as there is an ability to read behaviour which does not presuppose facility with propositional attitude psychology. This paper investigates behaviour reading and its relation to mirror neurons and mirroring motor cognition generally. It articulates and defends a single conjecture: that mirroring motor cognition plays a fundamental role in behaviour reading, that of discerning and providing representations of the smallest behavioural segments ("chunks"), which is exactly the role mirroring motor cognition plays in comprehending language according to the motor theory of speech perception (Liberman and Mattingly 1985; Liberman and Whalen 2000; Galantucci, Fowler and Turvey 2006). The first part of this conjecture has been endorsed before (Rizzolatti, Fogassi and Gallese 2001; Byrne 2003) but not developed in detail. The conjecture enables us to characterise the mechanisms linking mirroring motor cognition to action understanding and theory of mind; it supports refined claims about the relation between action understanding and language (cf. Rizzolatti and Arbib 1998); and it solves a fundamental difficulty, so far unaddressed, for accounts of behaviour reading.

The rest of this introduction summarises the more detailed discussion in the following sections.

There is a close analogy between perceiving speech and reading behaviour. In the case of speech, acoustic and visual stimuli are segmented into phonetic

primitives (chunks), and patterns among chunks are then discerned enabling individuals to identify words (J. R. Saffran, Newport and Aslin 1996) and clauses (Soderstrom, Nelson and Jusczyk 2005) independently of any semantic knowledge. Research on behaviour reading has shown that subjects identify word-and clause-like units in the flow of behaviour, and that the cues and learning mechanisms involved are comparable to those used in processing speech. This is not surprising: given that phones *are* behaviours (Liberman and Whalen 2000; Browman and Goldstein 1992), perceiving speech is a special case of reading behaviour.

Two prominent issues in speech perception are: which variations in stimuli cause them to be classified as the same or different phonetic primitives, and what is the nature of these primitives? By contrast parallel questions have barely been raised in the case of behaviour reading. As we will see, there is not even direct evidence that there are recurring segments, although detecting patterns in behaviours depends on identifying recurring segments. A key problem for accounts of behaviour reading, then, is to characterise the primitive behavioural chunks.

The first step to resolving this hypothesis may appear so obvious that it has been assumed without comment. Just as articulatory tract gestures are the basic units of speech perception and so of linguistic comprehension, so it seems plausible to hypothesise that the basic units of behaviour reading will be motor actions of some kind. I take this hypothesis to be correct but do not regard it as solving the problem for it immediately raises two further questions: which kind of motor actions are behavioural chunks, and how are they represented?

There is a complementary problem concerning the role of mirroring motor cognition in action observation. Observing action involves mirroring motor cognition in much the way that perceiving

speech does: as the action unfolds, the same motor commands are activated in the observer as in the agent (e.g. M. Gangitano, Mottaghy and Pascual-Leone 2001; Fadiga, Craighero, *et al.* 2002). In the case of speech, there is wide agreement about at least one purpose of mirroring motor cognition, for it enables one to identify phonemes or allophones reliably in the face of variation in acoustic and visual signals (Lieberman and Mattingly 1985). In this way, mirroring motor cognition plays an essential if modest role in linguistic communication. In the case of action observation, however, there is marked disagreement about the utility of mirroring motor cognition. Claims that mirroring motor cognition provide direct knowledge of others' intentions (e.g. Fogassi, Ferrari, *et al.* 2005) face insurmountable objections (Jacob and Jeannerod 2005; Borg 2007). On the other hand, weakening these claims so that mirroring motor cognition provides knowledge only of sequences of small-scale motor actions and their outcomes (e.g. Gallese, Keysers and Rizzolatti 2004) leaves its utility unclear because it is unclear why knowing the outcomes of small-scale motor actions matters. After all, the outcome of grasping a ball is just having a ball in your hand; action interpretation seems to be concerned with why the ball is there. In fact, recently some researchers have shifted emphasis from action understanding to suggestions that the utility of mirroring may consist in its allowing us to discover facts about objects rather than agents (Rotman, Troje, *et al.* 2006) or to disconfirm prior goal attributions rather than do identify goals (Csibra 2007). On the face of it, then, there is a problem about the function of mirroring motor cognition in action observation.

A natural conjecture connects behaviour reading with mirroring motor cognition. One function of mirroring motor cognition is to segment the flow of behaviour and to represent the fundamental chunks

(Rizzolatti, Fogassi and Gallese 2001; Byrne 2003). If this is correct, the role of mirroring motor cognition in understanding action is comparable to its role in understanding speech: it does not give us direct insight into intentional or semantic properties, but it does play a fundamental role in the ascription of those properties by enabling us to identify the units and patterns of behaviour that are vehicles of intention and meaning.

This conjecture about the function of mirroring motor cognition tells us two things about behaviour reading. First, the primitive chunks of behaviour are represented by motor commands. (This explains how it is possible to represent them without insight into the nature of the motor system.) Second, the primitive chunks of behaviour are the motor actions to which the mirroring motor system is sensitive. Accordingly we can lever the existing research on invariance in the responses of mirror neurons (e.g. Umiltà, Kohler, *et al.* 2001) and motor cognition (e.g. Massimo Gangitano, Mottaghy and Pascual-Leone 2004) to make predictions about the nature of the chunks on which behaviour reading builds.

In short, then, speech perception is a case of behaviour reading more generally and mirroring motor cognition plays the same role in all behaviour reading that it plays in speech perception: it enables us to segment and represent the primitive chunks of behaviour which are the building blocks for larger patterns corresponding to intentional actions.

The following sections support the claims made in this outline. Behaviour reading is given a selective but relatively detailed overview because no recent survey of this literature exists.

II. Behaviour reading

Behaviour reading involves segmenting a stream of bodily movement into chunks and discerning structural relations among chunks on the basis of statistical patterns and other cues. To illustrate, chunks may include reaching, tearing, poking, or orienting to an object or location; and patterns may include orienting to an object's raising the probability of reaching for it.

Behaviour reading is held to be useful or even necessary for recognising intentions (Newtson, Engquist and Bois 1977: 861; D. A. Baldwin, Baird, *et al.* 2001: 708). Behaviour reading may also have functions independently of a capacity to ascribe intentions: efficiently representing events (Kurby and Zacks 2008), discerning structure in actions in such a way as to identify their effects (Byrne 1999), and predicting either what others will do (Povinelli 2001) or when an event likely to be of interest will occur (Swallow and Zacks 2008: 121).

Behaviour reading is also thought to be important for development. Much as identifying phonemes, grouping them into word-like clusters and uncovering clause-like clusters in these patterns is held to "bootstrap" the development of linguistic comprehension (Werker and Yeung 2005), so also behaviour reading is held to be a steppingstone to understanding intentional action (Saylor, Baldwin, *et al.* 2007). This is plausible because (as we will see) behaviour reading identifies the structures to which ascriptions of belief, desire and intention are explanatory.

Segmenting chunks. Experimental research on behaviour reading was pioneered by Darren Newtson who showed adults a five minute silent video of a man performing everyday activities in an unbroken sequence, including filling in a questionnaire, putting a match into a bin and stamping out a cigarette on the

floor. Subjects were asked to identify the boundaries of fine-grained chunks that were “meaningful sequences of action” (1973). Newton found reliable intra- and interpersonal agreement on the location of such boundaries, and also that boundaries shift depending on context (see Newton, Engquist and Bois 1977: 847 for an overview; he calls the boundaries “breakpoints”). These results have been confirmed using indirect behavioural measures (Newton and Engquist 1976) and fMRI (Zacks, Braver, *et al.* 2001; Zacks, Swallow, *et al.* 2006). Follow-up work has shown that extracting chunks of behaviour occurs automatically and online, so may be a perceptual process (Zacks, Tversky and Iyer 2001). This suggests extracting chunks of behaviour may be compared to the perceptual segmentation of phonemes from acoustic and visual stimuli.

Newton hypothesised that chunks of behaviour serve as an intermediate representation for ascribing intention. This hypothesis is supported by the discovery that segmenting chunks has effects on narrative comprehension, recognition, learning and memory (see Kurby and Zacks 2008 for a review). Perhaps, then, comparison with speech can be extended: the role of phonemes in linguistic comprehension may provide a useful model for the role of the role of behaviour chunks in attributing beliefs, desires and intentions (D. Baldwin and Baird 2001).

One problem noted above concerns the nature and representation of these behavioural chunks. This problem will be discussed later; for now I take for granted that suitable chunks occur.

The usefulness of behaviour reading depends on the complexity of the patterns that can be discerned. A key distinction is between ‘serial’ and ‘hierarchical’ patterns. This distinction is analogous to that between finite state and phrase structure grammars

(Chomsky 1956), but can be understood informally for present purposes. Serial patterns involve relations among consecutive chunks of behaviour and so can be identified by relatively simple learning mechanisms that track statistical relations between pairs of consecutive chunks (Conway and Christiansen 2001). Some serial patterns arise from the fact that pursuit of a goal characteristically involves a sequence of adjacent chunks; for example, eating a hamburger involves grasping it and then repeatedly bringing it to the mouth. Hierarchical patterns involve relations among non-consecutive chunks. Such patterns sometimes occur because pursuit of a goal involves subgoals that can be performed in different orders, as when making a burger involves preparing the bun and fillings individually before bringing them together. Patterns in the behavioural chunks that constitute burger-making activities cannot be learnt only by tracking relations between pairs of consecutive chunks.

The nature and recognition of these patterns can be understood on the model of speech perception, where the organisation of phonemes into words counts as a sequential pattern (ignoring, for simplicity, complex variability in the allophones that realise phonemes) and organisation of words into clauses counts as an hierarchical pattern.

Sequential patterns. Consider the following sequence of mundane actions:

... push door, walk through frame, walk to desk, grasp paper, orient to pen, reach for pen, write on paper, fold paper, place pen, reach for headphones, push button, ...

Here the flow of movements is described in terms of phoneme-like chunks that infants and adults can readily identify (Saylor, Baldwin, *et al.* 2007). Infants (from around 9 months) and adults can also group

these chunks into word-like units such as the letter-writing sequence in the above example: these are chunks that seem to belong together and to be separate from previous and ensuing chunks (Saylor, Baldwin, *et al.* 2007). The boundaries of these units are generally “intention-relevant” (Saylor, Baldwin, *et al.* 2007), corresponding to the completion of a goal or subgoal.

In some cases, identifying these word-like behaviour units involves ascriptions of intention or purpose (Zacks 2004). However, ascribing intention is not generally necessary for isolating these units. There are also at least two types of non-intentional cue to their boundaries. First, commencement and completion of a goal or subgoal typically coincide with dramatic changes in the physical features of the movements (Zacks, Tversky and Iyer 2001). Baldwin and Baird express this idea graphically with the notion of a “ballistic trajectory that provides a temporal contour or ‘envelope’ demarcating one intentional act from the next” (D. Baldwin and Baird 2001). Research using schematic animations has shown that adults use a variety of movement features to group behavioural chunks into word-like units (Zacks 2004; Bridgette M. Hard, Tversky and Lang 2006). The second non-intentional cue is statistical. Chunks of behaviour that are all steps to a single goal or subgoal are more likely to occur in succession than chunks not so related; thus transitional probabilities in the sequence of chunks could in principle be used to identify intention-relevant units, much as phonemes can be grouped into words by means of tracking transitional probabilities (J. R. Saffran, Newport and Aslin 1996; Gómez and Gerken 2000). In fact, Baldwin and colleagues demonstrated that adults can learn to group small chunks of behaviour into larger word-like units on the basis of statistical features alone (2008). And since the statistical learning mechanism required for discerning such units is

automatic (Fiser and Aslin 2001), domain-general (Kirkham, Slemmer and Johnson 2002) and present in other species including monkeys (Hauser, Newport and Aslin 2001) and rats (Toro, Trobalç and n 2005), it seems plausible to suppose that sensitivity to intention-relevant units in sequences of behavioural chunks can be found in individuals that lack the ability to ascribe intentions or goals.

In short, relatively humble learning mechanisms enable chunks of behaviour to be grouped into word-like units on the basis of ballistic envelopes and transitional probabilities. Just as words are semantically relevant units yet can be discerned without knowledge of semantics, so these behavioural units have boundaries corresponding to the fulfilment of intentions but can be identified without knowledge of intention.

Note that discerning pattern in sequences of behaviour is not limited to intention-relevant units: behaviour reading may group behavioural chunks into units that have no clear relevance to a goal, as studies requiring subjects to segment behaviours using reversed animations and other manipulations show (Bridgette M. Hard, Tversky and Lang 2006; D. Baldwin, Andersson, *et al.* 2008; Swallow and Zacks 2008). Similarly, use of statistical cues to group phonemes into words does not yield only meaningful words. Discerning intention-*ir*relevant units may be undesirable overgeneration or it may have predictive value. The definition of behaviour-reading allows for the exploitation of any useful patterns, not just those induced by goals and intentions.

Hierarchical patterns. In discussing sequential patterns we were concerned with grouping together strings of consecutive behavioural chunks. But there are also patterns in behaviour connecting non-consecutive chunks. For example, making a burger involves

several steps whose order is only loosely constrained, where some of these steps can be omitted or replaced (veggie burger, hamburger) and where steps can be interspersed with irrelevant actions (answering the phone). Grouping together all and only the behavioural units involved in making a burger therefore involves discerning hierarchical structure. Clearly this is possible, but is it possible independently of ascribing intention?

Richard Byrne studied Rwandan mountain gorillas' preparation of nettles for eating, a complex task involving several sub- and sub-sub-goals (2003). He identified two types of cue that reveal structure in this behaviour. The first type of cue depends on comparing different occasions: repetitions, omissions, pauses and substitutions indicate boundaries of significant units, as do points where one-off (and therefore probably extraneous) behaviours are interwoven and the points where recovery from interruption happens smoothly. The second type of cue includes changes in motion features and pauses. I shall call these "prosodic" cues because they resemble cues such as changes in pitch and pauses used by infants to identify clause boundaries (e.g. Soderstrom, Nelson and Jusczyk 2005; Seidl and Cristi... 2008).

The existence of such cues shows that it is possible in principle to identify plan-induced hierarchical patterns in behaviours without knowledge of intention. Is there any evidence on whether these cues are in fact detected?

Several researchers who appear to be addressing this question are actually concerned only with identifying word-like units (Zacks, Tversky and Iyer 2001; Bridgette M. Hard, Tversky and Lang 2006; Bridgette Martin Hard, Lozano and Tversky 2006; Kurby and Zacks 2008). They label this 'hierarchical structure' because flowing motion is first being divided into phoneme-like behavioural chunks whose organisation then defines boundaries of larger word-

like units. We can thus regard the chunks as parts of larger units, which justifies the label 'hierarchical'. However, discerning word-like units only requires sequential *learning*—it only requires tracking associations between consecutive chunks. Sensitivity to these "hierarchical structures" does not imply that one can also detect patterns involving optional components or variable ordering, such as those involved in making a burger or preparing a nettle. Following Conway and Christiansen (2001), I therefore reserve the label 'hierarchical' for patterns involving relations among potentially non-consecutive chunks.

To my knowledge the question of learning genuinely hierarchal patterns in behaviour has not been directly addressed. We do know from research in other domains that hierarchical patterns can be learnt by adults (Newport and Aslin 2004) and infants from 12-months or earlier (J. Saffran, Hauser, *et al.* 2008). Learning hierarchical patterns differs from learning sequential patterns in that it requires attention (Cohen, Ivry and Keele 1990) and there appear to be constraints on the types of pattern that can be learnt (Newport and Aslin 2004).

Infants' ability to identify patterns in action may be boosted by 'motionese'. Brand and colleagues (2002) found that "mothers spontaneously modified their infant-directed actions in a number of ways ... a higher level of interactiveness, greater repetitiveness and movements that were larger in scale but reduced in complexity." These modifications may assist infants' efforts to detect hierarchical patterns in behaviour, amplifying and inserting prosodic cues to reduce infants' dependence on statistics.

There is uncertainty on whether non-humans are capable of learning hierarchical patterns of any kind. Evidence for an absence of hierarchical learning has been offered for cotton-top tamarins (Fitch and Hauser 2004) and rats (Toro, Trobalç and n 2005),

and there is no convincing evidence for such learning in nonhumans generally (Corballis 2007). Inability to learn hierarchical patterns would limit the power of behaviour reading to activities that are often uniformly realised (no variation in the steps taken) and performed without interruption.

In short, executing goals with sub-goals tends to create statistical and prosodic cues in the pattern of behaviours. Discerning these cues makes it possible to recover the structure of the behaviour without knowledge of intention. Such patterns can be discerned by humans, including infants, but possibly not by other animals. Learning these patterns is not automatic and requires attention.

III. What are behaviours?

Current research as overviewed in the previous section has yet to squarely face the question, What are behaviours? In answering this question we can focus on the phoneme-like chunks of behaviour, for more complex behavioural units are built from these. This question should be addressed from the point of view of behaviour readers: What do they represent when they represent behaviour?

This question has two related parts. One concerns factors with respect to which behaviour reading is invariant. To illustrate, reaching for a ball may involve quite different motions on different occasions. Do subjects classify different reaches as instances of a single chunk or as instances of unrelated chunks? The other part of the question is about the nature of these behaviours. Are they intentional actions, bodily movements or what? This is partly a philosophical issue: just as philosophers have attempted to say what intentional action is (e.g. Davidson 1963 [1980], 1971 [1980]; Bratman 1984), so there should be a philosophical account of non-intentional behavioural chunks. But this issue is not narrowly philosophical,

for it has implications for how behavioural chunks could be represented.

Our guiding hypothesis is that behaviour reading is possible independently of ascribing goals or intentions. This means that behaviour readers cannot represent chunks of behaviours as intentional actions or in any way that presupposes intentions or goals. A natural alternative is to suppose that behaviour readers represent chunks of behaviours as sequences of movements by bodies or body parts.

This alternative is unacceptable, however. The power of behaviour reading depends on extracting patterns, which requires a relatively small number of chunks that recur. Further, extracting patterns that are sometimes relevant to intentions requires some degree of invariance with respect to anatomical and postural differences between individual agents and environmental obstacles to action. The movements involved in achieving even the simplest goal, such as grasping an apple, vary from occasion to occasion and lack any such invariance.

Comparison with speech is useful here. Phonemes are realised by different allophones on different occasions (Davenport and Hannahs 1998: 96ff.), and allophones in turn are realised by various acoustic signals (Repp and Liberman 1987; Nygaard and Pisoni 1995: 72–5). That is, phonemes are separated from acoustic and visual signals by at least two layers of abstraction. Given the requirements of recurrence and intention-relevance on behavioural chunks, they must be similarly distant from movements. Indeed, variation in many of the same factors affect both speech and action: anatomy, posture, manner of execution (e.g. fast vs. slow), and environmental impediments. Context is a further factor: the way a phonetic gesture or behavioural chunk is realised depends on neighbouring chunks (Davenport and Hannahs 1998: 130ff.; Johnson-Frey, McCarty and Keen 2004). Invariance with respect to context and

the other factors requires layers of abstraction between movement and behavioural chunks. This is why the basic chunks used in behaviour reading cannot be movements.

So what is the nature of this abstraction, this analogue of a phoneme in behaviour reading, if it is neither goal-directed nor movement? So far no research has directly addressed this question.

Existing research does not directly address this question. In the case of speech, researchers have investigated when subjects classify different stimuli as instances of a single phoneme (e.g. Repp and Liberman 1987), giving rise to detailed claims about the nature of the primitive chunks (e.g. Browman and Goldstein 1992). Newton, whose work set the scene for contemporary research, did not address the parallel question about when different stimuli are treated as the same behavioural chunks. To my knowledge he never even gave an informal description of the fine-grained segments of behaviour his subjects identified. And his research, which used video, involved unvarying sequence of movements. He showed that inter- and intra-subject agreement on where the boundaries of chunks fall (that is, “breakpoints”) when subjects are given precisely the same movements. It is consistent with this finding to deny that relevantly similar boundaries would be identified in other movements that differ superficially but are attempts to achieve the same goals. In principle his findings might be explained by subjects’ relying on changes in motion features that do not result in categories applicable to different occasions. So not only do we not know what behavioural chunks are, there is no direct evidence that they even exist.

Recent research has generally ignored this lacuna (Loucks and Bladwin 2006: 240). The nature of the stimuli, which are invariably precisely repeated movements delivered by video or computer, tends to

obscure the importance of an ability to identify varying motions as the same behavioural category. Accordingly, research has focussed questions about which factors induce subjects to identify a boundary between chunks and not on whether these identifications give rise to a coherent set of categories. To illustrate, if I understand Zacks and colleagues' Event Segmentation Theory (EST), it works independently of whether there are recurring categories or whether behavioural segments are identified fresh in every case (Zacks, Speer, *et al.* 2007). This suggests these researchers have deliberately remained neutral on the issue of whether behaviour reading involves recurring segments. However, without recurring chunks, behaviour reading would have relatively predictive little power and depend on ascription of intention. (Indeed Newton or Zacks *et al.* may not have envisaged behaviour reading is possible independently of knowledge of intention.) The view that behaviour reading is possible independently of knowledge of intention requires recurring behavioural chunks, as indeed do most of the claims about its utility.

Dare Baldwin and colleagues have assumed the behavioural chunks (they call them "motion elements") include items naturally labelled "poke", "reach", "place", "drink" (D. Baldwin, Andersson, *et al.* 2008). In initial research on how such instances of such categories might be identified, they asked subjects to classify point-light displays of behaviour as same or different (Loucks and Bladwin 2006). This research was designed to answer the question "what sources of information [do] people extract from the motion stream to assist in categorizing and identifying actions?" (Loucks and Bladwin 2006: 240). Accordingly it does not directly address the issue of whether there are categories or what the categories are.

Finally, note that the methods used by Newton and subsequent research will not necessarily reveal the primitive behavioural chunks. If you ask ordinary subjects to segment speech, you may find breakpoints at word- and clause-boundaries without necessarily finding them at phonetic boundaries. Indeed, this method for measuring segmentation is unsuitable for discovering primitive units of speech perception because phonetic gestures overlap (Lieberman 1998). There is similar overlap in the execution of small-scale units of action. For example, in a sequence that involves reaching for and then grasping an object, some of the movements required for grasping occur simultaneously with reaching (M. Jeannerod, Arbib, *et al.* 1995). This makes it unlikely that measuring segmentation of behaviours by identifying breakpoints will reliably yield the smallest segments that are discerned in behaviour reading.

To sum up, I defined behaviour reading as segmenting a stream of bodily movement into chunks and identifying patterns among the chunks. These chunks are the primitive in the sense that behaviour readers recognise only behaviours which are chunks or combinations of chunks. Behavioural chunks are thus comparable to phonemes in speech behaviour. Given this definition it is not obvious that behaviour reading exists. While several researchers share the assumption that behaviour reading involves primitive chunks out of which everything else is built (e.g. Byrne 2003; D. Baldwin and Baird 2001), others do not (e.g. Zacks, Speer, *et al.* 2007). And as things stand there is no direct evidence on the existence of behavioural chunks. The reasons we have for believing in them are indirect. In particular, behaviour reading could not play the envisaged roles in intention reading, development and discerning the effects of complex behaviours unless behavioural chunks exist (see Section II).

Assuming they exist, several basic questions about behavioural chunks cannot be answered yet. First, what are the subjects of the behaviours: are they agents, body parts or something else? Second, are the behaviours in some sense directed to objects, as suggested by labels like "poke" and "reach"? Third, is there a fixed list of behavioural chunks (as with phonemes) or can novel chunks be identified ad hoc? Fourth, are behavioural chunks invariant with respect to anatomy, posture, environmental obstacles or context?

In the rest of this section I begin to answer these questions by describing a conjecture which may already be implicit in some work on behaviour reading: behavioural chunks are sub-intentional motor actions including object- and location-directed actions. Grasping a ball is a paradigm example of a behavioural chunk. As a preliminary, proper understanding of this conjecture requires distinguishing between goal-directed and object-directed actions (these are frequently conflated).

Object-directed isn't goal-directed. Some motor actions are object- or location-directed. Their specification involves an object whose properties control how the action is executed. For example, in grasping a ball various action parameters for action are set by properties of the ball including its size, shape and weight (M. Jeannerod, Arbib, *et al.* 1995).

Object-directed motor actions resemble goal-directed actions in that they can be unsuccessful (for example, you can fail to grasp an object). They also resemble goal-directed action in that the same motor action could have been executed differently (for example, if an object had been differently positioned, the grasp would have been executed differently).

Perhaps because of these points of resemblance theorists sometimes conflate object- and goal-directed

actions (e.g. Gergely, Nadasky, *et al.* 1995; Gallese and Goldman 1998; Pacherie and Dokic 2006). Informally it is sometimes appropriate to refer to objects as goals. But if we are concerned about action and the nature of our understanding of it, we must distinguish object- from goal-directed behaviour. This is particularly important given that behaviour reading can occur independent of knowledge of goals and intentions and does not rely on understanding these.

How, then, do goal- and object-directed actions differ? One fundamental point of difference is the nature of the representations involved. To represent goal-directed action you need to represent an end, for example the end of eating two bananas. This end may but need not involve reference to a specific objects, and it may but need not involve reference to basic actions such as reaching and grasping. By contrast, an object-directed action involves two distinct representations: a representation of an object and a representation of a motor action drawn from the agent's repertoire.

A second fundamental point of difference between goal-directed and object-directed actions concerns how the agent and action are related to the end or object. In the case of goal-directed action philosophers debate the exact nature of this relation (e.g. Taylor 1964; Wright 1976; Millikan 1993; Price 2001; Stout 1996). A simple counterfactual account is a useful starting point: for an end to be the goal of an agent's action is for it to be the case that had achieving that end required a different action, the agent would have taken that action instead (Bennett 1976: 39). This account needs refining in several ways to be useful in characterising everyday understanding of behaviour; clearly, agents will not normally do *anything* necessary to achieve their goals, and it is debatable whether the counterfactual element is appropriate. However, it is already sufficient to

distinguish goal- from object-directed action. For in the case of object-directed action, the relation between agent, action and object involves a motor command or schema. What makes one object rather than any other the object to which the motor action is directed is the fact that it is specified as the target in the motor cognition governing this action. This means that representing object-directed motor actions requires substantially different conceptual sophistication: it requires some understanding of motor cognition, but no understanding of goals and how they relate to action.

So goal- and object-directed actions differ both with respect to what is represented (ends vs. objects) and how these are related to the agent and action. It is possible for an individual to be able to represent object- but not goal-directed actions, and knowledge of which objects motor actions are directed to does not entail knowledge of an agent's goals in acting.

In favour of the conjecture. Several considerations favour the conjecture that behavioural chunks are motor actions, including object- and location-directed motor actions. First, motor actions are components of intentional action, so the conjecture makes it clear how identifying patterns of them in behaviour reading could feed into understanding intentional aspects of action. Second, motor actions follow rules independent of biomechanical constraints (Marc Jeannerod 1994). For example, in grasping there are rules governing when grip size expands and contracts (M. Jeannerod, Arbib, *et al.* 1995). The rule-bound nature of motor actions helps to explain how they could be recognised in moving bodies. Third, the object-directed nature of some motor actions plays a key role in behaviour reading. This is partly because it allows for the identification of patterns involving objects (for example, orienting to *an object* is often

followed by grasping *that object*), which is valuable because behaviours directed to the same object are likely to have interesting relations. Object-directedness is also valuable as a clue to intention, for the objects to which behaviours are directed often feature in the contents of the intentions agents are acting on. For these reasons, motor actions, particularly object-directed motor actions, are ideal as the phoneme-like chunks for behaviour reading.

In one respect motor actions may not be optimal for identifying patterns in behaviour. Motor cognition may involve representations that are fine-grained in some respects for behaviour reading, which would restrict our ability to identify patterns. To illustrate with a simplistic example, the motor system might in principle contain representations for different types of grasping action but lack a representation for grasp. Since for behaviour reading it may be advantageous to represent grasping in a less fine-grained way. This problem can be solved by introducing a further abstraction by analogy with the allophone-phoneme distinction. Object-directed motor actions may stand to behavioural chunks as allophones to phonemes. That is, chunks may be categories of motor action (for example, containing actions that differ only in being performed by different hands) or may be constituted from them by more complex rules. The key theoretical constraint governing the relation between basic motor actions and behavioural chunks is that every motor action that realises a chunk must have the same object. This ensures that the object-directed nature of the chunks flows from the nature of the motor actions that realise them.

The revised conjecture, then, is that behavioural chunks are *categories* of motor action, including object- and location-directed motor actions. Paradigm examples include reaching for a pen, grasping a cup and tearing a paper.

Neural correlates. The conjecture hinges on the claim that some cognitive processes are sensitive to object-directed motor actions. Is this claim plausible? In principle it is possible that cognitive processes track motion interpret it as intentional action without also tracking object-directed motor actions. However, single cell studies in the superior temporal sulci of macaque monkeys demonstrates that information about motor actions is present in the brain. Dave Perrett and co-workers found cells which fire selectively for object-directed motor actions including reaching, retrieving, manipulation, picking, tearing, presenting and holding (Perrett, Harries, *et al.* 1989: 102). Their sensitivity to these actions appears to be robust, for it can survive key parts of the motion and its object being occluded (Jellema and Perrett 2001). Intriguingly, follow-up work discovered cells sensitive to reaching whose firing is modulated by postural cues to attention such as head direction (Jellema, Baker, *et al.* 2000; see also Pelphrey and Morris 2006). This hints at the possibility that behavioural chunks may involve action in postures indicative of attention, thus increasing the probability that behavioural chunks are intention-relevant.

Perrett describes these neurons as responsive to actions under “goal-centred descriptions” and argues that such a description “provides a framework through which the visual system can achieve a rich understanding of the world which embodies causation and intentionality” (Perrett, Harries, *et al.* 1989: 110). In fact his findings demonstrate sensitivity to object-directed motor actions only. Indeed, his definition of “goal-centred” actually defines object- rather than goal-directed actions (Perrett, Harries, *et al.* 1989: 110).

These findings suggest that the superior temporal sulcus may be involved in behaviour reading and so, at

least indirectly, in theory of mind (cf. Saxe, Xiao, *et al.* 2004 for related fMRI studies in humans). For present purposes these findings are relevant as evidence for the less ambitious claim that some cognitive mechanisms track object-directed motor actions.

Predictions. As this conjecture is a generalisation of the claim that phonemes are categories of phonemic gestures (Liberman and Mattingly 1985; Liberman and Whalen 2000), a logic similar to that used to establish the claim about speech could also be used to test the more general conjecture. To illustrate, two simple predictions are first, that the cues subjects use to segment behavioural chunks will reliably pick out categories of motor action in circumstances where behaviour reading normally occurs; and, second, that subjects will treat different instances of the same category of motor action as equivalent behavioural chunks.

Two problems for the conjecture. Which motor actions are behavioural chunks, and how are these motor actions represented? I take these problems in turn.

Since motor actions can be arranged in a hierarchy, the conjecture calls for us to specify *which* motor actions constitute behavioural chunks. Given the requirements on behaviour reading and the parallel with speech, it seems likely that the relevant motor action will be complex, potentially involving several coordinated movements. But there may be several options consistent with this constraint. Relatedly, are the behavioural chunks an individual can identify only those motor actions in the individual’s own repertoire? Can they involve tools or are they only basic actions? Such questions are left open by the conjecture.

A second, more pressing problem is how the behavioural chunks are represented. To represent

intentional actions you have to understand what an intention is, because intentional actions are actions caused by intentions (e.g. Davidson 1978 [1980]). Similarly, to represent motor actions you have to understand motor cognition, because motor actions just are actions produced by motor commands. This generates a dilemma for the conjecture that behavioural chunks are motor actions. Either behaviour reading does not involve representing motor actions, or else individuals capable of behaviour reading understand what motor actions are, even if only implicitly. The second alternative is unattractive to the extent that there is no independent evidence for implicit understanding of motor actions. This forces us to accept the first alternative. How, then, is behaviour reading possible if the chunks are not represented?

(It is easy to miss the force of this second problem given the finding, mentioned above, that individual neurons fire in response to motor actions. Doesn’t this demonstrate that motor actions are represented? No, it shows that cells carry information about motor actions, which is defeasible evidence that some cognitive processes track motor actions. This does not entail that the individual represents those actions (Pacherie and Dokic 2006: §3).)

Solutions to both problems hinge on mirroring motor cognition, which is the topic of the next section. To anticipate, when seeing or hearing an action motor commands are activated in the observer corresponding to those activated in the agent. This pattern of activation in the observer enables her to segment the behaviour into chunks that are motor actions. She is thus able to track motor commands not by representing them but by activating them (Marc Jeannerod 2001: S104). It follows that the motor actions which constitute behavioural chunks are the motor actions which are mirrored. This enables us to lever research on which motor commands are

mirrored to specify which motor actions constitute behavioural chunks.

IV. On the role of mirroring motor cognition

Fadiga and colleagues recently studied evoked motor potentials of tongue muscles while subjects heard spoken words (2002). They found a significant increase in these when subjects heard words speaking which involves moving the tongue. This is evidence that hearing speech involves motor cognition matching that required for producing speech. But what is the role of this mirroring motor cognition?

Imagine trying to answer this question without any knowledge of linguistics. Some imaginary researchers might state that this mirroring motor cognition gives us direct access to the meanings of others' utterances (Gallese and Goldman 1998; Umiltà, Kohler, *et al.* 2001). In response philosophers might object that this is impossible on the grounds that meanings to do not reduced to motor actions and it is motor action rather than meaning that is mirrored. But without knowledge of linguistics the philosophers would either omit to explain the function of this mirroring motor cognition (Borg 2007) or else insist that it serves only to enable us to predict which sound someone will make next (Jacob 2008; Jacob and Jeannerod 2005). This would make the function of mirroring motor cognition in speech a mystery.

Bizarre as this imaginary debate sounds, it is a true reflection of debate on the function of mirror neurons in understanding action (see cited papers). The reason debate is so much further advanced in the case of speech is our knowledge of linguistics. Phonotactics, prosody, syntax and the rest provide a mechanism connecting mirroring cognition with linguistic understanding. The primary role of mirroring motor cognition in language is to segment phonetic gestures. Because we know the mechanisms

connecting phonetics to semantics, we understand how this humble service, although removed from understanding meaning, is vital for communication by language.

A primary function of mirroring motor cognition of action is to segment motion into behavioural chunks (Rizzolatti, Fogassi and Gallese 2001; Byrne 2003). Behaviour reading is the mechanism linking this humble function to an understanding of intention. That, anyway, is the claim of this section.

Background: mirror neurons. Contemporary research on mirroring motor cognition begins with Rizzolatti and colleagues' famous discovery of mirror neurons in the frontal cortex of pig-tailed macaque monkeys (Rizzolatti, Fadiga, *et al.* 1996). Follow up work using a variety indirect methods has established their existence in humans as well (Rizzolatti and Craighero 2004). Mirror neurons fire both when an individual observes an action as well as when the individual performs an action, where observation can be by seeing or hearing (in monkeys: Kohler, Keysers, *et al.* 2002; and in humans: Aziz-Zadeh, Iacoboni, *et al.* 2004). The actions mirror-neurons fire in response to are object-directed motor actions such as grasping, holding and tearing (Rizzolatti, Fogassi and Gallese 2001) and motor actions not involving objects such as lip protrusions in monkeys (Ferrari, Gallese, *et al.* 2003) and finger movements in humans (Iacoboni, Woods, *et al.* 1999). Some mirror neurons track fine-grained motor-actions, for example, actions involving a precision but not whole-handed grip; these neurons are called "strictly congruent". Others are "broadly congruent" and track more coarse-grained action types, for example grasping irrespective of grip (Rizzolatti and Craighero 2004). The sensitivity of mirror neurons to object-directed actions is robust, not requiring that the entire movement be visible

(Umiltà, Kohler, *et al.* 2001). Because research to date reveals that neurons in the superior temporal sulcus are responsive to a significantly wider range of motor actions than mirror neurons (particularly actions involving faces and walking), it is possible that mirror neurons do not exist for the full range of motor actions an individual can recognise (Keysers and Perrett 2004).

For present purposes, the key features of mirror neurons are: (1) they track motor actions, including object-directed motor actions; and (2) they track these actions irrespective of whether they are executed or observed. This strengthens the case for the claim that cognitive processes are concerned with object-directed motor actions, not just movements and intentional actions.

What do observing and performing actions have in common? The existence of mirror neurons shows that observing and executing actions have something in common. But what do they have in common? Several findings provide evidence, further, that some of the motor representations involved in planning and executing an action are also involved in simulating and observing that action. The most direct evidence for this claim comes from measurements of motor-evoked potentials generated by artificially stimulating the motor cortices of subjects observing videos of a ball being grasped. Gangitano and colleagues found motor-evoked potentials related to the very muscles used in performing the observed action at the very times those muscles were needed for the task (M. Gangitano, Mottaghy and Pascual-Leone 2001; Fadiga, Craighero and Olivier 2005). Relatedly, Calvo-Merino *et al.* investigated the causes of motor activation in subjects observing actions. They found that that what determines motor activation is the observers' competence to perform those actions rather than her visual familiarity with them (Calvo-Merino,

Grèzes, *et al.* 2006). This fits with the finding that performing an action can facilitate observational recognition of the same action (Miall, Stanley, *et al.* 2006).

What are the effects of motor cognition in the observers of action? First, observing others' actions can interfere with one's own actions, and the interference from observing one task while performing another is comparable to that involved in performing both tasks oneself (Sebanz, Knoblich and Prinz 2003). Second, subjects watching an actor stacking blocks directed their eyes to the blocks predictively in just the way that the actor herself would naturally do (Flanagan and Johansson 2003; Rotman, Troje, *et al.* 2006). Third, the difficulty of performing an action is reflected not only in the time it takes to execute an action but also in the time it takes to simulate the action, so that varying the difficulty of an action has the same effect on simulation as performance (Marc Jeannerod and Frak 1999), and this is true whether the action is simulated from a first- or third-person point of view (Anquetil and Jeannerod 2007). These findings are further evidence for the existence motor cognition in observers of an action corresponding to that required for planning an executing the action. They show that such motor cognition is automatic, occurring even when it hinders performance on a task (Sebanz, Knoblich and Prinz 2003), and that it has a range of cognitive and behavioural effects.

What use is this mirroring motor cognition? Objections to the claim that mirror neurons enable us to know other's goals and intentions (Borg 2007; Jacob and Jeannerod 2005) apply equally to mirroring motor cognition more generally. Perhaps in anticipation of such objections, some state that mirroring motor cognition enables us to identify motor actions but not intentions or intentional action (Gallese, Keysers and Rizzolatti 2004). The difficulty for this proposal is why we should care about motor

actions as such. These small-scale events rarely have effects of significance in themselves—that my competitors hand closes on a nut matters not in itself but because it amounts to stealing it from me. Similar difficulty attends the idea that mirroring motor cognition enables us one predict an agent's next motor action (e.g. Csibra 2007): such small-scale predictions are useful in themselves only when rapid intervention is possible.

A more promising account of the function of mirroring motor cognition involves behaviour reading. One function of mirroring motor cognition is to segment flowing movements into chunks matching motor actions. The value of this segmenting is not that recognising the motor actions or their effects is valuable in itself, but that it provides the starting point for the identification of larger-scale patterns in behaviour that support prediction, imitation, learning and ascription of mental states.

Solving the problem of how behaviour readers track motor actions without representing them. Recall from Section III the question of how motor actions are tracked for behaviour reading. Mirroring motor cognition answers this question, and it does so without requiring representational resources. By activating her own motor system, an individual can categorise observed motor actions without reliance on representational resources over and above those involved in acting. To illustrate with an analogy, suppose you want to talk about an action you have observed. One way to do this is with a representation: if you recognise the action, you could name or describe it. Another way to introduce the action into our conversation, which does not depend on knowing what the action is, is to reproduce it. Reproduction is not representation but it can sometimes serves similar purposes. My proposal, then, is that behaviour readers segment movements into motor actions not by

representing but by reproducing (that is, mirroring) motor commands (cf. Marc Jeannerod 2001: S104). That it has this function explains the importance of mirroring motor cognition for understanding action.

Does mirroring motor cognition depend on ascriptions of intention? I have claimed that behaviour reading depends on mirroring motor cognition, and that behaviour reading can be done independently of knowledge of goals or intentions. Against this proposal, Csibra and Jacob have independently hypothesised that the identification of motor actions in mirror cognition depends on knowledge of the agent's goals (Csibra 2007; Jacob 2008). While I reject this hypothesis, two concessions can be made. Knowledge of goals may have a top-down effect on mirroring motor cognition in some but not all cases, and behaviour not consistent with the existence of *any* goal may inhibit mirroring motor cognition. What cannot—and should not—be conceded is that mirroring another's motor cognition invariably depends on knowing (or guessing) the particular goal they are acting on. After all, in many experiments no information at all about a goal is given. Instead, a stripped down motor action is shown, for example a hand grasping a ball. The point is particularly clear in the case of speech, where the stimuli are utterances of words. While uttering a word is occasionally a goal-directed action (as in “birra (beer)!”), this is not plausible for most of the stimuli used for research (e.g. “serra (greenhouse)”) (Fadiga, Craighero, *et al.* 2002).

It is possible that Csibra or Jacob have conflated object- with goal-directed action. Perhaps they hold only that mirroring motor cognition depends on prior recognition of the object which is the target of an action, rather than on recognition of a goal as such. This weaker hypothesis does not conflict with my proposal that mirroring motor cognition can segment

streams of movement into object-directed motor actions independently of knowledge of intention.

Solving the 'which motor action' problem. A problem (the second of two) in characterising behaviour reading is how to discover which motor actions constitute the primitive chunks (see Section III). If the chunks are segmented by means of mirrored motor activations, one way to address this question is by asking which units are mirrored.

Massimo Gangitano and colleagues investigated this question by studying what happens when a grasping action takes an unexpected and unnatural course (2004). Their subjects observed videos in which a hand grasped a ball, except that in some conditions the hand closed half-way to the ball before re-opening. When the motor action was interrupted mid-way, motor cognition in the observer was significantly reduced or ceased. (As in (M. Gangitano, Mottaghy and Pascual-Leone 2001), this finding was obtained by measuring artificially stimulated motor-evoked potentials in observers.) The authors conclude that the plan for a whole motor action is mirrored and that this plan is not updated during an ongoing motor action. Accordingly we can conclude that the unit mirrored is at least as large as the plan for a motor action such as grasping.

It is possible that some mirroring motor cognition involves larger units. Some monkey mirror neurons are sensitive not just to whether a certain object-directed motion (grasping a piece of food) occurs but also to what object-directed motion follows (eating vs. placing) (Fogassi, Ferrari, *et al.* 2005; related results have been obtained in humans with fMRI Iacoboni, Molnar-Szakacs, *et al.* 2005). Because these authors also note that some mirror neurons are insensitive to context, these findings do not conflict with my hypothesis that behaviour reading both involves mirroring motor cognition and does not depend on

knowledge of intention. This hypothesis requires that some units smaller than intentions are mirrored, not that these are the largest units mirrored. However, it is worth noting the relevance of behaviour reading to interpreting the results of Fogassi *et al.* (and Iacoboni *et al.*). These researchers inferred that “this neuronal property allows the monkey to predict the goal of the observed action and, thus, to “read” the intention of the acting individual” (Fogassi, Ferrari, *et al.* 2005). This rich interpretation has been criticised (Borg 2007; Jacob 2008). An alternative interpretation is suggested by behaviour reading. As detailed above (see Section II), behaviour reading involves identifying word-like units of primitive chunks whose boundaries reliably correspond to fulfilment of goals. This process does not demand conceptual sophistication and is automatic, which makes it a candidate for realisation in the monkey mirroring system.

De Vignemont and Haggard (2008) claim that what is shared in mirroring motor cognition is an intention-in-action rather than a motor command. Their argument is based on the claim that mirroring motor cognition makes imitation possible and imitation without intention-in-action would not be mere imitation but mere emulation. Even if correct, this argument does not quite justify their conclusion because it shows only that intentions-in-action must be among the things mirrored and so does not exclude the possibility that motor commands are also mirrored. Their argument can also be questioned on the grounds that mirroring motor cognition may make insightful imitation possible only in conjunction with behaviour reading. This alternative view is defended in detail by Richard Byrne (1999; 2003).

To sum up, there is some evidence for the claim that the units of mirroring motor cognition include plans for motor actions such as grasping and placing. Given the hypothesis that segmentation in behaviour reading is achieved by means of mirroring motor

cognition, this is evidence that such plans define the primitive behavioural chunks.

Predictions. In this section I have explained the hypothesis that one function of mirroring motor cognition is the segmentation of bodily movements into chunks for behaviour reading. This hypothesis generates some predictions. First, behaviour reading will depend on one's own motor competence (this implies that behaviour reading of actions involving alien motor competence will involve greater difficulty, not that they will be altogether impossible). Second, behaviour reading will never identify units of behaviours whose boundaries are unrelated to whatever units are involved in mirroring motor cognition. Third, whether behaviour reading can be smoothly applied to actions involving tools depends on whether there is mirroring motor cognition for tool use.

The present hypothesis is a generalisation of the motor theory of speech perception, according to which phonetic gestures are identified by the simulation of motor commands for their production (Liberman and Mattingly 1985; Galantucci, Fowler and Turvey 2006). Accordingly, a further source of predictions for this hypothesis can be generated by generalising predictions used to test the motor theory of speech perception (e.g. Galantucci, Fowler and Turvey 2006).

V. Conclusion

If the motor theory of speech perception is correct, speech perception is behaviour reading in the special case where the behaviours read are phonetic gestures. This tells us something about the nature of behaviour reading: segmentation of behaviour depends on mirroring motor cognition and the primitive behavioural chunks are motor actions including

object-directed motor actions. It also tells us something about one function of mirroring motor cognition. It might provide us with an ability to identify motor actions (Gallese, Keysers and Rizzolatti 2004), to predict which motor action will be performed next (Jacob and Jeannerod 2005), or to verify ascriptions of intention (Csibra 2007), and it may facilitate interpersonal coordination (Pacherie and Dokic 2006; Knoblich and Sebanz 2006). But at least as important as any of these functions is its foundational role in behaviour reading and, thereby, theory of mind. This is the truth behind the claim that mirror neurons are part of a mechanism for understanding intentional actions.

References

- Anquetil, Thierry and Marc Jeannerod (2007), "Simulated actions in the first and in the third person perspectives share common representations". Brain Research, 1130:125-129
- Aziz-Zadeh, Lisa, Marco Iacoboni, Eran Zaidel, Stephen Wilson and John Mazziotta (2004), "Left hemisphere motor facilitation in response to manual action sounds". European Journal of Neuroscience, 19(9):2609-2612
- Baldwin, Dare A., Jodie A. Baird, Megan M. Saylor and M. Angela Clark (2001), "Infants Parse Dynamic Action". Child Development, 72(3):708-717
- Baldwin, Dare, Annika Andersson, Jenny Saffran and Meredith Meyer (2008), "Segmenting dynamic human action via statistical structure". Cognition, 106(3):1382-1407
- Baldwin, Dare and Jodie A. Baird (2001), "Discerning intentions in dynamic human action". Trends in Cognitive Sciences, 5(4):171-178
- Bennett, Jonathan (1976), *Linguistic Behaviour*. Cambridge: Cambridge University Press
- Biro, Szilvia and Alan M. Leslie (2007), "Infants' perception of goal-directed actions: development through cue-based bootstrapping". Developmental Science, 10(3):379-398
- Borg, Emma (2007), "If Mirror Neurons are the Answer, What was the Question?" Journal of Consciousness Studies, 14:5-19
- Brand, Rebecca J., Dare A. Baldwin and Leslie A. Ashburn (2002), "Evidence for 'motionese': modifications in mothers' infant-directed action". Developmental Science, 5(1):72-83
- Bratman, Michael (1984), "Two Faces of Intention". The Philosophical Review, 93(3):375-405
- Browman, C. P. and L. Goldstein (1992), "Articulatory phonology: an overview". Phonetica, 49(3-4):155-80
- Byrne, Richard W. (1999), "Imitation without intentionality. Using string parsing to copy the organization of behaviour". Animal Cognition, 2(2):63-72
- (2003), "Imitation as Behaviour Parsing". Philosophical Transactions: Biological Sciences, 358(1431):529-536
- Calvo-Merino, Beatriz, Julie Grèzes, Daniel E. Glaser, Richard E. Passingham and Patrick Haggard (2006), "Seeing or Doing? Influence of Visual and Motor Familiarity in Action Observation". Current Biology, 16(19):1905-1910
- Chomsky, N. (1956), "Three models for the description of language". Information Theory, IRE Transactions on, 2(3):113-124
- Cohen, Asher, Richard I. Ivry and Steven W. Keele (1990), "Attention and structure in sequence learning". Journal of Experimental Psychology: Learning, 16(1):17-30
- Conway, Christopher M. and Morten H. Christiansen (2001), "Sequential learning in non-human primates". Trends in Cognitive Sciences, 5(12):539-546
- Corballis, Michael C. (2007), "Recursion, Language, and Starlings". Cognitive Science: A Multidisciplinary Journal, 31(4):697-697
- Csibra, Gergely (2007), "Action mirroring and action interpretation: An alternative account". Sensorimotor Foundations of Higher Cognition. Attention and Performance XXII. Oxford University Press. Oxford:435-480
- Davenport, Mike and S. J. Hannahs (1998), *Introducing Phonetics and Phonology*. London: Arnold
- Davidson, Donald (1963 [1980]), "Actions, Reasons and Causes", in *Essays on Actions and Events*. Oxford: Oxford University Press.
- (1971 [1980]), "Agency", in *Essays on Actions and Events*. Oxford: Oxford University Press.
- (1978 [1980]), "Intending", in *Essays on Actions and Events*. Oxford: Oxford University Press.
- de Vignemont, F. and P. Haggard (2008), "Action observation and execution: What is shared?" Social Neuroscience, 99999(1):1-1
- Fadiga, Luciano, Laila Craighero, Giovanni Buccino and Giacomo Rizzolatti (2002), "Speech listening specifically modulates the excitability of tongue muscles: a TMS study". European Journal of Neuroscience, 15(2):399-402
- Fadiga, Luciano, Laila Craighero and Etienne Olivier (2005), "Human motor cortex excitability during the perception of others' action". Current Opinion in Neurobiology, 15(2):213-218
- Ferrari, Pier Francesco, Vittorio Gallese, Giacomo Rizzolatti and Leonardo Fogassi (2003), "Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex". European Journal of Neuroscience, 17(8):1703-1714
- Fiser, J. zsef and Richard N. Aslin (2001), "Unsupervised Statistical Learning of Higher-Order Spatial Structures From Visual Scenes". Psychological Science, 12(6):499-504
- Fitch, W. Tecumseh and Marc D. Hauser (2004), "Computational Constraints on Syntactic Processing in a Nonhuman Primate". Science, 303(5656):377-380
- Flanagan, J. Randall and Roland S. Johansson (2003), "Action plans used in action observation". Nature, 424(6950):769-771
- Fogassi, Leonardo, Pier Francesco Ferrari, Benno Gesierich, Stefano Rozzi, Fabian Chersi and Giacomo Rizzolatti (2005), "Parietal Lobe: From Action Organization to Intention Understanding". Science, 308(5722):662-667
- Galantucci, Bruno, Carol A. Fowler and M. T. Turvey (2006), "The motor theory of speech perception reviewed". Psychonomic Bulletin & Review, 13:361-377
- Gallese, Vittorio and Alvin Goldman (1998), "Mirror neurons and the simulation theory of mind-reading". Trends in Cognitive Sciences, 2(12):493-501
- Gallese, Vittorio, Christian Keysers and Giacomo Rizzolatti (2004), "A unifying view of the basis of social cognition". Trends in Cognitive Sciences, 8(9):396-403
- Gangitano, M., F. M. Mottaghy and A. Pascual-Leone (2001), "Phase-specific modulation of cortical motor output during movement observation". Neuroreport, 12(7):1489-92
- Gangitano, Massimo, Felix M. Mottaghy and Alvaro Pascual-Leone (2004), "Modulation of premotor mirror neuron activity during observation of unpredictable grasping

- movements". European Journal of Neuroscience, 20(8):2193-2202
- Gergely, György, Z. Nadasky, Gergely Csibra and S. Biro (1995), "Taking the Intentional Stance at 12 Months of Age". Cognition, 56:165-193
- Gómez, Rebecca and LouAnn Gerken (2000), "Infant artificial language learning and language acquisition". Trends in Cognitive Sciences, 4(5):178-186
- Hard, Bridgette M., Barbara Tversky and David S. Lang (2006), "Making Sense of Abstract Events: Building Event Schemas". Memory & Cognition, 34:1221-1235
- Hard, Bridgette Martin, Sandra C. Lozano and Barbara Tversky (2006), "Hierarchical Encoding of Behavior: Translating Perception Into Action". Journal of Experimental Psychology: General. Vol 135(4), 2006:588-608
- Hauser, Marc D., Elissa L. Newport and Richard N. Aslin (2001), "Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins". Cognition, 78(3):B53-B64-B53-B64
- Iacoboni, Marco, Istvan Molnar-Szakacs, Vittorio Gallese, Giovanni Buccino, John C. Mazziotta and Giacomo Rizzolatti (2005), "Grasping the Intentions of Others with One's Own Mirror Neuron System". PLoS Biology, 3(3):e79 EP --e79 EP
- Iacoboni, Marco, Roger P. Woods, Marcel Brass, Harold Bekkering, John C. Mazziotta and Giacomo Rizzolatti (1999), "Cortical Mechanisms of Human Imitation". Science, 286(5449):2526-2528
- Jacob, Pierre (2008), "What Do Mirror Neurons Contribute to Human Social Cognition?" Mind & Language, 23(2):190-223
- Jacob, Pierre and Marc Jeannerod (2005), "The motor theory of social cognition: a critique". Trends in Cognitive Sciences, 9(1):21-25
- Jeannerod, M., M. A. Arbib, G. Rizzolatti and H. Sakata (1995), "Grasping objects: the cortical mechanisms of visuomotor transformation". Trends in Neurosciences, 18(7):314-320
- Jeannerod, Marc (1994), "The Representing Brain: Neural Correlates of Motor Intention and Imagery". Behavioral and Brain Sciences, 17(2):187-245
- (2001), "Neural Simulation of Action: A Unifying Mechanism for Motor Cognition". NeuroImage, 14(1):S103-S109-S103-S109
- Jeannerod, Marc and Victor Frak (1999), "Mental imaging of motor activity in humans". Current Opinion in Neurobiology, 9(6):735-739
- Jellema, T., C. I. Baker, B. Wicker and D. I. Perrett (2000), "Neural Representation for the Perception of the Intentionality of Actions". Brain and Cognition, 44(2):280-302
- Jellema, T. and D. I. Perrett (2001), "Coding of visible and hidden actions". Attention and Performance XIX
- Johnson-Frey, Scott, Michael McCarty and Rachel Keen (2004), "Reaching beyond spatial perception: Effects of intended future actions on visually guided prehension". Visual Cognition, 11:371-399
- Keysers, Christian and David I. Perrett (2004), "Demystifying social cognition: a Hebbian perspective". Trends in Cognitive Sciences, 8(11):501-507
- Kirkham, Natasha Z., Jonathan A. Slemmer and Scott P. Johnson (2002), "Visual statistical learning in infancy: evidence for a domain general learning mechanism". Cognition, 83(2):B35-B42-B35-B42
- Knoblich, Guenther and Natalie Sebanz (2006), "The Social Nature of Perception and Action". Current Directions in Psychological Science, 15(3):99-104
- Kohler, Evelynne, Christian Keysers, M. Alessandra Umiltà, Leonardo Fogassi, Vittorio Gallese and Giacomo Rizzolatti (2002), "Hearing Sounds, Understanding Actions: Action Representation in Mirror Neurons". Science, 297(5582):846-848
- Kurby, Christopher A. and Jeffrey M. Zacks (2008), "Segmentation in the perception and memory of events". Trends in Cognitive Sciences, 12(2):72-79
- Lieberman, Alvin M. (1998), "When Theories of Speech Meet the Real World". Journal of Psycholinguistic Research, 27(2):111-122
- Lieberman, Alvin M. and Ignatius G. Mattingly (1985), "The Motor Theory of Speech Perception Revised". Cognition, 21(1):1-36
- Lieberman, Alvin M. and Doug H. Whalen (2000), "On the Relation of Speech to Language". Trends in Cognitive Sciences, 4(5):187-96
- Loucks, Jeffrey T. and Dare Bladwin (2006), "When Is a Grasp a Grasp? Characterizing Some Basic Components of Human Action Processing", in *Action Meets Word*: Oxford University Press.
- Miall, R. C., J. Stanley, S. Todhunter, C. Levick, S. Lindo and J. D. Miall (2006), "Performing hand actions assists the visual discrimination of similar hand postures". Neuropsychologia, 44(6):966-976
- Millikan, Ruth (1993), *White Queen Psychology and Other Essays for Alice*. Cambridge, Mass.: MIT Press
- Newport, Elissa L. and Richard N. Aslin (2004), "Learning at a distance I. Statistical learning of non-adjacent dependencies". Cognitive Psychology, 48(2):127-162
- Newton, Darren (1973), "Attribution and the unit of perception of ongoing behavior". Journal of Personality and Social Psychology. Vol. 28(1), 28(1):28-38
- Newton, Darren and Gretchen Engquist (1976), "The Perceptual Organization of Ongoing Behavior". Journal of Experimental Social Psychology, 12(5):436-50
- Newton, Darren, Gretchen A. Engquist and Joyce Bois (1977), "The objective basis of behavior units". Journal of Personality and Social Psychology. Vol. 35(12), 35(12):847-862
- Nygaard, Lynne C. and David B. Pisoni (1995), "Speech Perception: New Directions in Research And Theory", in J. Miller, L. and P. D. Eimas (eds.), *Speech, Language and Communication*. London: Academic Press.
- Pacherie, Elisabeth and Jérôme Dokic (2006), "From mirror neurons to joint actions". Cognitive Systems Research, 7(2-3):101-112
- Pelphrey, Kevin A. and James P. Morris (2006), "Brain Mechanisms for Interpreting the Actions of Others From Biological-Motion Cues". Current Directions in Psychological Science, 15(3):136-140
- Perrett, D. I., M. H. Harries, R. Bevan, S. Thomas, P. J. Benson, A. J. Mistlin, A. J. Chitty, J. K. Hietanen and J. E. Ortega (1989), "Frameworks of analysis for the neural representation of animate objects and actions". J Exp Biol, 146(1):87-113
- Povinelli, Daniel J. (2001), "On the Possibility of Detecting Intentions Prior to Understanding Them", in B. F. Malle, L. Moses and D. A. Baldwin (eds.), *Intentions and Intentionality*. Cambridge, MA: MIT Press.
- Povinelli, Daniel J. and Steve Giambrone (2001), "Reasoning about Beliefs: A Human Specialization?" Child Development, 72(3):691-5
- Price, Carolyn (2001), *Functions in Mind*. Oxford: Clarendon Press

- Repp, Bruno H. and Alvin M. Liberman (1987), "Phonetic Category Boundaries are Flexible", in S. Harnad (ed.) *Categorical Perception: The Groundwork of Cognition*. Cambridge: Cambridge University Press.
- Rizzolatti, Giacomo and Michael A. Arbib (1998), "Language within our grasp". *Trends in Neurosciences*, 21(5):188-194
- Rizzolatti, Giacomo and Laila Craighero (2004), "The Mirror-Neuron System". *Annual Review of Neuroscience*, 27:169-192
- Rizzolatti, Giacomo, Luciano Fadiga, Vittorio Gallese and Leonardo Fogassi (1996), "Premotor cortex and the recognition of motor actions". *Cognitive Brain Research*, 3(2):131-141
- Rizzolatti, Giacomo, Leonardo Fogassi and Vittorio Gallese (2001), "Neurophysiological mechanisms underlying the understanding and imitation of action". *Nat Rev Neurosci*, 2(9):661-670
- Rotman, Gerben, Nikolaus F. Troje, Roland S. Johansson and J. Randall Flanagan (2006), "Eye Movements When Observing Predictable and Unpredictable Actions". *J Neurophysiol*, 96(3):1358-1369
- Saffran, J. R., Elissa L. Newport and R. N. Aslin (1996), "Statistical Learning by 8-Month-Old Infants". *Science*, 274(5294):1926-8
- Saffran, Jenny, Marc Hauser, Rebecca Seibel, Joshua Kapfhammer, Fritz Tsao and Fiery Cushman (2008), "Grammatical pattern learning by human infants and cotton-top tamarin monkeys". *Cognition*, 107(2):479-500
- Saxe, R., D. K. Xiao, G. Kovacs, D. I. Perrett and N. Kanwisher (2004), "A region of right posterior superior temporal sulcus responds to observed intentional actions". *Neuropsychologia*, 42(11):1435-1446
- Saylor, Megan M., Dare A. Baldwin, Jodie A. Baird and Jennifer LaBounty (2007), "Infants' On-line Segmentation of Dynamic Human Action". *Journal of Cognition and Development*, 8(1):113-113
- Sebanz, Natalie, Guenther Knoblich and W. Prinz (2003), "Representing others' actions: just like one's own?" *Cognition*, 88(3):B11-B21
- Seidl, Amanda and Alejandrina Cristi... (2008), "Developmental changes in the weighting of prosodic cues". *Developmental Science*, 11(4):596-606
- Soderstrom, Melanie, Deborah G. Kemler Nelson and Peter W. Jusczyk (2005), "Six-month-olds recognize clauses embedded in different passages of fluent speech". *Infant Behavior and Development*, 28(1):87-94
- Stout, Rowland (1996), *Things That Happen Because They Should*. Oxford: Oxford University Press
- Swallow, Khena M. and Jeffrey M. Zacks (2008), "Sequences learned without awareness can orient attention during the perception of human activity". *Psychonomic Bulletin & Review*, 15:116-122
- Taylor, Charles (1964), *The Explanation of Behaviour*. London: Routledge
- Toro, Juan M., Trobalç and Josep B. n (2005), "Statistical Computations Over a Speech Stream in a Rodent". *Perception & Psychophysics*, 67:867-875
- Umiltà, M. A., E. Kohler, V. Gallese, L. Fogassi, L. Fadiga, C. Keysers and G. Rizzolatti (2001), "I Know What You Are Doing: A Neurophysiological Study". *Neuron*, 31(1):155-165
- Werker, Janet F. and Henry H. Yeung (2005), "Infant speech perception bootstraps word learning". *Trends in Cognitive Sciences*, 9(11):519-527
- Wright, Larry (1976), *Teleological Explanations*. Berkeley: University of California Press
- Zacks, Jeffrey M. (2004), "Using movement and intentions to understand simple events". *Cognitive Science*, 28(6):979-1008
- Zacks, Jeffrey M., Todd S. Braver, Margaret A. Sheridan, David I. Donaldson, Abraham Z. Snyder, John M. Ollinger, Randy L. Buckner and Marcus E. Raichle (2001), "Human brain activity time-locked to perceptual event boundaries". *Nat Neurosci*, 4(6):651-655
- Zacks, Jeffrey M., Nicole K. Speer, Khena M. Swallow, Todd S. Braver and Jeremy R. Reynolds (2007), "Event Perception: A Mind-Brain Perspective". *Psychological Bulletin*. Vol 133(2):273-293
- Zacks, Jeffrey M., Khena M. Swallow, Jean M. Vettel and Mark P. McAvoy (2006), "Visual motion and the neural correlates of event perception". *Brain Research*, 1076(1):150-162
- Zacks, Jeffrey M., Barbara Tversky and Gowri Iyer (2001), "Perceiving, remembering, and communicating structure in events". *Journal of Experimental Psychology: General*. Vol. 130(1), 130(1):29-58