Lang. Teach. (2014), **47.4**, 467–483 © Cambridge University Press 2013
doi:10.1017/S0261444813000153 First published online 3 April 2013

Plenary Speeches

Language learning in mindbodyworld: A sociocognitive approach to second language acquisition

Dwight Atkinson Purdue University
dwightatki@gmail.com

Based on recent research in cognitive science, interaction, and second language acquisition (SLA), I describe a sociocognitive approach to SLA. This approach adopts a NON-COGNITIVIST *view of cognition: Instead of an isolated computational process in which input is extracted from the environment and used to build elaborate internal knowledge representations, cognition is seen as* ADAPTIVE INTELLIGENCE*, enabling our close and sensitive* ALIGNMENT *to our ecosocial environment in order to survive in it. Mind, body, and world are thus functionally integrated from a sociocognitive perspective instead of radically separated.*

Learning plays a major part in this scenario: If environments are ever-changing, then adaptation to them is continuous. Learning is part of our natural ability to so adapt, while retaining traces of that adaptation in the integrated mind-body-world system. Viewed in this way, SLA is adaptation to/engagement with L2 environments.

INTERACTION *also plays a central role in sociocognitive SLA: We learn L2s through interacting with/in L2 environments. Founded on innate, universal skills which evolutionarily* PRECEDED *language and make it possible, interaction supports SLA at every turn. Having presented this argument, I illustrate it by analyzing a video clip of an EFL tutoring session, indicating various 'sociocognitive tools' for interactive alignment which undergird L2 development.*

1. Introduction: Alternative approaches to SLA

Let me begin with a shocking and outrageous statement – the kind sometimes used to provoke audiences at the beginning of talks: All the SLA theories we have, all the approaches to SLA you and I know, support, and possibly even cherish, are WRONG. They're wrong in THIS sense – no single theory can do justice to the dizzyingly complex and multifaceted phenomenon we call SLA. As a result, each theory provides at best a PIECE of a larger and much more complex puzzle.

This is, I take it, the main argument for ALTERNATIVE APPROACHES to SLA – SLA theories which come from a range of disciplinary, conceptual, and methodological directions, but which all share a fundamentally non-cognitivist, and therefore non-mainstream,

Revised version of an invited presentation made in the University of Wisconsin Language Institute Speaker Series: Language, Cognition, and Sociality. Madison, Wisconsin, USA, 17 October 2012.

https://doi.org/10.1017/S0261444813000153 Published online by Cambridge University Press

orientation (Atkinson 2011a). While non-cognitivist approaches have (arguably, at least) always had a place in the field, the 1990s attempt at 'theory culling' – limiting SLA theorizing to a highly cognitivist core (e.g. Beretta 1993) – provoked the development of a range of counter-theories, as captured in Lantolf's (1996) call for 'Letting all the flowers bloom!'

The alternative approach to SLA I'd like to describe today is the one I've been involved in developing: THE SOCIOCOGNITIVE APPROACH. As the name itself implies, its distinctive claim is that individual mind and ecosocial world aren't radically separated but rather functionally integrated, and that this functional integration offers one key to understanding SLA. Given the cognitivism pervading the SLA field (Atkinson 2011b), I realize that this claim may sound odd, if not nonsensical; my main aim here is thus to develop a plausible conceptual basis for the sociocognitive approach.

2. Interaction: Three views

To accomplish this task, let me begin indirectly: by briefly considering three views or versions of human interaction. Interaction is at the heart of a sociocognitive approach to SLA, but to get there I need to present two contrasting views of the phenomenon first.

2.1 Robot interaction

This first version of interaction emanates from monumental efforts in artificial intelligence to develop robots, or CHATBOTS, which emulate human interaction. The Holy Grail in these efforts is a chatbot that passes the Turing Test – that is, which leads human beings to believe they're communicating with another human. Without further comment, then, let me show you a short video of two chatbots conversing with EACH OTHER – not exactly what they're designed for, as just noted, but still a fairly good indicator of what kind of interactors they are. As you watch the video, I'd ask you simply to note what, in your judgment, is particularly humanlike and particularly nonhumanlike in this interaction. The video can be found at: www.youtube.com/watch?v=WnzlbyTZsQY- please watch it before continuing to read. (For further information about this interaction, see www.npr.org/2011/09/01/140124824/robot-to-robot-chat-yields-curious-conversation and http://creativemachines.cornell.edu/AI-vs-AI.)

Now that you've watched the video, here's my own comment: One thing that strikes me about this interaction is its largely non-cooperative character. The two participants – or should we call them contestants? – seem to be playing the linguistic version of competitive (versus cooperative) ping pong: They hit the ball at each other, but not in ways designed to make it easy to return. As cooperative conversationalists, then, these chatbots don't seem wildly successful.

https://doi.org/10.1017/S0261444813000153 Published online by Cambridge University Press

2.2 Interaction in mainstream SLA

The second version of interaction I'd like to briefly describe is the one taken by the INTERACTIONIST – or INPUT AND INTERACTION – school of SLA, regarded by many as the single most influential SLA approach currently on offer (Block 2003). This approach is highly cognitivist in that it treats the brain as an information-processing device, or cognitive computer. Here, interaction is viewed as a form of input PRE-processing – through input 'modification' or 'negotiation' via interaction, input becomes comprehensible and therefore available for further processing by the cognitive computer. This view also underlies the current focus on error feedback/correction in instructed SLA.

Gass's (1997: 3) influential model of the SLA process clearly indicates the role of interaction in this approach. Arranged vertically, the model has 'input' at its top and 'output' at its bottom, separated by dotted lines from the model's eight components arrayed sequentially in between. Of these, 'Negotiation –> NS [i.e. Native Speaker] Modification' is the only component that is not mind-internal, with the final component, 'Personality, Mode (Oral/Written), Situation', being a mixed case. Despite the centrality of the IDEA of interaction in this approach to SLA, then, it actually occupies a very small part of the overall model, the vast majority being mind-internal.

2.3 Interaction in a sociocognitive approach to SLA

It's not my intention to argue that the two versions of interaction just reviewed are wrong in any categorical sense; certainly, help from all quarters is needed in understanding something as complex as interaction or SLA. I would suggest, however, that the robot version seems to underestimate the depth and complexity of human interaction – two human beings designed by evolution (see below) to interact/talk WITH each other is not the same as two machines designed by humans to talk AT each other – while the mainstream SLA version may undersell interaction in L2 development by giving it too limited a role.

The third version of interaction I want to discuss, then, is one that treats it more complexly, and one that can give it a more robust role in SLA. Luckily, human interaction – including but not limited to linguistic interaction – is a topic of serious interest among academics these days, especially those investigating the roots of human sociality (e.g. Goodwin 2000, 2003; Enfield & Levinson 2006; Levinson 2006; Schegloff 2006; Tomasello 2008, 2009). I depend on their work in what follows.

There's little doubt that human interaction has been THE 'primordial site of [human] sociality' (Schegloff 2006: 70) over our species history, which began, by recent estimates, between two and eight million years ago. Schegloff notes that even as countless social structures (e.g. tribes, political and economic systems, and whole complex societies) have arisen and disappeared across human history, interaction as a social structure has continued largely unaffected. And until the advent of the telephone, at least, FACE-TO-FACE INTERACTION must have been THE main means by which new tribes, political systems, and complex societies were established. I'd like to make three points here: First, if we've been engaging in close,

https://doi.org/10.1017/S0261444813000153 Published online by Cambridge University Press

usually face-to-face interaction for from two to eight million years, then we've probably become quite good at it – particularly if such interaction comprises a highly evolved, largely innate set of skills. In comparison, most of humanity has been interacting by telephone for 50 years at most, and by computer for 10 years at most (although much of humanity is probably still not online) – thinking in this way, we get a sense of just how highly developed face-to-face interaction may be. Of course, this is an unfair comparison since telephone and computer communication rely largely on the same interactive abilities developed through face-to-face interaction over the ages, so they're not separate things.

Second, our high-level interactive abilities are probably the ultimate skill set separating us from other animals. Based on extensive research comparing apes with human infants, for instance, Michael Tomasello (e.g. 2008) has shown that even at three months infants display interactive abilities far outdistancing those of their primate cousins – abilities that can only be innate given how early they appear.

Third, language in this view is the crowning achievement of human interaction rather than its underlying basis – our sophisticated interactive abilities are the FOUNDATION ON WHICH language is built rather than vice-versa. We simply couldn't have developed symbolic communication without robust, preexisting interactive abilities, and language in use still relies on those abilities at every turn (Levinson 2006; Tomasello 2008).

The picture of interaction I want to adopt here, then, presents it as the fundamental basis for human sociality, and that includes language learning. In the words of the pioneer sociologist Max Weber, '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' (1922/1978: 30). Compared to the chatbot conversation shown earlier, interaction in this sense is like a cooperative (versus competitive) ping pong game, in which two or more partners coordinate or align their activities sensitively and ongoingly for their mutual benefit. And they do so in at least three ways.

First, the participants align COGNITIVELY, as interactants become intercognizers. That is, they use their profound, evolutionarily well-developed skills of SHARED INTENTIONALITY (Tomasello et al. 2005), THEORY OF MIND – i.e. moment-by-moment insight into what the other is thinking (Premack & Woodruff 1978), COMMON KNOWLEDGE – knowing what the other is thinking about what you're thinking (Lewis 1969), and cross-cognitive ANTICIPATION (Kinsbourne & Jordan 2009) whereby fluent social activity is enabled to 'go on' without perceptible breaks. The combined result of such abilities has various names: SHARED COGNITION and INTERSUBJECTIVITY are two of them.

Second, partners in interaction align AFFECTIVELY, minimally in the sense that the activity itself is presumably mutually beneficial, but typically in much deeper ways. Thus, research across fields reveals the many public signs of affective and affiliative alignment in cooperative human interaction, from enhanced eye contact (e.g. Kendon 1990), to rhythmic entrainment of body movement (e.g. Collins 2004), to mutual smiling (e.g. Chartrand & Bargh 1999), to linguistic accommodation (e.g. Ylanne-McEwen & Coupland 2000). Once again, we don't just affectively attune ourselves to the other in an ongoing, highly sensitive way, but additionally engage in 'attunement to the attunement of the other' (Rommetveit 1998: 360). The discovery of the HUMAN MIRROR NEURON SYSTEM, in which we neurally enact animate

https://doi.org/10.1017/S0261444813000153 Published online by Cambridge University Press

actions we are merely witness to (Rizzolatti & Sinigaglia 2008), may provide one powerful key to understanding how human beings align naturally in the affective domain. That is, I may literally 'feel your pain'.

Third, and finally, interactants align with each other PHYSICALLY in moment-to-moment, highly sensitive ways, with variations in alignment signifying varying roles in and degrees of commitment to the interaction. The natural history of an interaction can even be traced in terms of physical movements, the original coming together, for example, being marked by mutual presentation of a maximum of physical symptoms through a literally face-to-face engagement, often including touch. From these beginnings, an interaction might adopt an 'L-structure' (Kendon 1990) wherein interactants stand at right angles to each other, thereby allowing themselves physical and perceptual access to both their interlocutor and what (and who) else might be happening (or present) in the environment. Alternatively, the full face-to-face orientation might be maintained. Finally, as interactants prepare to withdraw, they make subtle adjustments of proximity and orientation to signal the approach of an ending, although such endings are typically negotiated and contingent on the other's acquiescence, usually at a subconscious level. But the main point I want to prepare the ground for here is simply this: Interaction underlies and supports human social behavior in profound and complex ways, and if language learning is a form of social (or, preferably, sociocognitive) behavior, then interaction supports language learning as well.

3. A sociocognitive approach to SLA

A sociocognitive approach to SLA is therefore one that sees social interaction/alignment as the main dynamic or 'engine' (Levinson 2006) of SLA. Our immensely powerful and evolutionarily refined interactive abilities, which precede language and make its effective use possible, also underlie and support the learning of additional languages. By bringing our behavior into attunement with that of social others – in the case of SLA with speakers of the language to be learned (or at least social artifacts produced by such speakers, like books or websites – see Sections 3.4 and 3.5 below) – SLA is enabled and supported. Without appeal to such formidable interactive abilities, the process of language acquisition in general, and SLA in particular, becomes much more mysterious and difficult to explain.

Gregg (1988) warned that describing SLA from a particular perspective without first specifying the meaning of 'L' (language) and 'A' (acquisition) assumed in that perspective is a haphazard endeavor. Heeding Gregg's warning but extending and modifying his list, I'd now like to answer five questions from a sociocognitive perspective: (1) What is COGNITION for? (2) What is ACQUISITION/LEARNING for? (3) What is LANGUAGE (and language learning) for? (4) Why do we learn SECOND/ADDITIONAL LANGUAGES? and (5) HOW do we learn second/additional languages? By answering these questions, I intend to provide at least a rough conceptual base for sociocognitive SLA, in each case focusing on the capitalized concept. The fifth and final question – concerning the 'HOW' of SLA – is in fact the area in which most sociocognitive SLA research has been conducted. I'll introduce an example from that research to conclude my talk.

https://doi.org/10.1017/S0261444813000153 Published online by Cambridge University Press

3.1 What is COGNITION for?

In a sociocognitive approach to SLA, human cognition is reconceptualized as ADAPTIVE INTELLIGENCE (A. Clark 1999; Atkinson 2012) – as having the major evolved purpose of promoting adaptation to our ecosocial environment. The argument goes as follows: Organisms which don't adapt to their environments typically die, and rather quickly; for this reason, most motile zoological organisms have at least proto-nervous systems. Single-cell organisms like the amoeba, for example, while lacking articulated nervous systems, have sensory organs that allow them to detect, on the one hand, danger (e.g. for amoebae, noxious substances in their aqueous environment) and, on the other, affordances (e.g. food sources). More advanced, multicellular organisms have correspondingly more advanced environmental sensing/nervous systems, with the 'top of the line' in the animal kingdom being the human nervous system, which is centrally based on our highly developed brain. Brains are where our millions of neural pathways converge, leading, at least theoretically, to the combination of every bit of sensory information with every other bit. The overall result is fine-grained, intimate, ongoing attunement to the ecosocial environment, at least where needed for survival.

While few if any cognitivists would deny that the brain has evolved for adaptive purposes, the cognitivist program has pushed biological adaptivity to the margins in an effort to develop the engineering-centered view that cognition operates like a digital computer (e.g. Boden 2006; Noë 2009). Cognitivists, in other words, treat cognition as the logic-based processing of abstract symbols and building of abstract representations, from which environmental actions and solutions (if any – see Chomsky's (1965) influential restriction of linguistic theory to the realm of competence) are then computed. Luckily, however, the ground has been shifting in the cognitive sciences: Instead of disembodied computation, biological/adaptive approaches to cognition emphasize a different skill set – one focusing on ENVIRONMENTAL PATTERN FINDING based on the organism's deep and interactive embedding in its environment. A. Clark (1997) states the argument in humorously memorable form: Humans are 'good at frisbee, bad at logic' (p. 60), where 'logic' denotes the computational representation and solution of logic-like problems in a cognitivist model of cognition, while 'frisbee' denotes complex pattern-finding abilities originally associated with evolutionarily crucial survival skills like hunting. Significantly, computational modeling of such skills has proven difficult, and a number of alternative cognitive theories, all emphasizing the ways in which minds orient bodies to adaptively align with the world, have arisen. Examples include 'motor resonance' enabled by the human mirror neuron system (Rizzolatti & Sinigaglia 2008); adaptive, intuitive 'rules of thumb' (Gigerenzer 2007); low-dimensional environmental sensing which 'use[s] the world as its own best model' (Brooks 1991: 139); and direct perception of environmental affordances (Gibson 1979). Such approaches vary in the degree to which they depend on internal representations, but Shanon's (in Glenberg 1997: 3) statement provides a general guide:

> It appears that the underlying substrate of mental activity is not a repertory of well-defined, well-structured abstract symbols, and that the workings of mind cannot be generally characterized as the computational manipulation of such symbols. Rather, the substrate in which mental activity takes place should be one that meets the following requirements: It should afford maximal sensitivity to unspecified dimensions and

https://doi.org/10.1017/S0261444813000153 Published online by Cambridge University Press

distinctions, it should be context-sensitive, and it should be embedded in the framework of the organism's action in the world.

3.2 What is LEARNING for?

Cognitivist approaches to learning seem to see it as the accumulation and structuring of knowledge in isolated cognitive space – a so-called competence view of learning (e.g. Long 1997). This view also seems to underlie a substantial part of our educational practice – knowledge is to be squirreled away in the mind and then tested. It further seems to account for what is perhaps the reigning methodology in SLA research: An experimental group is exposed to an artificially high dose of some linguistic feature or other, and then given a discrete-point test on that feature to see what the subjects retain. How this accords with a view of acquisition as a largely implicit, gradual, non-linear developmental process is usually left unclear (Doughty 2003).

The approach to learning adopted in sociocognitive SLA, on the other hand, is that, like cognition itself, learning is primarily adaptive – our formidable learning abilities have evolved to help us more sensitively and efficiently adapt to our environments. Learning can therefore be defined, at a first approximation, as the ability to adapt to our environment BETTER (than formerly) through progressive interaction with/experience in and alignment with that environment. As a spectacularly mobile and wide-ranging species, humans need to adapt to a spectacular range of environments, as well as to retain traces of that adaptation in the mind-body-world ecology. In Wenger's (1998) words, learning is therefore 'as much a part of our human nature as eating or sleeping' (p. 3).

3.3 What is LANGUAGE (and language learning) for?

Arguably THE major means by which humans adapt to and exist within their ecosocial worlds is language. While research on language evolution is inevitably speculative, this much seems clear: For language to have evolved as both/either part of the human genome (Hauser, Chomsky & Fitch 2002) and/or part of our human cultural heritage (Tomasello 2003), it must first have provided us with a selective adaptive advantage. While some cognitivists see the resulting phenomenon as having primarily an internal function – Chomsky (in Andor 2004), for example, believes that the primary function of language is to mediate cognition – the most obvious adaptive advantage would have been in SOCIALLY mediating the relationship with our natural world. In other words, language has provided human beings with a selective advantage over our predators and competitors by allowing us to work together more effectively in groups. How else could such weak, ecologically unspecialized creatures have survived and prospered in competition with predators and others which were stronger, faster, or better at climbing trees?

To put it differently, language is for social action. If so, then language learning means learning how to operate adaptively in our social environments, which have the primary role of mediating our relationship with nature. From this viewpoint, learning a natal language is

https://doi.org/10.1017/S0261444813000153 Published online by Cambridge University Press

not first and foremost learning linguistic forms, although these are clearly important. Hopper (1998) put it radically: 'Children do not seem to learn sentences, but rather. . .to adapt their behavior to increasingly complex surroundings. It is an enterprise that does not stop at the age of 6, but continues throughout a lifetime' (pp. 161–162). My own preferred version would be that children learn language, which certainly includes and is based on multiple formal systems, yet the purpose of such learning is to enable children not to become linguists, but to survive and prosper in the world. Even today, humans without language live exiguous existences, made possible in most cases only through massive social support. Until quite recently in our species history they must not have survived at all.

3.4 Why do we learn SECOND/ADDITIONAL LANGUAGES?

Second/additional languages have the same evolved function as first languages: They support adaptation, or social action-for-survival, but in L2 ENVIRONMENTS – settings in which the L2 is a primary medium of such action. From this viewpoint, human beings learn L2s naturally, at least to the degree needed to function adaptively in L2 worlds. This is why so much of the world's population is multilingual, at least to the level people need to be. The monolingual bias of SLA studies (e.g. Firth & Wagner 1997) has played a large role in supporting the apparent belief that L2 learning takes place primarily through formal education.

The paradigm example of an L2 environment is obviously a city or country where the L2 functions as the primary linguistic tool for adaptive action in everyday life. Just as obviously, however, there are many other possible L2 environments: Intense personal relationships, well-designed language classrooms, the internet, music, or an engrossing novel can all be rich and supportive environments for SLA. Interaction as I use the term should thus be extended beyond its face-to-face version to include a wide range of social activities of more mediated types. While face-to-face interaction may have special qualities that make it ideal for some kinds of learning, any social action or artifact that promotes interactive engagement with/in L2 environments potentially provides support for SLA.

3.5 How do we learn second/additional languages?

The correct answer to this question can only be 'in many ways'. As mentioned at the beginning of my talk, no theory can claim to be anything like a total account of the complex and multifaceted phenomenon known as SLA, and most are very partial indeed. In this sense, Cook & Seidlhofer's (1995: 4) vision of the complex nature of language must hold equally for SLA:

> [Language is] a genetic inheritance, a mathematical system, a social fact, an expression of individual identity, an expression of cultural identity, an outcome of dialogic interaction, a social semiotic, the intuitions of native speakers, a collection of memorized chunks, the sum of attested data, a rule-based discrete combinatory system, or an electrical activation in a distributed network. . .We do not have to choose. Language can be all of these things at once.

Within this scheme, my main claim today has been that interactive alignment, the means by which and processes through which individuals build moment-to-moment social relations

https://doi.org/10.1017/S0261444813000153 Published online by Cambridge University Press

and cooperative social action, is a major facilitator of SLA. If the primary evolved purpose of language is to promote cooperative activity by which we adapt to and function in our worldly environments, then how we come to perform such action in L2 settings should be prominently included in SLA theory.

Cognitivist approaches to SLA see it as a largely invisible process – once the data enter the cognitive processor there is no direct way to study what is going on (Doughty 2003). A sociocognitive approach diverges from this view in positing that the SLA process is at least partly visible. That is, as individuals interact in the L2, and mutually adapt, adjust, and align their behavior in the course of interaction, signs or tools of their alignment become publicly available. If, for instance, while you and I are interacting in your second/additional language, I use a word you don't know as I point to or otherwise indicate an object in the environment, you have a pretty sound basis on which to associate the word with the object, and therefore possibly to learn that word or at least some incremental aspect of it (see Churchill 2008 for a detailed case study of such learning). Although a comprehensive list of sociocognitive tools of this sort does not yet exist, the following may indicate some of the possibilities:

1. *Gaze following*: Human beings know intuitively that if they don't understand what someone is saying, and the latter is referring to the here-and-now environment, then a good way to figure it out is to follow the other's gaze. Infants certainly know this – they learn to 'word the world' in part by following their caregivers' gaze while the latter talks about the environment (Baldwin et al. 1996). Yu, Ballard & Aslin (2005) have experimentally demonstrated the facilitative effect of gaze-following on word learning and segmentation skills in SLA.
2. *Gesture*: As mentioned above, pointing can be a powerful public index of what in the environment is being referred to linguistically. Yet there are many kinds of gestures providing many in-the-world affordances for SLA (e.g. Churchill et al. 2010; Gullberg 2013; Smotrova & Lantolf in press). Mori & Hayashi (2006) studied the use of 'embodied completions' – gestures completing a turn at talk, as when someone says 'I want' and then points to the desired object – in ways which are sensitively calibrated to the linguistic proficiency of L2 speakers. Although acquisition per se was not demonstrated in this study, it certainly captures the semiotic richness and availability of in-the-world gestural support for SLA.
3. *Facial expression*: Among the richest sources of affective information are the faces of social others. We use them continuously to interpret meaning in interaction. Although no SLA research focusing specifically on facial expressions exists that I know of, it strikes me as a potentially significant area of study: Smiles often accompany linguistic expressions of affiliation, for instance, while frowns do the opposite. Even in the absence of visual data – for example during phone calls or when listening to the radio – 'smile voices' can be reliably identified and therefore used as indexes of meaning (Drahota, Costall & Reddy 2008).
4. *Body positioning*: In Goodwin's (2003) words, 'The positioning, actions, and orientation of the body in the environment are crucial to how participants understand what is happening and build action together' (p. 20). Although not yet widely studied, the role of bodily orientation in SLA is suggested by Seo & Koshik (2010), who identified distinctive

https://doi.org/10.1017/S0261444813000153 Published online by Cambridge University Press

head actions used by ESL tutors to prompt repair of non-targetlike utterances by their tutees. These head actions appear to be natural, automatic, and largely unconscious, and to be substantially the same ones used to signal momentary lack of comprehension in native speaker interaction.

5. *Physical objects and artifacts*: Aspects of the physical (but often humanly designed) environment provide rich semiotic support for language learning and use, frequently allowing the possibility of forms of interaction – and therefore learning opportunities – that could not otherwise occur. To give just one example from a very wide range of phenomena: Churchill et al. (2010; see Section 5 below for related work) showed how an English grammar review worksheet used by a Japanese junior high school student provided a suite of sociocognitive affordances for learning, thus acting as a 'cognitive technology' (A. Clark 2001: 240):

> This dense, multiliterate technology is likely to play several crucial roles in learning and teaching. Most generally, it reduces a complex and messy reality – the relatively less determinate uses of verb tenses in real-world (often spoken) communication – to a 'simplified world' – a determinate and highly learnable system of relations among verb forms, adverbials, and semantic functions, all presented in written form. In doing so, it helps make the target syntactic relations publicly visible and available. . . More specifically, the cognitive technology of the worksheet affords a systematic work routine: Segregation of like items into sections and graduated levels of difficulty within each section allow students to proceed through the items in a predictable order. Similarly, teachers and tutors can use the worksheet's elaborate [numerical, alphabetical, and spatial] indexing system. . .for pedagogical purposes – for example, to establish cross-cognitive attention and action, to select items that illustrate a grammar point particularly well, or to provide instructive contrasts. (p. 240)

6. *Theory of mind and common knowledge*: These concepts – popular in psychology and cognitive science in recent years – have a range of meanings and applications. By THEORY OF MIND I mean what Tomasello (2008) calls 'mindreading', and what Levinson (2006) describes as 'intention attribution: given the observed behavior, the [interactor] must be able to infer likely goals that would have motivated that behavior' (p. 48). In other words, based on our innate, highly refined social abilities, we are often able to predict what the other is thinking, at least in so far as how they might proceed in an interaction. Accompanying such skills is the ability to ascertain COMMON GROUND, or COMMON KNOWLEDGE (Lewis 1969; H. Clark 1996) – knowing what the other is thinking about what you may be thinking yourself. Levinson gives the example of two people getting separated in a department store, and then determining a common place to reunite based on such knowledge. How theory of mind and common knowledge might facilitate SLA is not yet known, but if minds are in fact designed this way it could certainly simplify SLA processes: If innate abilities to understand the other's likely actions in the course of interaction exist, then it might be substantially easier to determine how the language being used is related to those actions.

7. *Interaction skills*: Schegloff (2006) describes the rich repertoire of apparently universal abilities humans employ to conduct interaction in orderly, comprehensible ways. Turn-taking, sequential organization, and repair, for example, provide a flexible range of tools for building interaction both linguistically and non-verbally – the two typically function together as part of a larger communicative whole – that at the same time provides a predictable character to social interaction. Without such 'interaction-management' techniques, both speaking and learning to speak would be substantially more difficult, if

https://doi.org/10.1017/S0261444813000153 Published online by Cambridge University Press

not impossible. Conversation analysts (e.g. Kasper & Wagner 2011) have begun to show how these universal features of interaction support SLA.

8. *'Natural pedagogy'*: It is widely agreed, at least in theory, that humans are natural learners. Recent research under the rubric of NATURAL PEDAGOGY (Csibra & Gergely 2009), however, has established the possibility that we are also natural teachers. That is, we appear to have innate abilities to structure our social activities in ways that make them more transparent, parsable, and therefore learnable, and we appear to do so especially in the presence of 'learners' – typically younger and less experienced members of our social groups. The phenomenon of FOREIGNER TALK, which Zuengler (1991) interpreted as a form of sociolinguistic accommodation (and therefore alignment), is a notable case in point as regards SLA.

9. *Activity frameworks*: Most cooperative human activities have conventional structures. While these are in part culturally shaped, it is likely that they also have universal attributes. Service encounters, for instance, seem to share characteristics the world over: There are two parties – seller and buyer; the buyer usually initiates the transaction; the seller responds, and the specifics of the transaction are then transacted; the seller and buyer then complete the transaction as the goods are delivered and the compensation paid. Obviously, this is not a perfect template, yet the number of options in service encounters is probably quite limited. The behavioral predictability and shared goal orientation that such activity frameworks provide may constitute a rich base of environmental support for SLA.

10. *Language*: In the foregoing list, I have purposely downplayed the role of 'the tool of tools' (Wells 1999: 9) for sociocognitive alignment and interaction – language itself. This is because language and interaction are too often elided, rather than viewing more basic and long-lived forms of interaction as fundamentally UNDERLYING language use. Yet as Krashen (e.g. 1985) pointed out some years ago, one of the best ways to make hitherto unknown linguistic forms known/comprehensible is to embed them in already comprehensible language. While I have alluded to examples of this phenomenon in my talk today, I will briefly cite two explicitly here: Codeswitching and/or L1 use can often provide rich support at the right moment for SLA.

4. Alignment and interaction in SLA: An example

I would like to conclude my talk with an example from sociocognitive SLA research – by analyzing a short video of an English as a foreign language (EFL) tutoring session for evidence of alignment in interaction. This example first appeared in Atkinson et al. (2007). It involves two people: 'Ako', a Japanese junior high school student working on a 22-page English grammar review worksheet assigned by her school to prepare her for upcoming high school entrance exams; and 'Tomo', a highly experienced EFL teacher and Ako's aunt, acting here as her tutor.

The grammar structures reviewed in the worksheet are those Ako was supposed to have mastered over her two-and-a-half years studying English at the junior high school. Here, Ako and Tomo are nearing the end of the worksheet section on present perfect tense/aspect, and as the video clip begins, they are just starting to work on an item which requires translating a

https://doi.org/10.1017/S0261444813000153 Published online by Cambridge University Press

Japanese sentence – *Anata wa ima made ni eigo de tegami o kaita koto ga arimasuka* – into its English equivalent, *Have you ever written a letter in English?* The grammatical focus of this item is the 'Have you ever' present-perfect experience construction (e.g. *Have you ever climbed Mt. Fuji?*), plus the correct form of the main verb. A transcript of the interaction, followed by a key to the transcription conventions used, is presented below; the video clip itself is located at: www.youtube.com/watch?v=hJhCP8t0YFQ&feature=youtu.be

Please watch the video at least twice while following along on the transcript – getting oriented may itself require a certain amount of alignment if you aren't familiar with conversation analysis transcription conventions (e.g. Ochs, Schegloff & Thompson 1997: 461–465) or Japanese.

Transcript: Have you ever written a letter in English?

01 A: ((Reads first part of exercise item quickly under her breath))
02 > °Anata wa ima made ni eigo de tegami° < (2.0) Ima made ni =
[Lit: You (Top) until now in English letter] *Until now*
03 T: ((Softly shadowing A's volume, intonation, and gestures))
04 = °Un ima made°
Right, until now
05 (1.0)
06 A: [Ne:ba:
Never
07 T: [>Sakki no tsukaeba iin janai?<
Why not use the one you used before?
08 A: Ima made (.8) n- e:va ka
Until now *ne- ever?*
09 T: Have you ever toka nantoka =
Have you ever blank
10 A: = Have you eva writu, written =
Have you ever write, written
11 T: = Un.
Right.
12 A: ((Softly, then laughs as she writes on the worksheet))
13 °Mattete wasureru°
Wait, I'll forget

Transcription Conventions

,	Non-final/continuing intonation followed by pause
.	Falling intonation followed by pause
?	Rising intonation followed by pause
:	Vowel lengthening
((smiles))	'Stage directions', i.e., non-linguistic event descriptions
(0.6)	Pauses timed by 10ths of a second
=	'Latching', i.e., second speaker's turn begins without pause after first speaker's turn
[	Vertically aligned square brackets mark simultaneously occurring talk
[	
>**Yes**<	Angle brackets enclose talk which is faster than surrounding talk
°Yes°	Degree marks enclose talk which is quieter than surrounding talk

https://doi.org/10.1017/S0261444813000153 Published online by Cambridge University Press

In line 1, Ako reads the first part of the sentence to be translated quickly and quietly – a form of 'self-talk' or 'private speech' (Lantolf 2011). In so doing, she organizes her own cognition publicly by reading aloud. Beyond this fact – this 'externalized' form of cognition – what is particularly interesting here from a sociocognitive viewpoint is that Ako simultaneously organizes TOMO'S cognition – by 'externalizing' cognition into public space she also focuses Tomo's attention on the same part of the sentence she is focusing on, resulting in a classic example of cross-cognitively SHARED or JOINT ATTENTION (Enfield 2006). Next, Ako pauses for two seconds and then utters *ima made ni* (lit: 'until now', or 'ever'), the time adverbial in the part of the sentence she has just read. (It can be noted that Ako has been taught in school to focus on time adverbials to help her determine the correct verb tense in English.) By uttering the time adverbial, Ako once again focuses both her own and Tomo's cognition on the part of the exercise to be worked on, resulting in more focused joint attention.

Tomo then responds to Ako in line 4. In doing so she indexes at least five types of alignment: (1) She 'latches' Ako's utterance in line 2 – that is, there is no perceptible gap between the end of Ako's utterance and the beginning of Tomo's. This action already assumes a high level of alignment – the latcher must be able not only to predict when her interlocutor will stop speaking, but also what the interlocutor will say, since there is no time to pause and think once the utterance has concluded; (2) Tomo speaks more quietly than usual, mirroring Ako's low volume in line 2; (3) Tomo both confirms (*Un* 'Right') Ako's focus on the time adverbial and repeats it in partial form (*ima made*); (4) Tomo mirrors Ako's intonation pattern as she partly repeats her *ima made ni*, which followed a 1–3–2 intonation pattern, just as Tomo's *ima made* does here (note: This is hard to hear unless you use headphones); and (5) There is remarkable synchrony of body orientation between 0:05 and 0:08 in the video: Ako leans to her right and puts her right hand to her head, her left arm resting on the literacy artifact in front of her. This is followed momentarily by Tomo adopting virtually the same orientation. In these five ways, Tomo exhibits and produces a high level of interactional alignment, perhaps indicating 'we are on the same page' in the mutual activity they are engaged in.

In line 6, Ako makes her first attempt to translate the time adverbial, coming up with *never.* While this is not the correct answer, it is close. At precisely the same moment, Tomo suggests in line 7 a solution to the translation problem, saying in effect, 'Why don't you look back at an exercise item we completed earlier – it contains the same form so you can use it as a model'. In so doing, Tomo encourages a classic form of EXTENDED COGNITION (A. Clark & Chalmers 1998) – the offloading of cognition onto the environment. That is, she suggests to Ako that she need not work out the answer in isolated cognitive space since it already exists in the environment and is locatable there. The non-human environment therefore becomes an active part of the cognitive process.

In line 8, as Tomo searches for the already-completed item she has just encouraged Ako to use, Ako continues along her earlier course of action, attempting to translate the time adverbial. She begins by restating it in the same partial form Tomo used in line 4, pauses for about one second, makes a false start (perhaps partially restating her earlier guess, *ne:ba*), and then proposes the correct answer, *e:va ka* ('ever?'). In line 9, Tomo then appropriates Ako's answer and builds it directly into a fill-in-the-blank formula: *Have you ever toka nantoka* (where *toka nantoka* means 'blank'). In so doing, Tomo radically reorganizes Ako's cognitive task, turning a fairly complex sentence translation exercise into a much simpler fill-in-the-blank

https://doi.org/10.1017/S0261444813000153 Published online by Cambridge University Press

one, where the sole requirement is to supply the correct principal part of the verb. In fact, this has been a major focus of the verb-form exercises featured in the worksheet: to promote the accurate use of irregular verb forms. The main sociocognitive point here, however, is that Tomo's very active reframing of Ako's cognitive task is also an act of alignment and cross-cognition – by radically simplifying the task, Tomo radically reconstitutes Ako's cognition in terms of ease and efficiency of functioning, if not learning, as evidenced in Ako's response, as described just below.

In line 10, Ako immediately takes up Tomo's reframing and begins to fill in the blank with the appropriate verb form. Her first try, *writu* ('write'), is not the correct verb form, but it IS the correct verb, and so represents a partial solution. Ako then uses this form to 'step up' to the correct answer – proceeding in the same general order she used when actively learning the English irregular verbs, i.e. 'externalizing' them into the world according to the following pattern: *eat* –> *ate* –> *eaten; go* –> *went* –> *gone*; *write* –> *wrote* –> *written*. It is impossible to know with certainty, but perhaps Ako is re-enacting that same general pattern here, while skipping the second principal part. At any rate, Ako produces the correct answer in her second attempt: *written*.

Ako's gestures as she produces the correct verb form are also quite interesting from a sociocognitive viewpoint. Exactly as she produces the first, intermediate answer, she puts her left index finger on the worksheet, one of the two major sociocognitive affordances in her environment; then, as she produces the correct form, she points her right index finger at the other main sociocognitive affordance – Tomo. In so doing, she seems to dramatize the deeply shared and environmentally embedded cognitive activity she has been involved in – a task in which Tomo, Ako, and the worksheet have all functioned integratively to facilitate correct answering. Although this example cannot provide evidence for cross-time SLA, it certainly reveals MICROGENESIS, or moment-to-moment development (Lantolf 2000; see also Atkinson et al. 2007 for preliminary evidence of cross-time learning in accompanying data).

5. Conclusion

I have tried in this talk to provide a conceptual foundation for a sociocognitive approach to SLA. Clearly, this effort is in its early stages: Development is required at all points, from theory to methodology to empirical research findings. While more elaborate accounts are given in Atkinson (2011c) and Atkinson (2012), I hope I have conveyed a general sense of the approach here.

Acknowledgments

I would like to thank Nick Boke, Ghada Gherwash, Takako Nishino, Joshua Paiz, Tetyana Smotrova, Steve Thorne, and *Language Teaching*'s reviewers for their much appreciated comments on this manuscript and/or earlier versions of this talk.

https://doi.org/10.1017/S0261444813000153 Published online by Cambridge University Press

References

Andor, J. (2004). The master and his performance: An interview with Noam Chomsky. *Intercultural Pragmatics* 1, 93–111.

Atkinson, D. (ed.) (2011a). *Alternative approaches to second language acquisition*. Oxford: Routledge.

Atkinson, D. (2011b). Introduction: Cognitivism and second language acquisition. In D. Atkinson (ed.), 1–23.

Atkinson, D. (2011c). A sociocognitive approach to second language acquisition: How mind, body, and world work together in learning additional languages. In D. Atkinson (ed.), 143–166.

Atkinson, D. (2012). Adaptive intelligence and second language acquisition. *Applied Linguistics Review* 2, 211–232.

Atkinson, D., E. Churchill, T. Nishino & H. Okada (2007). Alignment and interaction in a sociocognitive approach to second language acquisition. *The Modern Language Journal* 91, 169–188.

Baldwin, D., E. Markman, B. Bill, R. Desjardins, J. Irwin & G. Tidball (1996). Infants' reliance on a social criterion for establishing word–object relations. *Child Development* 67.6, 3135–3153.

Beretta, A. (ed.) (1993). Special issue on theory construction in SLA. *Studies in Second Language Acquisition* 14.3.

Block, D. (2003). *The social turn in second language acquisition*. Washington, DC: Georgetown University Press.

Boden, M. (2006). *Mind as machine: A history of cognitive science*. Oxford: Clarendon Press.

Brooks, R. (1991). Intelligence without representation. *Artificial Intelligence* 37, 139–159.

Chartrand, T. & J. Bargh (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality & Social Psychology* 6, 893–910.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Churchill, E. (2008). A dynamic systems account of learning a word: From ecology to form-relations. *Applied Linguistics* 29, 339–358.

Churchill, E., T. Nishino, H. Okada & D. Atkinson (2010). Symbiotic gesture and the sociocognitive visibility of grammar. *The Modern Language Journal* 94, 234–253.

Clark, A. (1997). *Being there: Putting brain, body, and world together again*. Cambridge, MA: MIT Press.

Clark, A. (1999). Visual awareness and visuomotor action. In R. Nuñez & W. Freeman (eds.), *Reclaiming cognition: The primacy of action, intention, and emotion*. Bowling Green, OH: Imprint Academic, 1–18.

Clark, A. (2001). *Mindware: An introduction to the philosophy of cognitive science*. Oxford: Oxford University Press.

Clark, A. & D. Chalmers (1998). The extended mind. *Analysis* 58, 7–19.

Clark, H. (1996). *Using language*. Cambridge: Cambridge University Press.

Collins, R. (2004). *Interaction ritual chains*. Princeton, NJ: Princeton University Press.

Cook, G. & B. Seidlhofer (1995). An applied linguist in principle and practice. In G. Cook & B. Seidlhofer (eds.), *Principle and practice in applied linguistics: Essays in honour of H. G. Widdowson*. London: Oxford University Press, 1–25.

Csibra, G. & G. Gergely (2009). Natural pedagogy. *Trends in Cognitive Sciences* 13, 148–153.

Doughty, C. (2003). Instructed SLA: Constraints, compensation, and enhancement. In C. Doughty & M. Long (eds.), *Handbook of second language acquisition*. Malden, MA: Blackwell, 256–310.

Drahota, A., A. Costall & V. Reddy (2008). The vocal communication of different kinds of smile. *Speech Communication* 50, 278–287.

Enfield, N. (2006). Social consequences of common ground. In N. Enfield & S. Levinson (eds.), 399–430.

Enfield, N. & S. Levinson (eds.) (2006). *Roots of human sociality: Culture, cognition, and interaction*. London: Berg.

Firth, A. & J. Wagner (1997). On discourse, communication, and (some) fundamental concepts in SLA research. *The Modern Language Journal* 81, 285–300.

Gass, S. (1997). *Input, interaction, and the second language learner*. Mahwah, NJ: Erlbaum.

Gibson, J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.

Gigerenzer, G. (2007). *Adaptive thinking: Rationality in the real world*. Oxford: Oxford University Press.

Glenberg, A. (1997). What memory is for. *Behavioral & Brain Sciences* 20, 1–55.

Goodwin, C. (2000). Action and embodiment within situated human interaction. *Journal of Pragmatics* 32, 1489–1522.

https://doi.org/10.1017/S0261444813000153 Published online by Cambridge University Press

Goodwin, C. (2003). The body in action. In J. Coupland & R. Gwin (eds.), *Discourse, the body, and identity*. New York: Palgrave-Macmillan, 19–42.
Gregg, K. (1988). Second language acquisition theory: The case for a generative perspective. In S. Gass & J. Schachter (eds.), *Linguistic perspectives on second language acquisition*. New York: Cambridge University Press, 15–40.
Gullberg, M. (2013). *Gestures in second language acquisition*. London: Routledge.
Hauser, M., N. Chomsky & T. Fitch (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science* 298, 1569–1579.
Hopper, P. (1998). Emergent grammar. In M. Tomasello (ed.), *New psychology of language*. Mahwah, NJ: Erlbaum, 155–175.
Kasper, G. & J. Wagner (2011). A conversation-analytic approach to second language acquisition. In D. Atkinson (ed.), 117–142.
Kendon, A. (1990). *Conducting interaction: Patterns of behavior in focused encounters*. Cambridge: Cambridge University Press.
Kinsbourne, M. & J. S. Jordan (2009). Embodied anticipation: A neurodevelopmental interpretation. *Discourse Processes* 46, 103–26.
Krashen, S. (1985). *The input hypothesis: Issues and implications*. London: Longman.
Lantolf, J. (1996). SLA theory building: Letting all the flowers bloom! *Language Learning* 46, 713–749.
Lantolf, J. (2000). Introducing sociocultural theory. In J. Lantolf (ed.), *Sociocultural theory and second language learning*. Oxford: Oxford University Press, 1–26.
Lantolf, J. (2011). The sociocultural approach to second language acquisition: Sociocultural theory, second language acquisition, and artificial L2 development. In D. Atkinson (ed.), 24–47.
Levinson, S. (2006). On the human 'interaction engine'. In N. Enfield & S. Levinson (eds.), 39–69.
Lewis, D. (1969). *Convention: A philosophical study*. Cambridge, MA: Harvard University Press.
Long, M. (1997). Construct validity in SLA research: A response to Firth and Wagner. *The Modern Language Journal* 81, 318–323.
Mori, J. & M. Hayashi (2006). The achievement of interculturality through embodied completion: A study of interactions between first and second language speakers. *Applied Linguistics* 27, 195–219.
Noë, A. (2009). Extending our view of the mind: Review of A. Clark, Supersizing the mind. *Trends in Cognitive Sciences* 13, 237–238.
Ochs, E., E. Schegloff & S. Thompson (eds.) (1997). *Interaction and grammar*. Cambridge: Cambridge University Press.
Premack, D. & G. Woodruff (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences* 1, 515–526.
Rizzolatti, G. & C. Sinigaglia (2008). *Mirrors in the brain: How our minds share actions and emotions*. Oxford: Oxford University Press.
Rommetveit, R. (1998). Intersubjective attunement and linguistically mediated meaning in discourse. In S. Braten (ed.), *Intersubjective communication and emotion in early ontogeny*. Cambridge: Cambridge University Press, 15–40.
Schegloff, M. (2006). Interaction: The infrastructure for social institutions, the natural ecological niche for language, and the arena in which culture is enacted. In N. Enfield & S. Levinson (eds.), 40–96.
Seo, M.-S. & I. Koshik (2010). A conversation analytic study of gestures that engender repair in conversational tutoring. *Journal of Pragmatics* 42, 2219–2239.
Smotrova, T. & J. Lantolf (in press). The function of gesture in lexically focused L2 instructional conversations. *The Modern Language Journal* 97.2.
Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
Tomasello, M. (2008). *Origins of human communication*. Cambridge, MA: MIT Press.
Tomasello, M. (2009). *Why we cooperate*. Cambridge, MA: MIT Press.
Tomasello, M., M. Carpenter, J. Call, T. Behne & H. Moll (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral & Brain Sciences* 28.5, 675–691.
Weber, M. (1922/1978). The nature of social activity. In W. Runciman (ed.), *Weber: Selections in translation*. Cambridge: Cambridge University Press, 7–32.
Wells, G. (1999). *Dialogic inquiry: Toward a sociocultural practice and theory of education*. Cambridge: Cambridge University Press.
Wenger, E. (1998). *Communities of practice*. Cambridge: Cambridge University Press.

https://doi.org/10.1017/S0261444813000153 Published online by Cambridge University Press

Ylanne-McEwen, V. & N. Coupland (2000). Accommodation theory: A conceptual resource for intercultural sociolinguistics. In H. Spencer-Oatley (ed.), *Culturally speaking: Managing rapport through talk across cultures*. London: Continuum, 191–216.

Yu, C., D. Ballard & R. Aslin (2005). The role of embodied intention in early lexical acquisition. *Cognitive Science* 29, 961–1005.

Zuengler, J. (1991). Accommodation in native-nonnative interactions: Going beyond the 'what' to the 'why' in second-language research. In H. Giles, N. Coupland & J. Coupland (eds.), *Contexts of accommodation: Developments in applied sociolinguistics*. Cambridge: Cambridge University Press, 223–244.

DWIGHT ATKINSON is an applied linguist with a wide range of research and teaching interests: second language acquisition, second language writing, culture theory, postmodernism and poststructuralism, qualitative research approaches, discourse analysis, and English language education in India, among others. Recent publications include the edited volume *Alternative approaches to second language acquisition* (Routledge), and articles/entries in the *Routledge encyclopedia of second language acquisition*, *Routledge handbook of intercultural communication*, *Continuum companion to discourse analysis*, and *Applied Linguistics Review* (De Gruyter). Dwight is an associate professor in the Second Language Studies/ESL Program at Purdue University.

https://doi.org/10.1017/S0261444813000153 Published online by Cambridge University Press