



A Linguistically and Educationally Oriented Theory of Learning Based on Situated Meaning

James Paul Gee  
Mary Lou Fulton Presidential Professor  
Regents' Professor  
Arizona State University  
[james.gee@asu.edu](mailto:james.gee@asu.edu)

## INTRODUCTION

In this paper I lay out a linguistically-oriented theory of learning. The theory is based on the concept of “situated meaning” (Gee 2004, 2014a, 2015a), a linguistic notion that is related to the concept of embodied cognition or situated cognition in psychology (the two terms tend to mean the same thing, see Wilson 2002 for an overview). Unfortunately, there are unsettled problems around theories of embodied cognition (Hickok 2014). Claims about mirror neurons (neural systems that allow humans to re-enact or “mirror” what others’ are doing and, thereby, figure out what their intentions and goals might be, see: Rizzolatti & Craighero 2004) and the role of the motor system and motoric processing in understanding action and language (Glenberg & Gallese 2012) are popular today, but controversial.

In the theory I develop in this paper, I take *thinking and learning through experience* (Gee 1992, 2004, 2014b) as the core concept behind embodied cognition. I do not rely on theories about mirror neurons and motor responses. Rather, I rely on ideas about how humans use stored experiences (memories) to give meaning to language in context, a process that we will see is closely related to how humans give meaning to objects and actions in the world as well.

Saying that thinking and learning are based on experience does not say much. We need to know how the experiences we have had function in the mind/brain to fuel thinking, learning, and meaning, the latter of which I will argue is crucial for the first two.

The theory I develop here is linguistically-oriented because I believe the role of language in learning has not been considered seriously enough—or at least not in the right ways—in theories

of learning in psychology. This is so partly because linguistics and psychology are different disciplines and often exist in the typical academic silos. For me, as I will make clear, debates about embodied cognition are—or should be—really debates about how humans give meanings to words (and to other sorts of symbols) and to the world, as well (Bergen 2012; Gee 2015a).

The theory I develop here will also have a certain technology orientation. I will argue that just as digital computers once inspired psychologists as a source of possible metaphors and models of the mind (Gee 1992), so, too, today video games can serve as a fruitful source of metaphors and models for a situated or embodied view of human thinking, learning, and action in the world (Gee 2013; Gee & Hayes 2011). Indeed, I will argue that video games are a better source of such metaphors and models than were digital computers.

This paper is not a research report. It is the statement of a theory that is meant to help guide thinking and research about learning and the application of this thinking and research to teaching and learning inside and outside schools. That is, the theory is meant to be applicable to debates about and interventions in education. The theory represents my view of what the research “says” most convincingly when we combine work in psychology, linguistics, and technology connected to embodiment (for a small sampling: Bergen 2012; Barsalou 1999a, b; Brown, Collins, & Dugid 1989; Churchland, & Sejnowski 1992; Clancey 1997; Clark 1989, 1997, 2003; diSessa 2000; Glenberg 1997; Hickok 2014; Tomasello 1999; Vygotsky 1987; Wertsch 1998). I evade controversies that are not relevant to a theory of meaning and its role in learning.

Work on embodied cognition in particular has made the term “simulation” popular (Hickok 2014; Gallese 2007). The mind/brain is said to run simulations based on experiences we have had in the world and with media as a basis for thinking and understanding. In the theory I develop here, a mental simulation-like capacity is crucial. But the term “simulation” is too general to buy us much clarity. When we experience the world we experience it through all our senses and in terms of how we process this sensory information at different levels in the brain. We then use our mentally stored data from experience to run simulations in our mind. These simulations share properties with the original experiences, that is, that are rich in sensory information and processing.

The simulation-like mental capacity my theory requires is that the mind/brain uses multi-modal, iconic, perspectival, flexibly manipulable representations. Multi-modal means that mental representations meld images, sounds, feel, words, and other human sensory information. Iconic means that these representations have parts that can “look like” (feel like, sound like, smell like) parts in other representations in ways that allow us to readily recognize and use patterns and sub-patterns. The word “tree” does not resemble the word “bush” in any way, but a picture (even a stick drawing) of a tree and a bush show resemblances of parts. “Perspectival” means that the representation is “from a point of view” or perspective and that things are foregrounded and backgrounded in terms of this perspective (or interest). For example, we can think about—just as we can draw—a city from different perspectives (top-down, internal, from a side, far, near, etc.) and we can foreground or focus on certain things or places and background others.

“Manipulable” means that the parts of these representations can be flexibly arranged, rearranged,

and transformed “on the fly” (as we process in real time) and can represent fantasy as well as reality.

Thankfully we have one good example of a multi-modal iconic perspectival flexibly manipulable representation and that is a role-playing video game (RPG). In an RPG (Gee 2014b) we get sounds and images that are iconic in ways that make patterns of resemblance clear. A player can clearly see where and how different types of aliens in *Halo*, say, resemble each other more than they resemble Master Chief, who resembles, but is not exactly like a (normal) human. An RPG allows the game’s world to be seen from the perspective of an avatar or role that helps determine how the world ought to be seen and dealt with for effective action and the accomplishment of goals in a given context where a problem must be solved. And, the game’s software can freely combine sounds and images to create “real” and “unreal” things and change things quickly across time and in relation to what players do. Of course, words, dialogue, and social interactions (either among characters or real people in a multi-player game) are also part of the multi-modal package in video games (and, of course, words, dialogue, and interactions are things we sense as well).

So imagine the mind/brain works like an RPG. It can “run” multi-modal scenes (with sounds, images, feel, smells, words, and emotions in the case of humans) taking the perspective of self, other, or even an animal or object in and on that scene. The scenes or scenarios it creates and runs do not have to be “real”, but can be fantasies or “what if” scenarios (Tomasello 2014). I can imagine myself, someone else, or even a river, seeking the sea.

Note that the “machine” that runs the RPG—whether game machine or wetware in mind/brain—might well use procedures or algorithms stateable in symbolic or propositional formats to run its scenarios, but this does not mean that such procedures or algorithms are the right level for semantics, that is, for the representation and manipulation of meanings. I will argue that the right level for semantics is the level of mental RPG-like multi-modal iconic perspectival flexibly manipulable representations. Such representations are a type of simulation, but not all simulations are RPG-like multi-modal iconic perspectival flexibly manipulable representations.

I will be arguing, then, that humans think, understand, and mean things thanks to their capacity to, in a sense, play video games in their minds. That we have the capacity to imagine ourselves, others, or things acting in the real world or a fantasy world is clear from our everyday experiences as humans. This does raise the old problem of a homunculus in the mind, a little “person” who, inside our heads, sees what else is in there and makes decisions. There is no such central observer or command center in the brain, of course (Hood 2012; Macknit & Martinez-Conde 2010; Swaab 2014). Who, then, is playing the video games in our minds? When we imagine ourselves asking someone for a first date, who is doing the imagining?

It would take us too far afield to deal with this issue in any complete way here. All I need at this point for my theory is the feeling we humans all have that we can “see” ourselves acting in the world and can imagine ourselves (or others) acting in the world in certain ways and suffering certain consequences (Bergen 2012; Gee 2014; Kean 2014). Since we will here be dealing with semantics, that is, with meaning, we will be dealing not with reality *per se*, but with the ways in which humans give words and things significance and communicate meanings with each other.

I need a simple term for the mouthful of “RPG-like multi-modal iconic perspectival flexibly manipulable representations”. The term “simulation” is too general. I will use the term “scenario” for short, but keep in mind the term is meant to capture multi-modal simulations that have a perspective defined by a role being played in the scenario and where elements of the simulation can be mixed, matched, and transformed flexibly and “on the fly” to create “real” and “unreal” things and situations (for a formalization of situations or scenarios see: Barwise 1981, 1988; Barwise & Perry 1983). Notice by the way that movies do not have this capacity, while video games do. A movie cannot change a sequence based on how the viewer sees it, but a game can and does modify what comes next based on what a player has done. So, too, we can flexibly transform what happens next in our mental imaginings based on what the “avatar” in our mind has been imagined to do or feel. We can play out in our heads several different scenarios of an event we are getting ready for.

It is important to note though that mental scenarios are much more multi-modal than even video games, since they can involve every possible human sensation, not just sound and image. Furthermore, it is crucial to keep in mind that in both video games and in life, experience includes not just images and actions, but dialogue, social and verbal interactions, reading, and media representations. Indeed, it is not uncommon that humans sometimes cannot remember whether they experienced something themselves, heard about, read it, or saw it in a movie. We humans treat media as vicarious experience and use such experiences quite regularly to think and plan with (Reeves & Nass 1999).

Just like video game graphics, our mental scenarios can be of very different grain sizes or levels of detail. A video game can be an 8-bit representation that obscures most details in the service of basic shapes and dimensions. Other video games can be nearly fully realistic representations. And there can be levels of detail all along a continuum between these two extremes. So, too, the mind can run scenarios of different degrees of detail, realism, and abstraction (Marr 1982). An 8-bit representation in a game is close to being more symbolic than imagistic, but is still imagistic as well. And things could get even more abstract and we could represent people (e.g., guards) by dots, things (e.g., corridors, doors, rooms) by rectangles, and lines of sight (e.g., where the guards are looking) by cones, as happens in some map-like representationa in video games meant to help players plan strategies (Gee 2014b).

A scenario in the head is not just a low-level sensory representation (as in some theories of embodied cognition), but a representation at the level of affordances and opportunities for action or goal accomplishment (Gee 2014b; Gibson 1977, 1979; Glenberg 1997; Glenberg & Robertson 1999). The scenario displays scenes that are “edited” from a perspective, namely the perspective of an agent (self or other) who wants to accomplish goals (or an animal or thing that acts so as to cause things to happen and we humans often treat animals and things as agents/causers). Thus, the scenario foregrounds the aspects of things and actions that can afford the agent the opportunity to accomplish goals. Imagining a scenario where you have to give a toast will make the glasses on the table prominent as objects for banging on to get silence and spoons prominent in their capacity as things to bang with. Thinking of setting a fine table will highlight different properties of glasses and spoons. Representations at the level of affordances and opportunities

are, of course, fairly high level and certainly sub-served in the brain by a variety of different levels of neural architecture.

## **SITUATED MEANING**

Claiming that humans think, plan, and create meaning through mental scenarios or “video games” is in opposition to the view that humans (like a digital computer) think, plan, and mean primarily via language-like symbol systems that constitute a language of the mind (Dennett 1969; Fodor 1975; Pylyshyn 1984). Linguists have used the term “mentalese” for this “language in the mind” (Fodor 1975) and philosophers see this language as composed of “propositions” (Dennett 1969), while psychologists see it as composed of “procedures” (rule-like formulations, Pylyshyn 1984). In any case, the idea is that a thought like “The puppy is brown” would be represented in a string of (universal) symbols like: THERE IS A YOUNG DOG SUCH THAT IT IS THE COLOR BROWN, where the capitalized words are universal mental and logical symbols that are part of a mental apparatus that lies behind all human languages and thinking. While I will be arguing that humans often use scenarios (image-like representations) rather than language-like symbols, we will still need to consider the role of more abstract, symbolic processing below, as well.

Humans, like many other animals, can run scenarios in their minds to plan and get ready for action (Suddendorf 2013; Tomasello 2014). We can imagine what will happen if we pull one apple from the bottom of a stack of apples. In fact, it is basic to human action that we engage in what I will call the “cycle of reflective action” (CORA). In the CORA when we want to accomplish a goal, we think about several possible courses of action and their possible outcomes,

then act in the world where our action is a sort of probe to test our hypotheses about actions and outcomes. Then we think about whether the response from the world to our action (what happened) was good or bad from the perspective of our goals. Then we can think again about new actions and their possible outcomes, adjust our strategy, act again, and continue this way until we either accomplish our goals, change them, or give up (Gee 2013).

The CORA is crucial to human intelligence and survival and is a core capacity that schools and society need to “train up” for complex thinking, complex problem solving, and collaboration that can lead to collective actions and intelligence. As with many elements of intelligence, the CORA can be supplemented and expanded by external devices and technologies.

The CORA is important in any theory of learning for several reasons. First, there is ample evidence that humans learn best when they are having an experience where they need to take an action (or actions) whose accomplishment they care about in the sense that something is meaningfully “at stake” for them in accomplishing it (Cosmides 1989; Wason 1966, 1968). Humans learn best, too, when they have clear goals for taking actions whose outcomes really matter to them (Hattie & Yates 2013).

Second, the CORA is important for theories of learning in that it stresses the necessity that learners can evaluate or assess whether what they have done and how the world has reacted is good or bad for their goal and, in terms of this assessment, can figure out what to do next. This capacity is an “appreciative capacity”, the capacity to appreciate in effective ways the goal-based outcomes of action in a sequence of actions (Schon 1983). In Wittgenstein’s (1953) work, the

capacity to know what to do next is virtually definitive of whether or not someone knows a specific “language game” and its related “form of life”. Learners need to be taught (and have modeled for them) the norms or standards for what constitutes successful outcomes and good next moves in a given domain or for a given social group.

Third, learning from experience can be too complex for newcomers (Kirschner, Sweller, & Clark 2006). There can be too many things to pay attention to. Beginners need help knowing how to manage their attention, that is, they need help knowing what to pay attention to and how and why they should pay attention to it in their experiences. Thus, parents, teachers, mentors, and social groups often design experiences that are constrained in some respects and that are well mentored to facilitate learning. For example, a new birder (engaged in “birding” or bird-watching) is taken out in just the right environment to maximize chances of seeing a certain type of bird (e.g., shore birds) and helped to know where the best spots to search are (e.g., shore line for waders, further out to sea for gulls, terns, and pelicans). Here, again, technology can often help, for example, virtual worlds where experiences can be constrained, samples concentrated, and important things highlighted. Learners need help in managing—and learning how to manage for themselves—their attentional economies.

So here get a set of supplementary principles for learning from experience in effective ways, at least early on in a learning progression: 1) assure that there are actions learners want and need to take and clear goals (that, of course, they can rethink in some circumstances as they probe the world with actions); 2) help learners understand and develop the norms and standards for outcomes and next moves that constitute a good appreciative system; 3) design well-structured

and well-constrained experiences and help learners to manage their attention in these experiences (what to focus on).

## **LANGUAGE**

Some animals can without doubt engage in the CORA (Tomasello 2014). They can work out some possible actions and outcomes in their heads before acting and suffering adverse outcomes. Our fellow primates can most certainly do this. Furthermore, for many baby animals their early limitations serve as ways to constrain their experiences to manageable proportions (just as children's early limitations in speech production and reception constrain what aspects of the input can become intake to be processed). But humans are far superior in this regard and the reason this is so, I would argue, is because we have language. Indeed, if we want a theory of learning that differentiates between general primate learning and more specific forms of human learning, then language is surely a key, since humans alone have language in the sense that the term "natural language" is used in linguistics (Pinker 1994).

To see the crucial role of language in thinking and learning we need first to see how meaning works in language and the role of what I will call "situated meaning" (Gee 2004, 2014a, 2015a). There are two types of meaning a word can have. First, a word can have what I will call a "basic meaning" (other terms here that mean much the same thing: literal meaning, utterance-type meaning, lexical meaning, or dictionary meaning). Thus, a word like "cat", in terms of basic meaning, means a certain type of animal, a feline. This type (cat) is not specified in our heads, in terms of vernacular language, by its genotype, though it is so specified by biologists. Indeed, there are various theories of how people represent basic meanings in their heads. These theories

include various sorts of verbal definitions, sets of features, or prototype examples (Cruse 1986; Murphy 2010; Smith & Medin 1981). However we think basic meanings are stored in the head, they set the range of possible situated meanings a word can have and the parameters via which new situated meanings can be formed in new contexts. Basic meanings and their relationships in sentences or propositions are a plausible domain for language-like symbolic processing and are the domain of semantics proper (“semantics” in its narrow sense) in linguistics.

Situated meanings, on the other hand, are the contextually-derived meanings speakers and readers give words in actual contexts of use (Gee 2004, 2014a). Thus, to take a very simple example, we give different situated meanings to the word “coffee” in the examples below (Clark 1989):

1. The coffee spilled, go get a mop.
2. The coffee spilled, go get a broom.
3. The coffee spilled, go stack it again.
4. Big Coffee is as bad as Big Oil in the corporate world.
5. Her coffee skin gleamed in the moonlight.
6. The coffee ice-cream is delicious.
7. What the Americans call “coffee” isn’t really coffee (i.e., it is not strong enough).

All words take on nuanced meanings in context and they can take on new meanings in new contexts. Situated meaning is accomplished by active work by listeners and readers. Some of this work is routine and has been done many times in the past, some is less routine and more

innovative. For example, an utterance like “Any political system without strong protections for private property is not a democracy” requires a good deal of work to assign an appropriate situated meaning to the word “democracy”. Since the basic meaning of the word “democracy” is “a system where people vote”, it is not at all clear why a country with no private property could not have voting and, thus, in the basic sense, be a democracy. If you know Milton Friedman’s work and that of the school of economics he founded (“the Chicago School”), then you can readily give a situated meaning to the word “democracy” here based on Freeman’s theories or related theories (Klein 2007). If you don’t, you can’t. This example, by the way, makes clear how little basic meanings buy you when you need to understand complex language and complex ideas, a hallmark of education.

Another example: “Corporations are persons” is clearly false on any sense of basic meaning, but it is now enshrined in law in the United States thanks to the Supreme Court. Assigning the word “person” here a situated meaning requires knowing a good deal about U.S. history and politics (“person” here means “bearer of civil rights”, thus, in U.S. history, the 14<sup>th</sup> Amendment to the Constitution has been used in court many more times to protect the civil rights of corporations than it has been to protect the civil rights of African-Americans, for whom it was originally formulated, see: Hartmann 2010).

There is one quite clear way to see the role and importance of situated meaning—that is, the capacity to situate meanings based on one’s assessment of the context—and that is any technical manual. The first passage below is from the manual from a video game. The second passage is from a textbook. Any gamer will tell you that the language in the game manual is pretty useless

until you have actually played the game (Gee 2003, 2013). This is so because the language in the manual is about the game world and its images and actions. When you can substituted such images and actions for the words and phrases in the manual you can give them situated meanings in terms of which you know how to operate with them to accomplish goals. Otherwise, you must try to understand the manual by translating words not into images and actions, but into other words that act like definitions. I will call this a purely “verbal understanding” and will argue that a situated understanding is what leads to deep understanding and the ability to effectively use the language of the text for goal accomplishment. Verbal understanding alone are much less deep and useful (unless and until one has had lots and lots of experience in the relevant domain).

So, too, for the passage from the textbook. When I read this I have only a verbal understanding. I ask myself, in confusion, isn't “removal of weathered material” just a form of “erosion” (but they are treated as two separate things here)? What is a “transporting agent”? I would have thought the production of rock waste was a form of erosion, but again they are contrasted here and the production of rock waste is weathering, not erosion. Why are chemical changes not “transporting agents”? It would really help if I could “see” what these words and phrases mean, that is, I could situate their meaning through running scenarios (perspectival images and other sensations) about them in my mind. Clearly a verbal understanding might allow me to pass a multiple-choice test, but it will not lead to effective understanding and problem solving, that is, to action-oriented goal accomplishment in the world.

Your internal nano-processors keep a very detailed record of your condition, equipment and recent history. You can access this data at any time during play by hitting F1 to get

to the Inventory screen or F2 to get to the Goals/Notes screen. Once you have accessed your information screens, you can move between the screens by clicking on the tabs at the top of the screen. You can map other information screens to hotkeys using Settings, Keyboard/Mouse (p. 5). (Ion Storm, 2000, p. 5).

The destruction of a land surface by the combined effects of abrasion and removal of weathered material by transporting agents is called erosion. ... The production of rock waste by mechanical processes and chemical changes is called weathering. (Martin 1990, p. 93)

Based on what I said above about scenarios in the mind here is one theory about how situated meaning works: Humans have experiences in the world. They store these experiences in their minds, but in an edited fashion where they foreground or highlight the aspects of the experience that were relevant to their goals (an effect of how they paid attention to elements in the experience while having it). With more and more similar experiences, humans find patterns and sub-patterns in this experience and they organize their experiences in their minds, and how they recall their experiences, in terms of these patterns and sub-patterns (associations). Patterns and sub-patterns, after a lot of experience, can rise to a fair degree of abstraction and generalization, but, for the most part, abstraction grows bottom up from specific experiences initially viewed in fairly concrete and specific ways (diSessa 2000). When humans need to act they can run scenarios in their heads that flexibly combine and recombine elements of these patterns and sub-

patterns to imagine things from the perspective of themselves and their goals or from the perspective of other actors or even objects (e.g., imagining a river flowing to the sea).

In giving situated meanings to words in context, we associate the word with a scenario or a set of them. We “see” and “act” (and “feel”) in our mind’s eye and can do so vicariously for others as roles in our heads. Remember that these scenarios can be at various levels of abstraction, but must remain iconic enough for patterns and sub-patterns to be apparent, rearrangeable, and modifiable. So, for example, say you heard from a clerk in an ice-cream store: “The coffee spilled, so I quickly scooped it up off the floor before the boss saw me”. You may well give the word “coffee” here a situated meaning based on a scenario involving coffee ice-cream and ice-cream scoops. Of course, you could imagine a comedic scene in which the person is frantically attempting to scoop coffee as liquid with an ice-cream scoop. Situated meanings are sometimes said to be in the domain of pragmatics in linguistics or in the domain of semantics taken in a broader way than is sometimes typical in linguistics.

## **LEARNING**

So now we have the pieces in place to state a theory of learning. What do situated meanings and the ability to situate meaning have to do with learning? One key thing to realize here is that comprehension is a general faculty in the mind that applies to written language, speech, and understanding things and actions in the world (Biemiller 2003; Stanovich 2000). It is not specific to language or reading (the only mental capacity specific to reading for alphabetic scripts is decoding). Humans use the same comprehension skills (and the same parts of the brain) to understand the world, oral language, and written language. We use the same capacities to

understand what people say, write, and do. And the basic form of this understanding is the giving of situated meanings to words, things, and actions via scenarios.

Learning from experience, texts, or teachers means a change in our mind/brains in terms of what we know or what we can do. This change requires, in any and all cases, that we understand (comprehend) the world and language in ways that allow input of data to be intake (intake = data that we can understand and process) so that such a change can happen. We move, in learning, from experience that gives us input to what can be “taken in” from this input based on our current capacities to comprehend or understand in a contextually appropriate way. So by its very nature situated meaning is the core basis of learning. As we have seen purely verbal learning not tied to scenarios is much less good for understanding and problem solving, at least for beginners.

But now this raises the question of what the role of language is in learning, since the capacity to understand things in their actual contexts of situations is not germane just to language itself. We use the same capacity to understand things (e.g., glasses for banging or table setting) and actions (e.g., pointing as indicating or as threatening). There are two different senses of the term “language” that are relevant here. One sense is language as a communicational device composed of tools that help listeners and readers understand and situate meaning in the right ways (the way the speaker or author “intends”). On this view of language, we view grammar as “functional”, that is, grammatical constructions are viewed in terms of their communicational functions (Halliday 1978). An example would be restrictive relative clauses which function to help listeners or readers identify things in context. If I say “The woman you liked best for the job is back for her last interview” I am using the relative clause construction “the woman you liked

best” to signal to you that you already know the woman and what part of the context is most relevant for you to use to recall who she is and what she means in this conversation and upcoming situation. This sort of functional approach to language is essentially a way of talking about grammar as a set of tools to help listeners and readers situate meanings (and this what functional grammarians do, Halliday & Matthiessen 2013).

A second sense in which the term “language” can be used is language as a system, as grammar composed of morphological, lexical, syntactic, and semantic systems (semantics here narrowly meaning basic meanings or type meanings and logical form). In the system sense of “language” we care about how grammatical constructions and words create meaning systems that are not specific to context (Saussure 1916; Vygotsky 1987). For example, imperfective aspect (signaled by “-ing” as in “washing”) means an action is viewed as ongoing, continuous, or repeated (“John is/was washing his car”). Perfective aspect (signaled by the simple present or past tense as in “washes” or “washed”) means an action is viewed as a point in time (“John washes/washed his car”). There are other aspects and they make up a system of related variables. So, too, words make up systems of contrast, such as cow/beef; pig/pork; sheep/mutton; deer/venison; lamb/lamb; chicken/chicken; fish/fish. Here we see a system in which some terms have different words for a live animal and the animal as food, but some do not. However, the words that do (e.g., cow/beef) indicate to us that the words that don’t (e.g., chicken/chicken) have two different basic meanings. This lexical system organizes one aspect of reality. Another example: the words “heat” and “temperature” relate to each other differently in vernacular everyday language and thought than they do in chemistry. Everyday language equates “high temperature” with “hot”, but these terms are tied to quite different concepts in chemistry.

Language as system is a way of organizing experience in relational and categorical ways via grammatical constructions and labels, and everyday language works differently in many respects than does the language of chemistry. An example of how grammatical constructions can categorize experience is nominalization in English. English has a near mania for treating processes as if they were abstract things, for example: grow → growth; vary → variation; destroy → destruction. This tendency is taken even further in many branches of science (Gee 2015) where nominalizations are very frequent because something like “growth” seems more quantifiable than a process like growing (e.g., consider “Hornworm growth exhibits a significant amount of statistical variation”, note here: “growth” rather than “grow” and “variation” rather than “vary” in comparison to something like: “Hornworms vary a lot in how well they grow”).

Language as system raises a chicken-and-egg paradox when we consider language as system side by side with situated (contextually-derived) meaning. As Vygotsky (1987) argued long ago, language as system regiment experience. Language as system places things in categories and relationships. Different registers (different varieties of language) regiment experience differently. The language of physics cuts up and relates things and processes in the world quite differently than do any vernacular varieties of language and differently again than, say, the language of chemistry or biology does. The language of traditional grammar cuts up English much as if it were Latin (which is where traditional grammar came from), while the language of Systemic Functional Grammar (to take a more formal academic approach) cuts up English in terms of forms married to functions (Halliday & Matthiessen 2013) and the language of generative grammar (to take another more formal academic approach) cuts up English into

abstract structures that are, in part, universal across all languages (Chomsky 1995). Languages and registers tell us where to cut at the joints in naming and thinking about experience in different domains (e.g., animals, biology, birding, chemistry, basic shapes and colors, plants, botany and a nearly endless number of other domains). Different languages and registers cut reality at different joints.

But now we reach a paradox. I have argue that scenarios give words and phrases situated meanings based on our experiences in the world. But I have also argued that language regiments (organizes) experience. So, if experience gives meaning to language and language tells us how to regiment and cut up this experience, which comes first: language or experience? The answer is that they bootstrap each other. The literature on the acquisition of language and the learning of registers is clear here (diSessa 2000; Hoff 2013; Gee 2004). Humans start by giving things concrete, contextually-specific meanings (represented by scenarios in our heads). For example, children may first assume that only four-legged animals with fur are animals and only later generalize the term to fish, bird, and insects (and maybe never to humans). With more experience we find patterns and sub-patterns in our related experiences and move towards more abstraction and generality, but always ready to go concrete again in new experiences.

The way language learning works in moving from specific and concrete meanings to more general and abstract ones is true of other domains as well. Andy diSessa (2000) points out that algebra does not distinguish effectively “among motion ( $d = rt$ ), converting meters to inches ( $i = 39.37 \times m$ ), defining coordinates of a straight line ( $y = mx$ ), or a host of other conceptually varied situations.” They all just look alike. He goes on to point out that “[d]istinguishing these *contexts*

is critical in learning, although it is probably nearly irrelevant in fluid, routine work for experts,” who, of course, have already had many embodied experiences in using algebra for a variety of different purposes and applied to different sorts of specific contexts.

The problem for human learning, viewed as a form of pattern recognition, is how to know out of all the myriad of patterns and sub-patterns the ever-ready pattern-recognizing human mind can find (many of them spurious), which patterns and sub-patterns are “right” or useful. We need help knowing which patterns and sub-patterns to look for, pay attention to, retain, and use (Gee 1992; Kirschner, Sweller, & Clark 2006). Where does this help come from? Some such help comes from our genes, undoubtedly. But, mostly it comes from parents, teachers, mentors, and social and cultural groups that help us to monitor and focus our attention and help us to learn new forms of language as a guide to the appropriate patterns and sub-patterns to be found and used in a given domain (Gee 1992, 2015b; Tomasello 2014). The new language is not general and abstract initially for us learners, but it is more so for the mentors who can inscribe instances of that language in and on our experiences of the world and media. Mentors can point the way along the path from situated meanings to meanings as system and back again as a way to organize experience in more and more general ways (when we need to).

The basic categories and relations of a language as system (ones that mentors already know) serve as attractors for learners to gradually organize their growing experiences and scenarios around. A good example of this process is birding. New birders need to see all sorts of birds in all sorts of different contexts (ecologies). They need to come to be able replay all sorts of bird scenarios in their heads (from the perspective of themselves, other birders, or birds themselves).

The experiences they have had and the ways then can use these experiences in their heads allows them to give situated meanings (contextually appropriate meanings) to the bird talk they hear and the bird texts they read, as well as to the bird shapes, colors, and behaviors they see (e.g., quick sight of a white rump in a field of grass = Bobolink; quick flash of white rump in forest = Flicker, a type of woodpecker; begging behavior in one setting = request to be fed from juvenal bird; begging behavior in another setting = part of sexual foreplay and social bonding).

At same time, bird books and guides and bird talk—in the guise of “bird language” or birders’ “ways with words”—help new birders organize and regiment their experience in terms of what patterns and sub-patterns they begin to recognize and use, patterns and sub-patterns that come, more or less, to match the system inherent in the books, guides, and talk. Of course, learners can sometimes come to see things in their experience that make them question or attempt to change the language as system they are learning as part and parcel of their education.

Mentors (more advanced birders) and social groups (clubs) of birders serve as the intermediators between situated meanings (birding in environments in the world, in the field) and language as system, abstraction, and generality. Mentors see that situated meaning and systemic abstraction bootstrap each other. Language as function and language as system come together.

I am using the terms mentors and mentoring broadly here for the ways in which parents, teachers, more advanced peers, or different technologies can engage in the core aspects of instruction or teaching: telling, showing, modelling, resourcing, constraining experience, and designing interactional and participatory frameworks. A key to any such teaching—in or out of

school; in social groups, cultures or institutions; in workplaces or popular culture—is to deftly relate both language as function and language as system, on the one hand, to action in the world and scenarios in the mind, on the other.

Words and symbols need to be supplied “just in time” (small amounts that can be applied in context right away) or “on demand” (larger blocks of talk or texts when learners need them and are ready for them and know they are or ask for them) (Gee 2003, 2013). How to situate meanings for words and symbols in different contexts needs to be modelled, as does assessing whether things in action in context have worked “right” or not (Bereiter & Scardamalia 1993; Schon 1983).

Words as systems for organizing experience and finding the “right” patterns and sub-patterns in that experience need to be inscribed on experience and representations of experience (and graphs and diagrams can be representations of experience) (Wellington & Osborne 2001; Zwiers 2014). One powerful new technology for inscribing language on experience is augmented reality (Klopfer 2008). Augmented reality uses mobile devices to lay images, graphs, and words over real environments that learners can navigate and see in real time both on a screen and in the real world. With or without technological support, deep learning is a mediated process among experience, scenario building, language as function, and language as system, serving (and changing) the interests and needs of different domains, groups, cultures, and institutions.

## **EDUCATION**

The view of the mind, understanding, language, and learning I have sketched here has important conceptual and ethical implications for learning, teaching, and education. Think a minute about maps. Maps are fairly abstract representations (much like a video game with 8-bit graphics). Concrete experiences of the terrain the map maps allows us to better understand and more effectively use the map. Handing a map of a city to someone who has had lots of experience with that city and asking them to find locations efficiently and effectively is easier than handing the same map to someone who has little experience of the city. It is even worse for people who have had little experience of cities at all. When we have had experiences with the terrain a map maps we can then use the map for our purposes and we can use it to supplement, deepen, and extend our own abilities as we enter new experiences.

So, too, when learners have had lots of experience with the contexts, images, actions, and interactions that a type of talk or text is about, for example, a lecture on molecules or a text on geology, they can use the talk or text as a reference, guide, or tool to obtain their own goals, supplement and extend their understanding and problem solving abilities, and learn new things. Minus experience, though, the words are only language as system that seeks to regiment and organize experiences the learner has not yet had. The learner can only gain a relatively abstract understanding of the talk or text as a system of relationships of words to words and grammatical constructions to grammatical constructions, systems unhinged from scenarios that could exemplify them and show how the system is a guide for finding and using normative patterns and sub-patterns in actual practice, action, and goal accomplishment.

Situated meaning and language as system cannot bootstrap each other if either one is missing. Situated meaning without language as system means that learners cannot see the forest for the trees and language as system without situated meaning means that learners would not know a forest if it bit them.

Learners who do not have good mentorship, social guidance, and instruction to ensure that they have the right sorts of experiences in the right order and have experiences that are well designed to focus their attention and help them find and use the right patterns and sub-patterns are asocial. They may be very creative, but they may well have learned and applied many a spurious generalization or been overwhelmed by the wonder and myriad features present in any rich experience in the world or in media.

Let me make clear that the theory of learning I have sketched here, when it is seen as a theory of education, is not progressive in the sense that learning and education are all about learners' own goals and immersion in experience, nor is it traditional in the sense that learning and education are all about overt instruction, teachers' goals, and lots of overt telling out of contexts of application. The theory developed here says that learning involves both immersion in experiences where learners have goals they care about *and* instruction, talk, and texts that regiment that experience and socialize it towards the relevant patterns and sub-patterns of a domain or social group (e.g., biologists). Done the right way this dance between immersion and instruction can lead to creativity and innovation as well, because each of us brings something unique to our experiences in the world and because knowing the norms of a domain or social

group is necessary for each of us to be able to contribute our unique perspectives to those domains or groups.

The theory of learning I have given in this paper suggests several core principles of deep teaching and learning (Gee 2008):

1. Learners must learn how to situate meaning for things, words, and actions for any new domain they need to learn. This requires lots of experience in specific contexts. They also need feedback about how appropriate their situated meanings are.
2. Learners must have an action (or actions) in their learning experiences whose outcome they care about and they must have clear goals.
3. Learners must get well-designed and well-mentored experiences that constrain complexity and help them to know how to manage attention in the experience.
4. Learners must have mentorship to help them develop an appreciative system in terms of which they can make good judgments in assessing the outcomes of their actions toward goals and in deciding what to do next.
5. Learners need mentors who use and help learners recruit language as system to label, guide, and organize their experiences and to recognize and use fruitful patterns and sub-patterns in a domain (i.e., help them grow generality and abstraction).

The term “mentors” here means parents, teachers, more advanced peers, technologies, social groups, and institutions. Teaching and learning is a coordinated and well-timed dance among experience, language as function, language as system, and mentors. One of the key capacities of new technologies like simulations, augmented reality, video games, and various forms of social media is their capacity to allow us to design well-structured, well-constrained, and well-mentored experiences for learners.

Giving learners tests or assessment when they lack situated meanings and have only verbal meanings is unfair, especially when others have had such situated meanings based on rich, well-mentored experiences. Giving learners experience but with too little mentorship and language as system can make them peripheral participants who do not share enough norms and systematic thinking with others to make real progress and gain standing in a domain.

## REFERENCES

- Barsalou, L. W. (1999a). Language comprehension: Archival memory or preparation for situated action. *Discourse Processes* 28: 61–80.
- Barsalou, L. W. (1999b). Perceptual symbol systems. *Behavioral and Brain Sciences* 22: 57–660.
- Barwise, J. (1981). Scenes and other situations. *The Journal of Philosophy*, 78: 369–97.
- Barwise, J. (1988). *The situation in logic*. Stanford: CSLI.
- Barwise, J. & Perry, J. (1983). *Situations and attitudes*. Cambridge, MA: MIT Press.
- Bereiter, C. & Scardamalia, M. (1993). *Surpassing ourselves: An inquiry into the nature and implications of expertise*. Chicago: Open Court.
- Bergen, B. K. (2012). *Louder than words: The new science of how the mind makes meaning*. New York: Basic Books.
- Biemiller, A. (2003). Oral comprehension sets the ceiling on reading comprehension. *American Educator* 27.1: 23, 44.
- Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.
- Clancey, W. J. (1997). *Situated cognition: On human knowledge and computer representations*. Cambridge: Cambridge University Press.
- Clark, A. (1989). *Microcognition: Philosophy, cognitive science, and parallel distributed processing*. Cambridge: MIT Press.
- Clark, A. (1997). *Being there: Putting brain, body, and world together again*. Cambridge: MIT Press.

- Clark, A. (2003). *Natural-born cyborgs: Why minds and technologies are made to merge*. Oxford: Oxford University Press.
- Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason Selection task. *Cognition* 31:187-276.
- Cruse, D. A. (1986). *Lexical semantics*. Cambridge: Cambridge University Press.
- Dennett, D. C. (1969). *Content and consciousness*. London: Routledge.
- diSessa, A. A. 2000. *Changing minds: Computers, learning, and literacy*. Cambridge: MIT Press.
- Gallese, V. (2007). Before and below “theory of mind”: Embodied simulation and the neural correlates of social cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences* 362: 659-669.
- Fodor, J. A. (1975). *The language of thought*. Cambridge, MA: Harvard University Press.
- Gee, J. P. (1992). *The social mind: Language, ideology, and social practice*. New York: Bergin & Garvey.
- Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. Second Edition, 2007. New York: Palgrave/Macmillan.
- Gee, J. P. (2004). *Situated language and learning: A critique of traditional schooling*. London: Routledge.
- Gee, J. P. (2008). A sociocultural perspective on opportunity to learn. In P. A. Moss, D. C. Pullin, J. P. Gee, E. H. Haertel, L. J. Young, Eds., *Assessment, equity, and opportunity to learn*. Cambridge: Cambridge University Press, pp. 76-108.

- Gee, J. P. (2013). *Good video games + good learning: Collected essays on video games, learning, and literacy*. Second Edition. New York: Lang, 2013
- Gee, J. P. (2014a). *An introduction to discourse analysis: Theory and method*. Fourth Edition. London: Routledge.
- Gee, J. P. (2014b). *Unified discourse analysis: Language, reality, virtual worlds, and video games*. New York: Routledge.
- Gee, J. P. (2015). *Social linguistics and literacies: Ideology in Discourses*. Fifth Edition. London: Taylor and Francis.
- Gee, J. P. (2015). *Literacy and education*. New York: Routledge.
- Gee J. P. & Hayes, E. R. (2011). *Language and learning in the digital age*. London: Routledge.
- Gibson, J. J. (1977). The theory of affordances. In R. Shaw and J. Bransford, Eds., *Perceiving, acting, and knowing: Toward an ecological psychology*. Hillsdale, N.J.: Lawrence Erlbaum, pp. 67-82.
- Gibson, J. J. 1979. *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Glenberg, A. M. 1997. What is memory for? *Behavioral and Brain Sciences* 20: 1–55.
- Glenberg, A. M. and D. A. Robertson. 1999. Indexical understanding of instructions. *Discourse Processes* 28: 1–26.
- Glenberg, A. M., & Gallese, V. (2012). Action-based Language: A theory of language acquisition, comprehension, and production. *Cortex*, 48: 905-922.
- Halliday, M.A.K. (1978). *Language as a social semiotic*. London: Edward Arnold.
- Halliday, M.A.K. & Matthiessen, C. (2013). *Halliday's introduction to functional grammar*. Fifth Edition. London: Hodder Arnold.

- Hartmann, T. (2010). *Unequal protection: How corporations became “people”—and how you can fight back*. Second Edition. San Francisco, CA: Berrett-Koehler.
- Hattie, J. & Yates, G. (2013). *Visible learning and the science of how we learn*. London: Routledge.
- Hoff, E. (2013). *Language development*. Fifth Edition. Belmont, CA: Cengage Learning.
- Hood, B. (2012). *The self illusion: How the social brain creates identity*. Oxford: Oxford University Press.
- Ion Storm (2000). *Deus Ex Manual*. London: Eidos Interactive. Retrieved from <http://gamecontentgreen.yummy.net/deusedemo/doc/Deusexmanual.pdf>.
- Kean, S. (2014). *The tale of the dueling neurosurgeons: And other true stories of trauma, madness, affliction, and recovery that reveal the surprising history of the human brain*. New York: Hachette Books.
- Kirschner, P., Sweller, J., & Clark, R. E. (2006). Why minimally guided learning does not work: An analysis of the failure of discovery learning, problem-based learning, experiential learning, and inquiry-based learning. *Educational Psychologist*, 41: 75-86.
- Klein, N. (2007). *The shock doctrine: The rise of disaster capitalism*. New York: Henry Holt.
- Klopfer, E. (2008). *Augmented learning: Design and research of mobile educational games*. Cambridge, MA: MIT Press.
- Macknit, S. L. & Martinez-Conde, S. (2010). *Slights of mind: What the neuroscience of magic reveals about our everyday deceptions*. New York: Henry Holt.
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York: Freeman & Co.

- Martin, J. R. (1990). Literacy in science: Learning to handle text as technology. In F. Christe, Ed., *Literacy for a changing world*. Melbourne: Australian Council for Educational Research, pp. 79-117.
- Murphy, M. L. (2010). *Lexical meaning*. Cambridge: Cambridge University Press.
- Pinker, S. (1994). *The language instinct: How the mind creates language*. New York: William Marrow
- Pylyshyn, Z. W. (1984). *Computation and cognition*. Cambridge, MA: MIT.
- Reeves, B. and Nass, C. (1999). *The media equation: How people treat computers, television, and new media like real people and places*. New York: Cambridge University Press.
- Rizzolatti, G. & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience* 27: 169–192.
- Saussure, F. de (1916). *Cours de linguistique generale*. Paris: Payot. English translation published as *A course in general linguistics* (trans. by Wade Baskin). New York: McGraw-Hill, 1959.
- Schön, D.A. (1983). *The reflective practitioner: how professionals think in action*. New York: Basic Books.
- Smith, E. E. & Medin, D. L. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.
- Stanovich, Keith E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York: Guilford Press.
- Suddendorf, T. (2013). *The gap: The science of what separates us from other animals*. New York: Basic Books.

- Swaab, D. F. (2014). *We are our brains: A neurobiography of the brain, from the womb to Alzheimer's*. New York: Spiegel & Grau.
- Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge: Harvard University Press.
- Tomasello, M. (2014). *A natural history of human thinking*. Cambridge, MA: Harvard University Press.
- Vygotsky, (1987). *The collected works of L.S. Vygotsky, Volume 1: Problems of general psychology. Including the volume Thinking and speech*. Ed. by R. W. Rieber & A. S. Carton. New York: Plenum.
- Wason, P. (1966). Reasoning. In B. M. Foss, Ed., *New horizons in psychology*. Harmondsworth: Penguin, pp. 135-151.
- Wason, P. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology* 20: 273-281.
- Wellington, J. & Osborne, J. (2001). *Language and literacy in science education*. Buckingham, England: Open University Press.
- Wertsch, J. V. (1998). *Mind as action*. Oxford: Oxford University Press
- Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review* 9: 625–636.
- Wittgenstein, L. (1953). *Philosophical investigations*. Oxford: Blackwell Publishing.
- Zwiers, J. (2014). *Building academic language: Meeting common core standards across disciplines*. San Francisco, CA: Jossey-Bass.

