Accelerating semogenesis
an ecosocial approach to photography
Johannessen, Christian Mosbæk; Boeriis, Morten

Published in:
Visual Communication

DOI:
10.1177/1470357219887769

Publication date:
2019

Document version
Accepted manuscript

Citation for published version (APA):

Terms of use
This work is brought to you by the University of Southern Denmark through the SDU Research Portal. Unless otherwise specified it has been shared according to the terms for self-archiving. If no other license is stated, these terms apply:
• You may download this work for personal use only.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying this open access version

If you believe that this document breaches copyright please contact us providing details and we will investigate your claim. Please direct all enquiries to puresupport@bib.sdu.dk
Accelerating semogenesis.
An ecosocial approach to photography

Abstract

Michael Halliday and Christian Matthiessen’s term “semogenesis” refers to how meaning potentials are created through processes on many co-occurring time frames, most prominently those referred to as “phylogenesis”, “ontogenesis”, and “logogenesis”. The concept was originally infused with linguistic concern in an attempt to link an SFL account of the lexico-grammatical and semantic strata with lived experience. In this article, we rethink the concept in order to (1) broaden its scope to the concerns of multimodal studies, and (2) accommodate how digital technology impacts on our communication practices. We do so by discussing semogenesis from the vantage point of Lemke (1995) and Thibault’s (2004) “ecosocial semiotics”, a perspective that asks us to blend both sociological, technological, material and biological understandings of human activity. Taking digital photography as an example, we argue that digital media afford an acceleration of processes of multimodal semogenesis on all semogenetic time frames. Picking up the notion of “microgenesis”, a fourth, faster-than-logogenesis time frame, that serves as a placeholder for any process enabling logogenesis, we suggest that this acceleration is driven by the global scale introduction of digital technology. Through a discussion of select examples from the history of photography, specifically contrasting nascent photographic practice with contemporary photography, we propose that the development from camera to digital camera and the subsequent consolidation in recent decades of digital cameras into smartphones has had a profound impact, not only on practices of photography, but also on the processes of meaning making with photographic material.
1 Introduction

Consider this example of contemporary photographic communication: After a conference on multimodal semiotics, the authors of this article visited Cape Town's Table Mountain. On the panorama platform near the top, we decided to post a selfie with the panoramic view of Cape Town in the background. We posed, arm in arm, and one of us held out his phone, oriented horizontally, to accommodate both of us in the frame. He struggled to fit both our faces and the view in the picture. The other then suggested he held the phone up higher while, at the same time, indicating the direction with his hand. Because we each only had one free hand, one of us held the camera while the other activated the phone's shutter release button (see figure 1). A brief moment later, the picture appeared on the screen but, thinking we looked silly, we decided to take another picture. This was repeated a few times. Satisfied, we wrote a humoristic caption, added our location and tags, and posted the image to Facebook. Thirty seconds later comments started coming in from family and colleagues telling us the shot looked good and to enjoy our trip.

Although this event did in fact take place, we deliberately enacted it so we could frame the argument in this article and provide a point of reference for our discussion of factors that are relevant for understanding processes of change in the contemporary semiotic system of photography.
Figure 1. Line tracings of still images from a video shot during the selfie event on Cape Town's Table Mountain in December 2016. In the first frame an agreement is made on which section of the panorama to include. In the second frame, arm in arm, the men collaborate to get the framing right. Man One holds the camera phone in one outstretched arm, Man Two instructs him to hold it higher. In the third frame, the two men coordinate to take the picture.

2 Background and motivation

The concept of “semogenesis” (Halliday and Matthiessen, 1999; Martin, 1999; Martin and Dreyfus, 2015) refers to how meaning potentials are created through processes on many different simultaneous time scales. Standard accounts of semogenesis enumerate three such processes (Halliday and Matthiessen, 1999, p. 17): The phylogenetic, ontogenetic, and logogenetic scales. On Halliday and Matthiessen’s account, phylogensis describes how semiosis slowly evolves on a population or community scale. The ontogenetic time frame describes how semiotic skill develops and grows over an individual’s lifespan. Finally, the logogenetic time frame describes how semiosis unfolds between people in particular semiotic events. On such a view, the Table Mountain example can be regarded as a logogenetic event of unfolding photographic communication. It is underpinned by whatever level of photographic skill we had on that day, and the development of that skill over time is an ontogenetic process. And the continuing societal development of the photographic conventions we attend to, self-portraits, or selfies, is a phylogenetic process.

When Halliday and Matthiessen (1999) outlined semogenesis in Construing Meaning Through Experience, their aim was to demonstrate how the lexico-grammatical and semantic strata in an SFL account of language link to lived experience. Thus, their agenda was driven by linguistic concerns, and, presumably, they did not need to describe faster-than-logogenetic scales because human bodily and cognitive capabilities (as they were then understood to relate to language) had remained relatively stable over thousands of years and could be taken as a given.

We believe the concept of semogenesis has a lot to offer multimodal studies, especially as a framework for understanding the dynamics of conventionalization processes. However, in order to do so, it must be expanded to encompass current multimodal and digital practices of communication like smartphone driven photography and the networked image (Van House et. al., 2005; Rubinstein & Sluis, 2008). This requires us to think carefully about how digital technologies considered as tools extend and change our bodily and cognitive capabilities, as exemplified by Andy Clark and David Chalmers’ (1998) “The Extended Mind”, and also appreciate how a digitally extended person now has capabilities that were difficult to predict only a couple of decades ago (Gómez Cruz and Meyer, 2012). One has to consider a “post-human perspective” (Braidotti, 2013) where the body
extends beyond the flesh into a “larger consciousness encompassing inorganic materiality” (Blaagaard, 2013, p. 371), which is connected both organically and technologically to other beings (Hayles, 1999).

Thus, our aim with this article is to rethink the concept of semogenesis in a way that includes the extremely fast time scales on which digital systems work and examine whether, at least in theory, these may in fact accelerate the dynamics on slower logogenetic, ontogenetic, and phylogenetic scales. Or, more simply put, we aim to develop a framework that helps us think about how technology accelerates processes of individual development as well as changes in semiotic conventions. Our inspiration for doing this comes from Jay Lemke’s (1995) and Paul Thibault’s (2004b) writing on *ecosocial semiotics*, a perspective on meaning making that attempts to bring social semiotic and materialist views together. More specifically, we suggest Thibault’s (2015) notion of “microgenesis” as a fourth semogenic process, a placeholder for events that are faster than those usually considered under the label of “unfolding discourse”.

The conceptual basis for our argument will be the topic for section 3, specifically presenting Jay Lemke’s notion of ecosocial semiotics and James Gibson’s ecological psychology (Gibson, 1968, 1986 [1979]). In section 3.1 we discuss Halliday and Matthiessen’s (1999) model of semogenesis and suggest that adding a fourth, faster scale of microgenesis would be productive for using the model to think about digital photography. We then suggest a framework of three types of affordances (section 3.2) and three types of functional coupling loops (section 3.3) that can help us understand the accelerating semogenesis of photography from an ecosocial perspective. In section 4 we apply this framework in a discussion of photography as an ecosocial phenomenon, showing how advances in photographic technology have changed microgenetic affordances for logogenesis, and arguing that this accelerates semogenic processes going all the way up the scales to sociocultural evolution. The final section discusses implications for how we might consider learning to take pictures as either a process of enculturation or enskillment.

3 Conceptual basis: Ecosocial systems

In a 2015 interview, Jay Lemke said the following about future trends in social semiotics:

“One of them certainly is the move towards integrating a semiotic approach with an embodied approach, taking up the implications of a basically materialist model of communication and language, and taking them a step further to talk about what it means to have animate bodies, moving, touching, interacting, doing, while, speaking, gesturing, drawing and writing and so forth” (Lemke, 2015, p. 138, our brackets and italics).
This, we feel, is a fair description of the direction we wish to take in our thinking about photography. We have a background in multimodal social semiotics, but wish to move beyond describing semiotic conventions, as they are instantiated in multimodal texts, and think about the dynamics of how such conventions emerge and remain stable or change and eventually disappear. We believe that multimodal communication studies and cognitive sciences could benefit from closer proximity, and therefore we want to present our argument in a way that makes it compatible with current writing about embodied and extended cognition (for an overview see Menary, 2010).

This brings us to why we call our efforts ‘ecosocial’. The term was coined by Jay Lemke to describe a perspective that combines his background in theoretical physics and science education with his later pursuits in linguistics and social semiotics. “We need to understand how the general principles that govern complex, material, self-organizing systems become further specialized in the case of human communities, where physical activity depends on social meaning” (Lemke, 1995, p. 100). In doing so, he had to work out how to reconcile two very different senses of the word ‘system’ understood as ‘ecosystem’ and ‘semiotic system’:

There cannot be two systems here, changing according to separate laws, relatively independent of one another. There can only be one unitary ecosocial system, material and semiotic, with a single unified dynamics, described under two aspects, by two different sorts of culture specific discourses [science and semiotics]. (Lemke, 1993, p. 251, our bracket)

An “ecosystem” is a reciprocal energy/matter coupling of an organism, or a population of organisms, and its environment. From the perspective of ecosystems, a “semiotic system” is a population-scale pattern in the symbolic behaviour of human organisms. In other words, ecosocial semiosis denotes how humans couple with their environment, which including how they relate to other humans. Of course, thinking of a photographer as an organism misses almost everything that is interesting about photography, but an ecosocial perspective asks us to blend our sociological understanding of photographers as social agents with our biological understanding of humans as a peculiar kind of organism: One that is unique because of its habit of channelling energy that does not come from food through its agency.

As Lemke has predicted, multimodal scholars are becoming increasingly preoccupied with the material and embodied underpinnings of communication practices. We believe that there is fertile common ground to be found between multimodal studies and Gibsonian ecological psychology, which is why we have chosen the Gibsonian tradition as an ecological counterpoint to our social semiotics thinking. One ambition for ecological psychology is to develop a theory of behaviour,
perception, knowledge, and learning that does not set humans fundamentally apart from other animals, which, according to proponents of ecological psychology (Chemero 2011), happens when theories of human cognition posit human-specific, computer-like information processing modules in the brain (exemplified, for example, by the computational psychology of Jerry Fodor’s (1975) *The Language of Thought*). Rather, ecological psychology seeks to explain cognition on the basis of material action-perception couplings with our environment, something all animals have in common. What sets us apart from animals in the ecological psychological account is still very much an open question, but our evolutionary history of modifying action-perception couplings through increasingly complex tool use is one avenue of explanation that is being considered (Ingold, 1993).

Gibsonian terminology has already been introduced to multimodality literature, most notably in Anthony Baldry and Paul Thibault’s (2006) *Multimodal Transcription and Text Analysis*, and both ecological psychology and multimodal studies use the term ‘affordance’ to denote, roughly, the possibilities for action we are offered by qualities of our material environment. By juxtaposing the two we seek to develop the common ground between multimodal studies and post-human, extended cognitive science.

### 3.1 Microgenesis: A fourth, very fast semogenic time frame

What appears in our introduction to be a straightforward case of us using a smartphone to take a selfie on Table Mountain and share it on Facebook can be understood as a complex coupling of material and semiotic systems on many different temporal and spatial scales. For example, closer inspection may reveal that it took several tries, each an event in its own right, to get the desired look in the picture. In turn, each try involved events of grimacing, fumbling and moving the camera around to get the background and angle of the shot right. Ultimately, the grimacing, fumbling and moving around can be analysed as very fast sensorimotor contingencies. Of course, this way of breaking events into smaller events goes down the continuum of temporal scales, but ecosocial events fold into one another up the continuum as well: The individual sharing of a selfie is only an infinitesimal detail in an autobiographical trajectory of learning to communicate with photography or indeed in a decade-spanning global cultural-scale process of digitalization.

From an ecosocial perspective, semogenesis describes the processes by which semiotic parameters in ecosocial systems change over time. Halliday & Matthiessen (1999) distinguish phylogensis, ontogenesis and logogenesis as a heuristic, “[…] a guiding principle in the form of some model of the process by which meaning, and particular meanings, are created” (1999, p. 17). They are fully aware that infinitely many relevant temporal scales can be analytically distinguished,
but these three are “the ones that matter” (1999, p. 17) for their, primarily linguistic, concerns. In their words, going down the hierarchy, slower scales “provide environment for” (1999, p. 18) adjacent faster ones: Processes of phylogenetic evolution provide the environment in which ontogenetic development happens, which in turn provides the environment in which logogenetic exchanges happen. Conversely, going up, smaller scales “provide material for” (1999, p. 18) adjacent larger scales. Halliday and Matthiessen’s use of the word “material” is probably a metaphor, unless one is willing to accept a strong upward causal relation between logogeny, ontogeny and phylogeny.

Therein lies the foundation of our argument. The “material for” logogenetic events of spoken utterances is more literally material. It is the sensorimotor and cognitive contingencies of our nervous systems and vocal tracts. Presumably, this has been a relatively stable principle for however long it has taken for speech to evolve in our species, and the Halliday and Matthiessen account of semogenesis is concerned with the evolution of the language system, not the evolution of the human body or the psychological reality of cognitive processes. However, a strong trend in contemporary cognitive science is to regard tool use as the crux of human cognitive development (e.g. Hollan, Hutchins and Kirsh, 2000; Clark, 2008; Malafouris, 2013). We did not invent tools because we were smart; we became smart because we used tools. On this view, cognitive processes can be said to extend across the barrier of our skin and into our material environment. Along with neurological processes, processes in tools are co-constitutive of cognition and, strictly speaking, we are different as cognitive agents depending on whether we swing baseball bats or look through the viewfinders on cameras. If the thesis of extended cognition (Clark, 2008) stands to reason, the material and cognitive underpinnings of logogenesis have changed quite dramatically as a result of digitalization because the cyclic rate in microprocessors is faster than neuron firings. We argue that the parameters set for logogenesis by faster technology cause the rate of logogenetic events to increase.

In order for our conceptual framework to grasp this, a fourth semogenic time frame is needed, a placeholder for processes that are faster than those generally considered as "unfolding discourse". We take inspiration from Paul Thibault (2015) to call this semohistory microgenesis. It covers processes of "[...] development on a brief present-time scale of a percept, a thought, an object of imagination, or an expression" (Rosenthal, 2004, p. 221, italics in the original). The term was originally coined in 1956 by gestalt psychologist Heinz Werner, in order to provide "a genetic characterization of the structure and temporal dynamics of immediate experience, and, more
generally, of any psychological process” (2004, p. 221). However, even if the scope of the term is generally psychological, we take cues from extended cognition (Clark, 1998, 2008; Clark & Chalmers, 1998; Hollan, Hutchins, & Kirsh, 2000; Malafouris, 2013; Menary, 2010) to include processes and events in our tools and artefacts that are co-constitutive (along with psychological and neurological processes) of communication.

This idea is in fact already anticipated, if not made explicit, in the way multimodal studies has adopted the ecological psychological term *affordance*, which denotes “[...] the idea that different modes offer different potentials for making meaning [...] Modal affordances are connected to both a mode’s material and its social histories” (Jewitt, Bezemer, & O'Halloran, 2016). In other words, different materials afford different logogenesis. And more to the point: Faster materials afford faster logogenesis, which again provides faster material for faster ontogenesis and, ultimately, faster phylogenesis.

### 3.2 Affordances

The term *affordance* was originally coined by ecological psychologist James Gibson in his (1977) article "The Theory of Affordances" to refer to the functional relationship between animal and environment (Gibson, 1986 [1979], p. 127). The Gibsonian concept of affordance encompasses *every aspect* of that relationship, all the way down to the breathability of air and the consequences of gravity for the animal. From the vantage point of the Gibsonian tradition, the social semiotic multimodal concept of “modal affordance” describes a subset of affordances, those that pertain to the semiotic activities of humans. Gibsonian scholar Katherine Loveland (1991, p. 100) considers what the world affords humans that it does not afford other species and suggests three “[…] nonexclusive categories of affordances that are present in and which define the human environment”. Interestingly, her categories cut across the multimodal concept of “modal affordance”:

*Affordances for physical transactions with the environment*: This most general category of affordance is necessary for any kind of animal activity. Basic examples from the animal kingdom include eat-ability, breathe-ability, grasp-ability, lift-ability, push-ability, etc. For our purposes, for example, a camera can be too heavy to hold high overhead or too large for a single hand to reach all the controls. Part of “modal affordance” is its affordance for physical transactions with the materiality of the mode.
Specific, culturally selected affordances: This category reflects culturally preferred (but not strictly necessary) relations between agent and environment. For example, a compact digital camera is heavy enough to afford weigh-down-papers-on-desk-ability, but this is not how most people culturally prefer to use it. Similarly, greaseproof paper affords diffuse-ability when photographers use it as a makeshift softening filter for photo lamps, but its preferred use is for cooking. From the point of view of multimodal studies, we argue that modal affordances mainly belong to this category.

Social affordances. This category reflects the meaning of human activity for other humans (Reed, 1988), including the affordances of talk and writing, facial expressions, gesture, body postures and movements, tone of voice and the direction of gaze. This latter class seems to include what social semiotics refers to as meaning potential; the social affordances of semiotic products.

3.3 Functional coupling
In ecological psychology, the concept of affordance is closely related to the concept of functional coupling (Shaw, 2003). It is through continuous sensing and acting on the environment, both of which requires movement, that an animal directly picks up on affordances, for example whether a surface is stable enough to support its weight. The environment is dynamic, and it pushes back when the animal acts on it. In the jargon of ecological psychology, animals and environments are functionally coupled with one another through action-perception feedback loops. To further our aim our we have developed three heuristic categories of loops:

Sensorimotor coupling loop: The necessary basis for any kind of activity, including photography, is the sensorimotor coupling of human and environment. Anything a person does, whether resting, moving, eating, moving around a scene, or operating the control interface of a camera, is guided by a constant flow of functional information about affordances picked up while acting.

Protext coupling loop: Some, but not all the time, people engage in text production. All acts of text production necessarily begin at one time and end at another. In the interim there is no text yet, only unfinished protext. Whether an unfinished string of written words flowing from a pen, a still amorphous lump of clay in the hands of a sculptor, or indeed, a scene viewed through a camera’s viewfinder, every text producer is functionally coupled with this text-in-becoming, the thing he is experiencing as something he is producing. His actions are guided by information about social
affordances as they appear. This is crucial for our argument, because developments in technology have steadily improved how cameras afford access to protext. Simply put, it makes a difference if you have to wait for minutes or hours to see the effect of a choice you make in the relation between camera and scene.

Social coupling loop: People mostly, but not always, engage in text production out of a desire to interact with other people (Frosh, 2015). Social coupling is perceptually guided by the flow of information about social affordances. A special case of social affordance for a text producer is the reaction of other people to either protext (e.g. the art director’s feedback after a photo shoot) or finished text (e.g. friends’ reactions to holiday pictures).

4. The case of photographic communication

Having laid down our conceptual groundwork we will proceed to discuss the concrete example of taking pictures as an ecosocial phenomenon, which requires particular sensorimotor skill as well as the ability to exploit specific culturally selected and social affordances. As part of our discussion we will relate very generally to the history of photography to show how major technological breakthroughs over 200 years of photographic practice changed parameters for microgenesis.

Changing affordances for sensorimotor coupling with camera at the microgenetic scale

We cannot do justice here to the impact digital technology has had on how people design and interact with semiotic tools, but some of its implications must be briefly considered. In modern cameras digital sensors, microprocessors and memory have replaced moving parts and chemical processes, which makes them much smaller and cheaper to produce than their mechanical and chemical predecessors. They are routinely built into other devices such as mobile phones or smartphones (Cobley and Haeffner, 2009). Apple’s iPhone 6S, which could well have been used in our Table Mountain example, measures only 138.3/67.1/7.1 millimetres and weighs only 143 grams.1 It consolidates many devices; telephone, camera, modem, computer, gps, etc., into one unit, which affords nearly effortless portability.

But not only do smartphones afford portability, they have been accepted into our daily routines to the extent that we might need them at any moment and so keep them on our person all the time. Therefore, cameras are effortlessly available to us, no matter what activity we might

1 http://www.apple.com/dk/iphone-6s/specs/
otherwise engage in. We don’t need to actively think to bring a camera anywhere; it is always simply very closely at hand (Rubinstein & Sluish, 2008, p. 15).

Furthermore, consider also the affordances for grab-ability of most smartphones. They are neither too small nor too big to be grabbed, held and used in one hand, even outstretched from the body or held high overhead (see figures 2 and 3 below) without fatiguing us and, beyond occupying one free hand, they hardly constrain our freedom of movement at all.

These are important affordances for sensorimotor coupling with a smartphone in terms of portability, availability, and grab-ability. In order to fully appreciate how such affordances may impinge on processes of multimodal semogenesis, let us contrast it with a few key innovations in the roughly two century long history of photography.

Very early cameras such as Joseph Niépche’s 1826 heliograph and Louis-Jaques-Mandé Daguerre’s 1838 daguerreotype were based on the design of the portable camera obscura, which had already been in popular use for centuries by dilettante painters as an aid to capture the perspective, proportions and detail of a scene in accordance with the Renaissance ideal of naturalism. (Clarke, 1997, p. 12) In its most basic form, a camera obscura—literally a ‘dark room’—works on the same principle as the vertebrate eye. A small opening in one wall of a dark enclosure lets in an array of light from the outside. Due to light’s rectilinear propagation, the array is inversely projected onto the opposite wall. In mature technological applications of this principle, a simple optical lens is fitted in an opening on one side of a wooden box. The lens projects the scene before it via a mirror onto a semi-opaque ground glass viewing screen on which the artist can place a sheet of paper and trace outlines with a pencil. The principle remains relevant in the design of contemporary digital cameras, but everything else including materials has changed.

---

2 Of course, rhetorically strong as it may seem, a comparison of the 1838 daguerreotype with the facebook-enabled 2016 iPhone 6s, without regard for the many stepping-stones that have lead from one to the other, will be very reductive. Even if one only considers the physical affordances for portability, availability and grab-ability of the camera itself, innovations leading to metal, bakelite, and plastic housings, dry plates and celluloid film, not to mention Polaroid photography, as well as smaller and cheaper optics have continually reduced the price, size and weight of cameras. The reality of technological development is far less neat than we make it seem, but our aim is to provide a general conceptual basis for discussing the impact of technology on semogenic change rather than a faithful historical account of the cultural history of photography.
Consider the portability, availability and grab-ability of the daguerreotype, the world’s first commercially successful camera. Figure 4 illustrates a photographer composing a shot with a
similar dry-plate camera mounted on a wooden telescoping stand. Niépche and Daguerre’s innovation was to allow a copper plate coated with silver-iodide, which turns dark when exposed to light, to be slid in front of the ground glass viewing screen of a camera obscura. This allowed them to substitute an artist’s pencil tracing with a persisting trace of the information in the light array. The original daguerreotype camera consisted of two sliding cedar- and walnut wood boxes, fitted with a glass lens in a brass head, and with a pivoting brass shutter door. Along with the necessary fuming boxes, burners, portable dark room for developing and fixating pictures, as well as vials of necessary chemicals, the setup weighed 50 kilos. (Hannavy, 2013) Even for someone who could afford to buy such equipment, photography was not a casual undertaking. Simply moving the equipment around took considerable effort, and so the sensorimotor base for portability, availability and grab-ability afforded by daguerreotype cameras was vastly different from anything we associate with photography today, as illustrated in figure 4. Note how large and heavy the setup appears to be. Compare it with the 1936 Kodak Box 620 (figure 5), which weighs only 440 grams and is roughly the size of a tea tin, or the 770 gram 2014 Sony A7 mirror less digital camera (figures 6-10). The way in which the latter two cameras afford moving around the scene, lying down on the ground or getting up high on a chair, is very different from the wooden dry-plate camera.

Changing affordances for protext coupling at the microgenetic scale
We must also consider how a camera affords perceptual access to protext. Camera manufacturers have historically deployed two ways of presenting the text-in-becoming to photographers. On the one hand we have eye-level viewfinders, on the other we have viewing screens. The respective protext couplings they afford are different in important aspects.

When a photographer raises an eye-level viewfinder to the eye (figure 8 and 9), the camera changes status, transcending the border between agent and world. It stops being an object in the environment and becomes a temporary appendage to the agent’s head. It goes where his head goes, swivels and moves with the head like eyeglasses or binoculars. Like the blind man’s cane, it becomes an extension of the perceptive system (following Clark, 2008, p. 31), which richly augments the photographer’s sensorimotor activity with information about the protext. We wonder if conventions such as bird’s eye and frog’s eye view and ensuing meaning potential of power and vertical angle (Kress & Van Leeuwen, 1996, p. 146) would have emerged had it not been for what eye-level viewfinders afford photographers.
Conversely, when a photographer engages with protext using a *viewing screen* on the camera (such as the matte plate on a daguerreotype camera, a Rolleiflex Twin-Lens Reflex or the LCD pre-view screen of a digital camera or smart phone), he looks at a *distal object* in the world, which already has important text-like properties. For example, he can move around it and look away from it. A viewing screen is experienced as separate from his own perceptive system unlike a viewfinder.

In figures 2-7 above, the photographer views protext using *viewing screens* of different kinds; a smartphone LCD screen (2,3), a matte screen (4), optical viewing screen (5), and an LED screen on a digital mirror reflex camera (6, 7). In figure 8 and 9, however, he uses an *eye-level viewfinder*.

In order to illustrate differences in how different cameras afford perceptual access to protext, we have produced figures 10-15. They illustrate the viewing mechanisms of 6 historically significant cameras:

<table>
<thead>
<tr>
<th>Figure</th>
<th>Year</th>
<th>Make and model</th>
<th>Viewing mechanism</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>1900</td>
<td>E. Suter Basel</td>
<td>Viewing screen (ground glass back plate)</td>
</tr>
<tr>
<td>11</td>
<td>1931</td>
<td>Leica 1</td>
<td>Eye-level viewfinder (separate from lens)</td>
</tr>
<tr>
<td>12</td>
<td>1936</td>
<td>Kodak Box 620</td>
<td>Optical viewing screen (Separate from lens)</td>
</tr>
<tr>
<td>13</td>
<td>1962</td>
<td>Voightländer Vito Automatic 2</td>
<td>Eye-level viewfinder (separate from lens with framing guides)</td>
</tr>
<tr>
<td>14</td>
<td>1971</td>
<td>Asahi Pentax SP 500 SLR</td>
<td>Eye-level viewfinder (through-the-lens with split-image focus guide and exposure guage)</td>
</tr>
<tr>
<td>15</td>
<td>2014</td>
<td>Sony A7</td>
<td>Eye-level digital viewfinder (through-the-lens)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Viewing screen (through-the-lens)</td>
</tr>
</tbody>
</table>

We have taken pictures *through* the viewing devices of each camera using a setup in which the cameras were placed at the same distance and angle from a scene containing a small figurine on a cloth-covered table. Evidently, the amount of information about protext afforded by different viewing mechanisms varies greatly.

Consider again the workflow of a 19th Century daguerreotype photographer. On this type of camera, the scene was projected onto a ground glass viewing screen in order to afford focus adjustment and composition. As can be seen in figure 10, this gave the photographer a fair measure of perceptual access to the protext as a faint, upside-down image up until the point when the photographic plate was slotted into the camera. Beyond that point, the plate blocked the action-perception coupling of photographer and protext and the photographer could no longer evaluate the effects of his actions on the text-in-becoming. The loop remained broken until the plate had been
exposed, removed from the camera in a dark room, developed and fixated.\(^3\) Only then could the photographer evaluate the finished picture and learn whether his choices had produced an apt communicative artefact. It is noteworthy that, although the image on the daguerrotype back plate was in fact very good, protext differed from finished text by showing colour—the finished text would be in black and white.

Leading up to the present day’s smartphone LCD screen, which shows the photographer in real time precisely what comes through the lens and what the finished text will look like in a text-like way cf. above, perceptual access to protext worsened significantly before improving. Early eye-level viewfinders, such as the separate viewfinder of the Leica 1 (figure 11) and the small optical viewing screen of the inexpensive Kodak Brownie and Box cameras (figure 12) from the 1930ies, were little more than simple aiming devices. The photographer looked through crude mechanisms, separate (and slightly offset) from the lens of the camera, at a very small and poorly resolved protext with different framing and angle than the finished picture. In addition, the viewing screen of the Kodak camera had to be held at approximately 40 cm distance to bring the image on the screen into focus (see also figure 8).

Figures 13, 14 and 15 illustrate an interesting development for our understanding of protext coupling: In recognition of growing demands for reliably being able to anticipate the outcome of a shot, camera manufacturers developed ways to present meta-information (about lighting conditions, aperture settings, shutter speed, film sensitivity, etc.) to the photographer within the eye-level viewfinder. Figure 13 shows the framing guides etched into the glass of the viewfinder of a 1962 Voightländer Vito to suggest how the shot would be framed. Figure 14 illustrates the exposure guide needle and split-screen focus guide in the eye-level viewfinder of a 1971 Asahi Pentax mirror reflex camera. Still, even given these innovations, the protext afforded by the cameras like the Voightländer and Asahi Pentax was a far cry from the final text. There was, for example, no access to how photo-technical choices would impact on exposure and thus on the final product. Accurately anticipating the outcome of a shot still required specialized technical know-how on the part of the photographer.

Digital technology introduced a dramatic improvement of such camera features. When digital sensors and memory replaced mechanical shutter and film, perceptual access to protext displayed

\(^3\) According to British daguerrotype photographer Christopher Brenton West, development takes 1-10 minutes using mercury and 1-3 hours using non-toxic Bequerel technique. Fixation takes 1-2 minutes. (Personal Communication, May 3\(^{rd}\), 2016)
on digital viewing screens and eye-level viewfinders (essentially miniature computer screens) became possible (Rubinstein & Sluis 2008, p. 13). This enables the photographer to see what the camera’s sensor sees and evaluate the image before print or screening (Rubinstein & Sluis 2008, p. 12). Figure 15 illustrates the view through the digital viewfinder of a 2014 Sony α 7 mirrorless digital camera. Note how information about shooting mode, available memory, battery status, as well as shutter speed, aperture and film settings is displayed. The photographer is even presented with real-time statistics about pixel distribution (in histogram format), which is very helpful for composing the picture in terms of contrast and exposure.

Current smart phone microprocessors are powerful enough to calculate, on 30 frames per second, the effect on the picture of any relevant photo-technical choice including post-production, colour grading and digital effects. On a smart phone viewing screen in 2016, protext at any given moment equals the potential finished text. There is no longer any skilled guesswork involved in photography and the gap between the amateur and the professional photographer is narrowed (Rubinstein & Sluis 2008, p. 13). The preview screen is entextualized in a way that stands in a lawful relation to the photographer’s bodily movements (Frosh, 2015).

Thus, we argue that the ontogeny of learning to take pictures after the introduction of digital cameras is not so much driven forward by punctuated sequences of (i) taking a picture, (ii) waiting while it’s developed (iii) evaluating its aptness. Rather, it is more a process of becoming attentive to changing one’s bodily orientation towards a scene until the protext aptly reflects desired social affordances. Digital technology has brought protext coupling and sensorimotor coupling closer to one another than ever before in the history of photography.
Figure 10. Picture of a scene as projected onto the ground glass viewing plate (right) of a 1900 wooden dry-plate studio camera (left). Note the upside-down protext.

Figure 11. Picture taken through the separate viewfinder (right) of a 1931 Leica 1 (left).

Figure 12. Picture taken from approximately 40 cm. distance to bring the scene into focus, of the separate viewfinder (right) on a 1936 Kodak Box 620 (left).
Figure 13. Picture taken through the separate eye-level viewfinder (right) of a 1962 Voightländer Vito Automatic 2 (left). Note the framing guides etched into the glass of the viewfinder.

Figure 14. Picture taken through the eye-level viewfinder (right) of a 1971 Asahi Pentax SP 500 SLR (left). Note the split-image centre (for focus adjustment), visible as faint circles in the middle, and the exposure guide needle on the right side of the viewfinder.
Changing affordances for social coupling at the logogenetic scale

The final kind of functional coupling we must—briefly—address in our discussion of accelerating semogenesis is the one that enables the dialogical coordination of photographic interaction between the photographer and his peers. In contrast to sensorimotor coupling and protext coupling, which are examples of microgenetic events, social coupling loops are events at the logogenetic scale. In his (2004a) *Agency and Consciousness in Discourse*, Paul Thibault writes “Human agents who engage in dialogically coordinated linguistic or other semiotic interaction produce discourse, which is itself embedded in and is often the direct product of meaning-making activity”. (Thibault, 2004a, p. 32) This is an apt description of how we see logogenesis.

Because a photographic event is situated in a normative cultural practice of photography, its aim is to produce social affordances (meaning potentials) with a good fit for the situation. From this point of view, photographic skill is the ability to reliably produce apt social affordances.

We have already argued that digital technology has enabled an acceleration of microgenetic coupling loops by changing affordances for sensorimotor coupling and protext coupling. We now argue that the same is true for the social coupling of the photographer with his peers by changing affordances for the *distributability* (after "distribution" in Kress & Van Leeuwen, 2001, p. 21) of photographic texts.

Consider the extremely different material constraints on distributing a daguerreotype plate and Facebook selfie. A daguerreotype plate—or indeed any pre-digital photograph including negatives and paper prints—is a spatiotemporally local and materially tangible artefact. Is has inertia. For photographic dialogue to happen, the daguerreotype must be moved physically from one human agent to another. Energy is expended in a lawful relation to inertia when it is moved, and inertia thus constrains the distributability it affords. Further constraining distributability, early photos were one-of-a-kind, as the picture developed directly on the plate. It was vulnerable to fingerprints, scratches and subsequent exposure to light, and it was expensive. Most daguerreotype plates remained rarities in private possession, only displayed for limited amounts of time during social calls. They signalled high social status (Clarke, 1997, p. 103). Because very few people at the
beginning of photography came into contact with photos, logogenetic samplings were few and far apart. We suggest that the ecosocial system was yet not *photographically saturated* to a degree that could impact ontogeny or phylogeny. However, as photographic technology improved and became less expensive due to dryplates, celluloid film and, eventually, negatives that enabled reproduction, saturation in the ecosocial system increased and the status of photography changed. As digital cameras began to become more affordable in the early 1990s image sharing was still a rather slow process where the viewers would gather around a single computer screen because online sharing of digital photographs was slow, expensive and required specialized knowledge about website hosting as well as coding skills (Rubinstein & Sluis 2008, 12). This changed as network speeds increased and with the rise of social media as platform for distribution.

Of course, the present day distribution of a Facebook selfie also requires energy, but this is a very different situation. In terms of its affordance for physical transactions with the environment, a digital photo is intangible, non-local and has no inertia. It exists solely as digital information, traveling as electromagnetic pulses in an electrical medium that saturates a globe-encompassing network of conducting wires. Invoking a metaphor from biology, one might refer to this as an electromagnetic *substrate* for communicative interactivity. The energy expended in the distribution of a single digital photo is an infinitesimal blip of no material consequence against the background of the totality of the substrate.\(^4\) As the network’s node clock frequencies (the machine cycles per second of all servers, processors and microprocessors coupled to the Internet) and edge bandwidths (copper lines, fibre optics, wireless- and phone networks and satellite links) continue to increase, the distributability afforded by digital photography is nearly effortless in terms of the distance it can travel and the time it takes. Thus, one might argue, changing energetic constraints collapses spatial and temporal constraints. Snapchat enables people across the globe to have something close to a photographic here-and-now conversation.

5. Discussion

We ended our presentation of a fourth, fast semogenic process in section 3.1 by suggesting that faster material affords faster logogenesis, which again provides faster (metaphorical) material for faster ontogenesis and phylogenesis. Considered from a material, processual perspective, digital patterns travelling as pulses in an electromagnetic communication substrate is almost as fast as is

\(^4\) This changes, of course, when someone decides to print a photo, but that only happens to a fraction of the hundreds of millions of pictures taken every day.
possible (at least until quantum computing or something similar becomes a commercially viable platform technology).

From this view of semogenesis arises a question of significant importance to the potential usefulness of our argument. Can it help us be more precise, for example, in how we think about learning? Halliday and Matthiessen’s account of semogenesis outlines an interesting dialectic between bottom-up and top-down causation. On the one hand, slower scale processes “provide environment” for faster processes. This idea would be cashed out top-down as constraint in the parlance of ecological psychology. On the other hand, faster scale processes “provide material” for slower scale processes. Ecological theory would recognize this as emergence.

Halliday and Matthiessen’s original (1999:17) model of semogenesis is essentially presented as a sandwich with ontogenesis in the middle, flanked by phylogenesis on top and logogenesis at the bottom. The model makes a clear statement that ontogenesis, that is, individual processes of learning, is influenced from both sides.

From the perspective we have presented in this article, this begs the following question: Whether one’s thinking about learning privileges top-down or bottom-up? Or, as stated by Timothy Ingold, whether one understands learning as “[...] an internalization of collective representations or, in a word, enculturation” (Ingold, 1993, p. 463, original italics) or as “[...] a process of enskillment, in which learning is inseparable from doing and in which both are embedded in the context of a practical engagement in the world” (ibid.).?

On the former view, which—or so we suggest—underpins multimodal literacy studies (e.g. Jewitt & Kress, 2003), learning is learning to know, understood as a process of sensitization to phylogenetic patterns. On the latter view, which informs the ecological psychological concept of direct learning (Jacobs & Michaels, 2007), learning is learning to do. Our aim here is not to rule out either position, as an ecosocial perspective precisely recognizes the strengths of both in relation to the study of photographic communication, but our perspective in this discussion favours the bottom-up enskillment perspective.

By now, many scholars have brilliantly treated the rise of digital photography. For example, Cobley and Haeflner (2009), with a clearly emancipatory agenda, have asked whether the boom of digital cameras “[...] teach, facilitate or otherwise enhance visual literacy among the public. Do they enable command over imaging? Do they foster any sense of the world of representation beyond the act of taking domestic photographs and communicating on a very localized, personal level?” (2009, p. 125). They argue that digital cameras, through their ubiquity, afford a more self-reflexive attitude
towards media in general (2009, p. 125) and possibly a critical literacy that can inform “what people want from photographs [...] and its non-verbal coordinates: the required pose, proper lighting, colour and exposure, and mimicry of the situation”. In their article, Cobley and Haeffner clearly privilege a top-down view of ontogenesis: Learning is cashed out as enculturation, an process of attunement to sociocultural patterns of normativity, knowing which poses are “required” and what lighting, colour and exposure is considered “proper” by the community. But there is another side of the argument, which Cobley and Haeffner do not elaborate. The articulation of socially apt affordances, “proper” lighting, colour and exposure (or indeed subversive and improper counterparts) necessarily requires the photographer to first master the technology. In other words, he must have learnt how to do lighting, colour and exposure with a camera. Learning to do, especially achieving expertise in a craft, requires many task repetitions. Think, for example, of the thousands of hours of sensorimotor rote learning that is required to master a musical instrument. In the age of daguerrotype photography, a person had to master complex processes of chemically treating and developing photographic plates as well as composing, focusing and exposing a shot. As we have tried to demonstrate in section 4, in the course of technological development, more and more of those aspects of skill have been designed into the tools (for example, in the viewfinder, as framing devices for composition, exposure needles and histograms for controlling lighting and exposure) thus rendering skills with doing unnecessary for the photographer.

This development does not go unnoticed by Cobley and Haeffner, as their discussion of digital photography touches on “the mutability of the image at the point of ‘production’ (as opposed to ‘post-production’) in the touch-of-a-button effects that digital cameras offer” (2009, p. 123). However, a lot of interesting insights are compressed into that statement and deserve unpacking. For example, that a contemporary smart phone camera essentially models the entire craft of photography in such a way that photographers do not need any dark room skills, or skill with or knowledge of shutter speeds, aperture settings or the sensitivity to light of film. They simply need to know how to operate sliders and buttons in a graphic user interface, or, increasingly to point and click. At any given time, the protext in every respect anticipates the finished text. And, increasingly, the camera will suggest apt social affordances (as when people are highlighted on the screen, or when the shutter is automatically released upon detection of a smile). Rather than having to achieve mastery of lighting, colour and exposure through repetitive shooting, developing, and evaluating, exposure, contemporary photographers can concentrate on the enskillment of orienting bodily with the camera in relation to a scene until the protext is apt in terms of social affordances. Thus, we
argue, at least when understood bottom-up, faster microgenesis engenders faster logogenesis, and, in turn, faster ontogenesis.

But what of accelerating phylogenesis? Again, many scholars before us have treated the impact of digital photography on conventionalization. In a 2014 article in New York Magazine entitled “Art at Arm’s Length: A History of the Selfie” art critic Jerry Saltz argued wrote: “Genres arise relatively rarely. Portraiture is a genre. So is still-life, landscape, animal painting, history painting”. He argued that the selfie is a distinct sub-genre in portrait photography, characteristic by nearly always being

[… ] taken from within an arm’s length of the subject. For this reason the cropping and composition of selfies are very different from those of all preceding self-portraiture. There is a near constant visual presence of one of the photographer’s arms, typically the one holding the camera. Bad camera angles predominate, as the subject is nearly always off-center. The wide angle lens on most cell-phone exaggerates the depth of noses and chins, and the arm holding the camera often looks huge. (Over time, this distortion has become less noticeable. recall, however, the skewed look of the early cell-phone snap.) If both your hands are in the picture and it’s not a mirror shot, technically it’s not a selfie—it’s a portrait (2014, p. 71).

In our terms, the characteristics of the selfie genre can be understood in terms of the smart phone’s affordances for sensorimotor coupling at the microgenetic scale. But it is extremely interesting that the genre has taken root in popular culture in such a short time. Of course, Saltz’s observations are not strong empirical justification for what we propose, but they do indicate that something is happening, and that it is happening quickly. Saltz argues that the selfie phenomenon began around 1999 as self-portraits “taken with crude early webcams, showing weird coloration, hot spots, bizarre resolution” (Saltz, 2014, p. 74). Very quickly, the basic fact of the matter has become that “I’ve taken them […] You’ve taken them. So has almost everyone you know” (Saltz, 2014, p. 72). The sheer rate and number of logogenetic instances of selfie photography and consumption of it provides ample material for fast ontogenesis. People very quickly catch on and begin to take selfies that are influenced by other people’s selfies through fast social coupling, and the ecosocial system begins to settle into a new pattern at the phylogenetic scale.

The selfie we took and posted on Table Mountain was not a very good picture in its own right, but it conformed to conventions of what a selfie picture should be like. And a bit of struggling to fit both of us in the frame aside, it was easy to take it. Which, if one thinks about everything that happened during that complex event, we find remarkable.

Acknowledgements
Thanks to Christopher Brenton West (http://www.daguerrotypes.co.uk) for sharing his knowledge of and experience with daguerreotype photography. Thanks also to Bjarne Meldgaard, director at The Danish Museum of Photography (http://www.fotomuseum.dk), who gave access to the museum's collection of cameras and allowed us to handle them, take pictures of them and through their viewfinders and framing devices.

References


