

# Theories and Research of Second Language Acquisition

*“A strong foundation and a good bridge!”*

Susan Malone, SIL International (in response to the question: “What is needed for a successful mother tongue-based multilingual education program?”)

## **Introduction**

If the foundation of the non-dominant language learners’ mother tongue is not strong, then any “bridge” built on it will be in danger. Without the good bridge, the non-dominant language learners will **not** achieve their educational goals.

The focus of this paper is *not* on the strong foundation. That is assumed to be present. The purpose is to present a summary of the theoretical ideas that contribute to building the good bridge. In order to do that in a condensed way, we will skim and skip over an enormous amount of research and theories.

### Theories, Methods, Hypotheses

Pre-20 <sup>th</sup> Century	Early 1900s	1940s, 1950s	1960s, 1970s	1980s, 1990s	1990s - present
Grammar-translation	Audiolingual, Direct Method	Behaviorist, S-R-R	Universal Grammar, LAD	Information Processing Models	Social Interactionism
	<i>Bloomfield, Fries</i>	<i>Skinner</i>	<i>Chomsky, Krashen</i>	<i>Anderson, McLaughlin</i>	<i>Vygotsky, Snow</i>

The discussion that follows is in two parts: (1) representative 2LA theories and (2) representative 2LA methods.

## **Representative 2LA Theories over the Years**

The following theories represent serious thinking over the past century about the way a person acquires or learns a second language. (Keep in mind that much complex theorizing has been condensed here.)

### *Behaviorism*

In mid-Century behaviorist theory (as typified by the work of B.F. Skinner) impacted almost all areas of scientific investigation. The behaviorist theory (simplified) predicted that any human behavior could be learned through a process of stimulus, response, and positive or negative reinforcement (sometime

abbreviated as S-R-R).<sup>1</sup> The object was to make the desired behavior become a habit, performed spontaneously. According to behaviorists this process applies to the way people learn language as well as other human behavior.

### *Language Acquisition Device (LAD)*

Linguist Noam Chomsky theorized that the observable data for language acquisition did not favor a behaviorist approach. Children in every language and cultural community learn to understand and speak at a remarkably early age. They are able to produce proper and unique/novel language even in infancy but they are not merely mimicking language patterns they hear. They are creating language themselves. “Out of the mouth of babes oft times come gems!” is an ancient proverb, now with a modern theoretical base (as if it needed one).

Chomsky called this innate ability to acquire and use language a Language Acquisition Device (LAD). He argued that the stages of development that are required for children to develop their cognitive abilities in other areas do not apply to learning language. They are using language – arguably the most abstract symbolic system they will encounter in their lives – appropriately and expressively, if not maturely, by the age of 5. Chomsky’s *theory* led to an entirely new approach in the field of linguistics: generative phonology and transformational grammar. This new approach to linguistic analysis focused almost entirely on the abstract “deep structure” of individuals’ native language. Chomsky was not as concerned with the actual language being spoken as with the unspoken but understood rules that made the utterances appropriate, correct to native speakers. Nor did Chomsky study how people acquire a second language. However, other linguists applied his theories to 2LA issues (e.g., Krashen’s Monitor Model below).

### *Information Processing*

Other cognitive scientists disagreed with Chomsky’s LAD hypothesis. They considered the human being’s ability to acquire language as merely a component of highly complex cognitive structures. R. C. Anderson (1983) developed the Adaptive Control of Thought (ACT) model that has been influential in studies of cognitive development and 2LA. In Anderson’s theory, intelligence is simply the gathering together and fine-tuning of many small units of knowledge that in total produce complex thinking. “The whole of language learning ability is not more than the sum of its parts, but it has a lot of parts!”<sup>2</sup>

---

<sup>1</sup> For a brief explanation of behaviorism and 2LA, see the PowerPoint presentation: [www.bu.edu/linguistics/JG/course/lx400/.../lx400-2a-history.pdf](http://www.bu.edu/linguistics/JG/course/lx400/.../lx400-2a-history.pdf)

<sup>2</sup> Quoted in Yates, Kenneth A. (2007). Towards a taxonomy of cognitive task analysis methods: A search for cognition and task analysis interactions. Unpublished Doctoral Dissertation, University of Southern California, Los Angeles.

This theory is based on the hypothesis that all human knowledge can be divided into two irreducible kinds of representations: *declarative knowledge* and *procedural knowledge*. In brief, declarative knowledge is learned rapidly and is stored in long-term memory through *images* and *schemas*.<sup>3</sup> Procedural knowledge is more complicated and refers to the gradual process by which a person learns how to do something successfully. In summary:

*The ACT model is enormously complex. ... It is not possible to do justice to this complexity here. The central points to grasp are the theoretical claim that learning begins with declarative knowledge which slowly becomes proceduralized, and that the mechanism by which this takes place is practice.* Ellis (1994) p. 389.

### *Social Interaction and Sociocultural Theories*

Where Chomsky decided to focus on the deep structure and abstractions of language acquisition, another group of linguists decided to focus on how the role of language as it is actually spoken contributes to 2LA. Also, where Krashen argued that “comprehensible input” is the necessary and sufficient source of successful 2LA, other linguists argued that comprehensible input may be necessary but not sufficient.

Social Interactionists, like Gass (2002), focused on the language learning context and “how learners use their linguistic environment (in particular, conversational interactions) to build their knowledge of the second language” (p. 17).

Vygotsky emphasized the role of the social environment (that is, the other people who interact with the child) on children’s learning. If, as Vygotsky argues, everything is first learned socially (that is, from someone else), then what is the role of social interaction in the language acquisition process? Does the 2LA program allow ample opportunities for the learners to interact with others using the L2?

Swain (1990) argues that “comprehensible output” in meaningful conversations is also necessary for successful 2LA. When learners talk in the L2 they notice a “gap,” a difference, between their knowledge of the L2 and what they want to/need to say. Having noticed the gap, they are now predisposed to modify their L2 speech, to pay attention to the L2 structure or grammar. At that point they can begin to think about the language – a metalinguistic<sup>4</sup> activity – and begin to internalize the way the L2 works.

---

<sup>3</sup> The term “schemas” refers to Anderson’s contribution to Schema Theory as an explanation of cognitive development and learning happen.

<sup>4</sup> Metalinguistic refers to “thinking about language” in a similar sense that metacognition refers to “thinking about thinking.”

Her research has provided ample evidence for the notion that second language learners learn the L2 from each other and from the teacher in actual interactions in the L2 as they receive feedback from their partners.<sup>5</sup>

### ***Representative 2LA Methods***

#### ***Background***

In Western countries, in the 19<sup>th</sup> Century, the primary purpose of learning a second language in school was in order to read the classical literature of that language. The method for doing that was called “grammar-translation.” The method for learning included the use of the learners’ mother tongue. Vocabulary, grammatical structures, and whole texts from the target language were presented to the learners, explained in their mother tongue, memorized and then translated from the L2 to the L1.

Not surprisingly, this method was difficult to implement in a meaningful and interesting way for the learners. The methods below are, in part, an effort to overcome that approach to learning a second language.

Surprisingly, perhaps, grammar-translation methods were used in India at least until the 1960s and in China up to the present<sup>6</sup>

#### ***Audiolingual Approach***

Linguists in North America developed the audiolingual approach in the early 20<sup>th</sup> Century as part of their project to record and document all of the Native American languages still in use. Because the Native American cultures were primarily oral cultures, the linguists’ language learning emphasis naturally shifted from the written to the spoken word. And since mother tongue speakers of these languages rarely if ever gained the linguistic education needed to write up formal linguistic analyses of their language, the outside linguist himself/herself had to devise a method for learning and recording the language.<sup>7</sup>

At the same time, behaviorist theory began to impact almost all areas of scientific investigation. The behaviorist theory -simplified - predicts that any human

---

<sup>5</sup> See [http://www.celea.org.cn/2007/keynote/ppt/Merrill Swain.pdf](http://www.celea.org.cn/2007/keynote/ppt/Merrill%20Swain.pdf) for an excellent presentation of her “output hypothesis” and its relation to sociocultural learning theory. It is only moderately technical. See also, Harley, Allen, Cummins and Swain (1990) for an academic paper on the subject.

<sup>6</sup> Rao (1996) discusses the deep cultural and pedagogical adherence Chinese learners have to teacher-centered, text-centered language education based in Confucian philosophy. As a result, English-as-Foreign-Language teachers in China run into difficulties trying to train teachers in purely communicative, learner-centered, small group oriented 2LA lessons. He suggests a way of combining Communicative Language Teaching with grammar-translation.

<sup>7</sup> Audiolingual is probably the method by which most SIL language staff learned indigenous languages prior to the 1980s.

behavior can be learned and performed through a process of stimulus, response, and reinforcement (sometime abbreviated as S-R-R).<sup>8</sup>

The S-R-R approach prescribed a 2LA method that involved a lot of repetition of the L2 grammatical forms, especially those that differed from the L1. It was assumed that the forms that differed would be the most likely source of errors for the L2 learner. This hypothesis was later discredited.

In sum, the L2 is learned by many repetitions of the L2 grammatical patterns with positive reinforcement provided when reproduced correctly and negative reinforcement when reproduced incorrectly. Interference by the learners' L1 was considered to be the main source of L2 error. Therefore, its use in L2 learning was avoided as much as possible.

### ***The Natural Approach***

Steven Krashen developed the Monitor Model of 2LA in the 1970s using ideas developed by Chomsky. Briefly, his theory was built on five hypotheses:

1. There is a "natural order" for learning the structure of an L2.
2. There is a distinction between language acquisition and language learning.
3. There is a function for a "monitor" (or "editor") in the production of the L2. That is, in language acquisition, the learner acquires fluent, correct speech in a natural way and applies it to speech unconsciously. When certain grammar or spelling rules are consciously learned, then a "monitor" or "editor" appears (figuratively) and corrects speech that is not appropriate in L2 (as far as the learner knows).
4. The necessary and sufficient cause of 2LA is "comprehensible input." That is, not only is meaningful input needed in 2LA, it is *all* that is needed. If you constantly receive comprehensible input you will learn to understand and speak the L2.
5. An "affective filter" plays a critical role in how well and how fast an L2 is acquired. This refers to the emotions/feelings of the learner in the 2LA process. If the process causes anxiety and fear, then less – if any – L2 will be acquired/learned.

Krashen also predicted that natural 2LA would include a "silent phase" when the language learner is receiving comprehensible input and beginning to understand how the language but not yet confident enough to speak. This phase might be shorter or longer depending on the individual learner.

Jim Cummins' work supports Krashen's hypotheses in several areas. Cummins proposes that L2 language learners have access to a common underlying proficiency

---

<sup>8</sup> For a brief explanation of behaviorism and 2LA, see [www.bu.edu/linguistics/UG/course/lx400/.../lx400-2a-history.pdf](http://www.bu.edu/linguistics/UG/course/lx400/.../lx400-2a-history.pdf)

(CUP) with respect to language learning. This idea builds on Chomsky's notion of a LAD available to children learning their L1. It means that the learners possess a built-in ability to acquire a second language using the same language learning aptitude they display in acquiring their basic mastery L1 in 5 years or less. In short, the surface features of the L2 may differ from the learners' L1, but the underlying, subconscious knowledge of how language works is available for the learners' use.

Cummins also hypothesized two domains of language: basic interpersonal communication skill (BICS) and cognitive-academic language proficiency, each requiring a different acquisition timeframe: BICS, 1-2 years, CALP 5-7 years. Cummins also predicted that amount of context provided in the learning situation and the degree of difficulty of the learning task combine to decrease or increase the ease with which the concept, skill or attitude can be learned. He illustrated this in the following figure:

**Low Cognitive Demand**

High Context  
Low Cognitive Demand  
[EASIEST]

Low Context  
Low Cognitive Demand  
[HARDER]

**High Context**

**Low Context**

High Context  
High Cognitive Demand  
[HARDER]

Low Context  
High Cognitive Demand  
[HARDEST]

**High Cognitive Demand**

The kind of language abilities needed for developing BICS would be found primarily in quadrant 1, with some instances in quadrants 2 and 3. Quadrant 4, the low context/high cognitive demand level, is the domain of language ability needed for CALP.

Both in Krashen's and in Cummins' (2000) approaches, the learners' mother tongue plays a critical role. A well-developed L1 in all communicative domains – listening, speaking, reading and writing – lays the foundation for transfer of these skills to the L2. Longitudinal research of over 42,000 students in bilingual programs by Thomas and Collier (1997, 2002) confirmed that prediction. Their primary findings included this: that the most accurate predictor of success in end-of high school exams in the L2 [English] by English language learners was the number of years they had of education in their L1.

***The Communicative Approach***

This approach grew out of theories that the kind of second language the learners need to learn is the kind they will need for real-world, real-life situations. More a

way of teaching than a specific 2LA method, Communicative Language Teaching (CLT) is designed around a set of classroom principles.

- *An emphasis on learning to communicate through interaction in the target language.*
- *The introduction of authentic texts into the learning situation...*
- *An enhancement of the learner's own personal experiences as important contributing elements to classroom learning.*
- *An attempt to link classroom language learning with language activities outside the classroom.*<sup>9</sup>

The theoretical framework for CLT comes in part from Social Interactionist theory that views the learners' use of the L2 as the best source for learning the L2. In other words, as the learners speak the L2, they receive feedback and that process helps them learn the language. Different approaches have spun off from this perspective, e.g., Task-Based Language Teaching (TBLT) and Content-Based Instruction (CBI).

An obvious concern for practitioners in MTB MLE is the emphasis that CLT makes on preparing the learners for the L2 they will meet outside of the classroom. This presupposes that the learners live in an L2-dominant environment with ample opportunities to engage in real-life L2 conversations and other language-related activities. This is definitely *not* the case in many of the situations where MTB MLE practitioners work. L2 learners in these contexts do not encounter the L2 meaningfully in the communities in which they live. For them, the L2 is actually a foreign language (e.g., English in Ethiopia) or virtually a foreign language (e.g., Standard Thai in many minority language areas in Thailand). In those contexts, the L2 is not likely to be used at all in real-life interactions in the learners' communities.

### ***Educational Practice for Bridging from L1 to L2***

The above theories and hypotheses about 2LA do not all focus on the process of building a good bridge from the child's mother tongue (L1) to the school educational language (L2) in MTB MLE programs. This is true especially in places where the L2 is virtually a foreign language to the learners, where their classrooms are in under-resourced educational systems, and where the classroom teachers receive minimal, in any, training in teaching the L2 as a second or foreign language.

MTB MLE practitioners need to adapt methods of 2LA and Foreign Language (FL) strategies to their specific situations, especially when the L2 and/or FL are introduced in the early grades. For example, sociocultural theory supports the kind of peer tutoring in small groups that Swain (2007) in her "output hypothesis" and

---

<sup>9</sup> Nunan (1991). Communicative tasks and the language curriculum. *TESOL Quarterly* 25(3), 279-295.

the CLT approaches use. However, that process may not work in Grade 1 and Grade 2 classrooms.

*Platt and Troudi (1997) report that for some activities, elementary school children have problems scaffolding<sup>10</sup> their peers because, even though they may themselves possess the expertise to carry out a task, they are not sensitive to their classmates' ZPD<sup>11</sup>... (p. 107)*

Lack of cognitive maturity may also be why the use of 2LA techniques like TPR work better for students in Grades 1 and 2. TPR is also an instructional approach that can be implemented by teachers whose own control of the L2 or FL is not fluent.<sup>12</sup> Teachers are given facilitating roles in Communicative Language Teaching lessons, but cannot easily implement them given their low level of mastery of the L2.

Here are eight instructional techniques that have been incorporated into 2LA classes with good success.

- TPR
- Cooperative learning (small groups, pairs)
- Language Experience Approach
- Dialogue Journals
- Academic Language Scaffolding
- L1 support for 2LA
- Accessing learners' prior knowledge
- Culture studies

**TOTAL PHYSICAL RESPONSE (TPR).** This instructional technique allows for the children to receive comprehensible second language input without requiring them to speak. This honors the "silent period" experienced by L2 learners. It is interactive and non-threatening.<sup>13</sup> The form of language is somewhat restricted, initially using only command patterns. However, a good deal of variety can be included. The variety is necessary because a teacher may unconsciously repeat the same pattern of TPR activity that results in boring the children.

**COOPERATIVE LEARNING.** This approach is frequently used in task-based language teaching (TBLT) and features small group activities that promote positive interaction. Students benefit by seeing the learning strategies used by their classmates. In

---

<sup>10</sup> "Scaffolding" is "the provision of appropriate assistance to students in order that they may achieve what alone would have been too difficult for them. ..." <http://esl.fis.edu/teachers/fis/scaffold/page1.htm>

<sup>11</sup> ZPD = Zone of Proximal Development; Vygotsky's notion, closely related to "scaffolding". The "zone" is the point in time when helping a person do something they cannot do by themselves will likely enable them to learn to do it.

<sup>12</sup> A colleague in Africa related how she visited a secondary school and was introduced to the school's English teacher. He was not able to understand simple questions she asked, and when the questions were translated into his mother tongue, was not able to respond in English.

<sup>13</sup> This may apply more to children than adults. A colleague once told me that some adults learning English found this approach threatening when they did not understand the commands given to them to perform.



addition, cooperative learning facilitates face-to-face meaningful interactions in the L2. That is, the interactions are necessary to complete the assigned task. Meaningful tasks are the basis of this interaction.

**LANGUAGE EXPERIENCE APPROACH.** This instructional technique has likely been used during the learners' L1 literacy acquisition period. Teacher and students share an experience of some type and then discuss it in the L2. The teacher suggests that the children make up a written story/text of the experience. The teacher elicits a beginning sentence, and writes it on the chalkboard. This continues until 4-5 sentences have been written in the L2. The children are then encouraged to read what they have written, as a group, in pairs, individually. For L1, this is a literacy learning activity. In L2, this is an activity that promotes the learners' transfer of their literacy knowledge and skill to the L2 writing system.

**DIALOGUE JOURNALS.** This activity creates a written "conversation" between the learner and the teacher. Teacher and students decide on topics of personal or group interest. Students write about the topic in the L2 in their "journal" and the teacher responds, in their journal. This provides the teacher with an opportunity to evaluate what the learner has written, but also to model correct language and provide a non-threatening opportunity for students to communicate in writing with someone more proficient in the L2.<sup>14</sup>

**ACADEMIC LANGUAGE SCAFFOLDING.** This involves the process of providing the learner with L2 input that is a little bit beyond his or her current level. The teacher or more advanced classmate act as the "support" or "scaffold"<sup>15</sup> for the learner as she builds her L2 vocabulary and fluency, providing positive and negative feedback, often in the process of completing a shared task.

This activity grows out of Krashen's idea of "comprehensible input,  $i+1$ "<sup>16</sup> and Swain's emphasis on comprehensible output in the learner's "zone of proximal development." Again, this activity depends on some degree of proficiency in the L2 by the teacher and by classmates. Where the L2 is actually or virtually a foreign language, this activity may not be possible.

**L1 SUPPORT FOR L2 LEARNING.** Where approaches in the past – the Direct Method, L2 Immersion, Audiolingual – banned the learners' L1 from the 2LA classroom, more recent research suggests that the L1 can be a very useful medium for L2 learning. Butzkamm (2007) concludes

---

<sup>14</sup> This technique is, of course, less useful where the teachers' L2 proficiency is too limited. Also, in crowded classrooms, the teacher will need to schedule responding to 5-6 student journals per day. In the course of one month the teacher could write in each student's journal perhaps two times.

<sup>15</sup> "Scaffold" is a term used for the structures that provide ladders and supports for workers as they repair or construct a building (often made out of bamboo or pipes or tubes made from steel or plastic).

<sup>16</sup> In Krashen's "comprehensible input,  $i+1$ ," the "i" indicates what the learner can do/understand at the present and "+1" indicates the next step up that she can do with help from the teacher or classmate.

*The mother tongue is therefore the greatest asset people bring to the task of foreign language learning and provides an indispensable Language Acquisition Support System.*

CULTURE STUDIES. This is, of course, a logical result of the paragraph above. The children's prior knowledge that they bring to the 2LA learning experience is learned in their L1. Building on that knowledge makes sense from a learning theory perspective as well as from 2LA theory. Also, much of the learners' cultural knowledge is embedded<sup>17</sup> in their mother tongue. Second language learning also requires a transfer of L1 cultural knowledge to knowledge of the L2 culture. For example, a Patani Malay child learning to understand and speak Standard Thai as an L2 will also need to understand something of Thai culture. Likewise, there is a culture that goes along with learning to speak English.

### **Conclusion**

At the end of their book, *Second Language Learning Theories*, Mitchell and Myles (1998) addressed the issue of the increase of theories of second language learning without any consensus by applied linguists as to how all of these theories fit together. In truth, they don't. The authors conclude, wisely I think, that it is too early to try to discover a single theoretical framework that takes into consideration all of what is currently known and theorized about 2LA. In terms of MTB MLE, we can feel liberated, perhaps, to include 2LA curriculum and instruction that appears to be working well with the ethnic minority learners and that it has theoretical support somewhere. In the meantime, serious, credible research and evaluation of MTB MLE programs and students will help in developing useful 2LA theories.

Building the "good bridge" for MTB MLE is definitely a work in progress.

### **REFERENCES**

- Anderson, J. 1983. *Cognitive psychology and its implications*. Second edition. New York: Freeman.
- Butzkamm, W. 2007. Native language skills as a foundation for foreign language learning. In W. Kindermann (Ed.), *Transcending boundaries: Essays in honor of Gisela Hermann-Brennecke*, 71-85. Berlin: Lit Verlag.
- Cummins, J. 2000. *Language, power, and pedagogy: Bilingual children in the crossfire*. Clevedon: Multilingual Matters.
- Ellis, R. 1994. *The study of second language acquisition*. Oxford: Oxford University Press.
- Gass, S. 2002. An interactionist perspective on second language acquisition. In R. Kaplan (Ed.), *The Oxford handbook of applied linguistics*, 170-181. Oxford: Oxford University Press.

---

<sup>17</sup> The term "embedded" refers to the close relationship that exists between a particular culture and the language that its adherents speak.

- Harley, B, Allen, P., Cummins, J. and Swain, M. 1990. The development of second language proficiency. New York: Cambridge University Press.
- Mitchell, R. and Myles, F. 1998. *Second language learning theories*. London: Arnold.  
 BOOK REVIEW by Dennis Malone in *The Journal* 1(1), xx-xx, 2005, Salaya, Thailand: Mahidol University-Salaya.
- Nunan, D. 1991. Communicative tasks and the language curriculum. *TESOL Quarterly* 25(3), 279-295
- Platt, E. and Troudi, S. 1997. Mary and her teachers: a Grebo-speaking child's place in the mainstream classroom. *Modern Language Journal* 81, 28-49.
- Rao, Z. 1996. Reconciling communicative approaches to the teaching of English with traditional Chinese methods. *Research in the Teaching of English* 30(4), 458-471.
- Thomas, W. and Collier, V. 1997. *School effectiveness for language minority students*. Washington, DC: National Clearinghouse for Bilingual Education.
- Thomas, W. and Collier, V. 2002. *A national study of school effectiveness for language minority students' long-term academic achievement*. Santa Cruz, CA: Center for Research and Education, Diversity and Excellence.  
[http://www.crede.ucsc.edu/research/llaa/1.1\\_final.html](http://www.crede.ucsc.edu/research/llaa/1.1_final.html)
- Yates, K. A. 2007. Towards a taxonomy of cognitive task analysis methods: A search for cognition and task analysis interactions. Unpublished Doctoral Dissertation, University of Southern California, Los Angeles.