The sociocultural theory (SCT) approach to SLA (henceforth, SCT-L2) is grounded in the psychological theory of human consciousness proposed by L. S. Vygotsky. Although not developed specifically to explain SLA, SCT as a theory of human mental activity has much to offer regarding how individuals acquire and use languages beyond their first. Although some SCT researchers have examined bilingual acquisition, including issues relating to biliteracy, most SCT research within the field of SLA has concentrated on adult learners. Therefore the focus of the present chapter is on SCT-L2 research relating to adult SLA.

Overview
The central thread that runs through most SCT-L2 research since its inception (Frawley & Lantolf, 1985), and which marks it off from other SLA approaches, is its focus on if and how learners develop the ability to use the new language to mediate (i.e., regulate or control) their mental and communicative activity. To be sure, research concerning the Zone of Proximal Development (see below) has directly addressed acquisition, but even there development is understood not only in terms of target-like performance but also in terms of the quality and quantity of external mediation required. Seen as a whole, then, SCT-L2 research is distinguished from other SLA approaches by the fact that it places mediation, either by other or self, at the core of development and use.

Theoretical Principle(s)
SCT’s foundational principle is that “all specifically human psychological processes (so-called higher mental processes) are mediated by psychological tools such as
language, signs, and symbols” (Karpov & Hayward, 1998, p. 27). Mediation is the creation and use of artificial auxiliary means of acting—physically, socially, and mentally. In the physical world, auxiliary means, or tools, include shovels, hammers, bulldozers, dynamite, etc., all of which greatly enhance the human body’s power to shape the environment: It is much easier to dig a hole with a shovel than one’s hands. In the social and psychological worlds, our tools consist of symbols, e.g., numbers, graphs, models, drawings, and especially linguistic symbols. As with physical tools, the power of symbolic artifacts resides not in their structure but in their action potential. Thus, the physical structure of a shovel says little about its function. One must press it into service to discover its capacity to mediate digging action.

Similarly, the structure of language tells us little about its power to mediate our social/communicative and mental lives. Language’s power resides instead in its use value—its meaning-making capacity. Early on, children start to appropriate the symbolic tools of their culture through joint goal-directed activity with adults. The process continues throughout the school years and into adolescence. According to Karpov and Hayward (1998), SCT distinguishes two types of symbolic mediation: self-regulation—the ability to plan, monitor, check, and evaluate self-performance (p. 27); and concept-based regulation—resulting from the appropriation and internalization of cognitive tools needed for mediation in specific “subject-domains” (p. 28). As Vygotsky stated:

Man [sic] introduces artificial stimuli, signifies behavior, and with signs, acting externally, creates new connections in the brain. Together with assuming this, we shall tentatively introduce into our research a new regulatory principle of behavior, a new concept of determinacy of human reaction which consists of the fact that man creates connections in the brain from outside, controls the brain and through it, his own body.

(1997, p. 55)

Children’s early appropriation of language is implicit (i.e., beyond awareness) since the main function of interaction is not usually language learning but learning something else, including how to participate appropriately in social activities. Language serves as a symbolic artifact to facilitate such activities, but it is in and through these activities that language is appropriated (Wertsch, 2007, p. 185). Consequently, language remains largely invisible, at least if and until children enter school, where they are immersed in literacy activities. The effect of schooling is thus to make language highly visible and to enhance children’s capacity to consciously shape it to meet their communicative needs.1

Crucially, there is a close relationship between the social and psychological uses of language. In its communicative function, language entails interaction between “I” and “You.” Eventually, however, a new function emerges, in which the conversation becomes intrapsychological, i.e., between “I” and “Me”, where “I” formulates plans and makes decisions and “Me” (the counterpart of “You” in social conversation) evaluates, critiques and revises these as necessary before the plan’s
external deployment (Vocate, 1994). The “I–Me” conversation is generally referred to as *private speech*, a term coined by Flavell (1966) to replace Piaget’s “egocentric speech.”

To appreciate how the symbolic “I–Me” conversation serves to mediate behavior, consider Marx’s example, borrowed by Vygotsky (1997, p. 68), of the architect, who first works out the design of the building symbolically in blueprint form before beginning to build. The blueprints comprise the plan of action on the ideal plane. The completed blueprint is actualized, or objectified, through physical activity that gives rise to the physical edifice. Vygotsky reasoned that all intentional human behavior, mental or physical, entails a transition from the ideal symbolic plane to the concrete objective plane. This includes acts of speaking, which are just as material as buildings. In other words, speaking entails the realization of an ideal symbolic plan of action that is realized in vocally emitted sound waves moving through space.

**Research Methods**

Because SCT focuses on the formation of mediational ability through appropriating and internalizing symbolic artifacts, it is not very useful to study this ability once formed, as with competent users of a language. That is, it is difficult to observe mediation once it has been internalized. In reaction-time research, for instance, when participants are asked to push a button in response to some stimulus, the thinking process that underlies the behavior is not observable and must instead be inferred by the researchers (Vygotsky, 1978).

Since SCT holds that development originates in the integration of biologically endowed abilities with culturally organized artifacts that mediate thinking, research concentrating on fully formed, “fossilized” (Vygotsky, 1978, p. 68) processes cannot differentiate behavior arising from one or the other source. The solution, according to Vygotsky (1978), is to trace the relevant processes during their formation—as they still operate on the external plane. This approach is known as the *genetic method*—“genetic” not as in found in genes, but because it is historical (i.e., tracks change over time). Thus, child development researchers study mediation by presenting children with tasks beyond their current developmental level while simultaneously offering them potential mediating artifacts and observing whether and how they integrate these artifacts into the problem-situation. Vygotsky (1978, p. 74) called this “the functional method of double stimulation” because, in essence, the children were presented with two tasks: to solve a difficult problem beyond their current ability and also to figure out a way to use an external auxiliary artifact to help them solve the problem. In L2 development, this means studying how learners deploy the new language to regulate their behavior when confronted with communicatively or cognitively challenging tasks.

Examples of how this methodology functions in L2 research are provided below. For present purposes, however, consider the forbidden-colors task (Vygotsky, 1986). Participants of different ages are asked questions and instructed to avoid using a
specific color in their responses. Thus, participants might be asked to describe their home, with white being the forbidden color. Four-year-olds find it difficult to avoid using the forbidden term in such circumstances if their house is actually white. To help the children over this hurdle researchers provide pieces of differently colored paper as external mediational tools for thinking. However, four-year-olds cannot use the tool and continue producing the forbidden colors. But seven-year-olds use the paper, which they often place nearby to remind them of the forbidden color. Twelve-year-olds and adults have no need for external support since they can remind themselves on the internal plane which colors are forbidden. This experiment suggests that children gradually develop the ability, first, to use external mediation and, later, to internalize it.

**Supporting Findings**

SCT-L2 research reflects both ways of conceptualizing symbolic mediation: self-regulatory mediation and mediation provided by conceptual knowledge. It must be stressed that the distinction is only analytical—in normal activity the two aspects of mediation are inseparable. The first phase of SCT-L2 research focused on the self-regulatory function of L2 mediation, beginning with Frawley and Lantolf (1985). This line of research has been thoroughly reviewed (e.g., Lantolf & Beckett, 2009; Lantolf & Pavlenko, 1995; Lantolf & Thorne, 2006), so only its major findings will be treated here. The second phase of SCT-L2 research began with Negueruela (2003), when concept-based instruction (CBI) first attracted the attention of researchers. However, this does not mean that work on self-regulation has ceased: It continues unabated but with an expanded scope that now includes nonverbal components, in particular gesture. Since the latter research has not previously been reviewed in detail, it will be examined more closely here along with the CBI research. The first three subsections that follow address the research carried out on self-regulation, while the remainder of the chapter discusses the growing body of work on concept-based mediation and how this is developed through educational praxis.

**Mediation as Self-Regulation**

The focus of Frawley and Lantolf (1985), as with most research dealing with self-regulation, was not on the accuracy of learners’ speech but on how their performance manifested their ability to maintain and regain self-regulation, in this case on picture-sequence narration tasks. Thus, the intermediate speakers in this study, unlike the advanced speakers, frequently used progressive aspect to describe events (e.g., “Here the little boy is eating the ice-cream cone”), much as one would describe action in a photograph. The researchers interpreted this usage as indicating that the speakers did not control the task and therefore could not create a coherent narrative. Instead, they opted to do what they were capable of—describe isolated pictures/events. Other speakers used the past tense to narrate some story events (e.g., “The man took the little boy’s ice-cream cone.”) rather than the historical
present used by native speakers (NSs) and advanced L2 speakers. The researchers argued that this past-tense usage represented an attempt to regain self-regulation because past-tense morphology provides a kind of temporal distance from events, much like how standing back from a painting allows one to see the whole. Other studies (e.g., Appel & Lantolf, 1994; McCafferty, 1994) uncovered similar though by no means identical mediational L2 uses to carry out complex tasks.

Based on extensive research in Russia, Ushakova (1994) suggested that L2 learners are unlikely to develop the capacity to use the L2 to mediate mental functioning, even when they can use it in social interaction. She cast her conclusion in metaphorical terms: “[A] second language is looking into the windows cut out by the first language” (p. 154). A decade later, Centeno-Cortés and Jiménez-Jiménez (2004), using a more complex research design than in previous studies, again found that L2 speakers, including advanced speakers, were unable to use the language to mediate their online thinking during complex tasks. They reported that, even when able to sustain L2 private speech (i.e., self-speech as a mediational tool), speakers could not complete the tasks given. If, however, they switched to their L1 in its psychological function, they were much more likely to complete the tasks.

Coughlan and Duff’s (1994) important study was the first to consider L2 self-regulation from an activity-theory perspective. Activity theory is considered by many SCT researchers as a sub-theory of SCT. It argues that human behavior is determined by its motive, goal, and the material circumstances in which it is enacted (Lantolf & Thorne, 2006). Coughlan and Duff showed that L2 performance need not be consistent across tasks for single learners or across different learners for single tasks. They argued that performance depends greatly on the specific goals individuals have for speaking. Similarly, Lantolf and Ahmed (1989) explained the variation in one L2 user’s performance across three speaking tasks—picture story, interview, and free conversation—as shaped by the speaker’s communicative goals. Specifically, the learner produced more accurate language when seeking to comply with the assumed interests of the researchers—to elicit evidence of his L2 ability. However, when conversing on a particular topic of interest to him, he produced much more speech and longer turns than in the previous tasks, but at the same time his speech became formally less accurate. The researchers concluded that the learner’s accurate performance reflected other-regulation by the researchers, whereas his less accurate but more interesting and relevant performance exhibited his ability to self-regulate through the language. In other words, the learner’s accurate performance in the first task reflected his attempt to comply with what he perceived as the interests of the researchers—to perform accurately in the L2 regardless of the content of the message. His performance in the second task reflected his personal interest in the topic and the only way for him to fully express this was through L2 speech that was not formally accurate.

Two studies have investigated using private speech to internalize L2 features in classrooms. Ohta (2001) studied the “vicarious” (p. 56) responses of learners when eavesdropping on interactions between the teacher and fellow students. Lantolf and Yáñez-Prieto (2003) conducted a smaller-scale study but similarly found that
through private speech learners focused on those aspects of the target language they desired to learn. An especially interesting finding was that, through private speech, learners appeared to exhibit greater uptake of teacher recasts than reported elsewhere, where socially overt uptake was usually the focus.

**Zone of Proximal Development**

Earlier, I mentioned that children appropriate their community’s cultural artifacts via socialization processes organized by caregivers. Vygotsky (1978) discovered that during socialization caregivers usually behave toward children as if they were able to carry out cultural activities, including those involving language, which they could not actually carry out by themselves. An especially important socialization activity for preschool children is play, which allows them to behave beyond their chronological age. Accordingly, “play contains all developmental tendencies in a condensed form and is itself a major source of development,” because play “creates a zone of proximal development of the child” (p. 102). The Zone of Proximal Development (ZPD) is the activity in which instruction (i.e., socialization at home and formal teaching at school) and development “are interrelated from the child’s very first day of life” (p. 84). Karpov (2005) pointed out that the optimal type of play for promoting development involves adults or older peers serving as models for imitation and providing mediation for children.

In the ZPD, mediators do things *with* rather than *for* children. A simple, though powerful, example of physical mediation in the ZPD is provided by Fogel (1993), wherein a mother undertakes to transfer her infant, who has only partial muscle control, from prone to upright position. One way of doing so is simply to reposition the infant from one posture to the other. A second way, however, is for the mother to mediate the infant by taking the infant’s hands in her own and coaxing her to pull against them while pulling the infant up. This difference in options illustrates the ZPD concept: The infant cannot sit up on her own so, in the first case, the mother makes this the focal point of her own action. In the second case, the mother instead integrates the infant’s behavior with her own mature capacity for bodily action. She collaboratively engages the infant in sitting-up action, instilling some sense of successful agency in her. This is how the ZPD works—by achieving through collaborative mediation what is unachievable alone. It is important to appreciate that the mediator needs to be aware of or discover those capacities that are in the ZPD of the other. Thus, if the infant in the example had no muscle control whatsoever, it would have been useless for her mother to try to move her into a sitting position.

Aljaafreh and Lantolf (1994) represents the initial study on L2 development in the ZPD. The researchers documented changes in learner control over specific L2 grammatical features resulting from mediation negotiated between three learners and an ESL tutor. The important findings of this study include: (1) different learners may require qualitatively different (i.e., more implicit or explicit) types of mediation for the same grammatical feature; (2) single learners may require different types of
mediation for different features depending on their level of control over the feature; (3) mediation sometimes needs to be withheld to determine if learners have control over given features; and (4) development is determined not only by changes in learner performance but also shifts in mediation from more explicit to more implicit. A small-scale study by Nassaji and Swain (2000) showed that randomly provided mediation is less effective than mediation geared to a learner’s ZPD.

Unfortunately, no further significant empirical research on L2 development in the ZPD was published for about a decade. However, Dunn and Lantolf (1998) dealt with the misconception that the ZPD and Krashen’s i+1 were similar concepts. They argued that Krashen’s concept is grounded in a Piagetian perspective that assumes a common internal syllabus for interlanguage development across all learners provided they receive sufficient comprehensible input, while development in the ZPD differs for different learners depending on the quality of mediation negotiated with others.

Dynamic Assessment

With the completion of Poehner’s 2005 dissertation (published as a monograph in 2008), SCT-L2 researchers again began to investigate development in the ZPD. This time, however, a new concept, dynamic assessment, was introduced. Coined by Luria (1961), dynamic assessment (DA) is the systematic integration of the ZPD into educational praxis as the dialectical unity of instruction and assessment (Haywood & Lidz, 2007; Sternberg & Grigorenko, 2002). (For “educational praxis” and “dialectical unity of instruction,” see subsection entitled “Educational Praxis and Concept-based Instruction” below). DA’s underlying principle is that effective instruction requires not only assessment of what individuals or groups can accomplish alone, but also information on how learners react to instruction (i.e., mediation). The former only uncovers past development—it fails to consider potential future development. Because future development depends on mediation, responsiveness to instruction becomes an indispensable component of the assessment process.

Poehner (2008) described a four-month-long project on the oral ability of advanced university L2 French learners. Significantly, this project incorporated the important concept of transcendence (Poehner, 2007)—learners’ ability to appropriate and generalize mediation to new, more complex activities. Through close micro-genetic (i.e., moment-to-moment) analysis of learners’ speaking ability, Poehner documented how learners extended their gains from mediated interaction during recall of a scene from the movie, Nine Months, to recall of a more complex scene from The Pianist, to recall of a different genre—a passage from Voltaire’s Candide. Poehner also corroborated two important findings from Aljaafreh and Lantolf (1994): (1) development manifests not only through changes in learner performance but also through changes in type of mediation supporting learner performance; and (2) development is not uniform for all learners. Thus, different learners need different
types of mediation (from explicit to implicit) for the same L2 features, and single learners often require different forms of mediation for different L2 features.

In a four-month-long DA project on L2 listening comprehension, Ableeva (2010) asked intermediate university French students to recall authentic oral texts in which NSs compared American and French eating habits. Like Poehner, Ableeva included transcendence activities wherein learners recalled not just the original text but new and increasingly complex texts. She found statistically significant improvement in learner comprehension determined by number of propositions accurately recalled from NS texts. She also discovered an upper limit to text complexity even with extensive mediation: Although learners could extend their developing comprehension ability to more complex texts, including a French TV documentary on smoking in restaurants, they could not deal with a radio commercial for a restaurant chain delivered at an articulation rate typical of short radio ads and without the redundancy and pauses of the other texts. This is not surprising: Development is not a process whereby learners can master anything at any time. Even with effective mediation, development in the ZPD has an upper limit, but not one established a priori; rather it is determined through negotiated mediation between learners and others (Vygotsky, 1978). To be sure, most of the students had problems with the radio commercial even when receiving mediation, but this does not mean they can never comprehend such texts. Further focused instruction and experience with the language of commercial texts is likely to help. Similarly, we cannot expect someone who has mastered basic arithmetic to suddenly do calculus, regardless of how much mediation they receive. They must first receive explicit instruction in, and master, algebra. In both language and math, if the individual has no ZPD for the object of study, then mediation is useless.

Antón (2009) studied the use of DA in placement testing in an advanced university Spanish program. Her goal was to achieve more sensitive placement so that instruction was better attuned to student needs. Recall that the ZPD begins with actual ability based on independent performance but is oriented toward future development determined by learner responsiveness to mediation: Students with the same actual level of development do not necessarily project identical future development. Antón demonstrated this important feature of DA through analysis of mediation protocols from learners at the same ostensible proficiency level performing the same narrative task. For instance, in independent performance, two learners had similar problems sustaining coherent use of past-tense morphology, which in Spanish distinguishes perfective from imperfect aspect. Under mediation, however, their performance showed marked divergence: One learner not only improved his performance but also indicated his awareness of the precise nature of his difficulties during independent narration. The other learner showed little improvement under mediation, in fact requiring explicit mediation throughout the narration; nor did he indicate awareness of the nature of his problem. Clearly, such differential abilities require different types of instructional intervention. Antón argued that such ability differences are seldom manifested during independent performance on assessment tasks.
Lantolf and Poehner (2011) traced the integration of DA into a primary-level Spanish course. The instructor employed Lantolf and Poehner’s (2006) teacher’s guide on dynamic assessment, which explains its theoretical basis in the ZPD and presents case studies showing effective versus ineffective mediation. Based on this guide, the teacher adapted DA to her classroom environment, one in which DA-type instruction was possible for only 15 minutes per day.

More specifically, rather than follow the “interactionist” (Lantolf & Poehner, 2004) mediation-providing procedure employed by Poehner (2008) and Ableeava (2010), the instructor formulated a set of eight prompts arranged from most implicit (“Pause to give the student an opportunity to self-correct”) to most explicit (“Provide the correct pattern with an explanation”)—an approach labeled by Lantolf and Poehner (2004) “interventionist” DA. The latter’s advantage is that it permits quantitative comparisons across learners on single tasks, and within learners across tasks at different times. Its disadvantage is that it restricts mediation to predetermined prompts and therefore risks missing opportunities to maximally help students.

Analyzing instructional conversations from this classroom, Lantolf and Poehner (2011) traced the cross-time development of Spanish nominal concord in the performance of one student. They argued that the student’s struggle with nominal concord resulted in development deemed unlikely had the teacher provided immediate recasts instead of calibrated mediation promoting his struggle. Another interesting aspect of this study, as argued by Poehner (2009), was that the other students appeared to benefit from observing this student–teacher interaction. This indicates that instructors and students can operate within a group ZPD (cf. Guk & Kellogg, 2007).

**Concept-based Mediation**

So far, I have discussed mediation as a self-regulating process growing out of other-regulation in the ZPD. The second form of mediation central to SCT is mediation through concepts. Concepts are here understood as the meanings that cultures construct to make sense of the world. The most pervasive concepts are found in language, including lexical, figurative (as in metaphor, metonymy, and other tropes), and grammatical meanings, such as tense, aspect, mood, voice, and anaphora.

Vygotsky (1986, chaps. 5 & 6) distinguished two kinds of concepts: spontaneous (i.e., everyday) concepts and scientific concepts. The latter will be treated more fully below—suffice it to say here that, while spontaneous knowledge is usually appropriated indirectly during socialization, scientific knowledge is appropriated through “the intentional introduction of signs . . . designed and introduced by an external agent” such as a teacher, resulting in an often marked reorganization of activity (Wertsch, 2007, p. 185).

According to Vygotsky, spontaneous knowledge is derived through observing entities and events as they appear to our senses. As a result, some types of spontaneous knowledge are superficial and therefore incomplete or even erroneous. For example, children often consider whales fish since they have fins and live in
water. Even adult language betrays empirical origins—e.g., in English, the sun "rises," "sets," and "moves" through the sky, although such words depart from our scientific understanding of celestial motion. Vygotsky considered the language of children prior to schooling part of spontaneous knowledge. As mentioned earlier, Vygotsky argued that schooled literacy brings spontaneously acquired linguistic knowledge into consciousness, but there are limits to what learners can be reasonably expected to figure out of the complex features of a language in the amount of time normally allocated to organized language study—and this is where scientific knowledge makes its distinctive contribution. This is an important point, as will become clear later in discussing concept-based instruction.

Paradis (2009) and Ullman (2005) have recently proposed models of SLA that distinguish between declarative and procedural (or implicit) knowledge. In L1 acquisition, grammatical knowledge is internalized (i.e., proceduralized) non-consciously through socialization processes similar to those posited by Vygotsky. Lexical knowledge, on the other hand, although also acquired during socialization, is accessible to consciousness. The distinction between the two types of knowledge in both Paradis’s and Ullman’s models is captured by the distinction between procedural and declarative knowledge.\(^2\) One line of research within SCT-L2 has focused on the ability of L2 users to appropriate declarative knowledge of L2 lexical concepts. This research has also been framed within Slobin’s (2003) thinking-for-speaking model, as discussed in the following subsection.

**Thinking for Speaking**

SCT-L2 research informed by Slobin’s (2003) thinking-for-speaking (TFS) model has investigated whether learners can develop the ability to appropriate and think through meanings available in the L2, especially within the semantic domain of motion in event narratives. In particular, the research has considered the interface between speech and gesture as verbal and imagistic carriers of meaning, respectively. This research is premised on the assumption that, when speakers encode their thinking in language for communicative purposes, language shapes (or completes—Vygotsky, 1986) the thinking process itself. Borrowing Vygotsky’s (1986) concept of inner speech, McNeill (2005) argued that not only speech but also gesture shapes thinking. He maintained that speech and gesture form a dialectical unity, or growth point, where gesture is the imagistic and synthetic co-expression of what is represented symbolically and analytically in speech. Specifically, the growth point is the focus of a speaker’s attention, as made manifest in the synchronization of the stroke (or movement) of the gesture with a particular segment of speech.

The main semantic domain in which L2 TFS gesture research has been conducted is motion events. Talmy (2000) proposed a typological distinction among languages according to how they encode such events. Some languages pattern like English and highlight manner of motion encoded in verbs (e.g., *skip*, *trudge*, *sidle*, *scamper*, *creep*), with path of motion marked in a satellite phrase (e.g., *through the swamp, up the ladder, down the stairs*). Other languages pattern like Spanish and highlight
path of motion encoded in verbs (e.g., salir “to exit,” entrar “to enter,” subir “to get into”—as with a car), with manner encoded (if at all) in an adverb or participle (e.g., El barril salió del sótano flotando “The barrel left the basement floating”). This does not mean that languages like Spanish lack manner of motion verbs, but their inventory is restricted compared to English-like languages. Thus, while Spanish has the equivalent of jump, walk, or run, it has nothing like sidle, scamper, or trudge.

McNeill (2005) showed that, along with expressing motion events through speech, speakers simultaneously gesture to co-express movement. Thus, English speakers often express manner of motion with complex manner verbs while producing a synchronized gesture. In Spanish-like languages, on the other hand, gestures used in such contexts synchronize with path verbs. In addition, speakers can express manner through gesture even when not expressed in speech, an option not generally found in English-like languages. Thus, a Spanish speaker might say, “The barrel left from the basement,” and mark the manner of motion with a gesture. It should be noted that, unlike self-regulation and concept-based instruction, motion-event lexical knowledge is not (as far as I know) intentionally taught in classrooms. Thus, if learners acquire this knowledge it is likely through indirect processes.

The question regarding L2 development is whether learners can adopt L2 TFS patterns as manifested in speech–gesture growth points. If so, it would provide evidence of the formation of a new inner order whereby learners internalize and use completely new conceptual meanings—at least as regards motion events—to mediate their thinking process. Some research (e.g., Özyürek, 2002; Stam, 2006) shows that L1 Spanish and L1 Turkish learners of English shift to the English gesture–speech pattern for encoding path of motion in verb satellites. However, there is little evidence of pattern shifts for manner of motion. Negueruela, Lantolf, Jordan, and Gelabert (2004) found no evidence of such shifts in advanced L2 speakers of Spanish (L1 = English) or English (L1 = Spanish).

Choi and Lantolf (2008) encountered one highly experienced immigrant L2 English (L1 Korean) speaker who used one English manner verb synchronized with an appropriate manner gesture in narrating a cartoon story. None of their advanced Korean L2 (L1 English) immigrant speakers showed evidence of moving to a Korean pattern for marking manner. Korean is similar to Spanish in that it marks path of motion on verbs (although in Korean they are compound constructions) and manner of motion with an adverb, or a gesture in the absence of verbal co-expressivity. The L1 English speakers in both this study and Negueruela et al. (2004) tended to display lexical search difficulties when expressing manner in their L2s. The problem was that the languages in question (Korean, Spanish) do not have the complex manner verbs the speakers were searching for. In both studies, this behavior indicates that L2 speakers continue to rely on their L1 to mediate TFS activity.

Gullberg (in press) showed that L1 Dutch French learners are able to shift from Dutch to French patterns of gesture–speech integration when describing object
placement. The Dutch lexicon has verbs that mark placement of objects in space (e.g., Dutch equivalents of put, set, place, stand, lie). French, on the other hand, has few such verbs, preferring mettre (“put”) to describe such actions. In addition, Dutch speakers use a co-expressive gesture depicting the shape of the object (e.g., bowl, bottle, dish) moved, while French speakers use a co-expressive deictic gesture to indicate the landing site of the object (e.g., extended index finger pointing to the target site). The L2 French speakers in Gullberg’s study not only used the appropriate French verb but also shifted from the gesture shape of their L1 to the indexical gesture of the L2, thus manifesting a concomitant shift in TFS. Although Gullberg’s research adds to the evidence provided by Stam (2006) of a shift in TFS, the jury is still out with regard to a shift in manner of motion.

Educational Praxis and Concept-Based Instruction

Although symbolic mediation is the core concept of SCT, in laying the groundwork for his new psychology Vygotsky insisted that theory could no longer be separated from practice, as stated in the 11th Thesis of Marx’s Theses on Feuerbach: “The philosophers have only interpreted the world, in various ways; the point, however, is to change it” (1845/1978, p. 145, italics in original). Vygotsky explained as follows:

Previously theory was not dependent on practice; instead practice was the conclusion, the application, an excursion beyond the boundaries of science, an operation which lay outside science and came after science, which began after the scientific operation was considered completed. Success or failure had practically no effect on the fate of the theory . . . Now the situation is the opposite. Practice pervades the deepest foundations of the scientific operation and reforms it from beginning to end. Practice sets the tasks and serves as the supreme judge of theory, as its truth criterion. It dictates how to construct the concepts and how to formulate the laws.

(1926/2004, p. 304)

The dialectical unity of theory–practice reflected in this quote is referred to as praxis—material activity adapted to specific goals and informed by theory, while simultaneously testing those same theoretical principles (Sanchez Vasquez, 1977, p. 95). For Vygotsky, the highest test of a theory (in the theory–practice dialectic that is praxis) is practice. In this spirit, Stetsenko and Arievitch (2004, p. 78) have suggested that Kurt Lewin’s famous comment that “there is nothing more practical than a good theory” should be expanded to include the “mirror expression—that there is nothing more theoretically rich than a good practice” (p. 77). The challenge of a praxis-based approach was to create a psychology that would promote the development of new processes rather than continuing to focus on observing existing ones.

A praxis-based approach to SCT has been applied to a wide array of social domains, including workplaces (e.g., Engeström & Middleton, 2008), medical
settings (e.g., Luria, 1973), economic and political domains (e.g., Ratner, 2006), and above all education (e.g., Moll, 1990). One might plausibly argue that Aljaafreh and Lantolf’s (1994) work on L2 learning in the ZPD signaled the beginning of a praxis-based approach to L2 education, especially given Vygotsky’s (1978, p. 85) proposal that mediation in the ZPD is an important part of what school learning brings to developmental processes. However, there is another component of educational praxis that is just as important—mediation through scientific concepts. Both components are necessary for a full commitment to educational praxis. Therefore, in my view, the pedagogical project reported in Negueruela (2003) marks the true beginning of L2 praxis, in which the theory is not just a lens for observing learning processes but a means for making them happen. Following Negueruela, several completed or in-progress studies on L2 instruction bring scientific concepts to center stage as the unit of L2 instruction. Before looking at this research, however, it is necessary to discuss the nature of scientific concepts themselves and how they differ from knowledge created in the everyday world.

**Scientific versus Everyday Knowledge**

I have already introduced the notion of everyday/spontaneous knowledge in my discussion of concept-based mediation. While this knowledge generally operates below the level of full consciousness, and in the case of native-language grammatical knowledge remains inaccessible to conscious inspection, scientific knowledge is highly explicit and completely open to conscious analysis. Setting aside implicit grammatical knowledge for the moment, declarative everyday knowledge not only of vocabulary but also other types of encyclopedic knowledge can be superficial and is often erroneous or incomplete. Scientific knowledge (sometimes referred to as theoretical knowledge), on the other hand, represents “the generalizations of the experience of humankind that is fixed in science, understood in the broadest sense of the term to include both natural and social science as well as the humanities” (Karpov, 2003, p. 66). Scientific concepts are not only explicit, they are also domain-specific, and “aimed at selecting the essential characteristics of objects or events of a certain class and presenting these characteristics in the form of symbolic and graphic models” (p. 71). The power of scientific knowledge resides in its “generative” capacity to the extent that it is generalizable across diverse situations (Kozulin, 1998, p. 55).

To illustrate the distinction between everyday and scientific knowledge, consider Ratner’s (2006) example: Everyday knowledge is reflected in such utterances as, “The clothes dried because I hung them out in the sunshine.” According to Ratner, the relationship between sunshine and drying clothes here is “associative” rather than causal in a scientific sense given that one can replace “because” with “the descriptive term when with no change in meaning” (p. 161). Legitimate scientific explanation, on the other hand, provides specific reasons for a process along with
its underlying mechanisms: “The clothes absorb the light, which increases the kinetic energy of the water molecules in the wet clothing to the point that they overcome the adhesive forces that bind them to the clothes” (p. 161). This account explains why clothes dry in the presence of any source that increases the kinetic energy of water molecules; it likewise accounts for the failure of clothes to dry.

A general assumption of SLA research is that the acquisition process is psychologically uniform no matter where it occurs. Long, for example, asserted the following:

Change the social setting altogether (e.g., from street to classroom), or from a foreign to a second language environment and, as far as we know, the way the learner acquires does not change much either (as suggested, e.g., by comparisons of error types, developmental sequences, processing constraints, and other aspects of the acquisition process in and out of classrooms).

(2007, p. 145)

The Universal Acquisition Hypothesis (UAH)—my term—is grounded in the assumption that the basic mechanisms of acquisition are situated inside learners’ heads. As Long explained: “An eight-hour flight from a foreign to a second language environment does not alter a learner’s brain after all, so why should one expect any basic differences?” (2007, p. 145, italics in original). To my knowledge, the UAH has been challenged by only one SLA researcher, Tarone (2007), who presented evidence of a learner manifesting different acquisition sequences for English questions in the school versus home environment.

SCT agrees with Tarone. This is because a central mechanism of mental development is the mediation available in different sociocultural environments. As already mentioned, mediation is realized through social, largely communicative, interaction and the cultural concepts individuals have access to in different environments. Of course, concepts and social interaction are interdependent—social mediation employs culturally constructed concepts, either the everyday variety or the scientific type encountered in educational settings.

Paradis’s (2009) model of declarative and procedural knowledge (see above) makes an important claim regarding SLA that resonates deeply with L2 educational praxis. His essential claim is that, in the absence of intensive and extensive immersion, L2 learners are unlikely to develop implicit automatized competence (i.e., procedural knowledge) to anything like the degree of L1 learners. Instead, L2 learners, in particular those with primary and unique classroom L2 exposure, build up explicit/declarative knowledge, which through practice can result in “speeded-up controlled use” (p. 8). Through speeded-up declarative knowledge, learners can become quite fluent and proficient in meeting their communicative needs. In the following two subsections, I discuss SCT-L2 research that focuses on the intentional—or as Vygotsky put it, “artificial”—development of communicatively functional declarative knowledge.
Concept-based Instruction

As mentioned above, concept-based instruction (CBI) has as its centerpiece, or unit of instruction, systematic, explicit knowledge of the relevant features of the L2. Several studies have been conducted within this framework on L2 praxis, beginning with Negueruela (2003). As Negueruela’s work has already been discussed in the research literature, however, I will present an overview of a nearly completed dissertation that deals with Chinese (Mandarin) as a foreign language (Lai, 2011). To be sure, examples from Negueruela’s work—as well as Yáñez-Prieto’s (2008), also framed within CBI—will be included where relevant. Before discussing Lai’s research, however, it is necessary to consider in concrete terms CBI’s implementation in language classrooms.

The specific procedures followed in most L2 CBI to date were stipulated by the best-known pedagogical theorist of Vygotsky’s school, Piotr Gal’perin (see Haenen, 1996; Talyzina, 1981). Gal’perin’s program, known as Systemic-Theoretical Instruction, follows a specific sequence of instructional phases, which, as implemented in most recent L2 studies, are: systematic verbal explanation of the concept in the target language, including comparison with the L1 whenever feasible —> materialization of the concept —> communicative activities —> verbalization —> internalization.

The first phase, explanation, must be based on scientific knowledge of the concept under study. In my view, cognitive linguistics provides a potentially useful source of such knowledge for language instruction (Lantolf, in press). This is because cognitive linguistics foregrounds meaning and seeks to develop theoretical concepts that generalize across language domains. A second reason I believe cognitive linguistics is compatible with SCT is its robust use of visual models to depict linguistic concepts. This links the theory closely with Gal’perin’s (1970) first two phases of instruction—explanation and materialization, with the latter concept being based on the premise that the former alone is often problematic for learners. Thus, Gal’perin argued that students tend to memorize explanations as rules rather than understanding concepts well enough for them to inform and guide practical activity. He therefore proposed the second phase—materialization—in which the concept is represented visually as a model, graph, or other synthetic depiction. Gal’perin called the materialization of a concept a schema for the orienting basis of action (SCOBA). SCOBAs provide learners with resources that are then formulated as a plan of action in the third, or communication, phase of instruction. In language instruction, communication can involve a wide array of activities, ranging from tasks to scenarios (Di Pietro, 1987) to literature-based conversations and writing activities (Yáñez-Prieto, 2008) to service learning (Grabois, 2008), where language is used as a tool for accomplishing specific goals. The fourth phase, verbalization, is the point at which learners use language (i.e., engage in languaging—Swain, 2006) to both explain the concept to themselves and explain to themselves how they use it in specific communicative activities. In essence, this phase calls for production of speech in its psychological function with the intention of giving rise to the final, or internalization phase of the developmental process.
Learners’ reactions regarding the value of SCOBAs and their verbalization have been quite revealing, given that the procedures are not commonly employed in language programs. One student from Negueruela’s (2003) study made the following comment:

The charts are a grammar-figuring-out-guide that work better than the rules (like the rules for preterit and imperfect) that we had learned in Spanish 100. It was very helpful to see the concepts in a visual structure because the concept of grammar is a very structural concept, and being able to visualize it made it make much more sense.

(p. 453)

A student from Yáñez-Prieto’s (2008) study offered the following remarks on how SCOBAs impacted his thinking:

It’s kind of funny how you can have a grammar . . . grammatical structure actually tell a story. I’d not really noticed that or seen that before. I mean, the words are telling the story and the grammar is telling the story, which is kind of weird. Yeah, I’d never seen that before. Interesting.

(p. 267)

The SCOBAs not only made the language feature visible for this student, they also compelled him to think deeply about the connection between structure and meaning-making.

With regard to verbalization, the comment of a student from Negueruela’s (2003) study revealed that talking to himself contributed to his understanding, at the same time helping him gain confidence in speaking the new language:

Although sometimes recording myself speak was a bit awkward, I think it was overall extremely helpful. It made me more comfortable speaking and improvising, and it forced me to truly think about the grammar.

(p. 438)

A second student from the same course corroborated this perspective:

The activities that have helped me the most are the verbalization ones with the cassette tape player. I feel as though with verbalization exercises I not only improved my speaking, but also learned a lot of information about the indicative and subjunctive.

(p. 434)

Swain, Lapkin, Knouzi, Suzuki, and Brooks (2009) contended that verbalization is also effective in a collaborative format where students explain to each other rather than themselves a concept and the implementation of its SCOBA. Swain et al.’s
reasoning is based on Vygotsky’s argument that speech is reflexive and therefore can simultaneously serve to mediate others and self. It remains to be seen if private verbalization or social languaging is more effective in promoting internalization, or if they are equally effective.

CBI in Chinese Temporal Grammar

Lai (2011) analyzed the effects of a CBI program designed to develop the ability of L2 learners of Chinese to use temporal grammar. Without going into detail, Chinese does not mark tense morphologically but instead relies either on context or adverbial particles. One set of particles marks time on a horizontal dimension and another marks it vertically. Taking the horizontal dimension first, whereas English conceptualizes past time as behind and future time as in front of the speaker, Chinese does the opposite—the particle qian, “front,” designates past, while hou, “back,” designates future. Thus, the Chinese lian tian qian, “two days front,” is expressed in English as “two days ago,” while san nian hou, “three years back,” is rendered “after three years.” As regards the vertical dimension, Chinese also uses it to designate past and future. Thus, the particle shang, “up,” indicates immediate past, while xia, “down,” marks immediate future, with shang functioning similarly to English “last” and xia similarly to “next.” Sang xinqi, “up week,” translates into English as “last week” and Xia yi nian, “down one year,” is “next year.” Thus, the language distinguishes between distal and immediate past and future in accordance with the two different spatial dimensions.

Lai (2011) pointed out that, traditionally, Chinese-language pedagogy has not focused heavily on teaching temporal particles as tense markers. In some textbooks the particles are introduced as part of compound lexical items without indicating their tense-marking function. In other cases, brief rules of thumb for marking tense are given, accompanied by illustrative examples. As argued by myself and others (e.g., Lantolf, 2007; Negueruela, 2003), rules of thumb are much like everyday knowledge—incomplete, superficial, and/or erroneous, and as such they do not explicate the essence of the concept under study. Using cognitive linguistics as her theoretical foundation, Lai presented her students with a sophisticated yet useable explanation of particle-based tense-marking in Chinese using the SCOBA given in Figure 1.1. For present purposes I leave out the details and focus instead on the most important aspects of the explanation.

Lai (2011) implemented her instructional program over a five-day period in a first-year/beginning-level university Chinese course. Except for the special instruction on temporal grammar, the students followed the regular course syllabus. After instruction, the students were given a post-test in which they were asked to write a narrative based on a picture story. The same test was given to two other classes: a beginning-level class, which received temporal instruction according to the course textbook; and an intermediate-level class, which had received instruction on temporal grammar in their first and second years of study. The class receiving Systemic Theoretical Instruction performed significantly better (p < .05) than the
other first-year class, and no differently (in terms of statistical significance) from the intermediate class.

**Future Directions**

The current SCT-L2 research program clearly suggests fruitful areas for future research. In my view, four are particularly important. The first is research that implements dynamic assessment within group-wide ZPDs. The SCT-L2 work carried out so far on this topic, although limited in scope, has generated interesting results. In educational domains such as L1 literacy development, the classroom itself has been construed as a ZPD where instruction is organized according to principles of activity theory, especially by using a division of labor whereby texts are read collectively with specific tasks (e.g., preparing lists of potentially problematic words, reading between the lines) parceled out based on teacher assessment of student strengths and weaknesses (Cole, 1996). The notion of division of labor is a powerful one that should be fully exploited in future L2 educational praxis.

A second area for future research is extending concept-based instruction beyond grammar—to pragmatics, figurative language (a form of language much underappreciated in L2 pedagogy), and especially the interface of language and culture (see Agar, 1994). Thorne, Reinhardt, and Golombek (2008) conducted a pilot study implementing a CBI approach to teaching international teaching assistants at a U.S. university how to deploy directive constructions with appropriate mitigators and intensifiers when interacting with American students. This study is just a first step in exploring CBI’s potential for pragmatics instruction. In theory, any conceptual domain can be taught through CBI, but it requires a clear and systematic statement of the concept, an imaginative approach to formulating a coherent SCOBA, and the implementation of appropriate communicative activities.
A third area for future research focuses on the central claim of educational praxis, which, as pointed out above, restructures developmental processes. The approach to education advocated by Vygotsky contrasts with Piagetian perspectives whereby processes of development unfold smoothly and consistently for all learners regardless of the learning environment, including intentionally and systematically organized educational praxis. Piagetian models of SLA (e.g., Krashen, 1981; Pienemann, 1998) support common and invariant acquisition sequences based on an internal natural syllabus that operates across all learners and across all learning circumstances. Given that Vygotskian theory views the environment as a central mechanism in development, it challenges the existence of a natural syllabus. Thus, an interesting test of the two theoretical stances would be to conduct an instructional study designed to interfere with the natural syllabus. In my view, the research conducted within Pienemann’s (1998) processability model has produced the most robust and consistent results supporting a natural syllabus based on processing constraints. It has also assessed negatively the impact of instruction on the syllabus; however, in my view this research has not taken sufficient account of quality of instruction. A praxis approach to instruction as discussed in this chapter should be brought to bear on the theory of natural development. To this end, we are in the process of conducting an instructional project on word order in Chinese topicalized constructions, including temporal and locative adverbs as well as direct objects. Zhang (2008), conducted within Pienemann’s framework, found no instructional influence on the developmental sequence predicted by the processability model. The goal of our project is to test Zhang’s findings through CBI.

A final area where I believe fruitful SCT-L2 research can be carried out is within Slobin’s thinking-for-speaking framework, especially if speech–gesture synchronization, as defined by McNeill (2005), is included. The question of whether L2 users can employ their new language to mediate their thinking processes continues to be intriguing, and approaching the topic from the perspective of McNeill’s growth-point hypothesis has exceptional promise.

### Conclusion

In a recent paper, VanPatten (2010), a proponent of an influential SLA theory, addressed the relationship between SLA research and practice. He pointed out that teachers often assume a direct relationship between research and teaching, and as a consequence expect that “research should somehow improve instruction” (p. 30). VanPatten’s response was that such expectations may be inappropriate. He stated that, while SLA research “cannot speak to the day-to-day issues that confront” teachers (p. 29), it can help them “understand acquisition and thus inform instruction” by offering “insight into what the learning problems are” (p. 30), and that this in turn “might also lead to a better set of expectations regarding the interface between teaching and acquisition” (p. 36). Because of its commitment to educational praxis, SCT-L2 adopts a very different stance on the relationship between theory and practice. A particular divergence between VanPatten’s position and what has been
advocated here concerns the explicit teaching of scientific knowledge of L2 linguistic features. In VanPatten’s view, “it would be absurd to expect learners to grasp issues related to theoretical linguistics as part of their language learning experience” (p. 36). Instead, he believes that adjustments should be made to the input to help learners infer the appropriate uses of the features to be internalized. The approach espoused by not only SCT-L2 but also applied cognitive linguistics (De Knop & De Ryckere, 2008) dictates that learners should indeed be presented with explicit and theoretically sophisticated explanations of L2 features, visualized in a pedagogically relevant SCOBA. That is, features should be presented in a form students can use to guide their thinking and performance.

As indicated throughout this chapter, SCT-L2 is very much concerned with concrete classroom activity and its impact on learning. It argues for the pedagogical relevance of explicit and rigorous linguistic explanation, especially that derived from cognitive linguistics, and is devoted to discovering how to make learning happen through direct instruction. It is different from other theories of SLA in that it does not assume that acquisition is a universal process. On the contrary, because mediation is different in different sociocultural domains, development in these domains is expected to be psychologically different.

Notes

1 Tarone, Bigelow, and Hansen (2009) demonstrated how low levels of literacy impact L2 learning. One of their interesting findings is that such learners have reduced levels of awareness of L2 input.

2 The declarative/procedural model will become relevant later in my discussion of scientific knowledge and concept-based instruction.

3 Following Hulstijn (2003), Paradis allows for the possibility that in some circumstances L2 learners can use explicit knowledge to indirectly benefit the formation of implicit, automatized knowledge. I have argued elsewhere (e.g., Lantolf, 2007, 2008) that explicit knowledge can also become automatic procedural knowledge as a result of internalization, with internalization in SCT being the process whereby individuals appropriate and integrate cultural artifacts, making them their own (Wertsch, 1998). Paradis’s understanding of internalization differs from SCT’s in that he sees it as relevant to the formation of implicit knowledge only; in turn, only implicit knowledge, according to Paradis, is automatized. Hence, he distinguishes between automatized (i.e., implicit) knowledge and speeded-up explicit knowledge. To avoid terminological confusion, I will use Paradis’s “speeded-up” rather than “automatized explicit” knowledge in the following discussion. But because internalization is central to Vygotsky’s theory, I will continue to use this concept to refer to the process whereby any type of knowledge is appropriated and used for mediation of social or psychological behavior.

4 Proximal and distal time depends on the temporal framework in which one operates. Within the time frame of “weeks,” for example, proximal time would be either “last week” or “next week,” while distal time would be either “the week before last” or “the week after next.”

References


VanPatten, B. (2010). Some verbs are more perfect than others: Why learners have difficulty with *ser* and *estar* and what it means for instruction. *Hispania*, 93, 29–38.


