
ANDREA ZHOK

Università degli Studi di Milano

andrea.zhok@unimi.it

A PHENOMENOLOGICAL DISCUSSION OF EARLY IMITATION

abstract

The following pages are devoted to an attempt to examine Andrew Meltzoff's discoveries concerning neonatal imitation in the light of Husserl's discussion of Fremderfahrung. We criticise Meltzoff's explanatory model AIM (Active Intermodal Mapping), which is introduced to account for his empirical findings, for two main reasons. First, the AIM model does not seem to properly reconcile the vindication of the intermodal character of imitation with the idea that early imitation is based on organ identification: these two claims seem to be reconcilable only at the cost of sacrificing the active, non reflex-like character of imitation. Secondly, the account of AIM does not fit in with the ordinary first-person experience of adult imitation. In its stead we propose a different explanatory approach, which is consistent with a basic phenomenology of imitation and does not depend on organ identification, but on the "rhythmic resonance" of gestures.

keywords

Husserl, Meltzoff, imitation, cross-modal identity, rhythm

The question of intersubjectivity in Husserl's thought is one of the most important and controversial issues in phenomenological debate. At the same time, the question of intersubjectivity, with special reference to empathy and imitation, has been the object of cognitive science analyses, some of which have also received the attention of the general public. Our aim in the following pages will be to provide a discussion of a well-known scientific finding in the light of, and in sight of, a phenomenological approach to the question of intersubjectivity. There are remarkable methodological difficulties in allowing the naturalism embraced by cognitive sciences and the fierce antinaturalism of Husserlian phenomenology to interact. We will not try to properly settle here this intricate question, but we just intend the following analysis as a tentative exemplification of a phenomenological use of experimental results in psychology. More specifically, we will briefly re-interpret Andrew Meltzoff's experimental discoveries concerning neonatal imitation by providing them with a phenomenologically inspired interpretive framework, alternative to the one proposed by Meltzoff.

Husserl devotes great attention and thousands of manuscript pages to the problem of the foundation of intersubjectivity, and we cannot try to provide here any plausible resume of such formidable analysis. We will limit ourselves to recall the special role played by the question of *Fremderfahrung* (the experience of the Other). In one of the crucial theoretical steps of his analysis, Husserl tries to explore our primal intuition of the Other, understood as an *alter ego*, an other-subject-like-me. The leading question is: how can we recognise the Other, so that we can discern at the same time the Other's extraneousness and the Other's identity with the Ego that we ourselves are? According to Husserl we cannot have an immediate *primordial* access to the *alter ego*, because this would not allow for our intuition of the Other's alterity: we could not distinguish the Other from ourselves. Husserl tries to explain our intuition of the Other by resorting to the idea of an *apperception* supported by an *analogy* between our body and the Other's one (Hua I, 140-141). Apperception in Husserlian terms is an *association* which pre-delineates the completing traits of the perceived entity. Apperception is an association of *essential* and not merely psychological character. This means, among other things, that the "analogy" between our body and the Other's body cannot be traced back to a contingent external "resemblance". Husserl does claim

that there is a kind of resemblance at the roots of the constitution of the “pairing” (*Paarung*) that supports our identification of the Other as such, but the nature of such resemblance is clearer as to what it is not, than with regard to its positive traits. It is commonly acknowledged that Husserl’s treatment of the issue of *Fremderfahrung* remains partially unfulfilled. Our present purpose is to see if it is possible to draw from known empirical investigations some suggestions relevant to that Husserlian problem, while preserving the methodological constraints that phenomenology requires.

Over the last thirty years Andrew Meltzoff and colleagues have produced a series of pathbreaking studies, which have deeply influenced our understanding of children’s imitation. Meltzoff’s experiments have shown that a sort of imitative process, which we will call *proto-imitation*, can be found even in neonates few hours after birth; such imitation concerns also gestures, like tongue protrusion, which do not allow any visual coupling of the gesture to be imitated with the imitating one (in the absence of mirrors).

The most crucial experiments devised by Meltzoff *et al.* (1977; 1979; 1983) show what follows: after the display of some gestures (tongue protrusion, lips protrusion, mouth opening, and, to older infants, finger movements and side movements of the head) a *majority* of newborns were showing a *greater* amount of repetitions of the displayed gesture than of other gestures. For instance, in an experiment where the experimenter alternately performed sessions of tongue protrusions (TP) and mouth openings (MO), the infants were performing more often MO than TP in the MO session (on average 7.1 vs. 5.4), and they were more often performing TP than MO in the TP session (9.9 vs. 6.5) (Meltzoff and Moore 1983, 705). Also the duration of the infant’s gestures was correspondingly greater for the displayed ones than for the others (*ibid.* 706). These results may appear weak, because of their merely statistical consistency, but thenceforth they have been repeated many times, by different researchers and with further methodological caveats. It is important to note that the displayed gestures (TP and MO) belong to the ones that newborn infants *spontaneously* perform. Therefore, the relevant imitation does *not* produce the institution *ex novo* of an unfamiliar gesture after an exemplification. What happens is rather conceivable as the summoning of a *propriocepted* gesture by means of a *visually perceived* gesture. These experimental results must be qualified by the following limitations:

1. Most infants are always producing *some* (even wholly incongruous) responses and, in a minority of cases, even responses

akin to the alternative exemplification (e.g., TP responses to MO examples).

2. The responses of the newborns are not immediately emerging in an accomplished fashion, but are gradually approximating a clearer and clearer replication over the course of the experiment.

3. Many infants (more than 60%) do not respond at all or interrupt their responsive involvement, because of greater urges, in the course of the experiment.

As Meltzoff notices, the looseness in responding (1 and 3) and the apparent dependence on learning (2) are incompatible with responses governed by fixed action patterns or reflexes. Responses do not appear to be “ballistically” triggered from appropriate stimuli. Unlike reflexes like the palmar grasp reflex and the sucking one, the working of proto-imitation rather seems to show the character of a primitive *response of “sociability”*, requiring a relaxed situation and an *exploratory attitude*.

At this point Meltzoff tries to provide a causal explanatory account of the mentioned experimental outcomes. He does so by resorting to an explanatory model called *Active Intermodal Mapping (AIM)*, whose general traits are the following:

[I]mitation is a matching-to-target process. The active nature of the matching process is captured by the proprioceptive feedback loop. The loop allows infants' motor performance to be evaluated against the seen target and serves as a basis for correction. According to this view, the perceived and produced human acts are coded within a common (supramodal) framework which enables infants to detect equivalences between their own acts and ones they see. (Meltzoff and Moore 1997, 180)

These general traits of the *AIM* seem to be required by the descriptive basis of the experiment: the idea that proto-imitation is a “matching-to-target” process is supported by the fact that apparently proto-imitative acts progressively approximate to the exemplified ones. And the supramodal dimension to which proto-imitative acts must refer is required by the fact that apparently imitation can take place between gestures belonging to different sense modalities (e.g. vision and proprioception). As we will see, this part of Meltzoff's proposal is consistent with phenomenological descriptions. But Meltzoff goes beyond this explanatory level. He tries to discern the “core mechanism” of imitation by postulating a system of

“organ identification” active at birth. The idea is that the newborn is able to identify her own body parts as “corresponding” to the Other’s parts and limbs. This idea is introduced by Meltzoff in order to explain both the general transmodal ability to imitate and the appearance that babies sometimes start the imitating process just by activating the relevant body part (e.g. unspecific activation of the tongue in imitation of TP) (Meltzoff and Moore 1997, 183). And how is supposed such organ identification to take place? Meltzoff formulates two hypotheses. The first one is that organs are innately identified on the basis of their *form*, which we are evolutionarily predisposed to recognise. This hypothesis is suggested by the apparently analogous discovery that neonates are innately sensitive to human faces, that is, to general face-like configurations. This option is the one that Meltzoff prefers and is also an interpretation that has been later endorsed by Shaun Gallagher (2005) in support of his idea of “body schema”. An alternative option mentioned by Meltzoff is that organs could be identified through their unique spatiotemporal pattern of movement (“kinetic signature”): the ways in which tongues, fingers or arms respectively move have specific kinematic constraints that may be sufficient to identify each organ.

At this point, we have to observe in passing that the explanatory proposal that Meltzoff embraces (the first hypothesis) departs from any possible consistency with phenomenological analyses. The reason for this departure is the naturalistic assumption that in the last instance explanations must take the form of a reduction to spatiotemporal causes. Meltzoff rightly remarks that imitation must take place at a supramodal level (that is, at a level common to different sensuous modalities), but then he takes for granted that the roots of both modal and supramodal levels must be causally traced back to events in physical space (brain mechanisms). Husserlian phenomenology would object that spatial determinations as such are *not modally innocent*, since they are *constituted* with reference to specific modal acts (visual and tactile kinestheses). Insofar as we assume that spatiality and causality are beyond the scope of experiential constitution, we just enrol in the mainstream naturalistic club.

But this observation is wholly internal to the phenomenological horizon and is destined to sound unconvincing to naturalistically oriented ears. Let us therefore try to show the limits of Meltzoff’s explanatory proposal from within. Let us take his first option: here imitation as a matching-to-target should be guided by organ identification, which in his view implies the identification of *forms* and *relations between forms*. He calls such *innately*

cued responses “organ relation (OR) end states”: they are the propriocepted innate responses that correlate with visual configurations like tongue-to-lips, tongue-between lips, etc. (Meltzoff and Moore 1997, 184). Here however we have a problem: how can the identification of seen and propriocepted *organs* (or OR end states) be conceived if they are all to be considered “forms”? Such “forms” cannot be *visual* (spatial) forms, as the mentioned case of the neonatal recognition of faces might misleadingly suggest. But if they are not forms in a visual sense, how else should we understand their formal nature? If we look for *supramodal* roots of protoimitation, the model of the inborn sensitivity to facial traits is inapplicable. Such forms cannot unilaterally belong to any single *modal* dimension.

But could we not just say that some visual forms are simply able to prompt appropriate propriocepted responses? Could we not just suppose that unknown evolutionary developments led our organism to produce appropriate couplings between different modal percepts like seen OR end states and propriocepted OR end states, and that there is nothing more to ask? Of course we could, but it is not what we need either to account for Meltzoff’s findings or to account for the mature forms of imitation that we experience in the first person. Inborn immediate coupling of responses is pertinent to physiological reflexes, but could not support the *active* and *generative* character of imitation. There is no doubt, as Husserl has often argued, that we rely on countless passive responses that are not in the power of our conscious activity, but ordinary imitation is an activity guided by an attempt to approximate a target, and *such target cannot be the particular sensuous exemplification that we have in front of our eyes*. It must be a supramodal or transmodal type. Strangely enough, although Meltzoff is perfectly aware of the necessity to resort to a dimension of active representation in order to account for imitation, he does not see that an innate coupling of organs (or OR end states) would never provide such a dimension: even if the visual appearance of eyes, tongue or lips would elicit immediate activation of eyes, tongue or lips, this could not account for the imitation of gestures. Inborn organ identification of a kind may be an initial cue that eases early imitation, but cannot explain imitation and much the less can support the recognition of Otherness. In fact, it is precisely the focus on organ identification that seems to be misplaced. While it is well possible that some reflex responses contribute to an early bodily orientation in *some* newborns (recall that the majority of babies subjected to the experiments do not provide results, either positive or negative), immediate matching of seen and felt organs (or organ relations) cannot support the imitation of gestures.

Meltzoff's idea seems to be that a series of identifications of relative position between organs (OR end states) could work somehow like dots to be connected by pencil: if I proprioceptively know what corresponds to visual tongue-in-mouth and tongue-between-lips, then I could perform imitation as drawing a line between such end points. But this idea, in the absence of essential supplements, would be clearly inadequate: if the relevant "dots" are reflex responses, they *do not take place in any unitary representational space*, as dots on paper would do. This means that we have no guide telling us how to "fill the gap" between reflex responses (which, by assumption, display correlations between visual events and propriocepted ones), because such responses just do not belong to a common representational space: in the experiential "content" of my eye-blinking and of my knee-jerk respectively there is nothing that posits them in specific *spatial* relations, which can be attributed only by reflecting on my body image. Meltzoff, it must be noted, is apparently aware of this problem and thinks that such "connective tissue" between innately cued spots could be originally provided by "body babbling", by which he means the spontaneous repetitive motion of limbs and facial organs that babies produce even before birth. Body babbling should enable the subject to learn how to connect the propriocepted innate responses ("OR end states") with each other and with spontaneous muscular activity so that a kind of body map obtains. This reasonable hypothesis, however, shifts the weight of the argument from the role of innate responses to that of the proprioception of spontaneous muscular activity ("kinaestheses", Husserl would say).

Indeed, if the problem is the one of the transition from the perception of the Other's body to the proprioceptive awareness of my body, the idea of an innate mechanism producing a static mapping of visual positions to propriocepted positions does not do. This model would give priority to the replication of static configurations, which in Meltzoff's own experiments appear to be the exception, rather than the norm. But this model appears to be especially awkward if we consider our first-person experience of imitation, which is after all the ground from which we draw the intuition that neonates may be producing proto-imitative acts.

What happens when we, as adults, try to imitate a facial expression? Except for professionals who may train in front of a mirror, the majority of people are able to immediately grasp a sort of "overall style" of the expression and to reproduce it, even if they do not have any idea about how their face looks. Actually, if our imitative attempts are shown to us (they were filmed, say) we are often baffled by how our face looks. The same could be seen in the imitation of dance steps: uneasiness aside, we can easily imitate the general style of the dance, the rhythm of moves, their lightness or tension, etc. without

any specific awareness of the precise position of our feet, of the lines that we are following, etc. True enough, this primary level of imitation can lead to outcomes sadly distant from the model, as any amateur dancer can recognise with the merciless help of a mirror; nevertheless something essential *has been grasped* and can be *recognised* by others. It is important to observe that this stylistic or rhythmic factor can be recognised and transposed even *in the absence of any organ identification*. We can imitate even the joyful attitude of a tail-wagging dog, although no obvious matching of organs is available. We can imitate human, animal and even inanimate motions with limbs that do not belong to the model to be imitated. And our tentative dancing movements can spontaneously coordinate with musical features in an instance of supramodal “imitation”. It seems that there is an imbalance between the spontaneity with which mature imitation can relate to the rhythm/style of perceived motion and the difficulty in picturing to ourselves where our bodily parts and what their spatial relations are. In the following this “rhythmic” or “stylistic” dimension will be named *rythmòs*, after the ancient Greek term, which has an appropriately comprehensive meaning including ordinary rhythm, proportion and style of patterns; (for an extensive discussion on *rythmòs* see Zhok 2012, 123 et seq.).

Coming back to Meltzoff’s hypotheses, we should ask whether the second hypothesis that he formulates, without subscribing to it, is more satisfactory. This option implies that organ identification would be obtained by detecting not *forms* but “*kinetic signatures*”. This idea relies on the discovery of the peculiar human sensitivity to biological motions, whose multifarious features are recognised with spontaneity and surprising swiftness (Johansson 1973). This idea, however, does not seem to be adequately exploited by Meltzoff, who is pursuing organ identification as a necessary step towards early imitation. Indeed, it is conceivable that the “kinetic signature” of some gestures is characteristic enough to lead to the supramodal recognition of specific body parts. This is especially possible if we think that newborn infants have a limited repertoire of propriocepted gestures in their command and therefore do not need to perform many refined distinctions in order to find correspondences between seen and propriocepted gestures. The pre- and post-natal body babbling involves few stereotypies and this means that the child does not need a high level of discrimination in order to discern, for instance, MO from TP. Two problems, however, remain: first, this would not amount to *organ* identification, but to *gesture* identification (with incidental identification of the relevant organs), and second, if the correspondence is still read as an immediate elicitation of specific organ activation, then we remain closer to

reflex responses rather than to active imitation. Nevertheless, the analysis of early imitation in terms of a kinetic correspondence between gestures opens up a new and more promising view on imitation and its tie to empathy.

It must be noted that the “kinetic identity” between perceived and propriocepted gesture must not properly concern movement as it is ordinarily understood. Although in the third person (i.e., for objective representation) the gestures that I perform and experience in the first person can be described as motions (i.e. changes of position over time), they are not propriocepted by me as something taking place in objective space and time. This means that the propriocepted gesture does not have ordinary “motion” in common with the perceived gesture. But then what do they have in common? We must take seriously the demand that the dimension where imitative matching can take place be *supramodal*. The representational space where motions are placed is inescapably tied to *visual* aspects (motions are represented as events in an ideal visual field), whereas our disposition to “feel in resonance” with seen gestures does not belong to such representational space: we can spontaneously perceive the “similarity” between the lowering pitch of a whistle and the approaching trajectory of a falling object. Such similarity is no exterior congruence of sensuous traits. Seen motions, propriocepted muscular activations and unfolding sounds can have in common a “way of changing”, an instance of *rythmòs*, to which we must not attribute any particular sensuous aspect: it does not occupy either a visual space or a tactile space or any other sensuously qualified dimension, since its “substance” is rather a *temporal form*. It must be noted that a truly supramodal dimension is necessarily superindividual: it must be possibly valid for an indefinite plurality of sensuous instantiations. Furthermore, it should be stressed that *rythmòs* must also be a *motivating* dimension, in the sense that instances of *rythmòs* are apparently *salient* for the newborn perceiver, whose corresponding gestures are called forth.

A notion like a supramodal motivating *rythmòs* can account for the early imitative response better than any *organ identification*, which can be either derived from the correspondence between instances of *rythmòs* (as “kinetic signatures”), or attributed to inborn reflex-like cues inessential to the imitative process. From this point of view some of the results of Meltzoff’s experiments become more easily interpretable.

As we said above, many newborn babies did provide apparent responses to the exemplified gestures, even if they were the “wrong” answers. But, if we take organ identification to be explanatorily prior and dynamic gesture

correspondence to be a secondary achievement derived from the former, then responses that get the pertinent organs wrong (TP instead of MO) are not exactly what one would expect. On the contrary, if we reckon that at first babies respond to instances of *rythmòs* as such (primarily conceivable as biological motion), we can account for the “statistic” trend toward proper matching of gestures. The majority of babies respond just to *instances of rythmòs*, and the more the corresponding process unfolds, the better some of them refine the dynamic pattern of their responses, down to the appearance of imitation proper.

The character of motivating salience of the instances of *rythmòs* accounts also much better for the paradoxical nature of imitation and empathy, such that we experience ourselves at the same time as identical with and different from the Other. What is proximately felt, what is “mine”, is *rythmòs*, whereas its contingent modal realisations are irreducibly “other”, extraneous. What is in common is what enables communication, primarily emotional communication through the style of gestures (softness, regularity, abruptness, violence, etc. are immediately detectable) (see Stern 2000, 48 et seq.). What remains ineradicably different is the specific sensuous implementation of the gestures as well as the relevant character of reactivity or spontaneity of each implementation.

Finally, the notion of *rythmòs* is not a mere hypothesis consistent with, but superimposed to, phenomenological observations: *rythmòs* is intuitively available in a plurality of phenomena which go well beyond (early or mature) imitation. The idea that *imitation* could be the primary access leading to empathy with the Other and her “state of minds” seems now somewhat misleading. Our primal access to the Other does not go either through direct coupling to the Other’s body parts, or through direct access to the Other’s feelings (in this case we would *be* the Other). We have direct access to the Other’s *expressions* of feeling, which are embodied in specific instances of *rythmòs*. In other terms, we need not think that we first produce a more or less faithful copy of the Other’s behaviour, on the basis of a structural correspondence of bodies, and then gain a mediate access to the meaning of the Other’s gestures. *Rythmòs* in the gestures is primarily evident and is expressively intelligible even in the absence of the recognition of specific body correspondences: this is the reason why we need no mediate introduction to grasp the expressive motions of animated objects in cartoons (brooms, cars, haunted houses, etc.). The way of moving is enough to define the object’s “emotional dispositions” as well as its targets and spatial orientation, even in the absence of any clear

reference to eyes or limbs. This perspective could be critically applied to theoretical views like the one recently proposed by Gallagher (2005), where a body schema endowed with the same structure of the average body image is used as explanatory key for cross-modal and intersubjective relations (see Zhok 2012, 202 et seq.).

REFERENCES

- Gallagher, S. (2005), *How the Body Shapes the Mind*, Oxford: Oxford University Press, 2005;
- Husserl, E. (1973), *Cartesianische Meditationen und Pariser Vorträge*, Husserliana I, S. Strasser (ed.), Martinus Nijhoff, Den Haag (quoted as *Hua I*), 1973;
- Johansson, G. (1973), "Visual perception of biological motion and a model for its analysis", *Perception and Psychophysics*, 14, 1973, pp. 201-211;
- Meltzoff, A. N. and Moore, M. K. (1977), "Imitation of Facial and Manual Gestures by Human Neonates", *Science*, 198, 1977, pp. 75-78;
- Meltzoff, A.N. and Borton R. W. (1979), "Intermodal Matching by Human Neonates", *Nature*, 282, n. 5737, 1979, pp. 403-404;
- Meltzoff, A.N. and Moore M.K. (1983), "Newborn Infants Imitate Adult Facial Gestures", *Child Development*, 54, 1983, pp. 702-709;
- Meltzoff, A.N. and Moore M.K. (1997), "Explaining Facial Imitation: A Model", *Early Development and Parenting*, 6, 1997, pp. 179-192;
- Stern, D. (2000), *The Interpersonal World of the Infant. A View from Psychoanalysis and Developmental Psychology*, New York: Basic Books, 2000;
- Zhok, A. (2012), *La realtà e i suoi sensi. La costituzione fenomenologica della percezione e l'orizzonte del naturalismo*, Pisa: ETS, 2012.