The Neuropsychological Basis of Lexically-based Language Teaching

Siusana Kweldju
Universitas Negeri Malang

Abstract: This paper proposes a lexically-based approach for the teaching of English in Indonesia. To develop the approach, neurolinguistic and psycholinguistic research findings for L2 learning are reviewed because this approach has to consider how human mind stores, processes, recognizes and retrieves words. Neurobiologically, the learning of L1 is different from L2. Thus, the lexically-based approach gives more emphasis on chunking, the learning of formulaic phrases, and conscious awareness of learning.

Key words: lexically-based approach, neuro-psycholinguistic underpinnings

Indonesian students have little knowledge of English vocabulary. This is the main reason why they cannot optimize themselves in oral and written communication in English. To solve this problem, language teaching in Indonesia is ideally lexically-based (Kweldju, 2000). Language is obviously not seen as individual words, but also collocation and colligation or lexico-grammatical units. Lexically-based language teaching teaches an L2 through, among others, lexico-grammatical units. Conscious learning, repetition or memorization plays an important role, because according to research findings the learning of L2 vocabulary must be contrived. Although Krashen (1989) hypothesized that vocabulary was best acquired by guessing from context through the act of reading itself, most findings show that all meanings of words—except high-frequency words—cannot always be inferred from context or learned incidentally (Beheydt, 1987; Raptis, 1997). Learning words is not only for confronting unknown words for the first time and for knowing their meanings, but also for consolidation which covers memorizing, incidental learning and practicing (Mondria & Mondria-de Fries, 1994).

The purpose of this paper is to incorporate a degree of neurolinguistic and psycholinguistic underpinnings into lexically-based language teaching. Lexically-based approach must be developed to meet how human mind stores, recognizes and retrieves words. Words cannot be heaped up randomly in the mind, because random facts and figures must be extremely difficult to remember, while in reality enormous quantities of data have to be remembered and located so fast. When we consider the brain for facilitating language learning, cognition, emotion, perception and memory become intertwined, and consciousness-raising or drawing learners' attention to the formal properties of language becomes essential (Gray, 1990). Conscious learning ties together such related concepts as attention, short-term memory, control and series processing (Schmidt, 1990).

L2 vocabulary must be contrived because, as the first constraint, L2 learners often lack sufficient, highly contextualized input in L2. This often makes it extremely difficult for an L2 learner to extract and create semantic, syntactic, and morphological specifications about a word and integrate such information into the lexical entry of that word in the mind. The second constraint is the presence of an established conceptual/semantic system with an L1 lexicon, which is relied on by L2 learners and makes them less motivated to use contextual cues for meaning extraction (Jiang, 2000).

To be more specific, this paper reviews relevant neuropsychological studies to answer these three following questions: Is the learning of L2 really different from L1? Does the learning of L2 require conscious awareness of learning? What is the neurobiological support for improving one's L2 through memorizing stock phrases and lexico-grammatical units?

MENTAL LEXICON

Mental lexicon is the human mental dictionaries; the locations of
which are not as important as the links it makes. Unlike book dictionaries, words in the mind are not organized solely on the basis of alphabetical order. Words, including proper names and idiomatic expressions, cannot be heaped up without complex orderliness because of their large number and, yet, efficient and fast retrieval. Human memory is both flexible and extendable, provided that the information is structured. Native speakers can recognize a word of their language in 200 ms or less from its onset, that is, approximately one-fifth of a second from its beginning (Aitchison, 1994).

The mental lexicon is a simple tug-of-war between perception, production, and memory, and it overlaps with other aspects of cognition and language. In outline, the mental lexicon can be regarded as consisting of two major components: semantic-syntactic and phonological. Mental lexicon is also connected with a person's general ability to relate it with general knowledge and memory.

In the mental lexicon words are not separate items attached to one another in a fairly random way. They are clearly interdependent. In some cases it is difficult to understand a word without knowing the word around it: orange is best understood by relating it to red and yellow, or warm is between hot and cold. According to the tip of the tongue phenomena, tongue slips such as antidote for anecdote indicate that a speaker does not only consult his mental lexicon according to an alphabetical list. Instead, he looks for certain salient features which characterize it: word initial, word final, number of syllables, stress pattern, and chunked prefixes. These salient features are presumably prominent in storage, since they tend to be recalled even when the full form of the word cannot be retrieved (Aitchison & Straf, 1982).

NEUROBIOLOGICAL ACCOUNT FOR LEXICO-GRAMMATICAL UNITS

Chunking and the Importance of Collocation for Native-like Proficiency

Native-like production is not a matter of syntactic rule alone, or as creative as understood by Chomsky's tenet (Peters, 1983). It is characterized by collocation and fluency, which both depend on chunking. A chunk is a unit of memory organization. It is welded with other chunks to build a larger unit, leading to a hierarchical organization of memory.

Chunking appears to be a ubiquitous feature of human memory and forms the basis for an equally ubiquitous law of practice.

In fact, an important index of native-like competence is that the learner uses idioms fluently. Speaking natively is speaking idiomatically using frequent and familiar collocations, and the job of the language learner is to learn these familiar word sequences. Native-like proficiency also depends on conversational speech which is broken into fluent units of complete grammatical clauses of four to ten words, uttered at or faster than normal rates of articulation (Ellis, 1997). Collocation itself is well-stored in human brain, a brain-imaging machine can identify and reports that there are brain areas where collocates and nodes are located (Sylvester, 1995).

Chunks are Stored as Whole Words

Learning individual words is insufficient, because words are not only stored as morphemes and individual meanings, but also as parts of phrases, or as longer memorized chunks of speech, and are often retrieved from memory in these pre-assembled chunks (Nattinger & DeCarro, 1992). In line with this, words must also be learned in units or chunks, because they are not stored only according to their individual word meaning, but in the concatenation of either complex experiences or longer coherent chunks for the actual process of human communication. During this process the semantic representations of individual words are retrieved, and they are combined into an overall meaning of the entire sentence (Smith, 1978; Ericsson & Kintsch, 1995).

As mentioned above, chunking is part of the actual process of human communication. Words and lexicogrammatical units are both involved in perception and parsing procedures. In actual comprehension and learning to understand a language, for example, at least three types of mutually interactive procedures, rather than sequentially ordered or independent, come into play: pattern-recognition that encodes the words of a sentence, syntactic procedures that parse the sentence or parse the speech stream into chunks and make apparent the grammatical role of each word, and semantic procedures for interpretation (Smith, 1978). This is another reason why paying attention to recurring chunks is important in L2 learning (Ellis, 1997).
Psycholinguistically, language consists of units or chunks, which are perceived holistically and invariantly, without being discretely analysed. These units or chunks underlie human cognition. For example, torrents of speech come at us extremely quickly, sometimes exceeding 20 each second and we do not have the chance to perceive the sounds one by one (Wardhaugh, 1993). Psychologically, cognition arises from the interaction of procedural and declarative knowledge. Procedural knowledge is represented in units or knowledge structures called production rules, and declarative knowledge is represented in units called chunks (Anderson, 1996).

Units in mental dictionaries are organized in relation to real-world situation for real collocational use. While a book dictionary contains only a very small amount of data about the syntactic patterns of a word into which it can slot, mental lexicon stores any possible syntactic patterns of a word. Wide and main are both classified as adjectives in a book dictionary, but it might not tell us that "the road is wide" is correct, but not "the road is main." Both eat and resemble are classified as transitive verbs, but it does not tell us that whereas "a cow was eaten by my aunt" is possible, "a cow was resembled by my aunt" is not (Aitchison, 1994).

Lexicon and Syntax are Inseparable in Chunks

As mentioned above, according to experimental psycholinguistics grammar and lexicon are inseparable as syntactic-semantic entities or lexico-grammatical units in language production and comprehension. Psycholinguistically, these lexical semantic codes and their syntactic features are intertwined and this union is called lemmas (Kempen & Huijbers, 1983). It is lexicon that determines grammar and grammar is represented lexically. Thus, hypothetically, words and sentences are stored in the same memory system. An experiment (Miller & McKay, 1996), for example, shows that sentential factors—syntax and semantics—influence short-term memory within rapidly presented lists. Familiar two-word phrases at unpredictable positions in a list could be immediately recalled, while the same words as unrelated words in identical positions, without phrases, were poorly recalled. For example, the word night was easily recalled as part of the phrase night gown, but the word night as an unrelated word in (2) was poorly recalled. But the unrelated word mind in (1) and (2) were equally poorly recalled

1. good faith mind night gown film (phrases in italic)
2. people faith mind night hose film (unrelated word list)

The explanation is that language production begins with a preverbal message. This message triggers language specific processes that make up the formulation stage. During grammatical encoding at this stage, the system retrieves lexical-semantic representations that encode (or are associated with) syntactic information required for computing the hierarchically organized syntactic structures for a complete sentence. Bowers et al.'s experiment (1999) on language production using reaction time to words and pictures shows that syntax is lexically rather than conceptually represented. A syntactic feature that has some semantic grounding—namely, the count-mass marking in English and grammatical gender—is not incorporated into conceptual representations, but to the lexical representation of a word.

L1 Acquisition is Neurologically Different from L2 Acquisition

Traditionally the neurobiological specializations for language are associated with Wernicke's area and Broca's area, yet these are only the tip of the iceberg, because language subserve so many functions (e.g. cognition, memory) and because language is so intimately bound to sensory and motor systems. Even regions previously unsuspected of playing a role in language appear to link with language, such as the fusiform gyrus in both hemispheres, the anterior cingulated and the prefrontal cortex. Using imaging technique, coupled with results from excision and electrical stimulation studies, enables researchers to identify whole neural systems involved in language production and acquisition, and discover that these processes involve a variety of distinct yet interconnected structures. For example, the long term potentiation process originating in the hippocampus and related structure, seems to lead to the memory needed for vocabulary storage. In addition, rapid retrieval and use of grammatical knowledge may be viewed as a procedural skill, thus involving the procedural knowledge memory system of the cerebellum (Lem, 1992).

The two half-brains are not separate but intersect with each other. A number of studies, however, using blood-flow measures, dichotic listening
procedure, and unilateral sodium amytal tests, for examples, still show that generally in the most majority of people, the left hemisphere as the cortical representation of language skills still generally shows superiority in memorizing and recognizing vocabulary (Peng, 1985; Metcalfe et al., 1995). It does not mean, however, that the right hemisphere does not contribute to vocabulary skill. Lesser’s (1974) and Gainotti’s (1981) study respectively showed that the right-hemisphere damaged subjects also had marked difficulty on the semantic word association skills, and lower ability in semantic discrimination tests. Thus, it can be speculated that the right hemisphere is also responsible for vocabulary learning, but not for the syntactic interpretation of sentences.

In spite of that, in view of neurolinguistic and psycholinguistic research, the cerebral organization of language is not identical in monolingual and bilingual speakers, as the right hemisphere might play a much important role in L2 acquisition than it had been assumed, especially in early stages of L2 learning when learning uses more drills or formulaic utterances (Krafteschner, 1986). According to Obler (1981) this happens particularly in an informal setting of L2 outside the classroom when learners tend to rely more on content than function words, prosodic rather than syntactic cues, and linguistic information in context rather than in isolation (Krashen, 1977, Galloway & Krashen, 1980). Gifted L2 learners might also use a greater proportion of the right hemisphere for L2 learning and may, in consequence, suffer some visual spatial disability, a skill associated with the right hemisphere (O’Connor et al., 1994).

Right hemisphere functioning is generally associated with holistic processing, while serial or analytic processing occurs in the left hemisphere (Taylor, 1990). The right hemisphere is responsible for the storing and processing of formulaic speech. The routines and patterns which comprise formulaic speech are unanalyzed wholes and as such belong to the gestalt perception of the right hemisphere. Genesee (1982) also inferred that the right hemisphere involves more in informal than formal language acquisition, because the use of formulaic speech is more prominent in settings where natural language use is common. Seliger (1982) suggests that patterns which are initially processed in the right hemisphere can then be re-examined later in left hemisphere functioning. This is partly because the former stores more exact traces than the left hemisphere, which is more capable of mental manipulation (Metcalfe et al., 1995). However, if subsequent analysis by the left hemisphere does not take place, the learner will not be able to utilize the language forms that have been drilled in the construction of spontaneous, creative speech. This offers a neurolinguistic explanation of why drills do not appear to facilitate natural language use immediately.

In other words, the left and right cerebral hemispheres partake differently in language processing. It has been proposed that the right hemisphere in bilinguals may share language functions with the left hemisphere to a greater extent than is the case in unilinguals (Vaid, 1982). As long as L2 learning is concerned the different experiences can be determined by the manner, the stage and the age of learning an L2. For example, when Hopi schoolchildren listen to stories in Hopi and in English, the right hemisphere shows significantly desynchrony (implying greater involvement in the task) during the Hopi story, but not the English one.

Learning, in computer analogy, involves recording information in a specified location, and retrieval involves going back to that location and reading out the information stored there (Baddeley, 1990). Unlike human memory, however, computers are trivally easier to store information when a storage location is available. Forgetting is negligible, and retrieval is all-or-none. The speed of operation of the neurons within the brain is very much slower than that achieved by the components of computers. On the other hand, human memory can achieve feats beyond that of average computer and has major virtues in speed and flexibility of access, and in the extent to which it abstracts, yielding ready access to the essentials of complex prior experiences.

Human brain operates in parallel, so that it has the capacity for so-called graceful degradation. Although it loses a large numbers of brain cells as a result of normal ageing or as a result of brain damage, it can still continue to solve problems. Serial machines, however, are very prone to breakdown.

Based on multistore model there are three components of memory system. At the sensory register the process is pre-attentive or unconscious, but consciousness takes place at the following processes of short-term and long-term memory.
Connectionism: an Alternative Model for SLA that Considers All Subcomponents of the Brain

Modern thinking about language has been dominated by the views of Noam Chomsky, although alternative theories exist. Chomskyan framework starts with a complex system of rules and constraints that allows people to distinguish grammatical and ungrammatical forms, and applied linguists used it to explain how children acquire this grammar, and how it is used in producing and comprehending utterances. L2 acquisition research, then, has been much reduced to second grammar acquisition. In this framework, language processing takes place in the innate system of principles, pre-linguistic state, and parameters that guide the acquisition of any human language, known as Universal Grammar (UG) or Language Faculty (in the past known as LAD or Language Acquisition Device), which is unique to human, but neurobiological studies show that UG is unreal, as language processing can also take place both in the perceptual and cognitive subcomponents of human mind. There is no existence of a distinct LAD or UG (Jacobs & Schumann, 1992). Therefore, it is quite possible to maximize all these possible subcomponents for L2 learners’ ultimate success.

Due to the development of cognitive science, the nature of human language processing is understandable from two major developments outside the linguistic mainstream, i.e. connectionism by neuroscientists, and computer science for the creation of major archives of adult and child language corpora (Seidenberg, 1997). Almost all recent literature on the mental lexicon finds reference to connectionism because of the neurophysiological account that human brain is capable of massive parallel processing (Singleton, 1999).

Connectionism as a model for a memory system reaches the field of SLA for proposing a model of L2 processing which is alignable with neurophysiological reality. Only in this way psycholinguistics takes into account constraints from studies of the nervous system. Connectionism is focused on learning and memory. It does not assume innateness and is not compatible with UG, of which the principles and parameters are stated in terms of variables or symbols, which cannot be wired in to the connectionist network on the outset. In some of these models patterns are not stored. What is stored is the connection strengths between units or the simple processing elements that allow these patterns to be recreated.

The term connectionism is associated with the neurophysiological activity of the brain. During any brain activity, numerous brain cells are active to send signals to other neurons. Some signals are arousing or inhibits other units. The pattern of excitation and inhibition in the system as a whole determines what to remember. It is also called the parallel distributed processing because different portions of information are processed independently of one another on different levels. That is, one event or stimulus may trigger different processing of micro-operations or detailed operations. Therefore, this computed modeling of learning is via the gradual buildup of richly inter-networked association potentials as the neural level which are selectively and simultaneously activated in specific patterns to perform cognitive activities, including language production and comprehension. Connectionism is a computational model to embody the essential features of the information processes that take place in the brain. Computational approach deals with data structures and the processes that operate on them yield outputs from given inputs. Therefore, the validity of this model must be tested in computer programs (Gasser, 1990).

Connectionist theories attempt to explain behavioral phenomena in terms of large numbers of simple, neuron-like networks of processing units. Unlike grammatical theories describing and explaining the structure of language, connectionist approach explains how networks come to perform tasks such as comprehension and production. In fact, connectionists want to explain how humans use mental lexicon. When producing a word, they must pick the meaning before the sound, and when they recognize the words, the sounds come first. Connectionism develops a model which allows information rushing back and forth, while excites more and more
related words. The relevant get more and more excited, while the unwanted fade away. Different words require different levels of activation to produce. Higher frequency words are easy to trigger, while the lower ones are harder to arouse (Aitchison, 1994).

Connectionist framework differentiates between L1 processing and L2 processing. In L1 processing knowledge of language consists of generalizations made over linguistic pattern complexes (LPCs), each consisting of features of form (morphosyntactic, phonological) and content (semantic, pragmatic, contextual). In L2 processing L1 patterns may transfer to L2. Transfer is important in a connectionist model. The claim on this transfer is that the primitive of form and content are the same across languages. Therefore, the basic network units to represent input form and content are the same for L1 and L2. The L1 patterns also suffer interference from the L2 pattern. Though the L2 patterns are initially difficult for the network, they are not as difficult as the L1 patterns were when they were first presented to the network.

Human Memory, Cognition and Conscious Learning

Learning through memorization involves a conscious process. Although Krashen (1982, 1985) sharply demarcated the conscious and unconscious processes of language acquisition, the evidence from a wide variety of fields, including neuropsychology, psychotherapy, acquisition theory and pedagogy, and cognitive psychology, suggests that conscious processes cannot be sharply separated from unconscious processes (Munsell et al., 1988).

Krashen's point was much influenced by Chomsky's argument that rejects conscious learning in L1 acquisition. Chomsky's argument about the creative construction hypothesis, however, cannot be made for L2 acquisition. Although there are natural orders, they do not eliminate the possibility of a role for selective, voluntary attention. Psychologists, who argue that there is no learning without awareness, believe that implicit learning happens after the learner attends to structured stimuli. The resulting knowledge of this learning is encoded in the form of unconscious abstract representation. For example, L2 learners with explicit knowledge of grammatical items performed significantly better than learners without (Schmidt, 1990; Hulstijn & Hulstijn, 1984; Grigg, 1986).

There are numerous multistore models of memory. Kihlstrom (1984) suggests that in most multistore models, the terms consciousness, focal awareness, and short-term store are essentially equivalent. These models also claim that processing in short-term memory is necessary for permanent storage. Anything that is not processed in short-term memory is forever consigned to oblivion. Once in short-term memory, information that is not further encoded into long-term memory is also lost. If consciousness is equivalent to the short-term store, conscious awareness is essential.

Attention and perception refer to the immediate processing of the sensory experiences of organisms, the experiences of sight, sound, smell, taste and touch. Attention includes bottom-up gate-keeping operations which ensure that we are alerted to speech sounds and give us the initial ability we need to deal with them. Cognition only follows attention, although it is hard to say where one kind of processing ends and the other begins (Wardhaugh, 1993).

One of the anterior cingulate's primary functions is in attention. L2 learners also use conscious self-monitoring when he speaks, listens or writes. For example, the auditory signal must first be filtered from all other stimuli. Then it is processed to associate meaning with the sounds. Planning is made to respond to the auditory message. The appropriate brain areas are recruited to generate the response. Monitoring of the output is also necessary, particularly in non-fluent learners. This monitoring requires attention to be focused on the appropriate stimuli, a focusing which has been shown to be a function of the anterior cingulate.

Conscious Learning of Lexicon Alone is Insufficient for Word Retrieval and Production

In the mental lexicon receptive vocabulary is much larger than the productive one. The distance or the gap between receptive and productive vocabulary is a matter of the imperceptible and infinite degrees of knowledge. One's most elementary knowledge takes place at the first encounter with a word. At this point, he might only recognize the word in terms of its length. The word is still stored in the least complete way which does not allow the subject to either reproduce or produce the word. Having a higher degree of familiarly to a word means having not only the phonological, morphological, semantic, stylistic, and polysemous knowl-
knowledge of the word, but also the lexico-grammatical knowledge or the collo- 
cutional knowledge of a word.

The degree of knowledge alone is not the only factor for the process of 
recognition to manifest itself, or how a word is retrieved from the 
mental lexicon. In the process of recognition after the first apprehension, 
certain traces remain in the brain, but it is insufficient for retrieval. Only 
after repeated apprehension or several occurrences can a word be consid-
ered to be a part of one’s lexicon, though reproduction of the item would 
still be quite impossible.

Compared to production, however, recognizing words is a simpler 
process, as it only needs partial information instead of a complete form. It 
takes less information about a word to interpret than it does to generate it, 
primarily because the interpretation process requires only sufficient 
information about a word to distinguish it from all other possibilities (Melka, 1997). For example, one can recognize an item even if one pho-
neme is inaudible or has been replaced by a cough; if a word is presented 
in context he will not even detect that a phoneme is missing (Morton, 
experiments that word recognition is extremely fast (one fifth of a second); they also show that, in normal contexts (i.e. not an anomalous 
context), a word is located in the mental lexicon even before all the word 
have been heard (Melka, 1997).

In fact, knowing a word is not an all-or-nothing proposition. Al-
though the knowledge of a word has not reached a totally productive (P) 
stage, some aspects may have become productive, while others remain at 
the receptive (R) level. According to the tip-of-the-tongue phenomenon 
one may have trouble to produce a complete, precise form, but can often 
give detailed information on the number of syllables, the initial and final 
sound or letter and even the stress pattern of the word. When the word is 
finally given away to him, he can recognize it immediately (Melka, 1997).

Between R and P are imitation, reproduction, and comprehension. 
Reproduction is an active reconstitution of what has been read or heard. If 
reproduction is performed with assimilation of materials then reconstruc-
tion activates memory. If assimilation does not take place, the learner 
loses all contacts with the corresponding ideas when producing ready-
made verbal formula. At this stage, imitation and reproduction precedes

comprehension, which precedes production (Belyayev, 1963).

Words must be Learned Seriously

L2 vocabulary must be learned seriously. Otherwise, it can easily be-
come fossilized, because L1 lexical development is different from L2 
experimental development. In L1 lexical development, the task of vocabulary 
acquisition covers the acquisition of semantic and formal entities of the 
word, but L2 words are learned as formal entities only, because the 
meaning is provided. Therefore, at the first stage of learning, the meanings 
of L2 words and their grammatical information may become available 
through the activation of L2-L1 links. As one’s experience in L2 in-
creases, stronger associations are developed between L2 words and their 
L1 equivalent. Then, the activation of L2 word form and L1 lemma in-
formation may result in a strong and direct bond between an L2 word and 
the lemma of its L1 translation. A word may be considered to have 
reached the second stage of development when L1 lemma information is 
copied into its entry, although the L1 lemma information still mediates L2 
word processing. The full development of lexical competence is the L2 
integration stage, when the semantic, syntactic and morphological specifi-
cations of an L2 word are extracted from exposure and use, and integrated 
into the lexical entry.

L2 lexical competence may cease to develop at the second stage of 
learning, no matter how extensive the contextualized input is. According 
to Jiang (2000) it is because of L1 lemma mediation which allows any L2 
meaning and other information to be accessed through L1. Thus, the lan-
guage processor will be less motivated to pay attention to the input of L2 
for meaning extraction. As a result, the transition from L1 lemma media-
tion to L2 integration can be longer than expected. Then, the presence of 
L1 lemma information in the L2 lemma structure may prevent the estab-
ishment of L2 lemma information within the lexical entry. In other words, 
once the space is occupied by the L1 lemma information, it becomes very 
difficult for the L2 lemma information to get in. This speculation is sup-
ported by neurobiological evidence. Primary language acquisition (PLA) 
is both subjected to and responsible for the development of the nervous 
system or the maturational neurological changes. SLA is essentially re-
stricted to the integra}
must integrate itself into the existing neuronal framework of PLA, because the postnatal brain cannot any more reproduce neurons.

**SUMMARY**

One ideal approach for L2 learning in Indonesia is lexically-based, but this approach must be developed based on neuropsychological underpinnings. Reviewing the literature, this approach must consider the importance of lexico-grammatical units, as words are not only stored according to their individual word meanings, but also in units or chunks. Words must also be learned to its fullest degree when they want to be used actively. Since the cerebral organization of L1 is different from L2, L2 learning needs more conscious processes for learning formulaic phrases, an important key for fluency.

**REFERENCES**


