Imitation and conventional communication

Richard Moore

Abstract   To the extent that language is conventional, non-verbal individuals, including human infants, must participate in conventions in order to learn to use even simple utterances of words. This raises the question of which varieties of learning could make this possible. In this paper I defend Tomasello's (The cultural origins of human cognition. Harvard UP, Cambridge, 1999, Origins of human communication. MIT, Cambridge, 2008) claim that knowledge of linguistic conventions could be learned through imitation. This is possible because Lewisian accounts of convention have overstated what one must know to participate in conventions; and because the required knowledge could be learned imitatively. The imitation claim that I defend is consistent with what we know about both the proliferation of conventional behaviours in human children, who are skilful imitators, and the comparative absence of such behaviours in non-human great apes, who are poor at imitative learning.

Keywords   Communication · Conventions · Language development · Imitation · Social learning

Convention and social learning: a brief introduction

Some actions are intrinsically well suited to function as means towards the pursuit of the goals for which they are characteristically used. For example, when a chimpanzee uses a hammer and anvil to crack a coula nut, the weight of the hammer, and the force generated by the ape's directing the hammer onto the nut against the solid surface of the anvil cause the nut to crack. The physical properties

R. Moore (✉)
Department of Developmental and Comparative Psychology, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany
e-mail: richard_moore@eva.mpg.de

of the objects used, and their skilful manipulation by the chimpanzee, make for a solution to the problem of how to crack nuts that is superior to available alternatives. Because this action works by virtue of the intrinsic properties of objects and their skilful manipulation, it works whether or not it is known by anyone else; and the ability to perform it need not be learned from others. Thus, whether or not young chimpanzees do learn the art of nut-cracking from their peers, a lone ape working in isolation from all others could in principle figure out the relevant causal properties of the hammer and anvil and so teach itself how to crack nuts.[1]

In contrast to the use of tools like hammers, which possess physical properties that make them particularly suitable for the ends to which they are used as means, words are tools which are not intrinsically well suited to the tasks that they are used to perform. They mean what they do only because of the practices and activities of language users—and in particular, only because there exists a particular variety of social coordination, namely conventions, for using words (and combinations of words) with regular communicative functions. Conventions are necessary for word use because words are arbitrary. The speech acts for which they are used could have been performed just as well by countless arbitrarily different vehicles of meaning; and the same vehicles might just as well have been used for the performance of very different speech acts. For example, the phrase "Je voudrais un croissant, s'il vous plait," is an excellent tool for acquiring croissants in France. But it isn't intrinsically so: many different strings of sounds could have done the job just as well. The appropriateness of uttering one string of sounds rather than another is therefore to be explained only by the holding of contingent socio-cultural facts about the linguistic behaviour of the members of a community. We have reason to use words in the ways that we do only because others use words thus too.

Since the relationship between vehicles of meaning (e.g., words and sentences) and the purposes for which they are used is fixed by the conventional practices of language-users, in order to be able to use words in communicative interaction, speakers must be able to track and adopt these practices: to participate in conventions. An agent isolated from other language users could not learn the causal powers of words—that is, the changes in the world that their utterance can be used to bring about—in the way that she might learn to use a hammer to crack a nut. Consequently, when infants around the age of 14 months start to use conventional words in their communicative interactions, they must be engaging in social learning of some variety. This pushes us to answer the question:

(1) Which manner of social learning—for example, which cognitive abilities— would enable one to acquire knowledge of linguistic conventions?

Whatever socio-cognitive abilities are required for the learning of linguistic conventions, they cannot presuppose knowledge of language—since an explanation of how this becomes possible is what needs to be given. Furthermore, if these socio-cognitive abilities are to be exploited in an account of how individuals actually learn

---

[1] This is a conceptual claim. The empirical question of whether or not, and how, chimpanzees engage in social learning is a hotly contested subject. For relevant discussion see Boesch (1991, 2012), Tennie et al. (2009) and Whiten et al. (2009); and Section "An imitation-based account of convention learning satisfies desiderata (i)–(iii)".

to use linguistic conventions, they ought to be within the ken of those who do learn to use conventions in this way—not least children of around 14 months. If the same cognitive abilities could plausibly be attributed to our early hominin ancestors, answering (1) could also give us insight into (some aspects of) the proliferation of language in human pre-history.

Imitation and cultural transmission

One answer to question (1) that has been defended by Tomasello (1999, 2000, 2008) and Gergely and Csibra (2005) is that cultural knowledge, including knowledge of conventions, can be acquired through imitation. Imitation, which is one variety of non-verbal social learning, is the ability to observe and reproduce not just the intended goal of another's action, but the means pursued in order to achieve it (Tomasello and Carpenter 2005).

According to both Tomasello and Gergely and Csibra, it is humans' early developing facility for imitation that explains both the highly cultural nature of human life (characterised by the accretion of cultural artefacts like language and technology over successive generations), and the great ease with which human children are able to learn from others. Tomasello in particular attributes to imitation a foundational role both in children's language acquisition, and in phylogenetic development of language. Of its role in ontogeny, he goes so far as to claim that "initially, imitative learning is all that children do for linguistic constructions" (Tomasello 2000, p. 161). In conjunction with colleagues he also argues that it is the comparative absence of a facility for social learning in great apes, including an inability to imitate, that explains the relative lack of culture and language in their communities (Tennie et al. 2009).

It should be noted that the definition of imitation offered above has been challenged by Paulus (2011), on the grounds that that it begs the question about whether the neural mechanisms that support imitation implicate intention reading. If Tomasello and Carpenter aspired to make a claim about the neural mechanisms that support imitation, this objection would hold. However, their strategy is different: the mechanism that they propose is a *functional* one. A central strategy of Tomasello's work is to make functional claims about the cognitive abilities (for example, intention-reading and means-copying) that are necessary for aspects of social cognition, and to support these claims with empirical studies that testify to the presence or absence of these abilities in young children and non-human great apes. This strategy makes no claims about neural implementation. Nonetheless, it should be noted that some imitation researchers use the term 'imitation' in a more restricted way than Tomasello—to describe bodily movements performed by one agent that are similar to and causally connected to observed behaviours performed by another (e.g., Heyes 2001; Paulus 2011). In what follows, I adopt Tomasello's usage.

Tomasello's account of language development

On the Gricean account of linguistic meaning that Tomasello (2008) adopts, words and sentences have their meanings in virtue of the communicative intentions with

which speakers utter them. For an expression to have a meaning is just for it to be used paradigmatically for the expression of a certain type (or types) of communicative goal.[2] Children learn to use conventional vehicles of meaning—first words and holophrases; only later sentences—to fulfil their communicative goals because they identify the goals with which others speak, recognise that their words are uttered as a means to the fulfilment of these goals, and reproduce the same utterances in pursuit of analogous goals of their own. Thus two complementary cognitive abilities lie at the heart of Tomasello's account of early language-development: first, the ability to infer a speaker's communicative intentions, and second, the ability to imitate.[3] The same story is used in Tomasello's (2008) discussion of the propagation of gestural conventions among our early hominin ancestors.

While the syntactic aspects of linguistic communication—knowledge of how to construct new utterances out of combinations of words—remain essential to a full characterisation of language, with respect to both phylogeny and ontogeny Tomasello thinks syntax late developing: a response to functional pressures that promote precise and flexible communication. Thus, for example, he argues that very young children's production of language is not compositional in the manner of adult language (Tomasello 2000). Rather, phrases are learned and reproduced whole, as children imitate what they have heard. Only later is their constituent structure discerned and used to create new utterances. Thus, at least on this set of assumptions, an account of the spread of conventional communication remains more fundamental than one of the acquisition of syntax. Reflecting these assumptions, I develop here an account of the acquisition of linguistic conventions, but say nothing about the acquisition of syntax.

Evaluating Tomasello's claim

Tomasello's imitation claim is consistent both with evidence about the imitative abilities of young children, and with the requirement that at least some conventional activities be learnable without language. Empirical evidence has now established that even pre- verbal children are flexible and sophisticated imitators. Infants (i.e., children younger than 18 months) not only identify the goals in light of which their caregivers act and reproduce the means produced in pursuit of this goal, they also do so discriminatively. They reproduce actions judged to have been accidental less often than deliberately performed actions (Carpenter et al. 1998), and do not

---

[2] Millikan (2004) calls these goal-types 'memetic purposes', although unlike Tomasello she contrasts her position with Grice's. On both approaches, it's taken for granted that the meaning of an utterance can be specified in terms of the perlocutionary goals for which it is paradigmatically used. Searle (1969) criticises this position on the ground that speech acts are first and foremost illocutionary acts (intentions to be understood), and not perlocutionary ones (loosely, intentions to make things happen). However, with respect to language development, perlocutionary acts are clearly more fundamental: one could not intend to be understood as having requested an object unless one first grasped that performing that utterance was a way of obtaining that object.

[3] Some may object that infants could not engage in Gricean communication, because it implicates socio-cognitive abilities that are beyond them. I have argued that this view is mistaken, because motivated by an intellectualised reading of what Gricean communication requires (Moore submitted).

reproduce those in which no effect was discernible (Gergely and Csibra 2005, p. 475). Further evidence shows that infants as young as 12 months also imitate rationally—modifying their behaviour in light of their knowledge of the situational constraints on an actor's performance. When an agent's action is performed because some constraint prevents the performance of a more natural alternative, infants do not imitate the agent but instead reproduce the more obvious means to the goal (Gergely et al. 2002; Schwier et al. 2006). By contrast, where no such constraints are apparent, infants imitate an adult's action even when its fitness-for-task is not evident (Gergely and Csibra 2005).[4]

Given this evidence for the imitative abilities of pre-verbal infants, it's consistent with Tomasello's hypothesis that imitation could play a role in the learning of linguistic conventions. Nonetheless, the plausibility of his claim will also depend on a number of factors independent of empirical facts about the imitative abilities of infants. First among these is the question of whether or not the knowledge required for participating in conventions is of such a sort that it could be acquired through imitation. The answer to this question will depend on an answer to a second question:

(2) What must one know in order to participate in conventions, and linguistic conventions in particular?

In the remainder of this paper, I set out to answer both questions (1) and (2). I answer (2) first, because the philosophical analysis of the nature of conventions that I take as my starting point—the account developed by Lewis in his book *Convention*—characterises them as a sophisticated phenomena that make strong claims on the socio-cognitive abilities of those who participate in them. The demands of this account have consequently made it questionable whether pre-verbal children and non-human animals could even participate in conventions. Indeed, Lewis was explicit in his prediction that children could not, because they would fail to grasp the conventional nature of their linguistic activities, which he thought necessary for participation:

[The] convention of a language is… a regularity restricting one's production of, and response to, verbal utterances and inscriptions. Linguistic competence consists in part of a disposition to conform to that restriction with ease; and in part of an expectation that one's neighbours will be likewise disposed, with a recognition of their conformity as the [only] reason for one's own. No doubt a child or an idiot may conform without reason; if so, he is not party to the convention and his linguistic competence is incomplete. (Lewis 1969, p. 51)[5]

Laurence (1998) also argues that participation in Lewisian conventions could not be fundamental to language development, since "a variety of developmental considerations actually provide a relatively compelling argument against such accounts"

---

[4] The claim that infants can read intentions is not completely uncontroversial. Additionally, non-rational explanations of children's imitative abilities have also been proffered (Paulus et al. 2011).

[5] As will become apparent in due course, inserting that parenthetical "only" to Lewis's words here gives a better indication of the position that he defends.

(Laurence 1998, p. 202). Before a case can be made for the fact that knowledge of conventions could be acquired imitatively, I aim to give an account of convention that is consistent with intuitions that infants could participate in them.

Lewis's account of convention

Lewis's account constitutes a valuable starting point for discussion of how individuals could acquire knowledge of linguistic conventions because, in addition to its containing an account of what a convention is, it also yields a specification of what one would need to know in order to count as participating in a convention—and so Lewis's own answer to (2).

A deep insight that we inherit from Lewis is that conventions are (at least often) solutions to coordination problems. Coordination problems are situations of interdependent decision making, in which participants have goals that can be satisfied by activities (or 'coordination equilibria') that are equally good (i.e. only arbitrarily different), but where the particular action that is chosen by the participants is less important than the fact that they choose the same. If they are to fulfil their goals, members of a population finding themselves in coordination problematic situations must coordinate on which coordination equilibria they should pursue. Conventions constitute solutions to coordination problems because their existence gives rise to a precedent for choosing one possible action over others, and so provide participating individuals with a non-arbitrary reason for reproducing that action to the exclusion of others.

Lewis's original account gives something like the following analysis of convention:

A regularity R in the behaviour of members of a population P when they are agents in a recurrent situation S is a convention if and only if it is true that in any instance of S among members of P,

(1) almost everyone performs R;
(2) almost everyone expects almost everyone else to perform R;
(3) almost everyone performs R on condition that most others do, since S is (at least often) a coordination problem and near-uniform performance of R is a proper coordination equilibrium in S.
(4) (1)–(3) is common knowledge among the members of P.

Clauses (1)–(3) of this analysis are explained by the fact that, since conventional actions are not intrinsically superior means for attaining goals than arbitrarily different alternatives would be, participants in a convention have reason to perform R in S only where others do so too. The common knowledge clause (4) requires some further justification.

As Lewis introduces the term, common (or 'mutual') knowledge occurs when two or more individuals not only know that φ, but also know that each other knows that φ, and also that each other knows that each other knows that φ and so on. Such knowledge allows us to form higher order expectations about the activities of others, and to plan in light of these expectations. Lewis gives two reasons for thinking that

(4) is not just desirable for participation in conventions but necessary. First, it reflects the fact that conventions typically are common knowledge among the members of a community (Lewis 1969, p. 59); and second it blocks from being counted as conventional a counter-example that Lewis found unsatisfactory. Those working in naturalistic philosophy are unlikely to be persuaded by Lewis's arguments, and so I won't repeat them. However, a functionalist argument can be given that would show mutual knowledge of conventions to be necessary for their use, by virtue of its coordinating role. The central idea here is that where performance of a practice R is mutual knowledge among the members of a community, then individuals' performance of R on any given occasion will be insulated from higher-order doubts about the performance of others. I return to discussion of this argument in "Mutual knowledge is unnecessary for linguistic coordination". In the meantime, it suffices to say that this argument has been accepted by at least some of those who have adopted Lewis's account of convention—not least Tomasello (2008, implicitly; and in conversation).[6]

Responses to Lewis

As many commentators have observed, although Lewis's analysis makes it in principle possible for non-verbal creatures to participate in conventions, in practice the socio-cognitive requirements that he diagnoses as necessary for participation may be prohibitively demanding; perhaps sufficient to exclude from participation both non-verbal creatures, and some linguistic ones too. In part this is because the addition of clause (4) to the analysis makes it a requirement that the members of P in S have common knowledge that the R that they perform "satisfies the defining conditions for [being] a convention" (ibid., p.61)—i.e. knowledge of clauses (1)–(3). For any two members of P, M_1 and M_2, they would have such knowledge only if:

> M_1 knows that R satisfies the defining conditions for a convention (M_1kCR)
> M_2kCR
> M_1kM_2kCR
> M_2kM_1kCR
> M_1kM_2kM_1kCR
> M_2kM_1kM_1kCR

… and so on, with no principled upward limit on the number of knowledge iterations that could be inferred by M_1 and M_2 about what the other knows. Knowledge of even the first stage of this expansion might be prohibitively difficult to acquire, since it would require its being common knowledge not only that others perform and expect others to perform R in S, but also its being common knowledge that they perform R in S only because others do so too, since R is a solution to a coordination problem.

---

[6] While Tomasello appropriates Lewis's account in his discussion of language acquisition, he does not comment on his claim that children could not participate in conventions.

This might make the requirements on participating in a convention exorbitantly high. Lewis mitigates against this worry, at least in part, by conceding that knowledge of R's being C need only be potential—i.e. knowledge of "evidence from which we could reach the conclusion that any of our conventions meets the defining conditions for a convention" (*ibid*. p.63). In that case, none of the members of P need actually grasp the above expansion. They need only be able to infer it, if they "bothered to think hard enough" (*ibid*. p.165). However, even with this relaxation in place, Lewis's epistemic constraints might suffice to preclude the participation in conventions of many individuals. There may be some—including young children, great apes, and our early hominin ancestors—who would fail to grasp the nature of coordination problems even if they thought very hard about them, and for a very long time. Indeed, if the above expansion were infinitely large, and if each stage of it needed to be graspable for individuals to count as having potential knowledge, then no one would have such knowledge of the defining conditions of conventionality—since none of us can think hard enough to grasp that.[7]

Intuitively, the conclusion that young children and others couldn't participate in conventions is unsatisfying. After all, if words are paradigmatic case of conventions then verbal communication should be a paradigmatic case of participating in a convention; and even infant language-users should be counted as participants. Perhaps because of intuitions like this, critics of Lewis have long argued that he intellectualises the knowledge required for conventions.[8] One of Lewis's critics, Skyrms, has sought to show the epistemological extravagance of Lewis's account by demonstrating that in a system of dynamic replication, communities of organisms can evolve stable, arbitrary signalling systems that function perfectly well in the absence of common knowledge—and indeed in the absence of any knowledge at all (Skyrms 1996, pp. 93–94). However, in taking as his starting point the process of differential reproduction among successive generations of organisms in a competitive environment and not rational choice theory, Skyrms develops a model of reinforcement learning that is only minimally psychological and which functions independently of the choices of rational agents who lack a common language but who are nonetheless engaged in a real-time process of trying to coordinate their activities with the activities of others. This marks a substantial difference between Skyrms's and Lewis's approaches, since the latter explicitly sought to model the thought processes that would rationalise conformity to a convention. Therefore, while Skyrms's approach can explain how arbitrary signalling systems could, over several generations, emerge in communities of bacteria, it may not be the right approach by which to explain how conventions perpetuate among human communities. A wealth of empirical evidence now undermines the idea that human infants are mindless, and shows them to be capable of some very sophisticated cognitive feats (e.g., Tomasello 1999, 2008). Since we have good reason to believe

---

[7] Thanks to Kim Sterelny for this point.

[8] It's worth noting that many of these critics—including Binmore (2008), Burge (1975/2007), Millikan (2005) and Skyrms (1996)—have failed to appreciate that the knowledge Lewis requires for participation need only be potential. This failure may have motivated, at least in part, their criticisms.

that infants are minded, then it would be at least hasty to conclude that Skyrms's model is the right one for modelling their behaviour. There may be relatively simple processes of reasoning by which non-verbal but still rational creatures could coordinate with others, and that would better describe the 14-month-old infant's first attempts to use language to interact with its caregivers, or the attempts of early hominins to coordinate their big-game hunting activities. Were that the case, these thought processes could be modelled.

In what follows, by way of a reappraisal of Lewis's account of the knowledge that is necessary and sufficient for participating in conventions, I sketch out an account of knowledge that would—in principle—be sufficient to explain the participation of even non-verbal creatures in conventional activity. I present this account with several desiderata in mind. First and foremost, (1) the modelled knowledge should explain how—by virtue of what knowledge on the part of participants—conventions could fulfil their coordinative functions. Additionally, (2) this account should specify the knowledge required for coordination in a way that tells us something about what sorts of reasoning and interaction among the members of P might support their coordination. If the reasoning and interaction that is sufficient for coordination is such that it could be undertaken by even non-verbal creatures, then a third desideratum can also be met: (3) the need to explain how non-verbal creatures might come to use linguistic conventions. If these desiderata are met, we can learn something not only about a way in which young children might acquire the ability to use words communicatively, but also about a way in which conventional behaviours might have perpetuated among communities of our early hominin ancestors.

A functionalist account of conventional activity

Lewis identified a fundamental functional role of conventions when he recognised that they constitute solutions to coordination problems. Communication is one clear case when coordination with others is required. Since the strings of sounds that speakers produce in communication are arbitrary, if speaker and hearer are to communicate—at least, if they are to do so with ease[9]—they must agree on which sounds are being used in which ways. The existence of linguistic conventions allows speakers to coordinate their communicative goals by using words in an established way. In setting out a more austere account of what one must know to participate in conventions, I propose to start by taking this coordinative role to be fundamental. This marks a departure from Lewis's account, since he was concerned "to write into the definition all of the important features common to [the] examples" (1969, p. 59) by which he introduced the term 'convention'. Some of these features may be unnecessary for coordination.

---

[9] Davidson (1986) argues that shared knowledge of a language is unnecessary for linguistic communication and, moreover, that convention has no valuable role to play in explaining communication. In fact, though, he does not show that conventions are unnecessary for language use; only that where speakers can track others' deviations from conventional use, linguistic communication can persist.

Here it might be objected that not all conventions are solutions to coordination problems; and that these would be well characterised neither by Lewis's analysis, nor by any account of convention taking its coordinative function as a starting point (including Millikan 2005; Davis 2003). This is surely true: convention is a family resemblance concept, and there exist various phenomena that some would call 'conventional', but where there exist no coordination problems to which these practices constitute a solution (see Davis 2003, p. 225, for examples). This can be conceded without loss, though. In contrast to Lewis, the purpose here is not to provide an account of "convention in its full generality" (Lewis 1969, p. 3)—a task which may well be impossible—but only to say something about how conventions of a certain sort, namely linguistic conventions, could be acquired by non-verbal creatures. In this case, the remarks that follow should be taken to apply only to the sub-class of conventions that are coordination equilibria.

What would the members of a community P need to know in order to be able to use conventional tools to communicate? That is, what knowledge is necessary and sufficient for participation in a coordinating convention? I answer this question through a reappraisal of Lewis's conditions (1)–(4).

Lewis's (1) and (2) are necessary for participating in linguistic conventions

An obvious first point that nonetheless bears repeating is that in a coordination problem S, with many possible but arbitrarily different solutions $R_1 \ldots R_n$, coordination among the members of a population P would be achieved so long as they did the same thing. To borrow an example from Wittgenstein (1953/2000—the "complete primitive language" of the builders at PI§2),[10] consider a language consisting of four words only—'block', 'pillar', 'slab' and 'beam'—which are used to request the objects to which they respectively refer. The builders call out "Block!" when they want a block, "Slab!" when they want a slab, and so on. If enough builders used these words in these ways then those asking for blocks should reliably receive blocks, those requesting slabs, slabs, and so on. There might be individuals who sometimes handed out slabs when they knew they should hand out pillars, but here the system would break down for lack of cooperation and not any lack of understanding.

This point explicates the functionalist considerations that underscore the first clause of Lewis's analysis of convention. The justification for Lewis's clause (2) follows when we ask what would be required on any given occasion to explain a speaker's intentional performance of a speech act R_1, and not some arbitrarily different speech act R_2.

A fact about communication is that it requires uptake. That is, for a speaker's message to be communicated successfully, speaker and hearer must come to some consensus about what that message was. Speakers know this, and so tailor their messages to maximise the possibility of uptake—using words in ways that they expect others to understand. This tailoring isn't just desirable for communication. In fact, as Grice argued (1989, p. 98), it's necessary for acting with communicative

---

[10] Substituting building activities for hunting trips, the language of Wittgenstein's builders has just the cooperative and purposive character one might expect early languages to possess, and the minimal syntax.

intent that one expects that one's intention could be fulfilled. That's because one cannot act with an intention that one does not believe possible. Thus, for example, while one could reasonably expect that, by uttering the word "Slab!", one would be interpreted as requesting a slab, one couldn't ordinarily thereby expect to be understood as inquiring about the health of a friend, or asserting that it is raining. The existence of a convention for using the word 'slab' with a limited range of perlocutionary goals generates in language-users an expectation that, in the absence of stage setting that would suggest otherwise, anyone using this word would be using it in accord with these goals. Anyone using words in ways inconsistent with others' expectations could not thereby expect their communicative goals to be fulfilled; and so could not act with communicative intent.

Lewis's second clause can be seen as an expression of the idea that acting with communicative intent requires expecting that one's message could be inferred. In many this would take the form of having expectations about the tendencies of interlocutors to use words with certain communicative goals; or at least about the capacity of interlocutors to understand others' uses of those words. As Millikan (2005) notes, it's normal that language users understand the ways in which others use words—e.g., slang terms, and words from regional dialects—that they themselves would not use. One needn't use a word oneself to have expectations about how others use it.

Given the above, we have reason to think that clauses (1) and (2) of Lewis's analysis are, in some sense, required for conventional communication. However, they may not be quite right on certain points of detail. This is certainly the case for Lewis's requirement that "almost all" conform and expect others to perform a convention. As Ruth Millikan has objected, large-scale performance of conventions need not be necessary for them to fulfil their coordinating functions. She writes that even in the case of what she calls 'blind conventions'—which, like linguistic conventions, depend for their effectiveness on the conformity of members within a group—"only regular enough conformity is necessary to sustain them" (Millikan 2005, p. 14). This may fall short of even majority conformity, since if even a few people perform R in any instance of S reproducing R may nonetheless be worthwhile. For example, there may be pieces of language that are idiomatic, technical or antiquated and consequently used only by a few (Davis 2003). At least where one has some reasonable expectation that one's communicative goal could be understood it may often be worth using such vocabulary—especially since the cost of trying and failing is likely to be low. If an interlocutor fails to understand and responds only with a puzzled expression, the cost of reformulating one's utterance and trying again is minimal.

Given this, Lewis's requirement that "almost everyone" perform conventions is too strong, and the analysis of convention should be changed to reflect this. It's surely not necessary, though, to develop any detailed specification of what proportion of a population must participate in a convention in order to sustain it. Since in actual cases we often lack precise knowledge of who is participating in a convention, it will be advantageous to leave this specification loose.[11] A further advantage is that doing this also accommodates cases in which the performance of R in S by some members of P matters more than performance by others.

---

[11] Davis (2003, pp. 218–9) makes a similar claim.

Knowledge of (3) is unnecessary for use of linguistic conventions

Clause (3) is inserted into Lewis's analysis to distinguish conventional activities from non-conventional ones. Since, in the case of conventional coordination devices, arbitrarily different regularities could have been used to the same effect, participants should perform R only where others do too.[12] However, as others have previously noted (e.g., Burge 1975/2007; Binmore 2008) it's surely unnecessary for coordination that the participants in a convention know (3). The irrelevance of it to the coordinating function of language is made evident by the fact that there could be—and throughout history have been (e.g., Eco 1995, pp. 95–103)—communities of language-users who not only fail to recognise the arbitrariness of their language, but explicitly deny it too. The fact that individuals could succeed in communicating this denial shows that conventions perform the same coordinating role whether or not participants know (3).

Mutual knowledge is unnecessary for linguistic coordination

Where the coordinative function of conventions is taken to be central to their characterisation, the extent to which common knowledge is necessary for conventional activity will be determined by its coordinative role. Such a role might consist in protecting participants in a convention from higher-order doubts about the conformity of others. These doubts could potentially undermine the functioning of a convention, by causing individuals to deviate from it in anticipation that others might do the same. To give an example, I may know that it's standard practice in our community to use "Slab!" as a means of requesting slabs. However, if I have reason to believe that you think that "Block!" is the appropriate means for requesting slabs, then I may have reason to utter "Block!", and not "Slab!", when requesting slabs from you. Alternatively, suppose that I have only recently learned that "Slab!" is the means by which to request slabs; having previously thought that it was "Block!". If I think that you are ignorant of my having learned this, and so expect me to use "Block!" to request slabs, then I may also utter differently. Doubts like this can recur at potentially any order of reasoning about our interaction, and might suffice to rationalise one individual's deviation from a convention, leading to a potential failure of coordination. Mutual knowledge that participants perform R would secure against such failure, by ruling out any uncertainty that might exist between interlocutors about the tendency of others to conform. Mutual knowledge, then, can play a role in enabling conventional coordination. The question is whether it is necessary for it.

A first point to note is that in practice none of the above-described points of confusion would inevitably and irreversibly lead to a failure of coordination. When conducted in real time—and not, for example, by post—communicative interactions

---

[12] It may be that, once established, arbitrary conventions acquire further properties—grounded in ritual and custom, or in the inability of old dogs to learn new tricks—that differentiate them further from alternatives that would once have been considered equally good. Here the members of P might become attached to these ways of acting in S, such that they would no longer cease to perform a conventional activity when others did; and so would no longer act in conformity to (3). I note this possibility, as does Burge (1975/2007, p. 34), but won't discuss it further here.

provide a wealth of opportunities for anticipating and correcting such mistakes. For example, if you really doubted my understanding of "Slab!", you might take measures to double-check, by showing me a slab and giving me a quizzical look before making the effort to hand it over. In some cases, opportunities for anticipating mistakes won't present themselves and breakdowns may occur. However, these mistakes will themselves constitute opportunities for learning, and so facilitate coordination in the future. In this case, while mutual knowledge might be important for communication that is fast, efficient and largely error-free, for the most part a bumpier form of communication could persist in its absence. In empirical fact, when we interact with strangers (and even friends) we often lack mutual knowledge of the proper uses of the words that we use.

A further consideration is that any need for epistemic reassurance about others' tendency to participate in a convention arises only in conjunction with the ability to doubt that participation. Were participants cognitively incapable of entertaining doubts about the ways in which others use words, iterations of higher-order knowledge would be unnecessary for coordination and clause (4) would serve no coordinating function. Thus, for example, it may be that Wittgenstein's builders learned the functions of their words unreflectively. Until given reason to grasp that not all used the same words in the same ways, they would simply go about their business, calling out "Slab!" to request slabs, and so on. If no one broke with convention, their coordination would persist.[13] If communication did break down, it could probably be rectified, since the mistaken subject could point to the object he desired—the block that had been mistakenly called a 'slab'. Now the storekeeper might insist "Block!" and hand it over; and either the confused builder would learn his mistake or not. Either way the same breakdown could be resolved in the same way on future interactions. Higher orders of coordination need never be required, since breakdowns would either not occur, or be dealt with when they arose.

Appreciating this fact considerably reduces the meta-representational demands of communication, without thereby entailing that communicators lack insight into the grounds of their communicative interactions. It may be that learners of a language acquire only unreflective knowledge of a rule about the function of a particular coordinating device:

doing R in S is a way to achieve E.

In such circumstances, until they acquired further reasons to doubt this assumption, users of R should succeed in coordinating. Coordination might sometimes break down—but here participants might be expected to bump heads a couple of times, before hitting upon an alternative strategy (or giving up). Consequently, while mutual knowledge can facilitate coordination, it isn't necessary for participation in conventions.

---

[13] Kim Sterelny points out an analogy here with Paul Seabright's fascinating discussion of the role of 'tunnel vision' in the functioning of economic markets. These markets tend to run smoothly when manufacturers and purchasers unreflectively attend to their needs and take the operation of the market for granted. See Seabright (2010, chapter 1) for discussion.

The functional analysis

The account of convention, and of the knowledge required for participating in conventions, suggested by these revisions is as follows:

A regularity R in the behaviour of members of a population P when they are agents in a recurrent situation S is a convention if and only if it is true that in any instance of S among members of P,

(1) members of P perform R;
(2) members of P expect other members of P to perform R (or expect that others will recognise their performance of R);
(3) members of P perform R on condition that others do, since S is (at least often) a coordination problem and performance of R is a proper coordination equilibrium in S.

In this analysis, the removal of 'all' from the clauses reflects the fact that we may require only a vague notion of who participates in a convention to make performance of it worthwhile. The parenthetical remark in (2) reflects the fact that not all need do R themselves in order to have expectations about how others use R. Clause (3) distinguishes conventional actions from non-arbitrary coordinating functions, but imposes no epistemic requirements. It simply describes the fact that (absent factors like nostalgia) the members of P have no reason to perform R in S unless others do too—because alternatives R_1–R_n could have done the same job just as well. The only knowledge requirement on participating in a convention is the expectation that, among the members of a population P, R can be used in S to achieve E. In other words, the members of P need only grasp that R can be used with a certain tool function.

The requirements on what members of P would need to know to use an arbitrary action R to communicate in S have now been specified. For the purposes this paper, this amounts to having knowledge of a linguistic convention—and so constitutes an answer to questions (2) posed at the outset. With this answer in place, I turn now to answer question (1)—the question of how a situation could arise in which the members of P had this knowledge.

Imitation as a means of coordination

In imitation, there exists a variety of social learning that can be used to explain how a behaviour that could be learned only from others could be disseminated across a community such that members would perform an action R_1, to the exclusion of arbitrarily different actions $R_2 \ldots R_n$, and which yields knowledge sufficient for participation in linguistic conventions.[14] More importantly, imitative learning can be non- verbal, and is a skill at which pre-verbal children are adept.

---

[14] Lewis himself observes that one can participate in conventions by imitating others. However, he correctly emphasises that "we should not conclude that any regularity which originates or persists by imitation is therefore a convention" (1969, p. 118)—since imitated behaviours may not be arbitrary.

Tomasello and Carpenter (2005) define imitation as the ability to observe and act with not just the intended goal of another's action, but also the means employed in pursuit of that goal. A general characterisation of this process might be given as follows:

> In imitation, an imitator M_2 reproduces an action R in pursuit of some end E on the basis of having seen another individual M_1 perform R in the pursuit of E.

In any coordination problem S, E would be the goal achieved with coordination (for example the requesting of a slab) and R the means for achieving that goal (for example, an utterance of "Slab!"). The word 'action' is here used in a fairly loose sense, such that instances of performing an action could include both gesturing for an other and producing a vocal utterance.[15] The goals with which communicators act are perlocutionary goals. These are the responses that they intend to elicit from interlocutors by uttering, and are typically specified in terms of intentions to produce beliefs in others, or to have them perform actions (Grice 1989, chapter 6; Moore submitted).

Knowledge learned imitatively is sufficient for linguistic coordination

In a coordination-problematic situation S in which at least some members of P—call them P-innovators—already coordinate, other members of P—call them P-imitators—could observe that among P-innovators they could achieve their E-type goals by performing R. Thus they might learn that:

> doing R in S is a way to achieve E.

P-innovators might learn both the goals of others and the means used to pursue these goals observationally. For example, suppose a new member of the community observes that when the other builders want slabs, calling out "Slab!" leads to them being given a slab; and so on. Understanding of the desire with which the builders act is inferred from the fact that when they call out "Slab!" and receive slabs, the builders appear contented. By contrast, when a call of "Slab!" is occasionally met with a block or a beam, they look unhappy and repeat their original utterance until they get a slab. The new member of P might reasonably infer from her observations that calling out "Slab!", at least around other builders, is an appropriate way to get a slab.

This observation that doing R is a way to achieve E in S gives P-imitators in S a reason to reproduce R in pursuit of E. Any individual doing this would satisfy the first clause of the functional analysis elucidated above. Additionally, since knowledge of R as a means of achieving E was learned from others, this would also generate a concurrent expectation sufficient for (2)—that others, or at least one other, also perform(s) R in S. In the first instance, it may be that individual P-imitators do not have strong expectations about how widespread this practice is; perhaps they would expect only that those from

---

[15] On accounts of imitative learning that take it to consist in the copying of body movements (e.g., Heyes 2001; Byrne 2002; Paulus 2011), one could not imitate another's utterances of sounds, but only emulate it. Some take this to mean that all vocal copying, including human language copying, is really emulation learning. As previously noted, I use 'imitation' in a less restrictive sense.

whom R had been learned perform it. However, throughout a course of repeated interactions with others, this expectation could be expected to generalise. In time, where participants could track high order representations of others' knowledge, they may even acquire mutual knowledge of the performance of R in S—with all of the implications for flexible coordination that this implies. In the meantime, even in the early stages of learning, the existence of some R-users would give P-novices a reason to experiment with R in the company of strangers, since their existence would give novices reason to think it possible to achieve E in S by doing R—thereby satisfying (2). This would be particularly true when the novices did not know any other way to achieve E in S.

Imitative learning need not engender knowledge of (3)

While imitative learning is sufficient to generate expectations (1) and (2), it need not generate any knowledge that some R satisfies (3). That's because where a convention already exists, problems need not present themselves as coordination problems demanding one of potentially many arbitrarily different solutions. Often, when we observe others perform speech acts, the words that they utter are simply given to us as the appropriate tools with which to pursue their communicative goals. It need never occur to us that, whereas hammers have physical properties that make them intrinsically suitable for cracking nuts, no analogous property of our words makes them intrinsically well-suited to the tasks for which they are used. So long as imitative learners don't recognise that arbitrarily different speech acts might equally have been used for just the same ends, they would fail to have even potential knowledge of the conventionality of their utterances. They could coordinate their activities with those of their peers, though.

While it isn't necessary that one recognise the non-intrinsic suitability of a conventional action to use it appropriately, a further feature surely is necessary for its appropriate use to be learned. Where a coordinating tool acquires its functional properties through convention, those who learn it must be particularly attentive to the nature of the action that others perform—that is, to the means R that an observed agent employs in pursuit of her goal E. That's because in conventional coordination, unlike in the case of nut-cracking described at the outset, one couldn't (unless one was very lucky) figure out the appropriate means to a goal independently of copying the actions performed by others.

An imitation-based account of convention learning satisfies desiderata (i)–(iii)

In conjunction with the appeal to imitation, the model outlined provides us with some insight into a process of reasoning by which coordination could be achieved, by specifying an inference that would be sufficient to motivate any individual's performing R in S—in the form of the insight that doing R in S is a way to achieve E. Consequently, desiderata (i) and (ii) specified above are met. Since this use of R could be learned imitatively, success need not presuppose any understanding of language—fulfilling (iii). Given the comparatively modest cognition that is necessary and sufficient for conventional coordination on this account, it could

also plausibly serve to explain the dissemination of conventional practices among both young children and our hominin ancestors.

Central to the requirements of imitation here are two abilities. First, the ability to grasp the goal with which an other's action has been performed; and second, the ability to recreate precisely that observed action in pursuit of that goal. There may be cases in which both the actions that demonstrators perform and the intentions with which they act are sufficiently easy to discern that non-human animals could imitate. For example, some variety of imitative learning may be present in pigeons (Saggerson et al. 2005). In such cases, associative learning may play a role in establishing a connection between the means and the goal. Nonetheless, intended-goal recognition and means copying may not themselves be explicable by appeal only to associative processes. Indeed, a wealth of evidence shows that human children can grasp the communicative goals of others in cases that cannot be explained by existing accounts of associative learning (Baldwin and Moses 1994; Akhtar and Tomasello 1996; Akhtar et al. 1996). Furthermore, evidence from chimpanzees suggests that precise action copying is difficult even for some cognitively sophisticated animal species.

Imitative learning in chimpanzees

Empirical evidence supports the idea that the facility for imitation could be an important factor in explaining the propagation of arbitrary conventions in human communities. Evidence of such conventions in ape societies is scarce, at least in contrast to human communities, and it may be no coincidence that apes have shown little evidence of an aptitude for imitative learning. For example, in a recent study by Tennie et al. (2012), only one chimpanzee, Baluku, from a group of 15 was able to copy a familiar action performed by a trained conspecific demonstrator, Mawa, in exchange for a reward. Moreover, when the familiar action performed by Mawa was replaced by an action that was not already in Baluku's behavioural repertoire, he failed to reproduce the new behaviour at all. This suggests that while some chimpanzees may be inspired to reproduce familiar actions by watching others, they do not (or only rarely) learn new actions imitatively. Tennie, Call and Tomasello offer an explanation of why this may be:

> When chimpanzees observe someone using a tool, they tend to focus on the effect being produced in the environment, and they pay little attention to the actual bodily actions of the tool user. They then use their own behavioural strategies to produce the environmental effect. They thus reconstruct the product rather than copy the process leading to it. (Tennie et al. 2009, p. 2407)

In other words, chimpanzees are inattentive to, or simply fail to reproduce, the means by which the effect on the environment is produced.[16] Where the relation of a means to its goal is arbitrary, as in the case of conventional communication, it

---

[16] Others have also found that copying behaviour in children is more faithful then in chimpanzees. See Call et al. (2005), Horner and Whiten (2005) and Nielsen (2009), for relevant discussion; and Boesch (2012), for a contrasting view.

follows that chimpanzees should be poor learners—because (unless by chance) the proper performance of an arbitrary social tool could not be learned except by attending to its performance by others. Whereas for imitative learning, some entertaining of the following thought is required:

Performing A in S is a way of doing E,

what chimpanzees grasp, upon seeing another perform A in S and achieving E, may be only:

E can be achieved in S.

Through a process of skilful trial and error learning, they then discern their own techniques for performing A.

Here it may be objected that conventional behaviours do exist in communities of chimpanzees—because, for example, their gestures or vocalisations are arbitrary. However, the evidence for this claim is by no means straightforward. In fact, two views dominate the literature on ape vehicles of meaning, and neither implies conventionality. There is roughly uniform consensus that ape vocalisations are innate (Tomasello 2008). Additionally, on some accounts (Cartmill and Byrne 2010; Hobaiter and Byrne 2011), ape gestures are also thought to be part of their biological inheritance. However, innate gestures and vocalisations cannot be conventional, since they are used not only because other apes also use them, but because they are part of an inherited adaptation for communication —violating (3). On the alternative to the nativist view (Tomasello and Call 2007; Tomasello 2008), some ape gestures work by displaying naturally meaningful bodily states (like erections, emotional expressions, or piloerect fur) via the use of attention getting behaviours (like foot stamping), while others are ritualised forms of action schemas ('intention movement signals')—which thereby visually represent the messages they are used to communicate. On this account, the vehicles of communication would also not be conventional, since they would again violate (3) by virtue of being intrinsically well suited to the tasks for which they are used. In that case, neither of the existing accounts of the mechanisms of ape gestural communication implies conventionality. By contrast, there is no question that a great many of children's early uses of language are conventional in form.

Conclusions

In contrast to Lewis's account of convention, knowing how to use conventional language to coordinate with others does not require sophisticated meta-represen-tational abilities, or any insight into the conventional nature of one's activities. It requires only that speakers learn that words can be used with certain tool functions: that the utterance of a certain word (or gesture) can be used as a means to the achievement of a particular goal. Since imitative learning is sufficient for the acquisition of such knowledge, it may well play a fundamental role in children's language acquisition. Certainly, infants are good enough at imitation for this to be the case. It may also be that the proliferation of conventional behaviour in human

societies, but not in communities of non-human great apes, can be explained by the human facility for imitation (Nielsen 2009). While the development of a flexible facility for social learning, in the form of the ability and motivation to imitate peers, was certainly not the only transition that led to the emergence of language in our early hominin ancestors, it may nonetheless have been a significant one.

References

Akhtar N, Tomasello M (1996) Two-year-olds learn words for absent objects and actions. Br J Dev Psychol 14:79–93

Akhtar N, Carpenter M, Tomasello M (1996) The role of discourse novelty in early word learning. Child Dev 67:635–645

Baldwin D, Moses J (1994) Early understanding of referential intent and attentional focus: evidence from language and emotion. In: Lewis C, Mitchell P (eds) Children's early understanding of mind. Lawrence Erlbaum, Hove

Binmore K (2008) Do conventions need to be common knowledge? Topoi 27:17–27

Boesch C (1991) Teaching among wild chimpanzees. Anim Behav 41:530–532

Boesch C (2012) Wild cultures: a comparison between chimpanzee and human cultures. CUP, Cambridge

Burge T (1975/2007) On knowledge and convention. Philos Rev 84:249–255. Reprinted in Burge (2007) Foundations of mind. OUP, Oxford.

Byrne R (2002) Imitation of novel complex actions: what does the evidence from animals mean? Adv Study Behav 33:77–105

Call J, Carpenter M, Tomasello M (2005) Copying results and copying actions in the process of social learning: chimpanzees (*Pan troglodytes*) and human children (*Homo sapiens*). Anim Cogn 8:151–163

Carpenter M, Akhtar N, Tomasello M (1998) Fourteen- through 18-month-old infants differentially imitate intentional and accidental actions. Infant Behav Dev 21:315–330

Cartmill EA, Byrne RW (2010) Semantics of orangutan gesture: determining structure and meaning through form and use. Anim Cogn 13(6):793–804

Davidson D (1986) A nice derangement of epitaphs. Reprinted in Davidson (2005) Truth, language and history. OUP, Oxford

Davis W (2003) Meaning, expression and thought. CUP, Cambridge

Eco U (1995) The search for the perfect language. Basil Blackwell, Oxford

Gergely G, Csibra G (2005) The social construction of the cultural mind: imitative learning as a mechanism of human pedagogy. Interact Stud 6(3):463–481

Gergely G, Bekkering H, Király H (2002) Rational imitation in preverbal infants. Nature 415:755

Grice P (1989) Studies in the way of words. Harvard UP, London

Heyes C (2001) Causes and consequences of imitation. Trends Cogn Sci 5:253–261

Hobaiter C, Byrne RW (2011) The gestural repertoire of the wild chimpanzee. Anim Cogn 14:745–767

Horner V, Whiten A (2005) Causal knowledge and imitation/emulation switching in chimpanzees (*Pan troglodytes*) and children (*Homo sapiens*). Anim Cogn 8:164–181

Laurence S (1998) Convention-based semantics and the development of language. In: Carruthers P, Boucher J (eds) Language and thought. CUP, Cambridge, pp 201–207

Lewis D (1969) Convention. Blackwell, Oxford

Millikan RG (2004) Varieties of meaning: the 2002 Jean Nicod lectures. MIT, Cambridge

Millikan RG (2005) Language: a biological model. OUP, Oxford

Moore R (submitted) Cognising communicative intent

Nielsen M (2009).The imitative behaviour of children and chimpanzees: a window on the transmission of cultural traditions. Revue de Primatologie [online], 1, document 5. http://primatologie.revues.org/254

Paulus M (2011) Imitation in infancy: conceptual considerations. Theory Psychol 21:849–856

Paulus M, Hunnius S, Vissers M, Bekkering H (2011) Imitation in infancy: rational or motor resonance? Child Dev 82:1047–1057

Saggerson AL, George DN, Honey RC (2005) Imitative learning of stimulus-response and response-outcome associations in pigeons. J Exp Psychol Anim Behav Process 31:289–300

Schwier C, van Maanen C, Carpenter M, Tomasello M (2006) Rational imitation in 12-month-old infants. Infancy 10:303–311

Seabright P (2010) The company of strangers: a natural history of economic life, 2nd edn. Princeton UP, Oxford

Skyrms B (1996) Evolution of the social contract. CUP, Cambridge

Tennie C, Call J, Tomasello M (2009) Ratcheting up the ratchet: on the evolution of cumulative culture. Philos Trans R Soc Lond B Biol Sci 364:2405–2415

Tennie C, Call J, Tomasello M (2012) Untrained chimpanzees (Pan troglodytes schweinfurthii) fail to imitate novel actions. PLoS ONE 7:e4154.8

Tomasello M (1999) The cultural origins of human cognition. Harvard UP, Cambridge

Tomasello M (2000) The item-based nature of children's early syntactic development. Trends Cogn Sci 4(4):156–163

Tomasello M (2008) Origins of human communication. MIT, Cambridge

Tomasello M, Call J (eds) (2007) The gestural communication of apes and monkeys. Lawrence Erlbaum Publishers, London

Tomasello M, Carpenter M (2005) Imitation reading and imitative learning. In: Hurley SL, Chater N (eds) Perspectives on imitation, vol 2. MIT, London

Whiten A, McGuigan N, Marshall-Pescini S, Hopper L (2009) Emulation, imitation, over-imitation and the scope of culture for child and chimpanzee. Philos Trans R Soc Lond B Biol Sci 364:2417–2428

Wittgenstein L (1953/2000) Philosophical investigations, 3rd edn. Trans. G. Anscombe. Blackwell, Oxford