Extended, Embodied Cognition and Second Language Acquisition

DWIGHT ATKINSON

Department of English, Purdue University, Indiana, USA
E-mail: dwightatki@aol.com

A cognitivist approach to cognition has traditionally dominated second language acquisition (SLA) studies. In this article, I examine two alternative approaches—extended cognition and embodied cognition—for how they might help us conceptualize SLA. More specifically, I present: (i) summaries of extended and embodied cognition, followed by reasons why the two can be treated as a single, synthetic perspective; (ii) an approach to SLA grounded in an extended, embodied view of cognition—i.e. a sociocognitive approach—in three principles; and (iii) a naturally occurring example of extended, embodied cognition-for-SLA.

A highly specific view of cognition—cognitivism—has dominated second language acquisition (SLA) studies. As one SLA pioneer describes it:

Even in those early days, we believed that we were witnessing the birth of a new field—one that did not see language as behavior, one that no longer ignored the mind, one that put cognitivism squarely at the forefront of its explanations. As it turns out, it was a powerful birthright. It is fair to say that a cognitivist view has dominated the field ever since. (Larsen-Freeman 2007: 775)

Yet conceptions of cognition have changed radically over the past century. Most recently, ‘a sea change in research and theory’ (Semin and Cacioppo 2008: 140) has occurred—toward extended and embodied views of cognition. Extended cognition conceptualizes mind/brain as inextricably tied to the external environment, while embodied cognition views cognitive activity as grounded in bodily states and action. These two approaches are related because bodies link minds to the world—we experience, understand, and act on the world through our bodies. As a result, extended and embodied cognition are sometimes grouped together.

This article begins by briefly locating cognitivism in SLA studies. Next, extended and embodied views of cognition are reviewed and synthesized. An approach to SLA based on this synthesis—a sociocognitive approach—is then described in three principles. Finally, extended, embodied cognition-for-SLA is exemplified in a videotaped English as a Foreign Language (EFL) tutoring interaction.
COGNITION IN MAINSTREAM SLA STUDIES

Following Descartes (1637/1960), cognitivism views the mind/brain as the self-sufficient source of cognition. In SLA studies, its *locus classicus* is Sharwood Smith’s (1991) declaration that while social context is the icing in L2 learning, cognition is the cake. This perspective, which has dominated the field, conceptualizes SLA in terms of what learners *know*—their internalized linguistic competence. In Doughty and Long’s influential view:

> Much current SLA research and theorizing shares a strongly cognitive orientation... The focus is firmly on identifying the nature and sources of the underlying L2 knowledge system, and on explaining developmental success and failure. Performance data are inevitably the researchers’ mainstay, but understanding underlying competence, not the external verbal behavior that depends on that competence, is the ultimate goal. Researchers recognize that SLA takes place in a social context, of course, and accept that it can be influenced by that context... However, they also recognize that language learning, like any other learning, is ultimately a matter of change in an individual’s internal mental state. As such, research on SLA is increasingly viewed as a branch of cognitive science. (2003a: 4)

This is a paradigmatic statement of the mainstream, cognitivist view of SLA—one which recapitulates Chomsky’s radical severing of competence from performance and assumes the centrality of ‘representation and computation...[in] a pre-eminently cognitive, information-processing approach’ to SLA (Long and Doughty 2003: 867). It also suggests that SLA studies is a cognitive science.

EXTENDED COGNITION

Yet cognitive science is many things, and to many people. Philosophers Clark and Chalmers (1998) coined the term ‘extended cognition’ to describe their distinctive take on what is variously called situated cognition, distributed cognition, grounded cognition, and active externalism. Based substantially on Clark’s work, I see four propositions as underlying the extended cognition thesis.

Cognition is environmental

Instead of being the self-contained logical system posited by cognitivism (Boden 2006), cognition depends heavily on the external environment. Mundane examples abound: We wake, sleep, eat, and shift activities according to clocks and calendars; do math with calculators or on paper; leave our keys by the door on returning home so we can find them later on; refrigerate food in see-through containers; store cleaning supplies under the sink; arrange
books and papers alphabetically or by subject; talk problems out in conversation; make group decisions; consult shopping lists or scan store shelves to help us shop; and literacy pervades our lives. Even ‘higher-order’ cognitive activities depend on the environment, as Clark’s (2001: 142) account of his academic writing process suggests:

The brain supported some rereading of old texts, materials, and notes. While rereading these, it responded by generating a few fragmentary ideas and criticisms. These ideas and criticisms were then stored as more marks on paper, in margins, on computer discs, etc. The brain then played a role in reorganizing these data on clean sheets, adding new on-line reactions and ideas. The cycle of reading, responding, and external reorganization is repeated, again and again. Finally, there is a product...But this intellectual product owes a lot to those repeated loops out into the environment. Credit belongs to the embodied, embedded agent in the world. The naked biological brain is just a part (albeit a crucial and special part) of a spatially and temporally extended process, involving lots of extraneural operations, whose joint action creates the intellectual product. There is thus a real sense...in which...the ‘problem-solving engine’ is...the whole caboodle...: the brain and body operating within an environmental setting.

Without question, cognition does sometimes operate in less environmentally dependent ways. But that we so frequently rely on the world when cognizing is not incidental. I present reasons for cognition’s deep environmentalism below.

Clark and Chalmers (1998) propose a ‘parity principle’ (Wilson and Clark 2009) for determining the scope of extended cognition:

If, as we confront some task, a part of the world functions as a process which, were it done in the head, we would have no hesitation in recognizing as part of the cognitive process, then that part of the world is...part of the cognitive process. (71)

This claim is the heart of the extended cognition thesis: The environment is part of cognition itself, at least when fundamentally abetting onboard cognitive resources.

**Cognition is adaptive**

This proposition holds that the primary purpose of cognition is to facilitate adaptivity to an uncertain environment (Barsalou 2008). It has two parts. First, rather than having evolved to address the abstract, logic-like problems of cognitivism, cognition is first and foremost flexible, on-line adaptive intelligence (Wheeler 2005). Cognition thus subserves adaptive action: crossing the street safely; staying warm and dry in cold rain; or knowing the right time, place, and partners for business, political action, or romance.
The second part of this proposition concerns the understanding that real-time cognitive resources are extremely limited (Schmidt 2001). If so, but fluent cognition is crucial for survival, then these resources must be guarded jealously. We have therefore evolved breathtaking abilities to offload cognition onto the world, as formalized in Clark’s cleverly named 007 Principle:

In general, evolved creatures will neither store nor process information in costly ways when they can use the structure of their environment...as a convenient stand-in for the information-processing operations concerned. That is, know only as much as you need to know to get the job done. (1989: 64)

Thus, even where cognizers can solve problems using just the naked brain, it is usually cognitively cheaper and therefore more adaptive to bring in the environment. For instance, I could have memorized (with great effort) the talk this article is based on; instead, I used notes and Power Point to enhance my cognitive resources.

The environment is highly structured for cognitive activity

This proposition is the converse of the first two. It also has two parts. First, the natural environment is rich in cognitive support structures, sometimes due to human–environment coevolution. Consequently, ‘the world is [often] its own best model’ for cognition (Brooks 1990: 5). That is, contra cognitivism’s decontextualized internal world models or representations, the natural environment often guides cognition-for-action. Thus, before leaving home I usually peek outside: If it’s sunny I know to apply sunscreen, if it’s rainy to take an umbrella. Another example: I once received these cognition-for-worldly-direction-finding instructions: ‘Turn right at the pond,’ and ‘Ours is the house with the large pine trees.’

Second, human environments are often designed for fluid, low-dimensional cognitive action. Modern homes, for instance, have their light switches by doors, door knobs uniformly positioned, closets located just where stored items can be accessed when needed, and kitchens organized in highly memory-enhancing ways, thereby enabling fluent cognition-for-cooking-and-serving. Organizations of space and material thus profoundly afford the cognition-action complexes comprising everyday human behavior.

In ground-breaking studies, anthropologists have indicated how human-designed environments support cognition-for-action. Hutchins (2005), for example, suggests that a line of people waiting for a concert is cognized as a ‘blend’ of the perceptual universal of linear salience, internalized cultural notions like ‘first-come, first-served’ (1559), and real-world structure, i.e. the line itself. To use this line for worldly action—e.g. to know how many precede you in it—requires no elaborate internal representation: Just look.
Cognition is shared and distributed

The extended cognition thesis emphasizes non-human affordances over human co-cognition. Yet shared cognition contributes crucially to human activity. Everyday conversation, for instance, is so highly coordinated that ‘the notion of an utterance as the sole product of a speaker, or...a mind, could hardly have been entertained had real talk-in-interaction’ been a central focus of linguistics (Schegloff et al. 1996: 20). Consider latching, the very common practice of positioning one turn-at-talk after another without a gap. Latching requires not just the ability to predict precisely when the preceding speaker will stop, but also what she will say, since the following turn must be relevant (i.e. appropriate, contentful, and formally congruent—Sperber and Wilson 1986; Schegloff 2007) vis-à-vis the preceding one. Likewise, the phenomenon of one speaker completing another’s turn is predicated on high-level co-cognition; both this and the foregoing example can be explained only via profound sociocognitive anticipation (Kinsbourne and Jordan 2009)—the ability to know what others are about to do—and synchrony (Semin 2007)—the ability to adjust one’s own real-time behavior accordingly.

There is converging evidence that anticipation and synchrony are built into human sociocognition. Research on infants, for example, shows apparently innate anticipatory and synchronization abilities such as proto-turn-taking (Foster 1990), an activity which has serious consequences for caregiver–child bonding soon after birth. Infants likewise actively imitate their caregivers, and begin to track their gaze late in the first year. Levinson’s (2006) and Tomasello’s (2008) research on what the former terms the ‘interaction engine’—the suite of innate abilities enabling the close ‘intention-reading’ that underlies all human communication, as described in more detail below—further supports the idea that humans are highly attuned co-cognizers. The same is true for the human ‘mirror neuron system,’ also described below.

EMBODIED COGNITION

In conceptualizing thought as the manipulation of logical symbols, the leaders of the cognitive revolution forgot that cognitive processes are also biological. The result was functionalism—the doctrine that cognition is implementation-neutral (Boden 2006). But cognitive science took a biological turn in the 1990s, leading to extensive research on embodied cognition. I briefly summarize this research here.²

Cognition, perception, and motor action are integrated activities

Cognitivists modularize cognition, strictly dividing it from perception on the input side and motor action on the output side (Barsalou 2008). Recent
research, however, suggests the integrated nature of these domains. Thus, *mirror neurons*, a class of neurons in the cerebral cortex, activate not just when we perform motor actions, but also when *perceiving others* performing those same actions (Rizzolatti and Craighero 2004). By providing evidence for the shared neural coding of self- and other-action, mirror neurons appear to account for action-oriented understanding, imitative learning, and synchronized behavior.

Cognitive representations are embodied and action-oriented

Cognitivism holds that cognitive representations are abstract, symbolic, and amodal (Harnish 2002); embodied cognition provides three alternative (if overlapping) understandings: simulation; analogical representations; and image schemas. *Simulation* is the cognitive ‘reenactment of perceptual, motor, and introspective states acquired during experience’ (Barsalou 2008: 618). Extending one of Barsalou’s examples, the act of sitting in a chair is accompanied by the integrated, multimodal recording of sensory experience (e.g. of the chair’s texture), action (the neuromuscular activity of sitting), and introspection (e.g. expecting the chair to be comfortable). These representations are simultaneously combined with multimodal representations of past acts of sitting. Then, on next encountering a chair, this perception-action-introspection complex is activated, enabling comprehension that the object is a chair.

*Analogical representations* store patterns of how our bodies ‘mesh’ with the environment—these patterns incorporate environmental information (Glenberg 1997). Adapting from Glenberg, a path through the woods is remembered via the means it affords for passing from point A to B: space for our body between two trees here, an uneven walking surface, a sequence of stones across a stream there. Analogical representations are thus directly action-oriented, and summarize embodied experience.

Johnson and Rohrer (2007) described *image schemas* as:

recurrent patterns of bodily experience, ‘image’-like in that they preserve the topological structure of the perceptual whole, …[which are] operating dynamically in and across time, are realized as activation patterns…in and between topological neural maps, …link sensorimotor experience to conceptualization and language, and…afford ‘normal’ pattern completions that can serve as a basis for inference. (36)

Image schemas are therefore analog representations based on cumulative sensory experience ‘that help solve…adaptive problems…in complex physical environments’ (Gibbs 2006: 69). Even abstract thought and language are
cognized as image schemas: In ‘She launched into her spiel,’ ‘launched into’ is comprehended via conceptual metaphors based on the literal meaning of ‘launch’ and ‘into.’ Thus, all thought is embodied, no matter how abstract or figurative.

**Empirical evidence for embodied cognition**

There is substantial empirical evidence for embodied cognition, much of it experimental. Barsalou (2008), Gibbs (2006), Glenberg (1997), and Semin and Smith (2008) provide partial reviews. Major findings include:

- Embodied states affect and are affected by cognition (Barsalou 2008)
- Brain regions previously thought to be reserved for cognition are active during perception and motor action (Iverson and Thelen 1999)
- Neural mechanisms underlying cognition show embodiment effects (Rizzolatti and Craighero 2004)
- Memory shows embodiment effects (Glenberg 1997)
- Understanding objects depends on embodied experience with them (Carlson and Kenny 2005)
- Emotions enable cognition/understanding, e.g. brain damage causing loss of evaluative ability renders decision-making impossible (Damasio 1994)
- Bodily orientation affects cognition (Lempert and Kinsbourne 1982)
- Cognitive development depends on opportunities for embodied action (Smith 2005)
- Synchronized bodily states and action promote shared cognition (Semin and Cacioppo 2008)

**Embodied cognition and language**

Embodied cognition researchers study language processing as a core function of cognition. Barsalou (2008), Gibbs (2006), Glenberg (1997), Iverson and Thelen (1999), and Johnson and Rohrer (2007) review findings, which include:

- The brain’s ‘language areas’ activate during sensorimotor action (Bonda et al. 1994)
- Brain ‘motor areas’ activate during speech (Hauk et al. 2004)
- Verbalization of memory is facilitated when assuming original body posture during recall (Dijkstra et al. 2007)
- Linguistic tasks are facilitated when accompanied by action (Rieser et al. 1994)
- Descriptions of spatial associations are comprehended faster than those of spatial dissociations (Glenberg et al. 1987)
- Words with high ‘body-object interaction’ ratings are recognized faster than those without (Saikaluk et al. 2008)
- Speech and gesture emerge together in infancy (Iverson and Thelen 1999)
SYNTHESIZING EXTENDED AND EMBODIED APPROACHES TO COGNITION

Although treated separately so far, extended and embodied views of cognition are broadly compatible and sometimes grouped together. The unifying idea is that cognition is naked and amodal only when conceptualized abstractly. Studied biologically, cognition is embedded—both in body and world.

In a review entitled ‘Grounded Cognition,’ Barsalou (2008: 618–19) synthesizes embodied and extended views of cognition:

Conceptions of grounded cognition take many different forms… Some accounts… focus on the roles of the body in cognition, based on widespread findings that bodily states can cause cognitive states and be effects of them… Most accounts of grounded cognition, however, focus on the roles of simulation in cognition… Still other accounts… focus on situated action, social interaction, and the environment.

Similarly, in a chapter entitled ‘Embodied, Situated, and Distributed Cognition’, Clark (1998) describes these forms of cognition in a loosely united framework. That he retains the individual terms suggests disciplinary differences, while his synoptic treatment points toward synthesis. In fact, extended cognition has been studied largely by philosophers and anthropologists, and embodied cognition by psychologists, linguists, and neuroscientists. In this article, I adopt the view that extended and embodied cognition coalesce in a larger project—the study of cognition-for-worldly-action. At the same time, I retain separate terms in my formulation—extended, embodied cognition—to indicate that researchers often distinguish them.

EXTENDED, EMBODIED COGNITION AND SLA

Extended, embodied perspectives on cognition have entered the cognitive science mainstream, but are just beginning to influence SLA studies. I try to promote their development here, working within a sociocognitive approach to SLA (Atkinson 2002; Atkinson et al. 2007; Churchill et al. 2010). Specifically, I introduce three SLA principles based on extended, embodied cognition: (1) The Inseparability Principle: Mind, body, and world work together in learning/SLA; (2) The Learning-is-adaptive Principle: Learning/SLA facilitates survival and prosperity in complex environments; and (3) The Alignment Principle: A major engine of learning/SLA is alignment—the means by which we effect interaction.

The Inseparability Principle and SLA

To appreciate the inseparability principle, first consider four questions vis-à-vis Pictures 1–4, taken from naturally occurring interaction: (1) Who is the person in the picture?, (2) What is she doing?, (3) Where and when?, (4) Why? The aim of this
exercise is to suggest that SLA is more than just a cognitive input/restructuring/output process, and what some of that ‘more’ may be.  

Possible answers to the ‘who’ question here are: ‘A young woman’, ‘An adolescent’. The ‘what’ question might be answered: ‘She’s thinking’ or ‘She’s gazing at something.’ The remaining questions cannot be answered because the context has been removed. Like the cognitivist cognizer, this thinker is an isolated entity, cognizing for no apparent purpose in an ecosocial vacuum. Yet she differs in being embodied—she is not just the abstract mechanism described, for instance, in Corder’s SLA field-establishing paper: ‘The internal structure of the (language acquisition) device, i.e., the learner, has gone relatively unexplored’ (1967/1981: 12).

Picture 1

Possible answers to the ‘who’ question here are: ‘A young woman’, ‘An adolescent’. The ‘what’ question might be answered: ‘She’s thinking’ or ‘She’s gazing at something.’ The remaining questions cannot be answered because the context has been removed. Like the cognitivist cognizer, this thinker is an isolated entity, cognizing for no apparent purpose in an ecosocial vacuum. Yet she differs in being embodied—she is not just the abstract mechanism described, for instance, in Corder’s SLA field-establishing paper: ‘The internal structure of the (language acquisition) device, i.e., the learner, has gone relatively unexplored’ (1967/1981: 12).

Picture 2
Who: Based on her possible actions (see below) and age, this may be a student, perhaps in secondary school. What: We now see what she may be cognizing about: Papers lie before her, and a pen is barely visible in her right hand. She may thus be thinking in order to complete a school-sponsored assignment. Where and when: She is sitting at a table, perhaps in a classroom given the above. Why: Given the above, her cognition may be part of the larger sociocognitive action of completing a school-sponsored assignment, partly or conventionally in order to learn. Remove the assignment, and the nature of her cognition will likely change as well.

Who: We now see that the young woman has an older companion. Both perform similar activities—at least each has a literacy artifact before her and pen in hand. They even share their physical orientation: left hand/arm on artifact, right pen-holding hand to head, and head tilted right. Are they student and teacher? What: They appear to be studying together, the younger person completing an assignment, the older reading a book. Where and when: In a classroom? During class time? Why: No new information, but given the above, the older woman could be playing a guiding role.
Here, possible answers start to blend together. *Where/who/what:* They are in a home, sitting (perhaps) at a dining table. This suggests a family relationship. This in turn suggests that homework is being completed, the older woman assisting. *When:* If the younger woman is a student, perhaps it is a weekend or after school.

In fact, the younger woman (‘Ako’) is a Japanese junior high school student completing English grammar exercises to prepare for high school entrance exams. Her aunt (‘Tomo’) assists her as they sit at Tomo’s dining table. Ako has just read out a Japanese sentence, and repeated an adverbial as a place to start translating. Tomo has confirmed this choice by latching Ako’s utterance with an agreement marker and repeating the adverbial, closely shadowing Ako’s intonation. Now, as Ako ponders the correct translation, Tomo behaves in two seemingly contrasting ways: She looks at the book before her, having earlier told Ako she would help her when absolutely necessary but that otherwise she was on her own; and she aligns her body with Ako’s, seeming to signal ‘We’re in this together.’ One might even venture that the shared physical orientation suggests co-cognition—that the exercise answer will be highly co-constructed, as indeed it turned out to be (Atkinson et al. 2007).

This pictorial illustration has three implications for SLA. First, it suggests that people cognize/learn not just mentally, but in environments comprised of bodies, cognitive tools, social practices, and environmental features. If, as the inseparability principle argues, such contexts crucially affect cognition/learning, then they cannot be treated as optional extras. Goodwin (2003: 20) makes this point specifically regarding embodiment: ‘The positioning, actions, and orientation of the body in the environment are crucial to how participants understand what is happening and build action together.’ Regarding learning,
this suggests that: (i) Learning is more *discovering how to align with the world* than *extracting knowledge from it* (Ingold 2000); and (ii) By being environmentally embedded, knowledge/cognition is made public and thereby learnable (Goodwin 2003). Indeed, Gee (1995) makes the latter a general principle of learning:

The meanings of the parts of new systems, whether words, visual symbols, actions, or objects, must initially be rendered public and overt, so that the learner can see the connections between the signs and their interpretations. And this is done... by the ways in which words, actions, and social interaction are integrally intertwined. (p. 337)

The second implication of this illustration concerns the *quality* of cognition/learning (van Lier 2002). The personal relationships learning involves, its role in identity construction (Norton 2000), where and under whose sponsorship it occurs, and how it is embodied and enacted fundamentally influence its outcome. Unlike computers, humans don’t just process—they *find value and meaning*. This affects learners’ engagement with learning opportunities, including whether they engage at all.

The qualitative nature of learning is well known in L2 teaching, but has not played a major role in mainstream SLA studies. This is because cognitivism views cognition as computational and amodal, and because it employs research methods which neutralize the quality of learning experiences in order to put them on common quantitative scales. Researchers thus negate what good teachers know—that learning depends on its meaningfulness to learners themselves.

Third, if cognition/learning is complex and multimodal, as the illustration suggests, then it must be studied complexly and multimodally. Quoting Goodwin (2003) again, ‘Rather than being lodged in a single modality..., many forms of human action are built through the juxtaposition of quite diverse materials, including the actor’s body, the bodies of others, language, structure in the environment, and so on’ (22–3). An extended example of multimodal SLA is provided below, so I will not elaborate here.

**The learning-is-adaptive principle and SLA**

Given that cognition exists to support adaptive action, sociocognitive approaches to SLA posit that learning’s main purpose is to increase adaptive functionality. This seems commonsensical—what else could learning be? Yet mainstream learning theory assumes the opposite: that learning means *detaching knowledge from the world and internalizing it*.

As evidence that learning is adaptive, consider our finest learners: Children learn so quickly and well that some deny they are learning at all. Thus, for Chomsky, ‘It’s probably a mistake to even *use* the word ‘learning’ in connection with language acquisition’ (Equinox Films 1995). Yet children clearly do
have a serious need to learn: As infantile dependence decreases, they must survive in a complex world. And this world is largely social. Children therefore learn best what they need most to be treated as humans—their social groups’ practices, including language.

The learning-is-adaptive principle has four linked implications for learning/SLA: (1) Learning/SLA is relational: Learning-as-adaptive-behavior concerns how to relate—how to articulate with one’s environment. For SLA, this means that L2s are learned as social practices—practices employed by the language’s users to accomplish social action; (2) Learning/SLA is experiential, participatory, and guided: One learns to relate by relating—learning is experiential. Of course, SLA involves a wide range of activities and motivations, but its paradigm route and goal centrally involve communicating with L2 users to effect social action, i.e., participation in social situations for conventional social purposes. This is a far cry from seeing acquisition-for-internalization as route and goal of SLA. Yet learning only from raw experience is risky business: Learning-as-participation is therefore frequently guided—mediated or scaffolded by various actors and structures; (3) Learning/SLA is public: One crucial way learning is guided is by externalizing its object while focusing the learner’s attention (Schmidt 2001) on it. That is, guides focus learners on particular learner-world relationships as relationships to be learned. These need to be learned because, as learner-world relationships, their learning enhances learner-world alignment. Ingold (2000: 37) calls such learning ‘education of attention,’ describing how, as they move through the woods with their elders, novice indigenous hunters are ‘educated’ to attend to specific environmental cues as markers of the disposition of their prey. In Churchill et al. 2010, we show a similar process at work as Ako interactively learns relationships among (English-specific conceptions of) past, finished worldly action, English simple past-tense forms, and simple past-tense-cueing adverbials through Tomo’s repeated use of ‘symbiotic gestures’ (Goodwin 2003)—gestures which serve to align attention, embodied action, and the sociomaterial world; (4) Learning/SLA is aligning, and learning to align: This final point summarizes the preceding three: Learning is a process of alignment—of continuously and progressively fitting oneself to one’s environment, often with the help of guides. I expand on this point immediately below in describing the final principle of a sociocognitive approach to SLA.

The alignment principle and SLA

Atkinson et al. (2007) define alignment as ‘the complex means by which human beings effect coordinated interaction, and maintain that interaction in dynamically adaptive ways’ (169). Here, ‘interaction’ includes action both between human beings, and humans and their non-human environments.

Regarding human interaction, sociality has alignment at its core: ‘A social relationship may be said to exist when several people reciprocally adjust their behavior to each other with respect to the meaning…they give…it, and
when this reciprocal adjustment determines the form...it takes’ (Weber 1922/1978: 30). Levinson (2006) proposes that humans possess a formidable ‘interaction engine’—an evolutionarily evolved set of capabilities which make us fundamentally interactive and social. Language is central here, but is built on more basic interactive abilities: (i) theory of mind—the ability to determine others’ intentions and what they know of our own; (ii) inferential skills—the ability to infer meaning beyond what is expressed; (iii) cooperative skills—the ability to coordinate behavior in achieving common goals; and (iv) interactive skills—means of responding in real time to attempts to interact, e.g. turn-taking, gaze, and bodily orientation. All four skill sets are fully displayed, according to Levinson, in mundane social interaction.

Levinson’s theory powerfully informs the concept of alignment: All animals interact socially, but nowhere near the level of humans. Yet Levinson gives insufficient attention to the public nature of interaction—the great contributions others’ actions-within-contextual-frames (i.e. actions vis-à-vis the sociomaterial environment) make to understanding them. Thus, requesting salt while gazing at your dining companion and gesturing toward the salt shaker shows your meaning substantially in your environmentally framed actions.

Regarding human–environment interaction, our online ability to align our behavior with our surroundings is largely what keeps us alive. This includes learning from those surroundings—finding exploitable features that enhance adaptivity. We also engineer our environments for adaptive advantage, as previously described.

But what does alignment mean for SLA? First, it means having formidable pre-existing capacities for interacting without necessarily sharing a language. That is, all language learners have powerful interaction engines supporting their learning at every turn. Second, SLA itself is a process of alignment—of learning the ‘differences that make a difference’ (Bateson 1972: 459) in the L2 environment. This is what our guides teach us as we engage with the environment: A requisite idiom or formula here, a form-function relationship there—to do that, you need to say this. By trying to align with our environment—by learning to behave in ecosocially adaptive ways—we become ‘enskilled’ (Ingold 2000: 5). Third, as noted above, alignment has a public face: Our aligning/learning behaviors extend into the world, and as our guides facilitate and respond to them, alignment is further externalized.

Ultimately, sociocognitive approaches to SLA are based on this tripartite premise: (i) Mind, body, and world are in continuous processes of interactive alignment; (ii) These processes are partly public; and (iii) In being public, they are learnable. Thus, if cognition is the site of learning, it is extended, embodied cognition that makes learning possible, at least in part.

EXTENDED, EMBODIED COGNITION-FOR-SLA: AN EXAMPLE

In this section I describe a case of extended, embodied—i.e. sociocognitive—cognition-for-SLA. The example comes from the videotaped EFL tutoring...
session featured in Pictures 1–4, where the participants are Ako, a junior high school student, and her aunt Tomo, an experienced EFL teacher. They are completing a grammar-exercise worksheet in preparation for high school entrance exams.

Two main forms of interaction occur during this session: grammar-focused interaction, conducted mostly in Japanese and in a fairly serious manner; and interaction around but not strictly focusing on the worksheet—this tends toward playfulness and sometimes features English. At a finer grain of analysis, various participation frameworks (Goffman 1981)—fluid configurations of actors, actions, roles/identities, contexts, and artifacts—are constituted, as described below.

Despite its mundane appearance, the grammar worksheet itself is a dynamic actor in this interaction—a highly designed cognitive technology (Clark 2001) for producing, comprehending, and learning L2 forms (Churchill et al. 2010). Its effective deployment assumes considerable background, e.g. in its four writing and two numeric systems, outline-like format, complex sequential organization, and the highly socialized behaviors by which such technologies are conventionally used. The grammar worksheet and skills needed to deploy it are therefore part of an academic ‘form of life’ (Wittgenstein 1958).

Example

In this example, Ako and Tomo are completing the final item in an exercise section devoted to the present perfect tense, wherein the last two items involve converting present-perfect statements containing time adverbials into ‘how’ questions cued by the adverbial. Figure 1 shows the (completed) item; the parenthesized directions read, ‘Make a question about the underlined part.’

The transcript (Figure 2) begins with Ako quietly reading the exercise prompt aloud (line 1). By actively externalizing her language processing into the sociocognitive ‘problem space’ between herself and Tomo, Ako focuses their distributed cognition—their shared attention and problem-solving—on the task at hand. Then, after a pause, Tomo repeats the underlined adverbial, thus further focusing their shared cognition on this key part of the prompt (line 2). Note, too, the grammar exercise’s role in: (i) initiating cognition by providing its content; and (ii) organizing cognition, e.g. by highlighting the adverbial. As the episode begins, cognition is therefore thoroughly extended and ecological, circulating across and through Ako, Tomo, and the grammar worksheet-as-cognitive technology.

Figure 1: Exercise item
A: ((Reads exercise prompt with flat intonation)) "I have been busy for two weeks."

T: For two [weeks

A: [How long?]

T: [><Ok ok.< (1.5) "Saeteru." ((Withdraws gaze and opens book before her))

Ok, ok. You're sharp!

A: ((Chuckles while writing answer)) (2.0) "How long?"

T: ((Holds up pen, glances at it while clicking it shut, then shifts gaze back to worksheet))

Sakki to onaji

Same as last one

(1.8)

A: Havu

T: Havu

A: I ((Tilts head to right))

T: >Nanka nukete nai?< Isn't something missing?

A: Have you been busy? ((Writes on worksheet))

(2.0)

T: >Nanka nukete nai?<

Isn't something missing?

A: Been=

T: =Been un

Right

A: [(Been busy un. (3.5) ((Speaks with great feeling)) "For a long time nei? (2.8) Kotoeta no kizuita. ((L eans toward A, laughing slightly)) I have been busy for a long time." Obeen busy, right. For a long time, right? Did you realize that I answered? I have been busy for a long time.

A: ((Nods and chuckles while writing)) For a long time 'kaku no ka to omotta" I thought I was supposed to write 'for a long time.'

T: ((Laughingly)) >Chigau chigau chigau< How long have you been busy. ((Points to A with pen)) (3.8) Chittomo I have never been busy desho. (1.5) How long have you been busy. (5.5) ((Pokes A with pen)) "Ako-san."

No, no, no! How long have you been busy? 'Never ever, I have never been busy,' right? How long have you been busy, Ako-san?" (1.5)

A: ((Looking at exercise sheet)) I havu been busy for two/i/ weeks ((Shifts gaze to T and laughs energetically))

T: Nande nani de, for whatto. (5.5) How come [why


A: I am sleep

T: Sleep de, hai.

Ok, from sleeping

A: Yes [yes ((laughs))

T: ["Kekkoo de gozaimasu doozo" ((taps pen on next item on worksheet)

That is fine, please proceed.

Figure 2: Example transcript
In line 3, Ako overlaps Tomo’s repetition of the adverbial as she tentatively converts it into *How longu?* Here, the overlap suggests highly aligned intercognition, while the final rising intonation seeks cross-cognitive regulation—Tomo’s confirmation. That Ako receives this even as she speaks—Tomo overlaps with *Ok ok*—again suggests high-level alignment. Tomo then adds an under-the-breath compliment, *Saeteru* (‘You’re sharp!’ line 4), and Ako, chuckling, writes *How long* on the worksheet and orally repeats it. Simultaneously, Tomo shifts her gaze to the book she has just opened.

In line 6, Tomo returns her gaze to the worksheet and says *Sakki to onaji* (‘Same as last one’), referring to the immediately preceding, just-completed exercise item, whose answer reads *How often has he read this book?* Tomo thus invokes Clark’s 007 Principle (2001 and above): If the answer already exists in the environment, why waste cognitive resources puzzling it out? Ako assents by offering *havu*, thus importing the word order from the just-completed item while retaining the auxiliary verb inflection from the current one. Receiving Tomo’s confirmation (line 8), Ako tries the same procedure in her next guess, venturing *I* for the inverted subject pronoun. This time, however, Tomo responds more complexly: Starting with a confirmation (*Un*), she continues by suggesting *you* as more ‘natural’ than *I*—i.e. that *How long have I been busy?* is a low-frequency usage. She then concludes with *have you how long have you:.?* Through this response Tomo performs multiple sociocognitive tasks: (i) confirming the partial correctness of Ako’s answer; (ii) suggesting a better answer; (iii) justifying the better answer; (iv) combining this answer with Ako’s immediately preceding answer, yielding *have you*; (v) summarizing their progress to this point with *how long have you*, thus stabilizing their shared cognition and providing a sociocognitive base from which to proceed; and (vi) via the elongated vowel and final rising intonation of *you:.*, prompting Ako to proceed.

In line 11, Ako repeats the answer achieved to this point, *Have you:*, pauses briefly, and then ventures *busy?* Here, vowel lengthening, internal pausing, and final rising intonation mark a cognitive state of uncertainty. A fast-paced turn-sequence ensues: Tomo answers with *Nanka nukete nai?* (‘Isn’t something missing?’), Ako offers an assertive-sounding *been*, Tomo responds with a latched repetition and agreement marker (line 14), and the sequence concludes with in-unison production of the mutually constructed answer, *been busy*, to which Tomo adds *un* (‘right’). The highly synchronized and rhythmic nature of this exchange is hard to capture in words, but it is an important additional way in which language, thought, and action are deeply aligned here (see also lines 2–4 and 20–24).

At this point, the full answer to the exercise item has been negotiated, and Ako is busy writing it down. The interaction now undergoes a marked shift in tone and participation framework. This begins as Tomo, after a long turn-internal pause, answers the question they have constructed, *How long have you been busy?*, with a quiet but emotion-filled *For a long time ne?* (line 16). In this period, Tomo was desperately trying to balance work, graduate
school, mothering, household duties, and caring for a declining parent-in-law, so this statement sounds like a heartfelt comment on her life. Her subsequent actions support this interpretation: She leans toward Ako, and with a hint of laughter quietly asks, Do you realize that I answered? She then states the full form of her answer: I have been busy for a long time (line 16).

In this turn-at-talk, Tomo performs several actions relevant to sociocognitive SLA. First, she imbues the lifeless grammar item with real-life meaning. Second, she apparently tries to engage Ako in real-life communication, as suggested by the highly personal nature of her talk, her new and more intimate physical orientation, her use of the alignment-seeking particle ne (Maynard 1993), her subsequent behavior (see below), and by the fact that she does so elsewhere in the tutoring session (Atkinson et al. 2007: 180–2). Third, Tomo expresses her feelings in English—English is thus now being used to communicate rather than to produce grammatical forms. Fourth, she seems to abdicate the teacher/tutor role, becoming an interlocutor with a real need to talk. Larger implications of these points are discussed below.

Ako responds by nodding and chuckling as she finishes writing the answer, and then, eyes still on worksheet, leans back and replies, For a long time kaku no ka to omotta (‘I thought I was supposed to write ‘for a long time,’’ line 17). This elicits an animated Chigau chigau chigau (‘No, no, no!’) from Tomo, said quickly with laughter and tight shakes of her head. Ako responds with laughter as Tomo leans close and asks, How long have you been busy, bringing her pen into Ako’s field of vision above the worksheet and pointing at her with it. Tomo then pauses, apparently expecting an answer, but none comes. Next, Tomo extends her utterance by animatedly answering for Ako—Chittomo I have never been busy desho (‘‘Never ever, I have never been busy’’, right?’ line 18)—probably teasing Ako about her lazy reputation. She then pauses, again waiting for Ako to respond. This is made clear in Tomo’s subsequent actions: She restates the question while poking Ako’s arm with her pen, and then follows with Ako-san, a form of address notable for its distance-indexing formality given their close relationship. As such, it serves as the verbal counterpart of the pen-poke—a strong bid for Ako to respond.

Finally, after a longish pause and still looking at the worksheet, Ako replies I havu been busy for twoil weeks (line 19), gazes into the middle distance, and then shifts her gaze to Tomo with a flip of her hair while laughing energetically. Tomo responds by leaning away from Ako, laughing, and covering her face with one hand. As their laughter subsides, she queries Nande nani de, for whatto (‘Why? From what? For what?’ line 20), pauses, and finishes with How come why as she brings herself vertical and leans back toward Ako. Ako overlaps with I am sleep, Tomo responds Sleep de, hai. (‘Ok, from sleeping,’ line 22), and Ako laughingly confirms, Yes yes, just as Tomo shifts the participation framework by sitting up straight, overlapping with a quiet and super-polite Kekkoo de gozaimasu doozo (‘That is fine, please proceed,’ line 24), and tapping her pen on the worksheet’s next item.
Implications of example for extended, embodied SLA

Analyzing interaction in detail yields a radically different picture of learning activity than the mainstream/cognitivist version. Here, I introduce four implications of the foregoing analysis for SLA.

Engagement

In cognitivism, interaction merely conditions input for the information-processing at the heart of SLA (Gass 1998). Following Chomsky, language is thus seen as incidentally interactive and social, but fundamentally internal and mental. The present approach, by contrast, sees interaction as SLA’s basic purpose and ground: Instead of isolating language in cognitive space, we wear it on our sleeve, so to speak, because it helps us live in the world. Engagement is therefore a better metaphor than internalization for conceptualizing SLA sociocognitively.

There is ample evidence of engagement in the example. Consider lines 16–24, where Tomo refocuses the interaction on their own lives. Although Ako is slow to participate fully, Tomo nonetheless engages her in multiple ways: She introduces her own plight in line 16, then leans toward Ako, checking to see if she has noticed the participation framework shift. On receiving an ambiguous response, Tomo teasingly answers for Ako (line 18), gives her a pen-poke, and uses an honorific suffix that momentarily alters their relationship. Ako responds, apparently using the grammar exercise to help her, and a brief conversation ensues, with Ako sharing a detail from her own life (I am sleep) and/or joining in the game of being called lazy. Then, in line 24, Tomo shifts the participation framework—and the role/identity construals within it—back to grammar-exercise completion, using the Japanese politeness system again in concert with body orientation and gesture.

Interaction

From a cognitivist perspective, L2 learners are viewed in terms of what they lack—the full and fully internalized L2 system. A sociocognitive approach, by contrast, sees learners as already possessing enormous interactive capabilities—interaction engines, in Levinson’s (2006) terms. In this view, language is just the apex of the interaction pyramid, though no less important for that. Yet the mainstream cognitivist perspective, by focusing solely on language, makes SLA appear more mysterious and difficult than it may actually be.

The interaction engine is fully displayed in the example—both the grammar-exercise answer and ensuing conversation are highly co-constructed. Certainly, Tomo scaffolds the interaction in various ways, including by using Japanese, yet without pre-existing interactive abilities there would be little to scaffold. I am not suggesting that cognitivists deny such capacities, but they
receive scant attention regarding SLA. This is doubtless because interaction is viewed in a limited, instrumental way.

Visibility

For cognitivists, language learning is invisible (Doughty 2003). Yet if cognition occurs not just in but between people, and between people and their sociocognitive environments, then it is also in the world. People learn, from a sociocognitive perspective, by participating in extended cognition. When, in lines 3–11, Ako guesses How longu?, Tomo replies Ok ok. Saeteru...Sakki to onaji (‘Ok ok. You’re sharp... Same as last one.’), Ako ventures Havu, Tomo repeats it, Ako guesses I, Tomo counters with have you how long have you, and Ako responds Have you; busy?, we see a worldly process of co-cognition, irreducible to happenings in individual heads. Although impossible to prove from a cognitivist perspective, where evidence is represented by post-hoc artifacts, especially on tests of conscious knowledge (Doughty 2003), what we see in such sequences is learning-in-flight—the learner engaging with the object of learning in the public, sociocognitive world. This view has novel implications for second language learning and teaching (Atkinson et al. 2007: 183–5).

Alignment

The preceding implications are central aspects of the more general concept of alignment. As argued previously, alignment is the basis of social life, and beyond that, of adaptivity to the environment. Views of SLA which ignore alignment neglect a fundamental aspect of humanity, as well as reasons for which and actions by which a fundamental human activity—learning—must occur.

CONCLUSION

Bakhtin’s dramatic vision is close to the one assumed in this article:

To be means to be for the other, and through him [sic], for oneself. Man has no internal sovereign territory, he is all and always on the boundary; looking within himself, he looks in the eyes of the other or through the eyes of the other... I cannot do without the other; I cannot become myself without the other; I must find myself in the other, finding the other in me (in mutual reflection and perception). (1979, quoted in Wertsch 1998: 116)

For all its literariness, this passage captures the fundamental extendedness of human existence. For almost 400 years now, Descartes’ separation of mind from everything else has mesmerized Western thought, and cognitive science was one result. A self-professed Cartesian—Chomsky—is the foremost modern exponent of this view, but he was also part of a larger movement. Cognitive scientists have now begun to question their field’s founding premises, and a
richer, more complex interdiscipline is emerging. They increasingly find that understanding the mind/brain means studying it in the body, and understanding the embodied mind means studying it in the world; and this is simply because the mind is in the body and the world. If SLA studies is a cognitive science—or seriously desires to become one—shouldn’t it follow suit?

ACKNOWLEDGEMENTS

I would like to thank Paul Kei Matsuda, Jinju Nishino, and Yumi Takamiya for expert technical help, Takako Nishino, Hanako Okada, and Yumi Takamiya for expert language advice, and Paul Bruthiaux, Ulla Connor, Paul Kei Matsuda, Takako Nishino, Hanako Okada, Tetyana Smotrova, and Beril Tezeller Arik for thoughtful comments on earlier drafts of this article.

NOTES

1 There is currently debate over whether all forms of cognition are extended and embodied, or just some (e.g. Wilson 2002). In one sense, at least, the former is obviously true—all cognition takes place in human bodies embedded in sociomaterial worlds. If it is further argued, as here, that cognition is fundamentally adaptive, then even off-line cognition must be influenced by that fact—i.e. if cognition exists to help us adapt to/align with our worldly environments, then it is essentially extended and embodied. At the same time, cognitive extendedness and embodiment obviously vary in degree depending on the situation. Such flexibility is a hallmark of human cognition.

2 This summary (and associated terminology) is based on standard historical accounts of cognitive science (e.g. Boden 2006), embodied cognition (e.g. Barsalou 2008), and studies of embodied language use (e.g. Glenberg 1997).

3 The presentation of pictures in this subsection is inspired by Rogoff (2003).

4 In the version of these pictures reproduced for publication it is virtually impossible to see Ako’s pen, except perhaps in Picture 4.

5 I regard cognition and learning as broadly continuous, in the same way that connectionists do: Each act of understanding updates/augments the organism’s (socio)cognitive state so that there is continuous ecological adaptation/alignment (Atkinson et al. 2007).

6 This is one sense in which the present approach differs from not only cognitivist SLA studies but also sociocultural theory (e.g. Lantolf and Thorne 2006), with which it also has affinities. See Atkinson (2002: 537–38) for further discussion.

7 Although this approach shares an emphasis on the public nature of cognition with conversation analysis-for-SLA (Mori and Markee 2009), it differs in that the latter tends to reduce cognition to public performance, whereas I view cognition as operating in integrated sociocognitive space (Atkinson 2002; Atkinson et al. 2007).

8 The example is transcribed using the system developed by Gail Jefferson (Ochs et al. 1996: 461–5).

9 Tomo’s comment is marked as ‘emotion-filled’ by: (i) a ‘catch’ or sudden huskiness in her voice; (ii) the long turn-internal pauses bracketing this comment; (iii) slowness of speech; and (iv) quietness.

10 Different approaches to language acquisition have different ways of
defining ‘acquisition.’ Here, Bickerton’s (1981) criterion of first appearance of a linguistic feature in the learner’s speech might be more relevant.

REFERENCES


Reiser, J., A. Garing, and M. Young. 1994. ‘Imagery, action and young children’s spatial orientation: It’s not being there that counts, it’s what one has in mind,’ *Child Development* 45: 1043–56.


