12

Why humans and not apes: the social preconditions for the emergence of language

CHARLES WHITEHEAD

12.1 Introduction

In this Chapter, I aim to show that the preconditions for the emergence of language are both social and uniquely human, using a theory that I think fits what we know about human behaviour better than its more popular rivals.\(^\dagger\) Social mirror theory arose at the end of the 19th century as an explanation of self-consciousness (Dilthey 1883–1911; Baldwin 1894). It holds that 'mirrors in the mind depend on mirrors in society' (Whitehead 2001)—that is, social displays (such as facial expressions, music, pictures, and language) make thoughts and experiences public, so that, during childhood, they make us aware of thoughts and experiences in our own minds and in the minds of others. In this way, we become self-aware and other-aware at the same time, and self/other awareness is a single, indivisible phenomenon (henceforth 'social awareness' or 'intersubjectivity').

Social mirror theory complements a second theory, originally proposed by Émile Durkheim (1964 [1912]), that the emergence of language depended on ritual. This idea has been extended to become 'ritual/speech co-evolution theory' (Knight 1998). The theory further maintains that any attempt to explain language as an isolated trait is akin to explaining the emergence of the credit card without considering the preconditions on which credit cards depend—including commerce, money, banking, the digital computer, and the means to detect and punish fraud.

\(^\dagger\) All Whitehead references cited in this chapter (except 1995 and 2004) can be downloaded in draft or final form at www.socialmirrors.org. Go to 'About Charles Whitehead'.

\(^2\) The main theories are reviewed at www.socialmirrors.org. Go to 'Self/Other Consciousness' and click link to section 2 (at foot of page).
According to these two theories, in order to explain the origins of language, we must first explain the whole package of social displays and social awareness on which human ritual depends. In what follows, I will discuss the unique nature of human social displays, social awareness, and culture, and present a ‘play and display’ hypothesis to explain the evolution of human sociality and our ‘culture-ready brain’, without which language would not be possible.

12.2 The human difference

12.2.1 Social displays and social awareness

We humans have a formidable armamentarium of social displays. According to social mirror theory, these displays account for our exceptional social awareness. Although all apes have a rich repertoire of displays, I submit that humans have a broader range of displays—some unique to our species, and some which are more greatly elaborated than those of other apes.

Note that, in the context of social mirror theory, a social display is any kind of shared behaviour which makes thoughts, intentions, or experiences public. Hence it includes behaviours such as grooming and play, in addition to those more commonly regarded as ‘signals’. In animals with mirror neurones (neurones which fire when an individual performs an action and when the individual sees another perform the same action), even pragmatic actions—such as grasping a nut to eat it—can function as social displays which, according to most theories of social awareness, would generate a basic level of insight into desires and intentions.

We humans have three distinct kinds of social display—communication, play, and performance—which serve quite different functions (Whitehead 2001, 2003). Communication is generally goal-directed and aimed at influencing others (Krebs and Dawkins 1984). Play is quite different. It serves exploratory and developmental functions, but has no goal—it is pursued ‘just for fun’. Performance is different again. It is both communicative and playful, and can serve the additional functions of social grooming and entrainment,\(^3\) enabling two or more ‘selfish’ individuals to behave like one much bigger selfish individual.

Our three types of display come in at least three modes—implicit, mimetic, and conventional—which sustain different levels of social awareness (Burling 1993; Mitchell 1994). Implicit displays deal with emotions (e.g. laughter) and autonomic states (e.g. yawning). Mimetic displays, on the other hand, deal with perceptions, representing concrete things, actions, and actors by resemblance (Donald 1990). Conventional displays are culturally transmitted and their meanings rarely depend on resemblance to their referents. Mimetic and conventional modes, being voluntary

---

\(^3\) ‘Entrainment’ implies coordination of thought, mood, and/or behaviour between individuals.
and intentional, can be used to lie, and their utility depends on sufficient levels of social trust.

Table 12.1 shows the nine types of display with some illustrative examples which hopefully make these distinctions clear (for further explanation, see Whitehead 2001, 2010a, 2012).

Many animals have highly elaborated displays, but these are usually of a single type and in the implicit mode. Only humans have multiple sophisticated displays of all three types in all three modes, and all nine are uniquely developed in humans, including many displays which have no obvious precedent in other species. Even implicit gesture-calls are greatly elaborated in humans—for example, ‘the eloquence of the human face’ involves considerably more facial muscles than are found in other primates (Young 1992).
12.2.2 Cultural displays and cultural awareness

Claude Lévi-Strauss (1969) noted that human societies—in contrast to animal societies—are governed by culturally variable rules (involving a social contract, on which the utility of language depends: Austin 1962; Grice 1969; Searle 1969). Human rules are of two types—regulatory and constitutive, where the latter serve to establish institutional forms of social order (Rawles 1955; Searle 1995). Two fundamental institutions which structure all known human societies are formal systems of inflated kinship and extended reciprocity. Kinship and reciprocity are familiar terms to biologists—used to explain the evolution of altruism—yet biologists seldom consider cultural institutions when attempting to explain ‘strong reciprocity’ in humans. A case could be made for ‘strong empathy’ in humans, based on role-play and hypnotic suggestibility, but large-scale co-operation and out-group reciprocity surely depend on cultural systems.

Of course, you cannot have rules without social displays (such as stories, images, and ritual performances) to transmit them from generation to generation. But social displays are more fundamentally important than that—they are virtually constitutive of human culture. Although post-industrial societies have an elaborate material culture with obvious or supposed instrumental utility, even utilitarian objects usually have some aspect of design or display which is intended to influence people or say something about the owners or users of such artefacts. Émile Durkheim (1964 [1912]) coined the term ‘collective representations’ to refer to the shared displays through which human communities define their identity and assert their moral authority. Many of our assumptions about the human condition and the reality we inhabit—those which are passed down from our ancestors, and are so commonsensically taken-for-granted that they are seldom questioned even by scientists (Bourdieu 1972)—are ‘collective representations’. Human infants appear to show a spontaneous interest in cultural displays (Trevarthen 1995) and there can be little doubt that collective representations have profound influences on psychological factors such as personality, self/other-awareness, and visual perception, with concomitant effects on functional brain anatomy (Chiao et al. 2008; Turner and Whitehead 2008).

One point that many behavioural scientists miss is the theatrical nature of enculturated human behaviour—first noted by G. H. Mead (1934) and elaborated by Erving Goffman (1959), Victor Turner (1982), and others. Role-play and theatre are the most fundamental displays on which human cultures are founded. Today, people who spend 35 hours a week role-playing managers, lawyers, employees, etc. may then spend as many hours of vicarious role-play watching television. We live in a make-believe world, wholly believing in our pretence. Wholly-believed-in role-play is a widely accepted definition of the hypnotic state—hence Charles Tart’s (2009) notion of the consensual trance. Constitutive rules, institutions, and theatrical fictions implicate a relatively large role for invention in the emergence of human culture.
12.3 The developmental spiral

Infant development supports two predictions of social mirror theory: first, that social displays and social awareness (self/other-awareness and intersubjectivity) develop in tandem; and second, the development of displays in one mode precedes major transitions in social awareness, accompanied by the emergence of displays in higher modes.

In what follows I rely mainly on research by performative psychologists, because they specifically investigate intersubjectivity and social displays. Colwyn Trevarthen spent 40 years studying human babies and infants in various cultural settings, including video and audio recording of social interactions especially between infants and mothers, and detailed analyses of what he called ‘proto-conversational exchanges’ which prepare for language acquisition and enculturation (Trevarthen 1974). He concluded that infants are born with a ‘virtual other’ representation, and are innately motivated to seek, mirror, and interact with the motives and expressions of sympathetic others (Trevarthen 1998). Repudiating the copy-cat view of ‘simulation theory’ (as embraced by the Parma team who discovered mirror neurones), performative psychologists have shown that human infants are pro-active in their own socialization and enculturation. The input → processing → output model derived from computing simply cannot accommodate this pro-active creativity. ‘Cognitive science’, Trevarthen writes, ‘restricting the role of motives and emotions, puts childhood play and imagination behind bars’ (personal communication).

Infant interactions are playful, jocular, and often musical (Gratier and Trevarthen 2008). Contingent mirror play is the first in a long series of self-motivated (i.e. playful) activities. By three months, babies express melodic vocalizations coordinated with balletic limb movements (‘song-and-dance display’); and around six months use indexical gestures and engage in preverbal ‘clowning, tricks, and jokes’ (Dunn 1991; Trevarthen 1995; Reddy 2001). At nine months, infants begin to make marks on any surface—including their own bodies—whether with paint, jam, or faeces (Jennings 1990).

These developments in display are followed, at around nine months, by the first of two developmental watersheds which have no apparent parallel in apes. Simon Baron-Cohen (1995) calls this the onset of the ‘shared attention monitoring mechanism’, whereas Trevarthen (1995) calls it the perception of ‘the self as participant’. The child becomes aware of herself as a social player and realises that she and mother can share awareness of an object of interest. Trevarthen (1995) observed that the child—unlike the scientist—strives to understand the world ‘in active negotiation of creative imaginings that are valued for their human-made unreality’ (emphasis added). He notes that babies are born ‘hungry for culture’ and it would seem that they are pre-programmed to live in a cultural world of fantasy and fiction.

Three months later children begin to engage in projective pretend play (playing with representational toys), use iconic gestures, and utter the first words (Trevarthen 1995). Pretend play precedes the emergence, at around 18 months, of the self-
conscious ability of children to recognise themselves in mirrors (Gallup 1994), when they may show embarrassment or coyness at their reflection (Parker et al. 1994). Soon afterwards, they begin to use the pronouns ‘me’ and ‘mine’ (Lewis 1994).

At 24 months occurs the second distinctive watershed—the beginning of the ‘terrible twos’ (Lewis 1994) and a new sense of self perceived in terms of social value. Until this age, infants are relatively passive, allowing parents to dress, wash, or change them as they wish. But the terrible two-year-old asserts her newly discovered autonomy and self-worth by rebellion.

This is also the age of the ‘verbal explosion’ and the beginning of role-play. Around the same time, mark-making behaviour develops into swirling ‘mandalas’ and geometric designs (Jennings 1990). By 30 months, language rules are understood, the ‘mandalas’ turn into faces, and children begin to draw ideographic pictures, especially of people. Six months later, children can experience new emotions (and recognize them in others), notably those associated with self-value: shame, guilt, pride, contempt, and hubris (Parker et al. 1994). By four years (or earlier: Dunn 1991) normal children have acquired explicit ‘theory of mind’ (ToM)—the ability to infer epistemic mental states (knowing, thinking, believing, etc.) in themselves and in others (Baron-Cohen 1995). At this age it becomes possible to hypnotize children (Bliss 1986). Hypnosis involves mind influencing mind and may well depend on reflective ToM. It seems logical that early attempts to explore personhood and agency—through role-play and picture-making—should precede the emergence of these abilities.

By the age of six, enculturation has progressed to the point where children understand consensual morality (Parker et al. 1994), and are ready to play competitive ‘games-with-rules’ (Parker and Milbraith 1994). Hypnotic ability peaks a year later and thereafter declines, at least in westerners (Brown 1991). At this age, role-play can occasionally achieve hallucinatory force, and lonely children may create imaginary companions (dissociated autonomous personae: Bliss 1986). I infer that ‘theatre of mind’ (the ability to imagine social scenarios with ‘toy people’ who behave as though they have minds of their own) must be established at this age.

Around the time of puberty, children can experience stage-fright at the thought of performing in front of any audience, and not merely when ridiculed (Mitchell 1994). It takes the whole of childhood to develop principled morality (Parker et al. 1994) and perhaps longer to create full-blown economico-moral personae—the social roles we play in adult life.

Table 12.2 summarizes the above, showing the spiral relationship between social displays and social awareness. Major transitions are marked by diagonal arrows. Note the build-up of social displays prior to each major shift in social awareness, and the way major shifts coincide with transitions to higher modes of display. Language, however, seems to follow a precocious schedule, suggesting some hard-wired basis. In Section 12.4, I will argue that social displays and social awareness co-evolved in a similar spiral manner.
### Table 12.2. Co-development of social displays and social awareness

<table>
<thead>
<tr>
<th>Social awareness</th>
<th>Social displays</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Weeks</strong></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>Innate <strong>Implicit</strong></td>
</tr>
<tr>
<td></td>
<td>Contingent mirror play/Embodied play</td>
</tr>
<tr>
<td></td>
<td>Affective and autonomic gesture-calls</td>
</tr>
<tr>
<td>6–8</td>
<td>Proto-conversational exchanges</td>
</tr>
<tr>
<td><strong>Months</strong></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Self as participant <strong>Projective mimetic</strong></td>
</tr>
<tr>
<td>6</td>
<td>Making marks</td>
</tr>
<tr>
<td>9</td>
<td>Iconic gesture-calls/Pretend play</td>
</tr>
<tr>
<td>12</td>
<td>First words</td>
</tr>
<tr>
<td>18 (Mirror self-recognition)</td>
<td>Coyness, embarrassment</td>
</tr>
<tr>
<td>20</td>
<td>Personal pronouns</td>
</tr>
<tr>
<td><strong>Years</strong></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Self as value <strong>Introjective mimetic</strong></td>
</tr>
<tr>
<td></td>
<td>Role-play</td>
</tr>
<tr>
<td>6</td>
<td>Verbal explosion</td>
</tr>
<tr>
<td>11 (Theatre of mind)</td>
<td>Making pictures</td>
</tr>
<tr>
<td>11</td>
<td>Language rules</td>
</tr>
<tr>
<td></td>
<td>Expressions of self-value (pride, guilt)</td>
</tr>
<tr>
<td>6</td>
<td>Increasingly conventionalized <strong>Hypnotic ability</strong></td>
</tr>
<tr>
<td>7</td>
<td>Expressions of surprise</td>
</tr>
<tr>
<td>11</td>
<td>(Theatre of mind) <strong>Dressing up/Complex joint role-play</strong></td>
</tr>
<tr>
<td>11</td>
<td>Games-with-rules</td>
</tr>
<tr>
<td><strong>Economico-moral personae</strong></td>
<td>Peak in hypnotic ability</td>
</tr>
<tr>
<td><strong>Fully enculturated</strong></td>
<td>Conventional roles</td>
</tr>
<tr>
<td></td>
<td>Collective representations</td>
</tr>
<tr>
<td></td>
<td>Wealth displays</td>
</tr>
<tr>
<td></td>
<td>Collective deceptions</td>
</tr>
</tbody>
</table>
12.4 The evolutionary spiral

If we consider the types and modes of social displays illustrated in Table 12.1, it is clear that they cannot all have emerged simultaneously, and must have evolved in a logical order (Whitehead 2003). Communication has to come first because, in contrast to play and performance, it is a universal feature of all living organisms. All animals that play also communicate, and social play would be impossible without communication. If performance is a playful extension of communication, it follows that communication must be older than play, and play than performance. A similar argument applies to our modes of display—implicit signals are universal in higher animals, mimesis is rare, and both have to be in place before they can be conventionalized to generate modern human culture, including language.

The argument goes further if we assume that performance in one mode scaffolds the emergence of communication in a higher mode. Thus, song-and-dance generates the levels of social trust and social insight—the ability to put ourselves in others’ shoes—required for the later emergence of informative mimesis (Whitehead 2003). It also appears that dance may have pre-adapted the brain for advanced mimesis, since dance engages a subset of the brain areas involved in pretend play (Whitehead 2010b).

This parallels Durkheim’s (1964 [1912]) theory that ritual pantomime (mimetic performance) is essential to the emergence of language (conventional communication). What distinguishes language from animal communication, Durkheim reasoned, is displaced reference—the ability to refer to things not present in the here and now, but only in the mind of the speaker. How can we encrypt an intangible, he asked, unless it is first made public by ritual pantomime—a conventionalized drama that everyone understands, and that is ‘sacred’ in the sense that it carries the consensual authority of an entire community?

Building on Durkheim, proponents of ritual/speech co-evolution theory (Gellner 1989; Knight 1998; Rappaport 1999) have advanced a second argument which points to the same conclusion. Words are cheap and it is too easy to lie. If language could not be trusted, they argue, it would simply be ignored. Hence language depends on a ‘social contract’ and, in societies without police or judiciary, this can only be implemented through ritual and ritually constructed supernatural beliefs.

There is a third argument which also points to a ritual origin for language, though this was denied by its eccentric author, who deemed ritual unworthy of scientific interest. Lévi-Strauss (1950) was greatly intrigued by the curious fact that, in animistic societies, words such as mana, wakan, and orenda—commonly translated as ‘medicine’ or ‘sacred power’—also function as empty referents. That is, like ‘something’ in English, they can be used to denote anything new, strange, or for which no other word can be found. He inferred a single big-bang origin for both language and religion. In that primal creative moment, ‘the entire universe all at once became
significantly’ (Lévi-Strauss 1950: 60). In his view, the first utterance would refer to this cosmic significance and then, as later words were differentiated from this mother-of-all-words, the residual signifier would continue to refer to the prime mover in creation and everything not yet included in our referential system of meanings.

Anyone familiar with religious experience research, as pioneered by William James (1902) and Alister Hardy (1979), will recognize this ‘entire universe all at once became significant’ as a classic feature of spiritual experience. While 40 per cent of westerners have such experiences, perhaps once in their lifetimes, they are much more widespread and frequent in foraging communities (Bourguignon 1973). Indeed, it is difficult to see how words like mana, wakan, and orenda could have acquired their significance in any other way. It should not surprise us if the first utterance that others interpreted in a language-like way (Burling 1999; Tinbergen 1952) did originate in an ‘altered state’. Such states are an inevitable consequence of human dissociative or hypnotic abilities and, among hunter-gatherers, virtually all innovations derive from dreams, visions, ritual trances, or divine visitations (review: Whitehead 2011). Foraging communities accept innovations only because they are regarded as gifts from the supernatural realm.

Durkheim’s argument regarding the ritual origins of language has never, to my knowledge, been effectively countered, and it has been widely influential in social anthropology. If my extension of his argument is accepted—that performance in one mode scaffolds the emergence of communication in a higher mode—then we have the evolutionary spiral shown in Table 12.3. This predicts at least three major transitions during human evolution. The first would be the emergence of song-and-dance display; the second a major expansion of mimetic and pretend play abilities culminating in ritual pantomime; and this in turn would trigger the third—the emergence of modern culture, with the development of social rules, language, religion, and all the cultural arts.

According to the ‘play and display’ hypothesis (Whitehead 2003), the proliferation of social displays—especially play and performance—was a major factor in human brain expansion, generating a ‘culture ready brain’ capable of performing the first ritual pantomime and maintaining culture of modern type (Whitehead 2010b). Social displays must make demands on the brains of performers and observers alike. All forms of play and performance require multi-modal integration, and performance can involve high levels of acquired skill. Song-and-dance in particular demands fine timing precision (in gelada choral displays, voices are synchronized to within a

---

4 This idea was first proposed in my MSc dissertation (Whitehead 1995) and is based on Turner’s (1969) theory of anti-structure, and the observation that the transformative power of ritual depends on collapse or inversion of everyday categorical distinctions. Such anti-structure is equally characteristic of altered states of consciousness, including ritual trance. Parallel ideas occur in psychological theories of creativity which invoke playful as opposed to instrumental thought, e.g. ‘creative chaos’ (Cooley 1902) and ‘cognitive synergies’ (Apter 2008).
millisecond: Richman 1978). Finally, role-play requires a brain that can run at least two minds in parallel. In our dreams and daydreams, there may be several *dramatis personae* who apparently do have minds of their own—capable of saying and doing things which ‘we’ do not expect.

Accordingly, the first two postulated transitions would initiate phases of brain expansion. The third transition—the emergence of modern culture—would not be expected to lead to further expansion, because social cohesion maintained ‘from the outside’ by rule-bound formal systems would, if anything, reduce the selection pressure for individual playful and performative capacities.

Ontogenesis would not be expected to exactly recapitulate phylogenesis. However, if new levels of self-consciousness require the prior development of appropriate displays in human infancy, it seems unlikely that our ancestors evolved a different way of doing this. So the play-and-display hypothesis would further imply that song-and-dance display would generate a level of social awareness akin to the ‘self as participant’ in human infancy, projective mimesis would lead to self-value, and introjective mimesis to ‘theory of mind’. Finally, the emergence of culture with regulatory and constitutive rules would necessarily produce economico-moral personae.

I do not have space here to review the archaeological and fossil evidence relevant to changes in hominin behaviour and the brain. I have done so elsewhere (Whitehead 2003, 2008, 2010b) though further work needs to be done. I claim only that it is not inconsistent with the play-and-display hypothesis, though other hypotheses cannot be ruled out. However, in the next section, I review neuroimaging evidence that appears to support the evolutionary sequence proposed here.

<table>
<thead>
<tr>
<th>TABLE 12.3. Hypothetical evolutionary sequence of social displays (from Whitehead 2003)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Communication</strong></td>
</tr>
<tr>
<td><em>Implicit</em></td>
</tr>
<tr>
<td><em>Mimetic</em></td>
</tr>
<tr>
<td><em>Conventional</em></td>
</tr>
</tbody>
</table>
12.5 The culture-ready brain

With a few exceptions, social displays have been relatively neglected in neuroscience. For example, following the discovery of mirror neurones in grasping cortex, there have been more than forty brain imaging studies of grasping, hand-object manipulation, and tool use (Grèzes and Decety 2001). In contrast, there have been only four such studies of dance (Calvo-Merino et al. 2005, 2006; Cross et al. 2006; Brown et al. 2006), three of pretence (German et al. 2004; Whitehead et al. 2009; Smith et al. 2013), and only one of role-play (Whitehead 2003). Nevertheless, what we have learned about the brain basis of social interaction and displays is highly suggestive (for fuller detail and references see Whitehead 2010b, 2012).

12.5.1 Body movement and empathy

Figure 12.1 shows the presumed human motor mirroring system, based on the many studies of tool-use, object manipulation, etc. The figure shows the right hemisphere only. Praxic actions such as tool-use activate this system mainly in the left hemisphere, whereas expressive gestures recruit more bilateral activity. Two main areas are involved. The superior parietal lobule is a navigational area that maps space in body-centred coordinates, while the left frontal operculum includes Broca’s area, classically associated with motor sequencing for speech. However, the operculum is clearly involved in all kinds of motor actions including movements of all body parts. Just below this area—in the cleft dividing the frontal from the temporal lobe—lies the insula. The anterior insula and anterior cingulate cortex have been associated with a mirror system for facial expressions and empathy (review: Rizzolatti et al. 2006).

12.5.2 Dance

The mirror system for body movement, as one would expect, is activated by dance. Dance also activates lateral and medial areas in the parietal and temporal lobes (Figure 12.2). The inferior parietal lobule is a multimodal integration area in a pivotal position—surrounded on all sides by visual, auditory, somatosensory, and motor cortices. The temporal pole is likewise a multimodal area.

Curiously, auditory cortex is activated even when people are watching silent videos of dance—perhaps reflecting imagined music. This is the only primary sensory cortex which is expanded in human relative to chimp brains (Deacon 1992a).

12.5.3 Projective pretence

Studies of observing pretence showed activity in the main areas associated with dance, plus orbito-medial prefrontal cortices and the superior temporal sulcus (Figure 12.3). The medial prefrontal cortices include the anterior cingulate (associated with a mirror system for facial expressions) and other areas of major social
Superior parietal lobule
Prefrontal operculum
Superior

Figure 12.1 Right cerebral hemisphere showing motor mirroring areas.

Inferior parietal lobule
Posterior cingulate gyrus
Precuneus

Superior temporal auditory cortex
Temporal pole

Figure 12.2 Major cortical areas in which activation loci were associated with dance.

Temperoparietal area
Ventromedial prefrontal cortex

Superior temporal sulcus
Orbital prefrontal cortex

Figure 12.3 Major cortical areas in which activation loci were associated with projective pretence. (Source: Whitehead 2003)
importance. Note that four major language areas—Broca’s area, Wernicke’s area, the temporal pole, and the superior temporal sulcus—are all activated by pretend play. Three of those areas are also activated by dance, whereas only one is associated with body movements more generally.

12.5.4 Introjective pretence
The same expansive trend is apparent in studies of introjective pretence (role-play) and narrative (mental role-play) (Figure 12.4), implicating all previously mentioned areas with the addition of dorsolateral prefrontal cortex (Whitehead 2003; Mar 2004). Numerous authors assume functional continuity between role-play, narrative, and daydreaming (Mar 2004)—which are all ‘story-telling’ activities of one kind or another.

12.5.5 Discussion
It has not been established that all the cortices shown in these figures include mirror neurones, though one would expect them to be fundamentally important for social displays. In the light of Arbib’s (2002) suggestion that human brain enlargement involved the expansion and reduplication of mirror systems, with new systems then adapting to new functions, the figures do suggest a build-up of display systems, with more recently evolved displays ‘piggy-backing’ on and extending beyond older systems. The overlaps between dance, pretend play, and classical speech areas, might indicate that earlier forms of display likewise pre-adapted the brain for language.

The data further supports a link between displays and social awareness. The mirror system for facial expressions has obvious implications for empathy. The temporo-parietal areas and temporal poles—common to dance and pretence—together with

---

**Figure 12.4** Major cortical areas in which activation loci were associated with narrative and role-play.
the superior temporal sulcus and anterior prefrontal cortices—activated by pretence—have been identified as a ‘theory of mind’ network. Likewise, brain areas associated with role-play and narrative coincide with the so-called ‘default network’. This is a set of regions which have regularly been reported as more active during baseline ‘rest’ periods than during periods when volunteers have been engaged in laboratory tasks designed to investigate cognitive functions. ‘Default activity’ has often been interpreted as daydreaming—that is, ‘theatre of mind’—telling oneself stories or imagining social scenarios with two or more dramatis personae who appear to have minds of their own.

In sum, imaging research suggests that there may be up to six distinct mirror systems in modern human brains—each of which evolved around and extended beyond phylogenetically older ones. The first two systems, which we share with monkeys, would be those for reading body actions and affective signals. A possible third would be a song-and-dance system, a fourth a mimetic and pretend play system, a fifth a role-play system, and a sixth a language system.

12.6 Conclusion

I set out to show that the preconditions for the emergence of language are both social and uniquely human. Invoking social mirror theory, I argued that there are profound discontinuities between human and animal behaviour, social awareness, and culture. At the heart of these discontinuities lies our formidable armamentarium of social displays, which sustain our high levels of social awareness, are preconditions for the emergence of language, and are virtually constitutive of modern human culture. Our three types and three modes of display do have animal parallels, but we are unique in having all nine, and in having multiple displays in all nine—most of which are more sophisticated than any animal equivalents, and many of which (including language) are uniquely human.

The evidence I have presented is at least consistent with social mirror theory and its corollary, the play and display hypothesis. There is strong evidence of a spiral relationship between social displays and intersubjectivity during human development, and I have suggested that a similar spiral is likely to have occurred during hominin evolution. Finally, the presumed evolutionary sequence of displays is supported by brain imaging research.