

Evidence From the Phylogenetic Domain and the Use of L1 in English Language Teaching

Hywel Evans (Corresponding author)

Department of English, Tsuru University

Tsuru, Yamanashi, Japan

E-mail: veryserioso@gmail.com

Vahid Rafieyan

International College of Liberal Arts, Yamanashi Gakuin University

Kofu, Yamanashi, Japan

Natsue Hasebe

Department of English, Tsuru University

Tsuru, Yamanashi, Japan

Received: March 2, 2020

Accepted: March 16, 2020

Published: March 30, 2020

doi:10.5296/ijl.v12i2.16588

URL: <https://doi.org/10.5296/ijl.v12i2.16588>

Abstract

The theoretical foundations supporting adherence to naturalistic language learning approaches appear to have entirely collapsed, just as non-naturalistic approaches to language teaching such as translation and translanguaging have become increasingly respectable in English language teaching. In line with this, language and language learning are increasingly being understood as sociocultural phenomena, inevitably situated in local contexts, rather than in terms of manipulation of an abstract, universally-shared set of syntactic rules. This prompts a reevaluation of both traditional methods and innovation in specific sociocultural settings. We offer a review of some important developments in our understanding of language acquisition, with particular focus on evidence from the phylogenetic domain, often overlooked by language professionals, and suggest directions for fruitful investigation. In this

regard, we recommend attention first to methods that have stood the test of time in the sociocultural milieu, that are acknowledged as having value by both local teachers and students. As a notable example of these, we focus on the pedagogical use of L2 vocabulary items embedded in L1 text. We find that the claim made by Japanese native local teachers, that such L1 use is helpful in vocabulary learning, receives significant support from our experimental results.

Keywords: Evolution, First language, Language evolution, Universal grammar

1. Introduction

Local teaching methods, widely perceived as out of date, often receive hasty and harsh criticism from foreign teachers of second language. We urge healthy skepticism of such criticism. The critics themselves, often monolingual, may well have no clear idea what is really happening in the local teacher's classroom, as they generally lack the language skills to observe effectively. In the context of foreign teachers letting off steam, local methods tend to face widespread condemnation (Poole, 2005) as examples of eternal, annoyingly robust, pernicious practices, even though they might in fact be rather recent, reasonably successful innovations.

In this way, the so-called Grammar Translation method, a perennial and easy target for criticism, may have become little more than a procrustean bed for imagined ills in language teaching, reflecting little correspondence to classroom reality. While not denying that bad teachers may exist, irrespective of methodology, we suggest that there is a need to fairly investigate the efficacy of surviving local teaching techniques, even where these may appear at first glance to be unintuitive or even bizarre. In recent years, there has been a partial reevaluation of local methodology and particularly the use of first language (L1), with an array of researchers noting that use of L1 (Butzkamm, 2011; Butzkamm & Caldwell, 2009), translation (Cook, 2010), or translanguaging (Garcia & Wei, 2013) may be highly beneficial when conducted effectively. We urge the need to take a close look at what is going on and give local techniques a fair trial.

The belief in naturalistic approaches that outlaw the use of L1 may, of course, be self-serving. Foreign teachers and researchers who do not speak their students' language have no vested interest in validating its use in pedagogical contexts and may dismiss certain methods too easily. For true believers in naturalistic approaches, first language may come to be viewed as merely an interference to be overcome and a reason for student reticence (Krashen, 1981), rather than as a tool to be usefully exploited.

Blind faith in naturalistic methods has also gained considerable support from mainstream theoretical linguistics, which posits an abstract syntactic Language Faculty (LF) playing a vital role in language acquisition. LF, according to this view, is inextricably implicated in the subconscious acquisition of language, rendering it distinct from ordinary learning (Krashen, 1981). Therefore, language learning is often believed to be a special form of learning, fundamentally not amenable to ordinary learning strategies that might otherwise be considered common sense. While we, of course, do not deny that a great deal of language

learning is subconscious, the same can be said of all forms of learning, such as learning how to ride a bicycle or how to juggle (Draganski et al., 2004; Jones, 2004). We do not deny that language has a special status as the cultural tool *par excellence*, of course, but that does not imply in any compelling way that the use of L1, or other non-naturalistic teaching tools, should automatically be ruled out in language learning.

We review evidence here that belief in LF and, concomitantly, belief in the set of abstract syntactic rules of Universal Grammar (UG) that characterize LF, has become untenable, as historical evidence piles up, forcing a review of our understanding of cognitive evolution and revolution. This kind of new evidence, unfortunately, generally goes unchallenged in the fields of mainstream linguistics and second language acquisition research, and many language teaching professionals remain unaware of alternatives to the Chomskyan model (Goldberg, 2006; Michaelis, 2012). This prompts a reevaluation of non-naturalistic local methods in language learning, with reference to recent evidence related to human evolution. Here we investigate the rather startling use of L2 vocabulary items embedded in L1 texts as an aid to vocabulary learning. We find that the effectiveness of such methods receives reasonably strong and rather surprising statistical support. We conclude that we are certainly not justified in dismissing local methods as necessarily backward (Bax, 2003; Harmer, 2003). On the contrary, it seems that non-local educators may have much to learn by examining local methods that have stood the test of time.

1.1 Difficulties Attempting to Characterize UG

The conviction that language learning is fundamentally special remains the main justification for full adherence to naturalistic (All-English) teaching methods. Language acquisition is still widely believed to be different from other forms of learning. The assumption is that language acquisition depends on innate capacities that are not amenable to normal learning processes (Chomsky, 1988; Clark & Lappin, 2011; Cowie, 2008). Chomsky (1982) offers extremely convincing evidence that humans demonstrate subconscious awareness of hierarchical linguistic structure. However, there is no compelling evidence to support the belief that such an awareness is necessarily rooted in specifically syntactic abilities, situated in a discrete syntactic component in the brain. If the ability to construct and experience such linguistic hierarchy is based in more general cognitive abilities, which is the more natural assumption, then there is no reason to believe that language learning is fundamentally different from other forms of learning.

Roughly speaking, the UG view is that our LF is triggered by target language input, so resort to other forms of input is futile, at least for subconscious acquisition. Of course, no one would deny that appropriate target language input is important. We do not need to believe implicitly in UG to see that optimal use of the target language is indispensable and that overuse of L1 is a problem. However, when one looks at the way English is taught in the localities, and the way that second languages are taught in English-speaking countries, it becomes clear that an array of cultural tools, some of which may be rather surprising, is often employed to ease the process. We take the position that these need to be given a fair trial as they may benefit the native speaker teacher as well as locals.

Modern mainstream linguistics has been largely driven by attempts to characterize rules of universal grammar via analysis of putative levels of syntactic structure (Chomsky, 1965). Despite this effort, UG has defied attempts to pin it down over the last half century or so. In the face of wholesale failure to propose candidate rules for UG, the hypothesis has become an easy target for those who would dismiss it as unscientific due to its unfalsifiable nature (Sampson, 2005). From the end of the last century, the recursive operation Merge, the combination of two objects to form a new unit, was deemed to be fundamental to syntactic operations. Crucially, the ability to recursively combine categories accounts for hierarchy in linguistic constructions. However, an operation such as Merge seems far too general to be a serious candidate as an operation unique to human syntax, or even language, and is now seen as untenable even from within the Chomskyan citadel (Murphy, 2015).

Rather convincing evidence that some languages might not even have (infinite) recursion (Everett, 2008, 2012) is often discussed in this context. However, this constitutes a relatively minor problem (Nevins et al., 2009) for the UG hypothesis. Far more problematic, common sense dictates that recursion is by no means unique to language. In fact, recursion is obviously not a specifically linguistic operation. For example, the Fibonacci Sequence (Livio, 2003) is a classic example of recursion. Recursion clearly occurs in non-linguistic cognitive domains such as mathematics (Lakoff & Núñez, 2000) architecture, and art. Similarly, it is enormously difficult to believe, even given the greatest effort of will, that the appearance of recursion just about everywhere in nature is merely an illusion (Perez, 2010). Thus, it seems we can say that recursion certainly operates in language, but it is blindingly obvious that it is not language-specific.

Also, clear parallels with Merge in the cognitive and natural worlds are not difficult to find. Fauconnier and Turner (2002), for example, indicate that thought and language depends on the human ability to subconsciously blend various elements from a variety of situational sources, with cognitive results that can be predicted. A parallel argument is made by Hofstadter & Sander (2013). Indeed, a whole array of broadly cognitive (non-UG) linguistic investigation (Holyoak & Thagard, 1995; Lakoff, 1987; Lakoff & Johnson, 1980, 1999; Turner, 1996, 2001, 2015) employs conceptual combinations, by no means specific to syntax, as a fundamental explanatory mechanism for cognition itself. Indeed, this consensus is predated in the world of literature, with Koestler (1964) suggesting that a kind of conceptual blending (bisociative thinking) lies at the root of the potential creative ability shared by all creatures. Also, analysis of human cognitive development from archaeological records assumes the combination of existing objects and concepts in innovative ways (Shipton et al., 2013) in the development of early toolmaking skills. We might hypothesize that general conceptual blends are not merely *involved* in thinking. We might even suggest that these blending processes *are* thinking, and surmise that they do not originate with language, let alone syntax.

In approaches to grammar, Goldberg (2006) argues, from a profoundly anti-Chomskyan perspective, that knowledge of language is founded on the human ability to generalize, as compatible grammatical constructions (involving semantics as well as syntax) are combined. Similarly, non-UG unification grammars such as Sign-Based Construction Grammar (Boas &

Sag, 2012) and Head-Driven Phrase Structure Grammar (Pollard & Sag, 1994) rely on a structured blending of information, semantic and phonological, as well as syntactic, inherited from constituents to phrases. In this regard, it should be noted that non-Chomskyan theories of grammar happily employ hierarchical structure without the need to make the unsupported assumption that humans are endowed with UG.

Indeed, Merge may not even be considered an operation exclusive to the cognitive domain. John Stuart Mill (1930[1843]) pointed out that chemical combination of two substances produces a third substance with predictable structural properties, often very different from those of either of the two constituents. As we see entities combining literally everywhere in the known universe, not forgetting the passing on of biological traits from parents to offspring, the characterization of Merge as a specifically syntactic property cannot be maintained. Indeed, as we are now urged to understand UG as a system of thought (Berwick & Chomsky, 2016), it is no longer clear that even Chomsky honestly believes in UG as a specifically syntactic endowment. This makes a confrontation with, and reevaluation of, such matters among language teaching professionals particularly urgent.

As suggested above, even within the Chomskyan (1995) Minimalist domain, the UG characterization has faced considerable criticism. Murphy (2015) plausibly argues that the recursive application of Merge is not only inadequate as an explanation for linguistic phenomena but also that something like a recursive application of Merge, with hierarchical organization of communicative structure, is clearly observable in non-human communication, an observation, it should be noted, that is fully consistent with the human archaeological record. Murphy argues, instead, that a different operation, Label, is a better candidate as a unique human cognitive ability. Human beings certainly have a subconscious ability to intuit what kind of object results from combining two objects, which seems to be broadly parallel to the cognitive blending operations mentioned above. Of course, while this seems to be a promising line of argument for approaches to grammatical analysis, such a view would appear to be already implied in non-Chomskyan unification-based grammars, for example, that have highly structured feature sharing (phonological, syntactic, and semantic information) between constituents and phrases, with syntactic features treated as merely one feature complex inherited from heads. Therefore, it is difficult to find any compelling reason to assume that a subconscious, intuitive awareness of this kind of hierarchical structure (much of which is inescapably semantic in any case) must be unique to syntax, or even language in general.

With respect to second language learning, we obviously expect learners to possess cognitive abilities equal to the task of understanding and learning language. Also, certain activities may not require subconscious “acquisition” to be effective. In the experiment that follows, it is not clear that simple vocabulary learning involves the acquisition of complex hierarchical constructions in L2. However, it does seem extremely likely that the use of L1 involves use of such mental abilities and that this may lessen the cognitive burden and expedite the L2 learning process.

1.2 UG and the Cognitive Revolution

Even more problematic for the Chomskyan conception of UG as a unique syntactic endowment is, without exaggeration, the totality of the human evolutionary record, the fossil record of human remains and associated artefacts. Chomsky claims that UG entered human cognition as a biological endowment by way of random mutation involving a single individual between 50,000 and 100,000 years ago. Given this timing, one might perhaps propose that it was responsible for the Upper Paleolithic cognitive revolution (Berwick & Chomsky, 2016), at a time (roughly 40,000 years ago) when Homo Sapiens made an astonishing leap forward, most famously characterized by the production of unambiguously symbolic cave art.

As UG is assumed, by Chomsky's adherents, to function primarily as a system of thought, rather than as a tool for communication, it might not be entirely implausible to imagine that the UG endowment ushered in evolutionary advantages that enabled it to spread throughout the human population. This conveniently deals with the otherwise insurmountable problem that a purely syntactic endowment would have offered no evolutionary advantage to an individual in a world otherwise without syntax. However, the notion of UG as primarily a system of thought raises serious and obvious questions. First, are we seriously expected to believe that language did not evolve primarily to enable human beings to communicate? Second, why is it helpful to consider UG as a uniquely syntactic endowment *at all* if it just boils down to a more sophisticated means of thinking? After all, if Chomskians are claiming no more than that language ability is ultimately rooted in the evolution of more sophisticated levels of ordinary cognitive processes, there is nothing to distinguish them from their rivals. In fact, it amounts to an admission that those rivals are correct, so there is no meaningful disagreement or reason to believe in UG.

Even a casual reading of the human evolutionary record introduces a cascade of problems for the Chomskyan UG thesis. Assume that Homo Sapiens, anatomically modern humans, who embarked on an exodus from Africa 40,000 ~ 50,000 years ago, were benefitting from a UG-style syntactic endowment. This hypothesis might gain support from the view that Neanderthals (despite possessing bigger brains than Homo Sapiens) lacked figurative art and symbolic behavior (Pääbo, 2014; Berwick & Chomsky, 2016) and, therefore, remained untransformed ape-beings by comparison with their UG-enlightened cousins. However, recent research indicates that the Neanderthals underwent a gradual evolution to a much higher cognitive level than was previously assumed (Finlayson, 2019; Papagianni & Morse, 2013). It seems, from very recent and convincing research, that they *did*, in fact, produce symbolic art (Hoffmann et al., 2018) at a time (at least 64,000 years ago) before there could possibly have been any interbreeding between Homo Sapiens, putative beneficiaries of the UG endowment, and UG-impooverished Homo Neanderthalensis. This means that the previously inexplicably large-brained Neanderthals must have independently undergone a parallel development of symbolic behavior (not to mention speech organs that were by no means entirely unsuited to speech) in the absence of UG. If the Neanderthals managed symbolic behavior without the random emergence of UG and were speaking non-hierarchical

proto-languages (to make the least, albeit condescending, assumption), then there is not the slightest reason to believe that it was necessary for Homo Sapiens either.

In other words, the emergence of complex language in Homo Sapiens must have been gradual, evolving with improved general cognitive processes, over millions of years. This gradualist view is clearly the more commonsense first assumption. This perspective also forces us to see ourselves, not unnaturally, as part of an evolutionary continuum, owing a huge debt of gratitude to increasingly intelligent and resourceful hominins rather than as recipients of a UG blessing in a random singular event that, bizarrely, separates us from the rest of evolutionary history. Evans (2016) suggests that this renders modern humans an evolutionary curiosity, a perspective not shared outside those involved in language studies.

One would readily admit that the demise of the Neanderthals, either killed off or subsumed in the Homo Sapiens population by around 30,000 years ago, is quite likely attributable to the fact that they were somehow not as well adapted to symbolic behavior and speech as Homo Sapiens. However, an explanation for this does not require any belief in the random emergence of a syntactic endowment in Homo Sapiens. In fact, any appeal to relative levels of adaptation automatically invokes a gradualist explanation and immediately rules out the random mutation hypothesis. Language professionals, then, should be clear that the recent Neanderthal evidence makes the idea of UG as a sudden, recent singularity of human cognition entirely untenable. Instead, it seems much more likely that the Homo species were united in communicating verbally and using symbolic behavior but that only the most highly evolved Homo representatives survived in a highly competitive world.

One might also point out, in passing, that recent excavations of a relatively small-brained hominin species, indicate that Homo Naledi (Berger et al., 2017) may have been disposing of their dead or dying around 250,000 years ago, suggesting that they too had the ability to carry out abstract thought processes. In addition, evidence related to the Denisova hominins suggests that they might have been producing startlingly sophisticated technology around 50,000 years ago and, hence, were far more cognitively advanced than might be expected. Such potentially inconvenient evidence should at least be investigated to determine whether it holds any truth, rather than being reflexively dismissed or ignored. Whether or not we take such evidence seriously, the archaeological record is being extended with new findings all the time, particularly in the few years since Berwick and Chomsky (2016) proposed their UG-cognitive revolution hypothesis. By stark contrast, evidence for a unique syntactic mental component has not been forthcoming.

Second language learners find themselves in the predicament of attempting to develop a mental “grammar,” which will certainly be a largely subconscious process. However, it is a normal cognitive process, like other forms of learning, that is likely to be amenable to mediation by a variety of tools. In the experiment under discussion below, L1 is utilized in order to construct contextual information that may speed up the learning process. At the very least, given the evidence from the phylogenetic domain, there seems no reason to object to the use of L1 as necessarily interfering with the learning process in any significant way.

1.3 Physical Reality

A further objection to the Chomskyan random mutation hypothesis is that it ignores all inconvenient, yet indisputable, mental and physical development in the Homo genus in its emergence from the bonobo-like Australopithecus, who lived on this planet between around four and two million years ago. To borrow a term from Sociocultural Theory, the Chomskyan UG hypothesis ignores the phylogenetic domain in its entirety. Perhaps most shocking is that we are expected to be content to believe that true language only became possible, for no reason in particular, at least 100,000 years *after* both the human brain and the human speech organs were fully developed and had evolved to be entirely identical to modern humans. The random mutation thesis does not address these issues. For example, we must assume that the human jaw and teeth had been fully primed for speech articulation for 100,000 years, including in ways that worked counter to post-Neolithic eating habits (Blasi et al., 2019), yet we refuse to acknowledge that anyone was really speaking human languages according to our strict definition.

Even if we assume that humans were only managing linear combinations of words, they must still have managed to convey meaning, and we must assume that those individuals who were the most skillful exponents of this kind of communication were being selected by evolution. Otherwise, we have no explanation for the development of the speech organs. However, if that was the case, the random mutation argument collapses once more because there would have been an obvious evolutionary advantage accruing to individuals who were in possession of better speaking abilities. Thus, the emergence and control of speech organs in Homo Sapiens is clear evidence of gradual evolutionary forces at work. However, if we already know that these gradual evolutionary forces were at work, why do we need an anomaly such as the sudden appearance of a putatively syntactic LF?

This becomes particularly problematic for the UG hypothesis when one considers alternative accounts of how brain size and speech organs may have evolved the way they did among modern humans. For example, Wrangham (2009) offers compelling evidence that early humans (Homo Erectus and before) learned how to control fire and, soon after, to cook their food, roughly two million years ago. This is hypothesized as having had a huge evolutionary effect as it enabled hominins, over time, to get more nutrition from their food with less time spent chewing and, hence, less evolutionary pressure to maintain powerful ape-like jaws that would have been unhelpful when trying to communicate via speech.

As it was easier to get nutrition, this would have made evolutionary increases in brain size and decreases in gut, jaw and tooth size possible. Our ancestors could abandon life in the trees because they had fire to protect themselves from animals while they lived on the ground. As a result, these ancestral hominins were able to become fully bipedal, standing upright, and able to engage in hunting activities that contributed to division of labor and greater social complexity, including the development of family life. If increasingly sophisticated verbal communication was emerging at this much earlier stage, as is implied in Wrangham's account of increased general cognitive complexity and directly hypothesized by Everett (2017), we have a natural, gradualist evolutionary explanation for why our speech organs

evolved in the way they did. Complex language skills would then also be expected to have emerged gradually, as part of generally more sophisticated cognitive abilities, and those individuals who possessed such skills would have held an evolutionary advantage. Random mutation always plays a part in evolution, in this case the gradual emergence of improved cognition and greater control over speech organs adapted for speech, but our first assumption is not the sudden appearance of anomalous singularities. We are not entitled to reject Darwinian theory in explaining the evolution of language merely because no other animal species is in possession of language yet.

The gradualist view would be largely consistent with the hypothesis (Bramble & Lieberman, 2004) that humans evolved as great endurance hunters, taking advantage of upright stance, small guts, and free hands, to chase down prey in the days before the invention of sophisticated weapons. Control of fire enabled humans, over time, to shed body hair and develop more efficient bodily cooling mechanisms conducive to long-distance running. Competing with other animals (who were faster runners, had bigger and stronger teeth, sharper claws, and a better sense of smell, for example) our brainy, long-distance running, early ancestors, who would not have possessed sophisticated weaponry such as bows and arrows, may well have been under evolutionary pressure to become good at reading signs such as footprints and other disturbances to the environment left by much faster-sprinting prey, the proximity of which could only have been deduced from indexical signs (BBC Earth, 2009). One might also speculate that the best sign-readers would have achieved extremely high status in their communities, so evolution would have favored them.

It would not be entirely fanciful to imagine that such requirements would have contributed to a control over signs and, perhaps, symbolic hand-gestures that may have become useful in the silent tracking of prey. This may have been significant in the transition to the use of language. It is quite possible to imagine evolution favoring individuals who developed communicational abilities that contributed to social organization involving high-performance hunting and cooking activities. As we have a reasonable hypothesis, that fits with Darwinian theory, regarding how language ability may have evolved over millions of years, there is no need for an unfounded theory of UG as a cognitive anomaly that suddenly appeared spontaneously, in a single individual, for no reason, after large brain size and speech organs were fully developed.

Shipton (2010) argues that evidence from Acheulean tools indicates that unique human cognitive abilities evolved much more gradually, emerging in a propensity for imitation and shared intentionality (Tomasello, 1999; Tomasello & Rakoczy, 2003) with roots developing from at least 2 million years ago. Stout et al. (2008) suggest that increased sophistication in toolmaking ability from the Early Stone Age (from 2.6 to 0.25 million years ago) would likely have proceeded in tandem with the development of language ability and would probably have involved overlapping mental development. Shipton et al. (2013) argue that generativity, the ability to create new forms out of previously existing elements, is evident in Acheulean technology, and may have emerged with the evolution of improvements in working memory. There is an abundance of evidence to support the view that complex language skills evolved gradually as part of a natural evolution of general cognition in

sociocultural activity. By contrast, it is extremely difficult to find any compelling evidence to support the view that language emerged full-blown in a single individual, realized as abstract syntactic abilities that would have had absolutely no cognitive counterpart in human experience up until that time.

It seems reasonable to suggest that mainstream and applied linguistics needs to confront evidence related to human evolution, drawn from observation of the historical-cultural domain. Very often, language professionals know nothing of linguistics beyond Chomskyan theory. Monolingual teachers are likely to find UG-inspired methodology, that does not sanction the use of L1, attractive. There may be political forces favoring monolingual teaching methods. Publishers prefer all-English textbooks with global appeal. Therefore, it is worth emphasizing that language learning itself should be guided more by ideas gleaned from the local sociocultural domain and, therefore, be more concerned with local teaching methods that occasionally allow the use of L1.

1.4 UG and Language Learning

In the absence of a theory of UG, which clearly lacks convincing supporting evidence, the arguments in favor of full adherence to naturalistic methods is severely weakened. It also severely weakens the case for blind belief in UG (black box) research in SLA, often based on deficient evidence, as the only “scientific” option (Long, 2007) even where evidence for UG is entirely lacking. As cultural evolution is inevitably situated at the center of our understanding of language and language learning, we need to examine the efficacy of what may be the surprising use of cultural tools, in their sociocultural context. Regarding such an investigation, we must certainly acknowledge that language is a special cultural tool and that L1 is, indeed, the most salient element of local culture. In line with Vygotsky (1980) we might propose that language and culture are mutually constructing and are not usefully separated. However, once belief in UG is abandoned, we can certainly hypothesize that language learning may, in fact, be amenable to mediation by cultural tools in normal ways and that L1, the most important cultural tool but still an aspect of cultural learning, may be helpful in learning L2.

2. The Experiment

This clearly suggests the need for an investigation of the use of L1 as a useful tool, mediating L2 learning in specific contexts. In response to this, we looked at specific forms of L1 use recommended by one Japanese teacher/researcher.

Ikeda (2009) claims that the burden of L2 vocabulary learning is eased by effective use of L1. Specifically, he claims that target L2 words embedded in L1 text will be learned more easily. We decided to test this rather startling claim that vocabulary learning proceeds more effectively where target L2 words are provided with an easily-accessible context.

Therefore, the research questions are:

Which type of vocabulary instruction (L2 vocabulary embedded in L1 text vocabulary instruction versus traditional vocabulary instruction) is more effective in the development of language learners' lexical knowledge?

Which type of vocabulary instruction (L2 vocabulary embedded in L1 text vocabulary instruction versus traditional vocabulary instruction) is more effective in the sustainability of lexical knowledge gained by language learners?

Is there any relationship between language learners' level of language proficiency and their level of lexical knowledge?

Accordingly, the null hypotheses are:

There is no significant difference between the effectiveness of type of vocabulary instruction (L2 vocabulary embedded in L1 text vocabulary instruction versus traditional vocabulary instruction) in the development of language learners' lexical knowledge.

There is no significant difference between the effectiveness of type of vocabulary instruction (L2 vocabulary embedded in L1 text vocabulary instruction versus traditional vocabulary instruction) in the sustainability of lexical knowledge gained by language learners.

There is no significant relationship between language learners' level of language proficiency and their level of lexical knowledge.

3. Method

3.1 Participants

Participants in the study were a total of 140 university freshmen in the Department of English at Tsuru University, Japan. The first language of all participants was Japanese. The English proficiency level for TOEIC ranged from 300 to 760 (the average 512.3 standard variance 93.38), CEFR equivalent level *A2 (elementary)* to *B1 (intermediate)*.

Participants were grouped in four separate classes from A to D. The A class was the upper level, composed of 31 students, with a TOEIC range above 590, the average being 664.03. Class B, C, D were determined alphabetically, with TOEIC scores ranging from 275 to 585 the average being 480. The participants in each class were randomly divided into two groups, one as the experimental group, and the other as the control group. Participants remained in the same group through each session of the experiment.

3.2 Instruments

Materials used for intervention sessions consisted of vocabulary items of a high order of difficulty, selected from *Eiken* Grade 1 textbooks, introduced in reading practice materials, supported by voice recordings. The two conditions were administered via L1-embedded reading texts and L2-only reading texts. Target items were written in bold and in red in either case. For both conditions, the meaning of the target items and phonetic representations were provided on the back of the sheets given to students.

The data collection instruments consisted of paper tests. Three tests of equal level of difficulty were developed for the study: one to be used prior to the intervention as pre-test, one to be used immediately following the intervention as immediate post-test, and one to be used a period of time following intervention as delayed post-test. Each test contained 20 vocabulary items. Tests involved matching Japanese meanings with randomly ordered target items or writing the meaning of target items in Japanese. Correct scores gleaned from these tests, conducted at the relevant stage, were recorded for each student.

3.3 Procedure

The study followed an experimental design. At the beginning of the first semester of the academic year 2019/2020 through April and May, the vocabulary test used as pre-test was administered to all participants of the study. The students in the four classes considered for the experiment were randomly assigned to one of the instructional methods: 70 students were assigned to L2 vocabulary embedded in L1 text vocabulary instruction (hereafter referred to as experimental group) and the other 70 students were assigned to traditional vocabulary instruction (hereafter referred to as control group). The instruction was conducted for two sessions a week, each intervention spanning 30 minutes, over a period of four weeks, comprising a total of 8 sessions of instruction in general. The intervention for the experimental group was in the form of reading comprehension tasks in Japanese with target vocabulary embedded in English and being highlighted whereas the intervention for the control group was in the form of reading comprehension tasks in English with target vocabulary being highlighted. Then respectively one week following the intervention in early May and four weeks following the intervention in late May, the vocabulary test used as immediate post-test and the vocabulary test used as delayed post-test were administered to all participants.

3.4 Data Analysis

The performance of language learners on vocabulary tests used as pre-test, immediate post-test, and delayed post-test were analyzed by giving 1 mark to each correct answer and 0 mark to each incorrect answer. As there were twenty items on each test, each participant could get a score ranging from 0 to 20 on each of the tests.

To compare the performance of language learners in experimental and control groups, the mean and standard deviation of performance of language learners in both groups on pre-test, immediate post-test, and delayed post-test were computed.

Mixed between-within subjects analysis of variance was then employed to determine whether the different types of instruction had resulted in any learning gains in language learners from pre-test to immediate post-test and whether they sustained the gains after a period of time. Mixed between-within subjects analysis of variance allows combining between-subjects and within-subjects variables in one analysis (Pallant, 2013). In this respect, both the effectiveness of intervention in general (within-subjects effect) and the effectiveness of specific type of intervention, that is, L2 vocabulary embedded in L1 text vocabulary instruction and traditional vocabulary instruction (between-subjects effect) was considered.

Partial eta squared was then used to examine the level of effect of intervention for both within-subjects and between-subjects categories. Partial eta squared can range from 0 to 1 and represents the proportion of variance in the dependent variable (lexical knowledge) that is explained by the independent variable (type of intervention) (Pallant, 2013). Cohen (1988) proposed a set of guidelines to interpret the values of partial eta squared. The guidelines for interpreting the values of eta squared, as proposed by Cohen (1988), are presented in Table 1. Finally, the graphical presentation of the performance of language learners in both experimental and control groups on the pre-test, immediate post-test, and delayed post-test was provided.

Table 1. Interpretation of partial eta squared values

Value	Effect Size
0.01	Small Effect
0.06	Moderate Effect
0.14	Large Effect

The relationship between language learners' language proficiency and their lexical knowledge was assessed through correlating language learners' TOEIC scores and their performance on the immediate post-test using Pearson product-moment correlation coefficient (r). Pearson product-moment correlation coefficient (r) is used to measure the degree and the direction of the linear relationship between two variables (Gravetter & Wallnau, 2013). The value of Pearson product-moment correlation coefficient (r) can range from -1 to $+1$. The positive and negative signs indicate whether there is a positive correlation (as one variable increases, the other variable increases as well) or a negative correlation (as one variable increases, the other variable decreases). The size of the value, regardless of the sign, provides an indication of the strength of the relationship. Values of closer to $+1$ or -1 are indicative of higher correlation between the two variables, whereas values of closer to 0 are indicative of a lower correlation (Pallant, 2013). Cohen (1988) suggests a set of guidelines to interpret the values between 0 and 1 . The guidelines, which have been presented in Table 2, apply whether or not there is a positive or negative sign out the front of the correlation value (r value).

Table 2. Strength of relationship

r Value	Interpretation
0.10 – 0.29	Small Correlation
0.30 – 0.49	Medium Correlation
0.50 – 1.00	Large Correlation

The squared correlation (r^2), called the coefficient of determination, was then used to measure the proportion of variability in lexical knowledge that can be determined from its relationship with language proficiency. Squared correlation would give a value ranging from 0 to 1 . Cohen (1988) has also suggested a set of guidelines to interpret the values of squared correlation. The criterion for interpreting the value of squared correlation (r^2), as proposed by Cohen (1988), has been presented in Table 3.

Table 3. Percentage of variance explained, r^2

r^2 Value	Interpretation
0.01	Small Correlation
0.09	Medium Correlation
0.25	Large Correlation

4. Findings & Discussion

4.1 Findings

Effect of L2 Vocabulary Embedded in L1 Text versus Traditional Vocabulary Instruction on the Development of Lexical Knowledge.

Table 4 presents the results of the descriptive analysis of performance of language learners on pre-test, immediate post-test, and delayed post-test. The descriptive analysis presented in the table consists of the number of participants for each type of instruction as well as the mean and standard deviation obtained for the performance of each group of participants on pre-test, immediate post-test, and delayed post-test. According to the descriptive analysis of the data, the mean score for the performance of language learners in both experimental and control groups on lexical knowledge enhanced in immediate post-test and delayed post-test. However, the mean score obtained by language learners in the experimental group was higher than the mean score obtained by language learners in the control group immediately following the intervention and within a period of four weeks following the intervention, although both groups experienced a decline in performance within this post-intervention period. The mean score by itself, however, does not show whether the difference among the three tests and between the two groups is considered significant or not. To determine whether the difference among mean scores obtained by the two groups over the three tests is significantly different from one another or not, the results of the analysis of mixed between-within subjects analysis of variance need to be observed.

Table 4. Descriptive analysis of language learners' performance on vocabulary tests

	Type of		Std. Deviation	N
	Instruction	Mean		
Pre-test	L2-Embedding	0.04	0.204	70
	Traditional	0.13	0.563	70
	Total	0.09	0.424	140
Post-test	L2-Embedding	14.87	5.130	70
	Traditional	12.31	5.358	70
	Total	13.59	5.382	140
Delayed Post-test	L2-Embedding	5.77	3.868	70
	Traditional	5.14	3.935	70
	Total	5.46	3.901	140

Table 5 presents the results of the main effect for within-subjects variable (Test: pre-test, immediate post-test, delayed post-test). To explore the main effect for within-subjects

variable, the value of Wilks' Lambda and the associated probability value given in the column labeled Significance (Sig.) needs to be considered. All of the multivariate tests yield the same result. However, the most commonly reported statistic is Wilks' Lambda (Pallant, 2013). A significance value of above 0.05 ($p > 0.05$) for Wilks' Lambda indicates a non-significant effect whereas a significance value of equal to or less than 0.05 ($p \leq 0.05$) is indicative of a significant effect (Gravetter & Wallnau, 2013). In the data obtained in the current study, the value for Wilks' Lambda for Test is 0.12, with a significance value of 0.00 (which really means $p < 0.05$). Because the probability value is less than 0.05, there is a statistically significant effect for Test. This suggests that there was a change in lexical knowledge across the three different tests. The main effect for Test was significant.

 Table 5. Multivariate tests^a

Effect	Value	F	Hypothesis		Sig.	Partial		Observed Power ^c	
			df	Error df		Eta Squared	Noncent. Parameter		
Test	Pillai's Trace	0.87	474.28 ^b	2.00	137.00	0.00	0.87	948.57	1.00
	Wilks' Lambda	0.12	474.28 ^b	2.00	137.00	0.00	0.87	948.57	1.00
	Hotelling's Trace	6.92	474.28 ^b	2.00	137.00	0.00	0.87	948.57	1.00
	Roy's Largest Root	6.92	474.28 ^b	2.00	137.00	0.00	0.87	948.57	1.00
Test * Group	Pillai's Trace	0.06	4.92 ^b	2.00	137.00	0.01	0.06	9.84	0.80
	Wilks' Lambda	0.93	4.92 ^b	2.00	137.00	0.01	0.06	9.84	0.80
	Hotelling's Trace	0.07	4.92 ^b	2.00	137.00	0.01	0.06	9.84	0.80
	Roy's Largest Root	0.07	4.92 ^b	2.00	137.00	0.01	0.06	9.84	0.80

a. Design: Intercept + Group

Within Subjects Design: Test

b. Exact statistic

c. Computed using alpha =

Although a statistically significant difference among language learners' lexical knowledge on different tests (pre-test, immediate post-test, and delayed post-test) was found, the effect size of this result also needs to be considered in order to determine the exact size of this difference. In this regard, the value of interest is partial eta squared. The value of partial eta squared obtained for Test in this study is 0.87. Using the commonly used guidelines proposed by

Cohen (1988), this result suggests a very large effect size. Expressed as a percentage, 87% of variance in the change in language learners' lexical knowledge is explained by the treatment they received between the three tests.

Now that the within-subjects effects have been explored, the main effect for between-subjects variable (type of instruction: L2 vocabulary embedded in L1 text vocabulary instruction versus traditional vocabulary instruction) needs to be considered. The results that need to be considered are in Table 6. In this respect, the significance value across the row labeled Group (variable name for the type of instruction) should be considered. A significance value of above 0.05 ($p > 0.05$) for Group indicates a non-significant effect whereas a significance value of equal to or less than 0.05 ($p \leq 0.05$) is indicative of a significant effect (Gravetter & Wallnau, 2013). In the data obtained in the current study, the value for Group is 0.03. This is less than the alpha level of 0.05, so the main effect for Group is significant. There was a significant difference in lexical knowledge for the two groups (those who were instructed based on L2 vocabulary embedded in L1 text method and those who were instructed based on traditional method).

Table 6. Tests of between-subjects effects

Source	Type III		Mean Square	F	Sig.	Partial		Observed Power ^a
	Sum of Squares	df				Eta Squared	Noncent. Parameter	
Intercept	17088.19	1	17088.19	736.15	0.00	0.84	736.15	1.00
Group	112.11	1	112.11	4.83	0.03	0.03	4.83	0.58
Error	3203.35	138	23.21					

a. Computed using alpha =

Although a statistically significant difference between the performance of language learners in the two groups (experimental group and control group) was found, the effect size of this result also needs to be considered to be able to determine the exact size of this difference. In this regard, the value of interest is again partial eta squared. The value of partial eta squared obtained for Group in this study is 0.03 which, according to the guidelines proposed by Cohen (1988), indicates a small effect size. Expressed as a percentage, 3 percent of variance in the performance of language learners on the vocabulary tests is explained by the type of treatment they received.

The graphical presentation of language learners' lexical knowledge for both experimental and control groups in pre-test, immediate post-test, and delayed post-test has been depicted in Fig. 1. As Fig. 1 shows, although language learners in both groups had a better performance on immediate post-test, language learners in the experimental group had a slightly better performance than their counterparts in control group immediately following the intervention. Both groups then experienced a decline within a period of time following the intervention, with a non-significant difference.

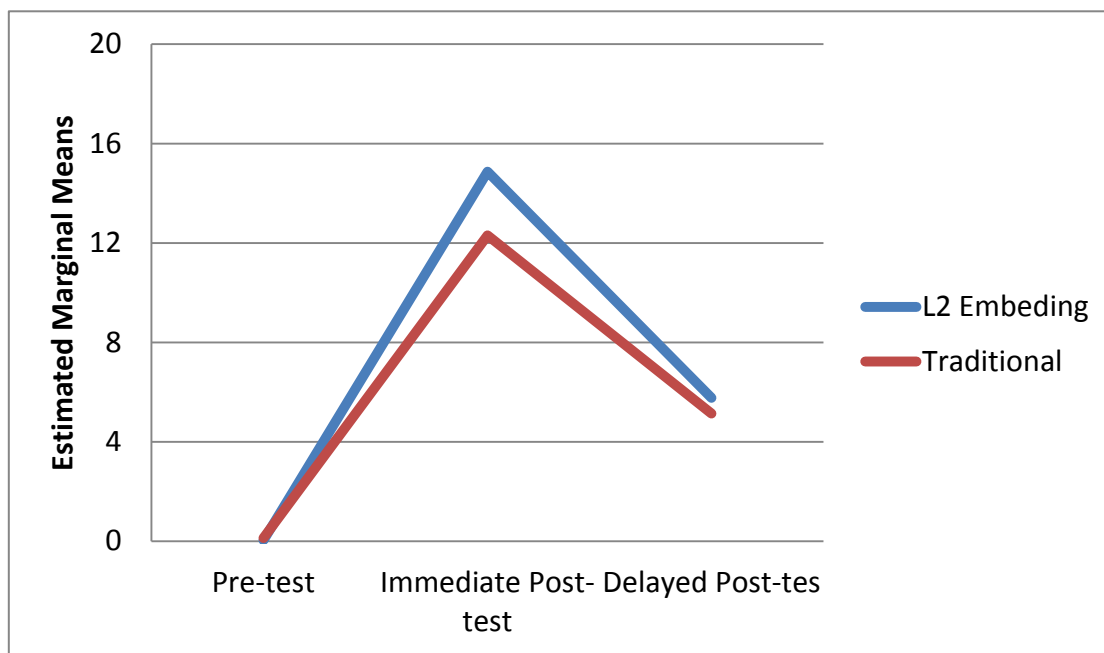


Figure 1. Language learners’ performance on pre-test, immediate post-test, and delayed post-test

Relationship between Language Proficiency and Lexical Knowledge.

Table 8 presents the results of Pearson product-moment correlation coefficient (r) analysis for language learners’ level of language proficiency and their lexical knowledge. The first thing to consider in correlation analysis is the direction of the relationship between the two variables (language proficiency and lexical knowledge). The data shows that there is a positive relationship between the two variables, that is, as language proficiency increases so too does lexical knowledge. The second thing to consider in correlation analysis is the size of the value of the correlation coefficient. This value indicates the strength of the relationship between the two variables (language proficiency and lexical knowledge). The value of correlation coefficient obtained in the analysis of Pearson product-moment correlation coefficient (r) is 0.20 which according to the guidelines proposed by Cohen (1988) to interpret the values of correlation coefficient suggests a weak relationship between language proficiency and lexical knowledge.

Table 8. Relationship between lexical knowledge and language proficiency

		Language Proficiency	Lexical Knowledge
Language Proficiency	Pearson Correlation	1	0.20*
	Sig. (2-tailed)		0.01
	N	140	140
Lexical Knowledge	Pearson Correlation	0.20*	1
	Sig. (2-tailed)	0.01	
	N	140	140

*. Correlation is significant at the 0.05 level (2-tailed).

To get an idea of how much variance the two variables (language proficiency and lexical knowledge) share, the coefficient of determination was calculated. This can be obtained by squaring the correlation value. The coefficient of determination for the obtained correlation analysis is $r^2 = (0.20)^2 = 0.04$ which according to the guidelines proposed by Cohen (1988) to interpret the values of coefficient of determination suggests a small correlation coefficient. To convert the value of coefficient of determination to ‘percentage of variance’, it was multiplied by 100, that is, $r^2 = (0.20)^2 \times 100 = 4$. This suggests that language proficiency helps to explain only 4% of the variance in language learners’ lexical knowledge. The graphical presentation of the relationship between language proficiency and lexical knowledge has been depicted in Figure 2.

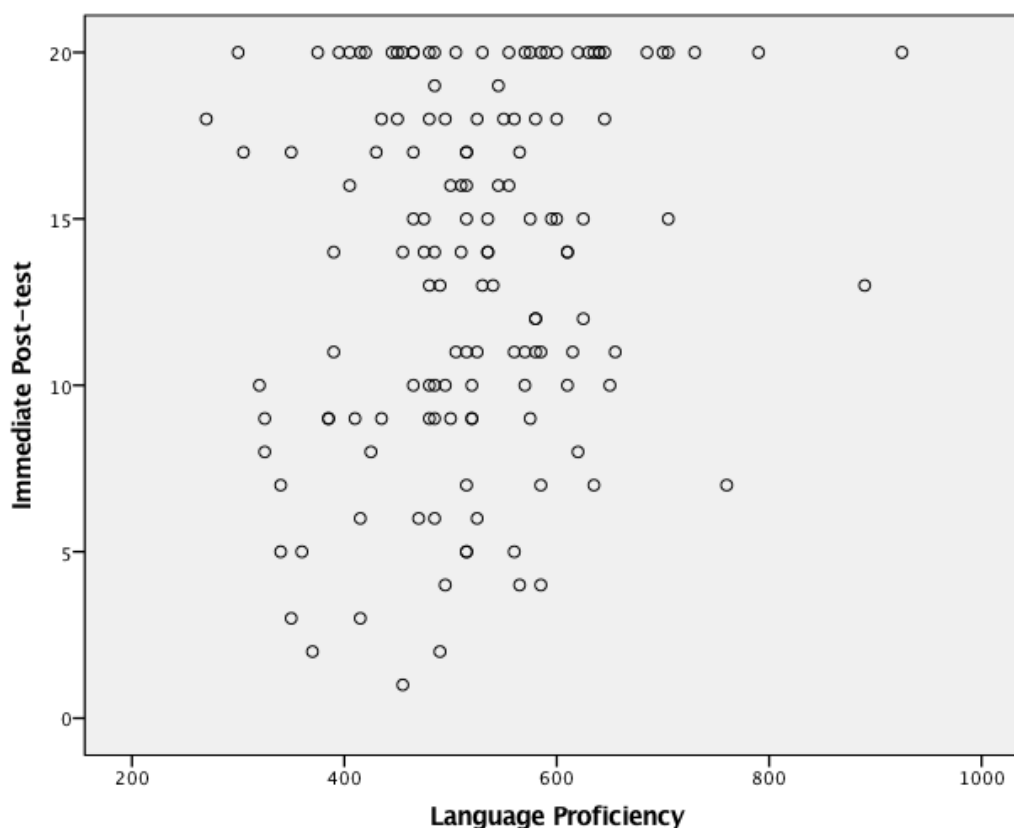


Figure 2. Relationship between language proficiency and lexical knowledge

4.2 Discussion

The study compared the effectiveness of two methods of vocabulary instruction including L2 vocabulary embedded in L1 text vocabulary instruction and traditional vocabulary instruction on the immediate and sustainable development of language learners’ lexical knowledge. The study also investigated the relationship between language proficiency and lexical knowledge. The study found that although both methods of vocabulary instruction are effective in the immediate and sustainable development of language learners’ lexical knowledge, L2 vocabulary embedded in L1 text vocabulary instruction led to a significantly higher performance. The study also found that although language learners tend to lose part of the

obtained knowledge over time, they could manage to sustain part of the obtained knowledge. In the current study, language learners who went through the L2 vocabulary embedded in L1 text vocabulary instruction outperformed the language learners who learned target vocabulary through traditional method of vocabulary instruction. This higher performance was evident in both immediate and delayed post-tests. Therefore, the first and second null hypotheses of the study which state that there is no significant difference between the effectiveness of type of vocabulary instruction (L2 vocabulary embedded in L1 text vocabulary instruction versus traditional vocabulary instruction) in the development of language learners' lexical knowledge and that there is no significant difference between the effectiveness of type of vocabulary instruction (L2 vocabulary embedded in L1 text vocabulary instruction versus traditional vocabulary instruction) in the sustainability of language learners' obtained lexical knowledge are rejected.

The study also found, incidentally, that language proficiency is a surprisingly weak predictor of lexical knowledge and being at a higher level of language proficiency is a small indicator of a higher lexical knowledge. In the current study, language learners who were at higher level of language proficiency performed slightly better on the vocabulary tests than language learners at lower level of language proficiency. Therefore, the third null hypothesis of the study which states that there is no significant relationship between language learners' level of language proficiency and their level of lexical knowledge is also rejected.

The findings obtained in the current study can be explained through the fact that language is a cultural tool rather than an abstract syntactic object, situated in some abstract space, that may only be triggered by use of the target language. While language is indeed a special tool, ultimately inseparable from culture, with culture and language mutually constructing each other, we still expect language learning to be amenable to normal support via cultural tools. This includes the use of L1, the most salient element of local culture, in mediating learning.

Lantolf (2000) and Lantolf and Poehner (2014) are considered the seminal works on sociocultural theory in SLA. They strongly urge a closing of the divide between theory and practice such that we see language acquisition taking place, in social activity, and in specific sociocultural contexts. Human language ability also developed gradually, in specific cultural activities, as a communicative tool. The evidence is clear that it did not suddenly enter the human evolutionary record as an anomaly, for no reason, before language even existed. Sociocultural theory tells us that mental development is mediated via the use of cultural tools. In the phylogenetic domain, language is likely to have featured as the cultural tool par excellence, in gradual mental development over a period in excess of 2 million years, making other forms of mental development possible (Stout et al., 2008). One might mention that it is odd that these seminal works in sociocultural SLA make virtually no reference to L1 as a mediating tool in language learning. The fact that the use of L1 is not given much attention even in sociocultural approaches to language learning must be indicative of continued bias among native speaker researchers and teachers working in the Anglosphere. For a theory of mental development that relies on the use of cultural tools to have virtually no mention of L1 suggests that far greater attention to this problem is needed from researchers in the localities. There are significant forces working against the use of L1 in second language learning and

there is good reason to believe that much work needs to be done in order to address this unfortunate bias that tends to work against local teachers. Theories of human language ability have been over-influenced by a belief in an abstract language faculty that is hypothesized as running directly counter to the generally accepted story of evolution in the sociocultural domain. Language is a cultural tool, and L1 may be effective in aiding in the acquisition of L2. In a world without UG, in which L1 may be considered the cultural tool par excellence, the above serves as an early step in the direction of a new reality in SLA, one that is consistent with the true place of language in human evolution.

5. Conclusion

The study found that although language learners in both groups experienced a significantly sharp increase in their lexical knowledge immediately following instruction, language learners in the experimental group had a significantly better performance than language learners in the control group. The study also found that although language learners in both groups managed to retain part of the obtained lexical knowledge over a four-week period following instruction, language learners in the experimental group retained a significantly higher rate of obtained knowledge than language learners in the control group. The study also found that language proficiency is a weak predictor of lexical knowledge and being at a higher level of language proficiency is a small indicator of a higher lexical knowledge. Given the startling effectiveness of local practices involving the use of L1, there is good reason to believe that a reevaluation of linguistic theory and theories of language learning are overdue. In particular, the place of sociocultural context in language learning requires urgent attention. The experiment offers strong support for local teaching approaches that utilize L1 as a cultural tool in L2 learning. Local methods that have stood the test of time should certainly be investigated, while it is important for non-local teachers to keep an open mind regarding the effectiveness of such approaches. These findings have important consequences for sociocultural approaches to second language acquisition. Language is a special cultural tool, but language learning, like all forms of mental development, takes place in culture, and L1 may play a useful part in aiding this developmental process. Language teachers, it seems, should not be afraid of exploiting learners' linguistic experience and expertise in their own language (Widdowson, 2003) in order to facilitate learning of L2.

References

- Bax, S. (2003). The end of CLT: A context approach to language teaching. *English Language Teaching Journal*, 57(3), 278-287.
- BBC Earth. (2009). *Human Mammal, Human Hunter Attenborough Life of Mammals BBC*. Retrieved from https://www.youtube.com/watch?v=826HMLoiE_o
- Berwick, R. C., & Chomsky, N. (2016). *Why only us: language and evolution*. Cambridge, Mass: MIT Press.
- Boas, H. C., & Sag, I. A. (2012). *Sign-based construction grammar*. Stanford, California: CSLI Publications.

- Berger, L. R., Hawks, J., Dirks, P. H. G. M., Elliott, M., & Roberts, E. M. (2017). *Homo Naledi and Pleistocene hominin evolution in subequatorial Africa*. Retrieved from <https://elifesciences.org/articles/24234#digest>
- Blasi, D. E., Moran, S., Moisik, S. R., Widmer, P., Dediu, D., & Bickel, B. (2019). Human sound systems are shaped by post-Neolithic changes in bite configuration. *Science*, 363, 6432.
- Bramble, D. M., & Lieberman, D. E. (2004). Endurance running and the evolution of homo. *Nature*, 432, 345-352.
- Butzkamm, W. (2011). Why make them crawl if they can walk? Teaching with mother tongue support. *RELC Journal*, 42(3), 379-391.
- Butzkamm, W., & Caldwell, J. A. W. (2009). *The bilingual reform: A paradigm shift in foreign language teaching*. Tübingen: Narr.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.
- Chomsky, N. (1982). Some concepts and consequences of the theory of government and Binding. *Linguistic Inquiry Monograph*, 6. MIT Press.
- Chomsky, N. (1988). *Language and problems of knowledge*. Cambridge, Massachusetts: MIT Press.
- Chomsky, N. (1995). *The minimalist program*. Cambridge, Mass: The MIT Press.
- Clark, A., & Lappin, S. (2011). *Linguistic nativism and the poverty of the stimulus*. Wiley.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cook, G. (2010). *Translation in language teaching: An argument for reassessment*. Oxford: Oxford University Press.
- Cowie, F. (2008). Innateness and language. *Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford.
- Draganski, B., Gaser, C., Busch, V., Schuierer, G., Bogdahn, U., & May, A. (2004). Neuroplasticity: Changes in grey matter induced by training. *Nature*, 427, 311-312.
- Evans, V. (2016). *Why only us: The language paradox*. New Scientist. Retrieved February 24, 2016, from <https://www.newscientist.com/article/2078294-why-only-us-the-language-paradox/>
- Everett, D. L. (2008). *Don't sleep, there are snakes: Life and language in the Amazonian jungle*. New York: Pantheon Books.
- Everett, D. L. (2012). *Language: The cultural tool*. New York: Pantheon Books.
- Everett, D. L. (2017). *How language began: The story of humanity's greatest invention*. Liveright (W.W. Norton).

- Fauconnier, G., & Turner, M. (2002). *The way we think: Conceptual blending and the mind's hidden complexities*. Basic Books.
- Finlayson, C. (2019). *The smart Neanderthal: Bird-catching, cave art, and the cognitive revolution*. Oxford: Oxford University Press.
- Garcia, O., & Wei, L. (2013). *Translanguaging: Language, bilingualism and education*. Palgrave Pivot.
- Goldberg, A. (2006). *Constructions at work: The nature of generalizations in language*. Oxford: Oxford University Press.
- Gravetter, F. J., & Wallnau, L. B. (2013). *Statistics for the behavioral sciences*. Belmont, CA: Wadsworth Publishing.
- Harmer, J. (2003). Popular culture, methods, and context. *English Language Teaching Journal*, 57(3), 288-294.
- Hoffman, D. L., Standish, C. D., Garc á-Diez, M. P., Pettitt, P. B., Milton, J. A. J., Zilh ão, J., ... Pike, A. W. G. (2018). U-Th dating of carbonate crusts reveals Neanderthal origin of Iberian cave art. *Science* 23, 359(6379), 912-915.
- Hofstadter, D., & Sander, E. (2013). *Surfaces and essences: Analogy as the fuel and fire of thinking*. Basic Books.
- Holyoak, K. J., & Thagard, P. (1995). *Mental leaps: Analogy in creative thought*. Cambridge, MA: MIT Press.
- Ikeda, K. (2009). *Eitango, kousureba hayaku oboerareru! English vocabulary. Do this and you can learn more quickly!*. Nihon Jitsugyou Shuppansha.
- Jones, R. (2004). Juggling boosts the brain. *Nature Reviews Neuroscience*, 5, 170.
- Koestler, A. (1964). *The act of creation*. Last Century Media.
- Krashen, S. D. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon Press.
- Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. The University of Chicago Press.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: The University of Chicago Press.
- Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. Basic Books.
- Lakoff, G., & Núñez, R. E. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. Basic Books.
- Lantolf, J. P. (2000). *Sociocultural theory and second language learning*. Oxford: Oxford University Press.

- Lantolf, J. P., & Poehner, M. E. (2014). *Sociocultural theory and the pedagogical imperative in L2 education: Vygotskian praxis and the research/practice divide*. London: Routledge.
- Livio, M. (2003). *The golden ratio: The story of phi, the world's most astonishing number*. Broadway Books.
- Long, M. H. (2007). *Problems in SLA*. Lawrence Erlbaum Associates Publishers.
- Michaelis, L. A. (2012). Making the case for construction grammar. In H. C. Boas, & I. A. Sag (Eds.), *Sign-based construction grammar* (pp.31-68). CSLI Publications.
- Mill, J. S. (1930 [1843]). *A system of logic ratiocinative and inductive*. London: Longmans, Green and Co.
- Murphy, E. (2015). Label, cognomes, and cyclic computation: an ethological perspective. *Frontiers in Psychology*, 6, 1-18.
- Nevins, A. (2009). Evidence and argumentation: a reply to Everett. *Language*, 85(3), 671-681.
- Pääbo, S. (2014). *Neanderthal man: In search of lost genomes*. Basic Books.
- Pallant, J. (2013). *SPSS survival manual: A step by step guide to data analysis using spss program* (5th ed.). Australia: Allen & Unwin.
- Papagianni, D., & Morse, A. (2013). *The Neanderthals rediscovered: How modern science is rewriting their story*. Thames & Hudson.
- Perez, J.-C. (2010). Codon populations in single-stranded whole human genome DNA are fractal and fine-tuned by the golden ratio 1.618. *Interdiscip Sci*, 2(4), 373.
- Pollard, C., & Sag, I. A. (1994). *Head-driven phrase structure grammar*. Chicago: The University of Chicago Press.
- Poole, G. S. (2005). Reform of the university English language teaching curriculum in Japan: A case study. In J. S. Eades, R. Goodman, & Y. Hada (Eds.), *The 'big bang' in Japanese higher education* (pp. 242-273). Melbourne: Trans Pacific Press.
- Sampson, G. (2005). *The "language instinct" debate*. Continuum International Publishing Group.
- Shipton, C. (2010). Imitation and shared intentionality in the Acheulean. *Cambridge Archaeological Journal*, 20(2), 197-210.
- Shipton, C., Clarkson, C., Pal, J. N., Jones, S. C., Roberts, R. G., Harris, C., ... Petraglia, M. D. (2013). Generativity, hierarchical action and recursion in the technology of the Acheulean to Middle Palaeolithic transition: A perspective from Patpara, the Son Valley, India. *Journal of Human Evolution*, 65(2), 93-108.
- Stout, D., Toth, N., Schick, K., & Chaminade, T. (2008). Neural correlates of Early Stone Age toolmaking: technology, language and cognition in human evolution. *Philosophical Transactions of the Royal Society*, 363, 1939-1949.

- Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.
- Tomasello, M., & Rakoczy, H. (2003). What makes human cognition unique? From individual to shared collective intentionality. *Mind and Language*, 18, 121- 47.
- Turner, M. (1996). *The literary mind: The origins of thought and language*. Oxford: Oxford University Press.
- Turner, M. (2001). *Cognitive dimensions of social science*. Oxford: Oxford University Press.
- Turner, M. (2015). *The origin of ideas: Blending, creativity, and the human spark*. Oxford: Oxford University Press.
- Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Widdowson, H. G. (2003). *Defining issues in English language teaching*. Oxford: Oxford University Press.
- Wrangham, R. (2009). *Catching fire: how cooking made us human*. Basic Books.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>)