A Thesis

entitled

Exploring the Chinese Room:
Parallel Sensory Input in Second Language Learning

by

Yifei Xin

Submitted to the Graduate Faculty as partial fulfillment of the requirements for the
Master of Arts Degree in English with a Concentration in ESL

_________________________________________
Dr. Douglas Coleman, Committee Chair

_________________________________________
Dr. Anchung Cheng, Committee Member

_________________________________________
Dr. Stephen Christman, Committee Member

_________________________________________
Dr. Patricia Komuniecki, Dean
College of Graduate Studies

The University of Toledo

May 2012
Language input in SLA is commonly defined in terms of what Saussure (1959) calls the “objects of language”, and this assumption has been widespread because of Chomsky’s theoretical framework for language learning, first put forward in the 1960’s. The idea that input in language learning is target language (TL) has been dominant in the field of Second Language Acquisition (SLA) that it heavily influences not only theoretical research but also teaching practices. However, what is commonly perceived as language – what people hear, read or see (in the case of sign language) does not carry meanings and cannot contribute to language learning. In other words, language is not physically real. As Klein (1986) shows, “what makes learning possible is the information received in parallel to the linguistic input in the narrower sense [generated by speech]” (p. 44). He illustrates this by using a Chinese Room analogy: if we put a person in a room and exposed him/her to nothing but recorded Chinese speech, he/she will not learn no matter how long the experiment lasts (Klein, 1986, p. 44). This study aims to empirically test this thought experiment, by providing two groups of subjects with two different treatments – one being speech and text, which is commonly perceived as language, and the other one being parallel sensory input, which is the sound of speech in parallel with
other sensory experiences. The proposed outcome of the study is that people cannot learn grammar, or rather the appropriate order of relevant speech articulation motor tasks, in a target language just by hearing the sound of speech and reading the text in the language to be learned. Thus, the study sets out to prove that the relevant input for language learning consists of the sound of speech in parallel with other sensory experiences, and what people learn is the ability to communicate.
To Dr. Coleman, for all the office hours I have taken from you,
and for all the guidance you have given me.
Acknowledgements

I would like to express my gratitude to all who have encouraged and supported me throughout my academic career.

My greatest appreciation goes to my thesis and graduate advisor, Dr. Coleman, without whose wise and patient guidance I would not be able to grow with and through the process of this thesis. Not only has he supported my growth as a scholar, he also challenged me to look at language learning and teaching from a real world perspective, which has changed my view towards the field of Second Language Acquisition entirely.

I would also like to thank Dr. Cheng for her kind support and insightful advice through the writing of this thesis, and Dr. Christman for his critical comments and constant push for better clarity.

I especially need to thank my mother, Yan Xin, who supported me unconditionally throughout my academic career, encouraged me in difficult times and provided guidance when I needed them.

Finally, I would like to say thank you to everyone else who has given me the support that is needed in order to complete this endeavor.
# Table of Contents

Abstract iii  
Acknowledgements vi  
Table of Contents vii  
List of Figures ix  

**I. Chapter 1. Input in Second Language Learning**  
A. Mainstream Assumptions about Input and Language Learning 1  
B. Why Language Input Alone Cannot Contribute to Language Learning 7  
C. Language Learning Device 13  
D. The True Nature of Input in Language Learning 17  
E. Hypothesis 21  

**II. Chapter 2. Methodology**  
A. Input Design 23  
B. Assessment Design 27  
C. Subjects 30  
D. Procedure 30  
E. Data Collection 31  

**III. Chapter 3. Data Analysis**  

**III. Chapter 4. Discussion**  
A. Implications 36  
B. Limitations 38  
C. Conclusion Remarks 40
References

Appendices

A. Script and Translation for Audio Used by Actors
   45
B. A Complete List of Storyboard Images
   48
C. Test Items
   53
D. Coding Key to Response Type
   54
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>Chomsky’s assumed language learning process.</td>
<td>3</td>
</tr>
<tr>
<td>Figure 2</td>
<td>Pies.</td>
<td>12</td>
</tr>
<tr>
<td>Figure 3</td>
<td>A neuron.</td>
<td>17</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Sample script for input design.</td>
<td>25</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Input design for experimental group and control Group.</td>
<td>27</td>
</tr>
<tr>
<td>Figure 6</td>
<td>Sample test items.</td>
<td>28</td>
</tr>
<tr>
<td>Figure 7</td>
<td>Sample test item breakdowns.</td>
<td>29</td>
</tr>
<tr>
<td>Figure 8</td>
<td>Sample coding key for both sense and form.</td>
<td>32</td>
</tr>
</tbody>
</table>
Chapter One

Input in Second Language Learning

Second Language Acquisition (SLA) is commonly described as the process of internalization of new knowledge, retention of information or skill, gained through either formal instruction or informal interaction with the target language (TL). Many factors are assumed to influence this learning process, both internal and external. The internal factors include mental disposition of a learner, the ability of comprehending the communication process or the aptitude. The external factors include the frequency, duration and quality of the target language exposure – in other words the input.

Mainstream Assumptions about Input and Language Learning

Traditionally, the underlying assumption about input in the field of SLA is that “input constitutes the language to which the learner is exposed” (Ellis, 1986, p. 299). In other words, input is the target language that “comes from a variety of sources including the language that the learner hears (e.g. in the classroom by the teacher, outside of the classroom by speakers of the second language), reads (in textbooks, in other reading materials), or sees, in the case of a signed language” (Gass and Selinker, 2008, p. 31). This assumption is that input exists in terms of what Saussure (1959) calls the “objects of language”, such as words or morphemes arranged in hierarchical relations in sentences possessing certain grammatical structures and so on (Coleman, 2011a, p. 1). Such an underlying assumption became widespread not only in the field of First Language Acquisition but also Second Language Acquisition after Chomsky adopted the concept and incorporated into his extremely influential theoretical framework of language.
Learning in the 1960’s, which included theories of Language Learning Device (LLD), generative grammar, and later on Universal Grammar (UG).

According to Chomsky (1964), input is considered to be “primary linguistic data”, which is defined as “well-formed sentences” (p. 26). He argues that the process of language learning cannot be separated from the exposure to “primary linguistic data”, and the mechanism of language learning formulates from innate processes (Chomsky, 1964). He notes that if a human baby and a kitten are both exposed to exactly the same linguistic data, the human baby will be able to acquire the ability to understand and produce language, while the kitten will never acquire either ability (Chomsky, 1975, p. 171). Chomsky (1964) attributes this relevant capacity human has and the kitten lacks to an innate device in the brain, which assimilates language and its varying complexities – also known as the Language Learning Device. Chomsky (1975) makes the argument that "all children share the same internal constraints which characterize narrowly the grammar they are going to construct" (p. 98). According to him, there is “no reason for supposing the mental world to be an exception, since we live in a biological world” (Chomsky, 1975, p. 94). In other words, under Chomsky’s assumptions, language learning is ultimately a process of receiving input, also known as “primary linguistic data”, stimulating and processing such input via the Language Learning Device, and producing output, which is generative grammar. Chomsky’s (1964) initial model of the Language Learning Device is shown in Fig. 1.
Gold (1967) criticizes this assumed language learning process from an logical point of view, proving that a generative grammar is essentially unlearnable given the assumption that the input is “primary linguistic data”, or well-formed sentences in the language. He points out that both positive evidence, which is a set of grammatical sentences the language learner has access to, and negative evidence, which is the evidence available to the language learner about what is not grammatical, are needed for such a process given Chomsky’s assumption. Thus, it can be inferred that learners need to be presented with both positive and negative evidence for these particular grammatical patterns, meaning that they need to hear sentences that are both “right” and “wrong”.

On the basis of supposed lack of adequate data, Chomsky (1975) responds to Gold’s (1967) criticism, by proposing the concept of Universal Grammar (UG). He accepts the premise that language patterns are unlearnable, and assumes that generative grammar is not learned but innate. The basic structure of his argument, as summarized by Cook (1993), is that since a native speaker of a language knows something about the language and this language feature is provably unlearnable given only the primary linguistic data, it can be concluded that the language feature is not learned from experience and thus the language feature must be built-in the mind (pp. 207-208). In other words, UG is presented by Chomsky (1975) as a solution to the generative grammar is unlearnable argument and is defined as “a characterization of the genetically
determined language faculty” (p. 3). It is commonly assumed that the Language Learning Device itself can be seen as synonymous with the language faculty, i.e., Universal Grammar (Cook, 1996, p. 81).

It should be noted that Chomsky (1964) also distinguishes language competence from language performance. Linguistic competence is usually defined as “knowledge of a language however it happens to be represented [and] this knowledge might correspond to a psychogrammar – defined here as the grammatical information used in parsing and producing a language” (Carroll, 2001, p. 24). As Chomsky (1965) puts it, competence is “the speaker/hearer’s knowledge of his language”, whereas performance is the actual use of language in concrete situations (p. 4). Chomsky (1964, 1965, 1971, 1975, 2002 and 2005) explicitly points out in many occasions that the output produced according to the Language Learning Device model is linguistic competence – the linguistic faculty of possessing a grammar consisting of principles, parameters and lexicon, and it is a limited set of rules for organizing language, out of which people develop their language abilities later on. It can be inferred that according to this distinction, people who receive a certain amount of “primary linguistic data” will have the ability of processing what is commonly considered to be grammatical structures without necessarily having the ability of actually communicating through the target language.

This framework about input and language learning has been widely accepted by linguists in the field of Second Language Acquisition. Cook (1993) researches extensively about the basic structural argument of the Language Learning Device and UG, and supports the point that language learning should be understood in terms of “how the [learner’s] mind turns the language input it encounters into a grammar by using its
built-in capabilities” (p. 1). He explains that “children hear a number of sentences said by their parents and other caretakers – the ‘primary linguistic data’; they process these in some fashion within their black box [brain], and they acquire linguistic competence in the language, i.e., a ‘generative grammar’” (Cook, 1996, p. 79). Through his understanding, the grammar contains the appropriate setting for the parameters of an assumed syntactic pattern, and it has thousands of lexical entries specifying how each word can behave in the sentence (Cook, 1996, p. 81).

Carroll (2001) holds a similar point of view. She accepts the basic assumptions about Language Learning Device and supports Chomsky’s viewpoint of UG, by arguing that language learning is “input driven” in the sense that learning mechanisms are triggered when parsing of an input fails (p. 68). UG, in other words, must provide the learner with information as to what a possible trigger can be (Carroll, 2001, p. 70). In other words, as Coleman (2011b) puts it, “[Carroll believes that] input has no role in language learning per se… It may function as a selectional trigger to what is already known, but it has no other possible role” (p. 6). This is also the position taken by many proponents of Chomsky’s theory (Coleman, 2011b, p. 6).

Krashen (1983, 1985) generally accepts Chomsky’s framework, yet differentiates himself from Chomsky by proposing the concept of comprehensible input. On the one hand, he explicitly points out that “input is the essential environmental ingredient”, and the acquirer “does not simply acquire what he hears – here is a significant contribution of the internal language processor (Chomsky’s Language Acquisition Device: LAD)” (Krashen, 1985, p. 2). He believes that LAD (Krashen’s version of Language Learning Device) generates possible rules according to innate procedures (p. 3). On the other hand,
he further argues that input needs to be made comprehensible, with the help of context, which includes “extra-linguistic information, our knowledge of the world, and previously acquired linguistic competence” (Krashen, 1985, pp. 2-3). The argument made here seems self-contradictory, yet Krashen (1985) believes that comprehensible input does not obey Chomsky’s definition of input, since input is essentially still the “primary linguistic data” and what makes input comprehensible is not part of the input. Perhaps the quote below can best express Krashen’s (1985) point of view on this subject.

The extensive evidence for the Input Hypothesis, reviewed below, supports Chomsky’s position, and extends it to second-language acquisition. We may see individual variation ‘on the surface’ – different sources of comprehensible input, different strategies for obtaining input, different messages, and of course different languages – and this variation may be of practical concern. But deep down, the ‘mental organ’ for language (Chomsky 1975) produces one basic product, a human language, in one influential way. (Krashen, 1985, p. 3)

The assumptions that input is language or objects of the language are also widely supported by most researchers. Gass and Selinker’s (2008) influential work in SLA puts UG at the core of the learning process, and claim that language input consists of objects of language. According to Gass’ diagram, input is defined as “a body of second language data” (Gass and Selinker, 2008, p. 400). Saleemi (1992) provides an overview for several versions of Chomsky’s paradigm, all of which seem to “share the assumption that the data for language leaning – also referred to as ‘the available evidence’ (p. 3), or
‘environmental input’ (p. 10) – consists of language” (as cited in Coleman, 2011a, p. 2). Morgan (1986) presents an alternative explanation, which essentially is still Chomsky’s view, that not only input consists of “primary linguistic data”, it also contains “bracketing information about the hierarchical structures present in it (as cited in Coleman, 2011a, p. 2). Ellis (2008) explicitly points out that input is language, and it serves as the data that learners use to determine the rules of the target language. The definition of input provided by VanPatten (1996) limits to meaning-bearing input, which refers to “language that the learner hears or sees that is used to communicate a message” (p. 6).

From reviewing the literature, it is clear that two underlying assumptions are made in terms of input and language learning: (1) input is the target language; (2) language learning is the processes of receiving input, and constructing knowledge of the grammatical structures of the target language, with the help of the innate abilities humans possess. However, upon a closer examination, both of these assumptions have fundamental flaws and it should be made clear that such input alone cannot contribute to language learning.

**Why Language Input Alone Cannot Contribute to Language Learning**

In order to explain why the generally assumed input alone cannot contribute to language learning, we first need to understand what language is. From early Sumerians to ancient Greeks, language has been regarded as an inherent part of the human existence that lends itself to study yet eludes an unequivocal definition (Postica, 2006, p. 1). “For centuries, schools of thought have re-defined ‘language’ to best fit the intellectual framework of their time” (Postica, 2006, p. 2). Thus, language has been considered as any of the following.
[Language is] something material, something behavioural, something mental, something biological, something abstract, something social or something cultural – to mention some of the stuffs language has been supposed to consist of. As for shapes, language is thrust down one’s throat as a thing, pushed as process, promoted as a procedure, auctioned as action, flogged as form, sold as system and marketed as means. Which bring us to diversity in design, conceptions of language being available in many a mode: nominalist, conceptualist, realist, obscurantist, eclecticist and so on. And each conception of language has its own finish. As for origin: conceptions of language come with all kinds of credentials, tagged as Aristotelian, Platonic, Cartesian, Humboldtian, Saussurian, Bloomfieldian, Sapirean, Wittgensteinian, Chomskyan, and so on and so forth. (Botha, 1992, p. xii)

Early grammarians began the formal study of language by systematically classifying sounds into consonants, vowels, and word classes such as nouns and verbs, dividing sounds into phonemes, formulating morphological and syntactical rules and so on (Bloomfield, 1914). In the 18th century, the use of the comparative linguistic method was first introduced. Bloomfield (1914) believes that through the use of comparative linguistic method, “the scientific study of language was broadened from one particular language to languages in general” (p. 310). Early in the 20th century, Saussure introduced the idea of language as a “semantic code”, which stresses examining language as a static system of interconnected units (Clarke, 1990, p. 143). He thus brought the shift from diachronic to synchronic analysis, meaning linguistic phenomena was viewed as
developments through time instead of at one point in time, as well as introduced several basic dimensions of semiotic analysis that are still important today (Clarke, 1990). In the 1960’s, Chomsky (1964, 1965, 1968) formulated the generative theory of language, which viewed the basic form of language as a set of syntactic rules that were universal for all humans.

Over the centuries, the study of language thus has been developing into a “science” that concerns itself with all aspects relating to language, and examines it from all of the theoretical viewpoints described above. However, as Yngve (1996) reviews in his book *From Grammar to Science*,

The whole point of modern science is to increase our knowledge of the natural world through a program of observation, experiment, and the development and testing of theories, where the criterion of acceptance of observational facts is their public reproducibility when questioned, and the criterion of acceptance of theories is their ability to pass appropriate tests against observational evidence when challenged. These criteria can be applied only in investigations in the physical domain, the domain of natural sciences, where observation by the sense is possible. (pp. 21-22)

By adopting the scientific method of studying, most linguists, thus, look at language as a real world observable property, with no exception in the field of SLA. The physical reality of language is assumed to be what people hear, read or see (i.e., in the case of sign language or lip reading) and one of the properties of such the reality is that language is meaning bearing, which is why it can be analyzed and studied. This is also the
fundamental reason why input, defined as the target language, seemingly can contribute to language learning.

However, language does not carry either meanings or grammatical structures that convey meaning from one person to another. In searching for the physical reality of what people hear, a presumable object of study in linguistics, Yngve (1996) provides the following examples.

A runner on the women’s Olympic team is at the starting mark. A gun is discharged and she hears the noise. With as short a delay as possible she starts to run. In this case the equipment in her head has been set to interpret the noise from the starting gun as a starting signal and to act on it as quickly as possible.

…

A schoolboy is about to start in a race. The starter says, “Go!” The boy starts to run with as short a delay as possible. Here the boy is set to start when he hears a different sounding noise.

…

After dinner the boy and his sister are considering playing a board game. “What game should we play?” he asks. “Go!” she says. Here the equipment in the boy’s head is set to interpret the noise as the name of a game, like checkers or chess. He does not start running with as short a delay as possible. (Yngve, 1996, pp. 3-4)

Through these examples, two observations can be made: different noises can be
interpreted in the same way, and the same noises can be interpreted differently (Yngve, 1996, pp. 3-4). The physical property of what people hear in all three scenarios is the acoustics quality of the sound waves; however, the interpretation of such speech sounds depends not only on what people hear, but also on the “equipment in the head of the speaker or hearer and the momentary state of the speaker or hearer and how the state changes reflecting the tasks in which the noise plays a part” (Yngve, 1996, p. 4). Sound waves, which can be scientifically observed and measured, clearly exist in the physical domain, and it is also clear that the interpretation of such sound is purely subjective. Therefore, Yngve (1996) concludes, “the common conception of utterances as having grammatical structure and carrying meaning conceals false assumptions from the tradition that will not stand up to even elementary scientific examination and are in fact incorrect” (p. 4).

A similar argument can be made in terms of written text, another form of language. In one example given by Coleman (2011d), the written word “pies” can be interpreted in several different ways depends on the reader’s interpretation (see Fig. 2), and he explains,

> [s]ince you are reading this (the word pies) in English, you probably will say “picture B” without hesitation. But if I were to ask someone else a similar question in Spanish, the person would respond by indicating picture C, and if I asked a similar question of a person who understood Polish, the response would be for that person to point at picture A. When we perceive words, phrases, and the grammatical structure of a sentence in marks on paper, we are actually projecting properties of ourselves outward
onto the marks. If we treat those marks as if they were an object having properties on its own, we are making a serious error. (pp. 8-9)

![Figure 2 – Pies (Coleman, 2011d, p. 8)](image)

Similar to speech sounds, the physical property of written texts is the ink on paper. It has a certain mass, density, flexibility, and adhered to its surface there are lings of graphite with traces of some other substances mixed in (Coleman, 2011a, p. 4). There is no such thing in the real world as a text that carries linguistic segmentation, or structure of any sort, whether in terms of phonemes, words, sentences, or any other of the constructs (Yngve, 1996, p. 9). We project our subjective interpretation onto the marks, and confuse our interpretation with the physical reality.

In short, language, perceived as “a natural phenomenon, the object of a science, a type of faculty, a kind of module, a type of object, a type of stuff, a type of system, as voluntary behavior, as something used, as something taught and learned, as having learned elements, as having patterns, as something spoken, heard, and learned, as something processed, as something organized and structured, as something produced and comprehended, and as data” (Yngve, 1996, p. 10), exists in the logical domain and does not have physical existence. It arises in our subjective experience because of our
individual properties and our point of views (Coleman, 2011a, p. 5). As Yngve (1996) puts it,

[Saussure] even exclaimed that the illusion [that] things [are] naturally given in language is profound. The illusion certainly is profound: many linguists still appear to believe that the objects of language exist in nature and are thus given in advance and appropriate candidates for scientific study. But Saussure was right. It is an illusion. (p. 30)

Thus, it is absurd to assume that “primary linguistic data” alone—defined as well-formed sentences (Chomsky, 1964, p. 26) can contribute to language learning, since the physical reality of language does not carry any meanings or what is considered as grammatical structures.

Language Learning Device

Now, what about the innate ability, the build-in structure, or the “mental organ” that processes input, enables input to arrive at the grammar of a language from the available data in the available time? Chomsky (1964, 1972) argues for the existence of Language Learning Device in details in a black box problem metaphor—something goes into a black box, something comes out of it; by looking at what goes in and what comes out, it is possible to arrive at some understanding of the process concealed inside the black box itself. He (as cited in Cook, 1996) goes on to explain that,

[s]uppose we see barley and empty bottles going in one door of a distillery, bottles of Scotch whisky coming out of the other; we can deduce what is going on inside by working out what must be done to the barley to
get whisky. Given a detailed analysis of the whisky and of the barley we could deduce the process through which one is transformed into the other.

Children hear a number of sentences said by their parents and other caretakers – ‘the primary linguistic data’; they process these in some fashion within their black box, called the Language Acquisition Device (LAD), and they acquire linguistic competence in the language, i.e. a ‘generative grammar’. We can deduce what is going on inside the child’s LAD by careful examination and comparison of the language input that goes in – the material out of which language is constructed – and the knowledge of language that comes out – the grammar. (p. 79)

However, first of all, I should point out when children interact with their parents and other caretakers, in extension to all human communicative interactions, what “goes into the black box” is not just well-formed sentences, since it has been made clear through previous analysis that there are no sentences, only sound of speech that do not carry either structure or meaning. Along with the sound of speech, people also perceive the context that is associated to them, which is the key factor in making the speech sound comprehensible. For instance, a mother walks into the living room, sees a broken vase on the floor and says, “Who broke the vase”? The child, who broke the vase earlier, hears the sound of speech, sees the mother pointing at the broken vase on the floor and giving an angry facial expression, and senses anger from the tone of the mother’s voice based on prior experiences, thus, says “I’m sorry mom”. In this case, the child takes in not only the sound of speech, but also the context, facial expressions, gestures, and the tone of voice.

As Deacon (1997) says, “Learning is not any one general process. Learning always
occurs in a particular context, involving particular senses and types of motor actions, and certain ways of organizing the information involved” (p. 48).

Now, we can take a look at the existence of Language Learning Device more closely. We speak of the body in terms of organs – the heart, the lungs, the liver, and so on. Chomsky (1980) believes that “we may usefully think of the language faculty, the number faculty, and others as systems of mental organs as well, analogous to the heart or the visual system or the system of motor coordination and planning” (p. 29). He argues that on the one hand “the theory of language is simply a part of human psychology that is concerned with one particular ‘mental organ’”, which processes human language (Chomsky, 2002, p. 36); and on the other that “the study of language falls naturally within human biology” (Chomsky, 2002, p. 123). Proponents of Chomsky’s theory believe that this language organ is physically present among other mental organs and should be described in biological, as well as psychological terms (Cook, 1996, p. 32), even though the precise physical location and form of Language Learning Device are yet unknown. As far as researchers in the neurobiology and psychology are concerned, there is no identifiable area of the brain functioning as the “language organ” and “parts of the brain which are involved in speech and the understanding of speech are also involved in other activities” (Lieberman, 2000). Lieberman (2000) also points out that, data suggest neuronal processes “are not genetically specified [as Chomsky’s theory claims], but are instead acquired phenotypically in a manner similar to the neural bases of various aspects of motor control” (p. 38).

Instead of assuming the existence of a “mental organ”, we can look at another more established explanation of “what is happening in the black box”: learning involves
changes in an individual’s neurology. “It is now generally accepted that such changes do not involve any significant physical rearrangement of neurons; rather it is via the modification of connections among neurons that learning occurs” (Coleman, 2011b, p. 29). Learning, at a basic physiological level, involves a complex “construction” process in which initially undifferentiated sensory impressions, properties of the developing organism’s brain and the organism’s unsuccessful behaviors interact in a dynamic and changing environment (Lawson, 2003, p. 5). The relevant neurons of the central nervous system, interneurons, are multipolar and form associations with other neurons, and thus, when a electrical impulse is triggered in the affected neuron depending on the overall excitatory and inhibitory effects of adjacent neurons (Hole, 1993, p. 347-350). To be more specific, as Coleman (2011b) points out,

[b]etween one neuron and the next, an impulse is one-way: from the exon of the first to a dendrite or the neuronal body of the second. The electrical impulse travels toward [down] a neuron’s axon, which typically ends in a set of synaptic knobs. When the impulse arrives at a synaptic knob, the synaptic vesicles contained there release a neurotransmitter which diffuses across the synaptic cleft and reacts with specific receptors of the postsynaptic membrane. (p. 30)

Thus, learning can be defined as the changes in linguistic properties of the communicating individual or their underlying neurological bases, and can be explained at a physiological base in one simple way: it is a biochemical modification a synaptic strengths potentials (Lawson, 2003). Fig. 3 shows the structure of a neuron.
Since we can explain “what happens in the black box” at a physiological level more scientifically, there is no need to assume the function of Language Learning Device, especially when is absolutely no empirical nor scientific evidence supporting the existence of it.

**The True Nature of Input in Language Learning**

Krashen (1983, 1985) is one of the few who have recognized that input cannot merely consist of “primary linguistic data”. In his Input Hypothesis, he explicitly points out that “we are able to understand language containing unacquired grammar with the help of context, which includes extra-linguistic information, our knowledge of the world, and previously acquired linguistic competence” (Krashen, 1985, p. 2). His viewpoint is that input itself is not sufficient, and additional contextual information is necessary for SLA because they are what make input comprehensible, even though Krashen still supports Chomsky’s theory about input and language learning and never denies the existence of LAD and the mainstream assumptions about input.
Saville-Troike (1985) holds a similar point of view. She argues that input contain not only “information about the phonological, grammatical, and lexical nature of the language”, but also “cultural information” (p. 52). Pinker (1994), a proponent of Chomsky’s theory agrees, saying that “through speech input is necessary for speech development, a mere soundtrack is not sufficient” (p. 278). Wong-Fillmore (1985) makes a point in terms of a teacher who is trying to explain a word to students in the target language. As Coleman (2011b) cites it,

[the example [given by Wong-Fillmore] concerns a teacher trying to explain the English word weak to a class of Spanish speakers. She provides a rather verbose explanation, which “it is doubtful that they could have made much sense out of” (Wong-Fillmore p. 35). Although what the teacher said in English she then translated into Spanish, the students could not possibly hold the English in mind until after understanding the translation, and while giving the explanation in English “she did nothing that allowed them to figure out what her words meant”. (p. 36)

While Yngve (1996) never comments explicitly on the definition of input, his model of energy flow channels in a way provides us with a basis for describing input in Hard Science Linguistics (HSL). According to him, input can consist of energy flow within an interaction, including gestures, facial expressions, objects in a relevant setting and the setting itself, and anything that the person says or does relevant to that linkage, which includes any and all concurrent events relevant to the linkage (Yngve, 1996).
Therefore, we can draw a conclusion that input is anything a participant in a linkage
experiences through the senses related to the communicative event.

In 1986, Klein provides a logical explanation for the issue once for all, pointing
out that speech, or “linguistic input in the narrower sense” cannot be input. He explains
this by using his version of Chinese Room analogy.

Suppose you were locked in a room and were continually exposed to the
sound of Chinese coming from a loudspeaker; however long the
experiment continued, you would not end up speaking Chinese; at most,
you might know the parts of the phonological system. (Klein, 1986, p. 44)

The implication of this thought experiment is that the sound of speech, which is thought
of as “language” by most people, do not carry either meanings or any form of
grammatical structures. Klein (1986) goes on and gives the following example.

Suppose you are a Japanese visitor and you happen to be in Germany
without knowing a single word of German. You are having breakfast in
your hotel with a group of Germans. One of the Germans turns to you and
produces a sequence of speech sounds like this:


Klein (1986) points out here that linguistic input in the narrower sense (the sound waves)
will not lead to language learning, since to the Japanese speaker, what is being heard is
only a string of noise that does not make any sense. However, as cited in Coleman
(2011a),
For example, suppose the speaker looks at you, raises his eyebrows, glances down at the table toward the salt and pepper shakers, gestures toward them and then speaks, holding out an open hand afterwards? For the Japanese speaker, there probably is too much new here for any language learning to occur. He/she might be perplexed as to which thing the German wanted, but would certainly be aware that the speaker was making a request. He/she might offer both, and let the German choose. If you know English, however, you would be likely to guess that one part of the sequence, [szalt], refers to salt, and would simply pass the salt: there would be both comprehension and learning. (p. 7)

The same reasoning process can also lead to the conclusion that in classroom teaching, if a teacher shows the saltshaker and provide the same communicative context along with the phonological information, there would also be both comprehension and learning. As Klein (1986) concludes,

What makes learning possible is the information in parallel to the linguistic input in the narrower sense (the sound waves): the learner must know who is speaking to whom, when and where, he must be able to watch the accompanying body language (gesture, facial expression, etc.) and he must note the reactions of the listener. Eventually he should be able to establish a relationship between identifiable segments of the sound stream and particular segments of the parallel information. (p. 44)

In the real world, people communicate through numerous ways, the most
particular way being through what is commonly called the use of language (Postica, 2006, p. 3). In other words, it's commonly accepted that human possess language ability, through which messages can be exchanged and meanings can be conveyed. Researchers, therefore, discuss human communication through the perspective of the reality of language. However, as previously discussed, the assumed properties of language – meanings and grammatical structures, do not exist in the physical domain. The physical reality of such assumed properties exist in the form of human communication (Yngve, 1996). Speech (sound waves), “who is speaking to whom”, “when and where”, accompanying body languages, “the reactions of the listener” are all parts of a communicative channel that exist in the real world (Yngve, 1996, Coleman, 2005). All of sensory experiences involved in such communicative channel contribute to language learning, and thus should be part of input. Coleman (2005) provides this definition of input, showing that input includes all of the parallel sensory events occurring during a given communicative event.

**Hypothesis**

In the *Chinese Room* thought experiment, Klein (1986) clearly indicates that people who are exposed to the sound of Chinese will not end up speaking Chinese. However, according to the mainstream assumptions, by being exposed to “primary linguistic data”, people will gain linguistic competence – the ability to perceive the underlying assumed grammatical patterns of the target language. In my thesis, I intend to empirically test this assumption with the help of the *Chinese Room* analogy. Since according to Chomsky’s (1975) theory, language competence and language performance should be clearly separated, I will also accept this premise by testing only grammatical
pattern recognition instead of meaning comprehension. The first group of subjects in my study will be given the “primary linguistic data input” – speech and text, and the second group of subjects will be given parallel sensory input – sound of speech in parallel with other sensory experiences that can be associated with specific elements of the speech.

Hₐ: subjects who are exposed to the parallel sensory input will perform better on the grammatical pattern recognition test than subjects who are given the primary linguistic data input.

Hₒ: subjects who are exposed to the parallel sensory input will not perform better or will perform equally on the grammatical pattern recognition test in compare to the subjects who are given the primary linguistic data input.
Chapter 2

Methodology

Input Design

In order to empirically carry out Klein’ s (1986) thought experiment, it was imperative to design two sets of input – one for the control group and the other one for the experimental group. The input for the control group should simulate the Chinese Room contexts, in which the subjects would be exposed solely to what Chomsky (1964) claims to be language input – “primary linguistic data”, which contains well-formed sentences. The experimental group, on the other hand, would receive parallel sensory input, as Klein points out that “what makes learning possible is the information received in parallel to the linguistic input in the narrower sense (the sound waves [generated by speech])” (p. 44).

With this in mind, a research team from Applied Linguistics I class, including undergraduate student Samuel Johnston, graduate student Sultana Nargis and me, first started to design a short dialog as the base of the input. After discussing and reviewing each of the designs from all the team members, we decided to go with the version I drafted and I revised it into a final draft based on the suggestions given by Sam, Nargis and Dr. Douglas Coleman, who was the supervisor of the research team. The dialog featured a common restaurant scene among four characters: three customers and one waiter. The three customers come in a restaurant, get seated, and the waiter takes drink order and food order from the customers. The difficulty of the dialog was kept to an elementary level, and the duration of the conversation is around three minutes,
considering the length of the learning exposure during the experiment, the form of input and the capacity of human short-term memory.

The form of the language was a key factor in this study because it was essential to find a language that was seldom spoken and that had few people that had any level of proficiency within the community. This was troublesome because “a significant portion of the student body at The University of Toledo were native speakers of languages other than English” (Ziegler, 2007, p. 19). The native speakers of English in the community were also likely to have foreign language exposures through secondary education. It could be inferred that the majority of the potential subjects of the study possessed a certain level of the skill of communicating through a non-native language. In order to avoid the confounding variable that the subjects tested might be familiar with the input to which they were to be exposed, we adopted an artificial language called Térus, designed by Douglas W. Coleman of The University of Toledo. Térus was created by systematically rotating the place of articulation of Polish consonants while preserving a simplified syntactical structure based very loosely on Polish – “a method inspired by Saussure’s analogy of the chess game” (Postica, 2006, p. 34). For instance, [t] became [k] and [k] became [p]; thus, Polish [tak] became Térus [kap] with [a] remained unchanged (p. 34). Térus had been used several times in previous studies (Postica, 2006; Sun & Bubalo, 2011), and the method of creating an artificial language had inspired other researchers as well (Ziegler, 2007). Therefore, the dialog, which I originally drafted in English, was translated into Térus with minor adjustments (by Coleman). A portion of the script used to develop the lessons can be seen in Fig. 4.
Additionally, the script was very carefully designed to show the physical relationship of the objects and actors. Therefore, basic meanings of the speech could be shown and observed in terms of the gestures of the actors, their interactions with each other, and their interactions with the objects in the communicative setting. For instance, when saying “Ekē shkē gra tamēta” (“here is table for you”), the waiter would be facing the three customers, indicating the table while making sweeping gestures at the three customers. Moreover, referential speech was frequently used in the dialogue, such as the combination of Ekē trok/ Ekē trokēt (here is/here are) and the coordinating objects.

Finally, the dialog contained well-formed sentences, which according to Chomsky (1964) is the essential element of “primary linguistic data”. In addition to the form of the sentences, we also emphasized on several grammatical features by repeating these patterns throughout the dialogue in order to reinforce learning experiences, such as the gender of nouns, “thanks” as a refusal or acceptance, singular/plural issues, and so on. It was important to do so because the objective of the experiment was to assess subjects’
ability to recognize what are considered to be grammatical patterns by most linguists. See Appendix A for a complete script of the dialog used for input development.

After we finished designing the base of the input – the dialog, the research team members each took one of the character roles and recorded the lines individually. Professor Douglas W. Coleman then edited the files into parts of the final dialog. In order to show the physical relationships of the characters and objects in the setting, Dr. Coleman designed a storyboard using 3D software to make images to fit his translation of the script I drafted. The storyboard contained a series of images that simulated the real world communicative event. Each image demonstrated information about the characters, their interactions, the objects they were interacting with and the setting in which the communicative event took place. The audio file was then captioned into the storyboard, along with the script in callout form. A complete list of storyboard images is shown in Appendix B.

Therefore, the primary linguistic data group (control group) would be given a PowerPoint presentation, through which they would hear the lines and see the text in the callouts (See Fig. 5). Subjects in the parallel sensory experience group would also hear the lines and see the text callouts, but they would also see the events pictured in a storyboard sequence as they unfolded (see Fig. 5)
Assessment Design

Since the duration of the input exposure was less than ten minutes and the difficulty of the input was kept to an elementary level, it was vital to design an assessment method that not only matched the level of difficulty but also served the purpose of examining the recognition of patterns. After carefully considering many possibilities, the research team decided to use a multiple-choice test as the assessment tool. Each of the team members undertook the responsibility of designing 3 to 4 test items that featured one specific syntactic and morphological pattern in English according to the original English dialog. Coleman then selected and translated the test items with
adjustments to the final *Térus* dialog. The final test contained ten items in total, which was designed to test the basic understanding of the purported syntactic and morphological patterns. The questions are in the form of sentence completion: subjects would choose an answer to complete the given sentence. It should be noted that we designed the items according to the entire scope of the script, rather than simply taking the original sentences, to make sure that subjects would not be tested on their ability to memorize the sentences. In fact, no item could be correctly answered based solely on a subject's memory of the dialog. In order to choose the formally correct response, the subject would need to learn something more general. The test strategy was based on that used by Postica (2006), also is also described in Postica and Coleman (2006). See Fig. 6 for sample test items. A complete list of test questions is given in Appendix C.

<table>
<thead>
<tr>
<th>1. Vukan, _____</th>
<th>2. Tso josto _____?</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. tamio</td>
<td>a. jasuk</td>
</tr>
<tr>
<td>b. zegio</td>
<td>b. zumio</td>
</tr>
<tr>
<td>c. kalkēt</td>
<td>c. tutsan</td>
</tr>
<tr>
<td>d. tam</td>
<td>d. vukan</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3. Ekē _____</th>
<th>4. «Espesal Burger»? Kap, ____</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. splamp koka</td>
<td>a. tutsan</td>
</tr>
<tr>
<td>b. splampio shkē</td>
<td>b. zuuk</td>
</tr>
<tr>
<td>c. splamp tama</td>
<td>c. tlesuk</td>
</tr>
<tr>
<td>d. splamp zegio</td>
<td>d. rēdan</td>
</tr>
</tbody>
</table>

Figure 6 – Sample Test Items

Since the parallel sensory experience group would be given the context, it was safe to assume that subjects in this group would perform better in terms of meaning comprehension. Thus, it was important for us to ensure that pattern recognition would be
fairly tested without the interference of comprehension issues. However, this was troublesome because traditionally speaking, multiple-choice question has one correct answer, meaning that this answer will be accurate in terms of both meaning and order, which is considered to be grammatically correct by most linguists. Therefore, subjects who had the advantage to the context of the dialog, thus, would perform better on both. In order to solve this issue, the choices given in the test followed a non-traditional approach. The four choices provided for each question were designed following a fixed pattern: (1) appropriate form ("grammatically correct") and sensible (correct meaning), (2) inappropriate form and nonsensical, (3) appropriate form and sensible, and (4) inappropriate form and nonsensical. In this way, meaning comprehension was separated from pattern recognition (Fig. 7).

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Vukan, _____ .</td>
<td></td>
</tr>
<tr>
<td>a. tamio (inappropriate form and sensible)</td>
<td>b. zegio (inappropriate form and nonsensical)</td>
</tr>
<tr>
<td>c. kalkēt (appropriate form and nonsensical)</td>
<td>d. tam (appropriate form and sensible)</td>
</tr>
</tbody>
</table>

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Tso josto _____ ?</td>
<td></td>
</tr>
<tr>
<td>a. jasuk (appropriate form and sensible)</td>
<td>b. zumio (inappropriate form and nonsensical)</td>
</tr>
<tr>
<td>c. tutsan (inappropriate form and sensible)</td>
<td>d. vukan (appropriate form and nonsensical)</td>
</tr>
</tbody>
</table>

Figure 7 – Sample Test Item Breakdowns
Additionally, a mini-survey would be given to all subjects at the end of the test. The survey included questions about the subject’s native language background and foreign languages exposure. The survey can be seen in Appendix C.

**Subjects**

The participants in this study were undergraduate students from the University of Toledo, both male and female. Intact groups were used instead of random selection due to the size of the sample needed for the experiment. Students from five Composition II classes and two native speaker of English Composition I classes during spring semester 2012 were selected to participate in the study – a total of 147 subjects. There were a total of 71 subjects for the parallel sensory experience group and 76 subjects for the primary linguistic data group.

The rationale behind this selection was that a significant number of students enrolled in Composition classes were college freshmen or sophomores, which means the age range of the subjects was likely to be more balanced. Moreover, both Composition I and II were compulsory courses. Students’ major of study was more evenly distributed in such classes in comparison to other classes, where students might have the advantage or disadvantage due to the nature of the study. For instance, students from a linguistic class might perform better compare to students from an engineering class, which would ultimately affect the outcome of the study. In addition, the gender of the subjects from a compulsory class was also likely to be more evenly distributed, whereas a certain field of study may attract more male students (i.e. engineering, computer science) or female students (i.e. literature, theater) in particular.

**Procedure**
After the initial contact with the instructors of the seven classes, I set up specific dates with each of the instructor. All the subjects were orally briefed about the experiment and given the consent forms. The participation was voluntary; thus, only the participating subjects signed the consent forms prior to the experiment. They were then given either the primary linguistic data input or the parallel sensory input. The selection of the treatment was semi-random: a specific non-random treatment was only given to a class based on the consideration of the distribution of the sample size for each group as the data collection proceed.

After the initial briefing, the researcher was present but did not have any further direct interaction with the subjects during the treatment, in order to avoid any potential confounding variables. The instructions were shown on the PowerPoint slides, through which the process that the subjects were to follow during the experiment and the test was explained explicitly. It should also be mentioned that for both groups, the dialog was played twice automatically in order to reinforce the input exposure. Immediately after given the treatment, the subjects were given the test and five minutes to complete both the test and the survey. The total duration of the experiment was around fifteen minutes for both groups.

**Data Collection**

The University of Toledo Composition II classes had students whose native language was not English as well, whereas the participating Composition I classes were exclusively for native speakers of English. Therefore, based on the survey collected, the results from 12 ESL (English as a second language) students were excluded from data analysis based on the consideration of language bias, in terms of comprehending the
content of the consent form or instructions given during the experiment. I also excluded 4 other test results because there were incomplete. Therefore, a total of 147 subjects participated in the experiment and 131 effective tests were used for data analysis, including 68 subjects from the primary linguistic data group and 63 subjects from the parallel sensory experience group.

The tests were then scored twice according the coding key (see Fig. 9). Each test was scored for two factors: (1) form, meaning what is usually considered to be grammatical structures and (2) sense, which referred to meaning comprehension. When a subject chose either of the answers that have the accurate form, he/she would accumulate 1 point for the first score and similarly either of the answers that had the correct meaning would receive 1 point for the sense score. Since there were ten question items in total, each test received two scores each on the scale of 0-10. See Appendix D for the complete coding key sheet.

<table>
<thead>
<tr>
<th>1. Vukan, _____</th>
<th>2. Tso josto _____?</th>
</tr>
</thead>
<tbody>
<tr>
<td>0/1 tamio</td>
<td>1/1 jasuk</td>
</tr>
<tr>
<td>0/0 zegio</td>
<td>0/0 zumio</td>
</tr>
<tr>
<td>1/0 kalkēt</td>
<td>0/1 tutsan</td>
</tr>
<tr>
<td>1/1 tam</td>
<td>1/0 vukan</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Ekē _____</td>
<td>4. «Espesal Burger»? Kap, _____</td>
</tr>
<tr>
<td>1/1 splamp koka</td>
<td>1/0 tutsan</td>
</tr>
<tr>
<td>0/0 splampio shkē</td>
<td>0/0 zuuk</td>
</tr>
<tr>
<td>1/0 splamp tama</td>
<td>0/1 tlesuk</td>
</tr>
<tr>
<td>0/1 splamp zegio</td>
<td>1/1 rēdan</td>
</tr>
</tbody>
</table>

Figure 8 – Sample Coding Key for both Sense and Form
Chapter 3

Data Analysis

After the data were collected and scored, they were entered into SPSS (Statistical Package for the Social Sciences) for further analysis. The initial outlook of the data in SPSS contained one grouping variable – parallel sensory experience group was coded as E (experimental group) and the primary linguistic data group was coded as C (control group). Since each subject received two scores, the scores were coded into two variables – one on sense (meaning comprehension) and one on form (pattern recognition).

Since the data was not certain to be interval level, the data was analyzed through a frequency test to determine if the skewness and kurtosis of the two groups was within -2 and +2 bounds. The results from the descriptive statistical analysis showed that the skewness of -.031 for the sense scores and -.056 for the form scores were within the bounds, respectively. Similarly, the results for the sense score sand form scores showed that the kurtosis of -.439 and -.602 were within the bounds as well, respectively.

Next, I ran a one-sample Kolmogorov-Smirnov (K-S) test in order to examine the normality of distribution. If the $p$-value was greater than .05, the test distribution would be considered as normal, meaning and the data could be treated as being at the interval level. The $p$-value for the K-S tests performed on the sense scores was .016, which was not greater than .05, indicating that the data was not normally distributed and it could not be treated as being at the interval level. Likewise, the $p$-value for the K-S test performed on the form scores was .011 respectively, suggesting that the data was not normally distributed and it could not be treated as being at the interval level.
Based on the nominal nature of the data, and the number of groups, I compared the median scores to see if there was significant difference between the two groups, by using the Mann-Whitney test. The test results showed that the $p$-value for sense and form scores were both $< 0.001$, indicating that there was a significant difference in terms of the central tendencies of the test scores from the two groups. The mean rank of the parallel sensory experience group’s test scores on form was 82.06 ($N=63$), whereas the mean rank of the primary linguistic data groups’ test scores on form was 51.13 ($N=68$). Thus, I was able to reject the null hypothesis, and safely concluded that the experimental group who received the parallel sensory input performed significantly better in comparison to the subjects who received primary linguistic data input. On the sense test, the mean rank for the parallel sensory experience group was 78.77 ($N=63$), and for the primary linguistic data group was 54.17 ($N=68$), which means that as expected, the parallel sensory experience group also performed significantly better on the sense test.

Since I was also interested in finding out whether or not is there was any learning occurred for the subjects who received the primary linguistic data input, I decided to run a second set of analysis to compare the test results to chance. As mentioned previously in chapter two, each of the test items was designed in a way that there is fifty percent of the chance to get a correct answer for both sense and form scores. Thus, if there was learning occurred during the experiment, the majority of the subjects from a group would receive a score higher than 5, since there were 10 test items in total. Similarly, if there were no learning occurred, the majority of the subjects would receive a score around 5. I compared the sense and form scores of the two groups against a fifty percent probability using the Binomial test. In the result for form scores, the parallel sensory experience
group performed significantly above chance (70% > chance, \( p = .002 \)); the primary linguistic data group did not (72% <= chance, \( p < .001 \)). That is, subjects who received the primary linguistic data did not learn (what is usually perceived as) grammatical structures, but subjects who received the parallel sensory input did learn. The parallel sensory experience group also performed above chance on the sense score (67% > chance, \( p = 0.11 \)); however, the primary linguistic data group did not (\( p = .545 \)), meaning they performed no differently from chance.
Chapter 4

Discussion

Implications

Based on the statistical analysis done in the previous chapter, we can draw the following conclusions regarding input and language learning. First, the subjects who were exposed solely to speech and text, which is considered to be “primary linguistic data”, did not show any results of learning. On the other hand, parallel sensory input seems to contribute significantly to language learning, even when the exposure was only less than 10 minutes in total. Second, the subjects who received parallel sensory input not only performed better in terms of meaning comprehension, but also in terms of (what is generally considered as) grammatical pattern recognition.

The results of this study first make us rethink the very definition of input in the field of SLA. The idea that input is the target language has guided researcher in their theoretical inquiry of how language is learned, and directed teachers in their practical application of how language is taught. However, speech and text clearly do not carry any meaning or grammatical structures from the perspective of the listeners or readers who have no knowledge of the target language, which can be confirmed from the results of the study. As Klein (1986) hypothesizes, what people perceive in the Chinese Room are only meaningless speech sounds. Likewise, subjects in this study who receive the so-called “primary linguistic data” only heard speech sounds and read texts that did make any sense to them. Given the same speech and text, along with images that simulates the communicative events by showing the contexts, body language, the interaction among speakers and the interactions between speakers and objects, subjects performed better on
both meaning comprehension and pattern recognition. This empirically confirms Klein’s (1986) thought experiment, and allows us to conclude that input, in fact, is the sound of speech (no the abstraction of language) in parallel with other sensory experiences.

Another theoretical implication of the study concerns the existence of Language Learning Device. Despite the different interpretations of Chomsky’s theory, as reviewed in Chapter one, most proponents of Chomsky agree that Language Learning Device is the essential element in processing “primary linguistic data”. This kind of innate ability is assumed to be universal among all humans and is what makes learning possible (Chomsky, 1964). Under this assumption, the two groups of subjects in my study, who were exposed to the exact same “primary linguistic data”, should have performed equally in terms of understanding the “underlying structure” of language. Additionally, both groups should also have shown signs of learning given the same amount of “primary linguistic data”. However, through the results of the experiment, it is clear that the primary linguistic data group first did not perform equally in terms of pattern recognition, and second showed no signs of learning.

Since input is all parallel sensory events occurring during a given communicative event (Coleman, 2005, p. 120), the use of gestures, pictures, and real physical objects in a classroom has vital value because they simulate the real world communicative event. Many of the current teaching practices in the field of SLA do stress the use of multi-sensory input, known as the input-based approaches, which includes methods such as the Natural Approach (Terrell, 1977), Total Physical Response (Asher, 1969), or Reading and Storytelling (Ray & Seely, 2002). However, these methodologies often take the use of objects or pictures as a supplementary part of the teaching practices, which is usually
inadequate in providing sufficient associations between meaning and speech sound. In addition, “they come out of a tradition that sees language as the object of study, which carries over to individual teachers, whole methods, and materials design” (Crist, 2008). Teachers are very likely to use pictures and objects as a sort of prototypical example, or pictures and objects that are not directly linked to the articulation. For instance, Terrell (1982) describes a typical Natural Approach class as such: when the teacher says “the book”, he/she touches the book; then he/she touches the table while saying “is the on the table”. The problem here is that the object used to associate with the articulation is only one example of a table. The students did not receive effective and sufficient amount of input to form the perceived definition of a table in the target language, which will creates confusion when they see a coffee table, dining table, or a desk. Thus, simply providing multi-sensory input is also not sufficient in a classroom. Sensory experiences, in parallel with the sound of speech is true nature of input, and breadth and limits of association, meaning positive and negative examples, must be included in the presentation of input in the classroom if students are expected to learn to communicate effectively (Coleman, 2011c).

Limitations

Although the results of this study are significant and do support my hypothesis, there are several possible flaws in the design of input as well as the data collection process.

First of all, the amount of input and the difficulty of the test items were factors that could not be accurately decided. Considering the nature and form of the experiment, the duration of the dialog was restrained to around three minutes, which provided 33
well-formed sentences in total. A longer dialog would certainly provide more exposure of
the input for the subjects, which would give them more learning experiences. However, it
would also give the subjects more information to process, which might exceed the human
short-term memory capacity and thus hinders the result the test. There was no accurate
way of deciding exactly how much input should be provided in a given amount of time,
thus the research team took a reasonable educated estimate. As for the assessment tool,
the difficulty of the test was not empirically measured. Since an artificial language is
used, none of the subjects had prior knowledge of this language. For beginner learners,
the test items should also be kept to a beginning level. The research team took this into
consideration when designing the test. However, when I was scoring the tests, I noticed
that a significant number of subjects scored higher on several test items, indicating that
these test items might be easier than other ones. Therefore, the complexity of the test
items, in addition to the short input exposure time, can definitely affect the final outcome
of the study. This is something that future research can address – to find a balance
between the length of the experiment and the difficulty of the test items.

I noticed another potential flaw of the study during the data collection process.
The treatments given to the subjects did not require any active involvement on their part.
The subjects simply watched the PowerPoint slides while listening to the audio track.
Based on my observation, it is possible that a certain number of subjects in the study did
not attend fully to the contents presented in the slides during the experiment. In addition,
the input given parallel sensory experience group includes images that simulates the
communicative event, which is seemingly more interesting compare to the primary
linguistic data input. Therefore, it can be inferred that subjects from the parallel sensory
experience group paid more attention overall in comparison to the speech and text group. Future research may consider address this motivational factor, by also providing images for the primary linguistic data group, but making sure the images do not provide any parallel sensory experiences (Coleman, personal communication, March 30, 2012).

**Conclusion Remarks**

The assumption of input being target language has been dominant in the field of SLA for many decades. I believe that the results of the study successfully supported the idea that when people observe or being involved in communicative behaviors, and when people interact with other people, object or the environment, the parallel sensory experiences stimulate the physiological processes in our brain, and enables us to gain the ability to communicate through the target language. Thus, there is a need to redevelop language education away from the false assumption of language and bring real world communication into the classroom instruction.
References


# Appendix A

## Script and Translation for Audio Used by Actors

The setting is the area around a table in "Restorant Mwo Tledron", to be shown as a set of computer-generated images in storyboard form. Various details, such as a table with a drink menu display, place settings, salt/pepper shakers, etc. are provided to facilitate the realism of the interactions of the actors.

<table>
<thead>
<tr>
<th></th>
<th>Script</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>W:</td>
<td><strong>Vukan panēt. Ekē shkē gra tamēta.</strong>&lt;br&gt;W gestures welcome, indicates table, makes sweeping gesture at C1/C2/C3</td>
<td>welcome-1stPerSgPres you-PL. here-is table for you-PL-A</td>
</tr>
<tr>
<td></td>
<td><strong>Ekēs kalkēt gra tamēta.</strong>&lt;br&gt;[W hands over menus]</td>
<td>here-are menu-PL for you-PL-A</td>
</tr>
<tr>
<td>C1/2/3:</td>
<td><strong>Gzampēwan.</strong>&lt;br&gt;[C1/2/3 smile, speaking as they receive menus]</td>
<td>thank-1stPerSgPres = thanks</td>
</tr>
<tr>
<td>W:</td>
<td><strong>Dalgzo tlesan.</strong>&lt;br&gt;[W nods and smiles]</td>
<td>very beg-1stPerSgPres = you're welcome</td>
</tr>
<tr>
<td></td>
<td><strong>Tsu gazatsan zegio?</strong>&lt;br&gt;[W offers glass of water from tray, looking at C1]</td>
<td>QUES I give water-IO</td>
</tr>
<tr>
<td>C1:</td>
<td><strong>Kap, tlesan.</strong>&lt;br&gt;[nodding, indicating a place for the water]</td>
<td>yes, beg-1stPerSgPres</td>
</tr>
<tr>
<td>C2:</td>
<td><strong>Kap, tlesan splampio zega.</strong>&lt;br&gt;[nodding, indicating place for the water]</td>
<td>yes, beg-1stPerSgPres glass-IO water-A</td>
</tr>
<tr>
<td>C3:</td>
<td><strong>Gzampēwan.</strong>&lt;br&gt;[shaking head, hand in refusal gesture]</td>
<td>thank-1stPerSgPres = no</td>
</tr>
<tr>
<td>W:</td>
<td><strong>Ekēs splampēt zega.</strong>&lt;br&gt;[W serves water to C1/2]</td>
<td>here-are glasses water-A</td>
</tr>
<tr>
<td></td>
<td><strong>Tso josto tutsuk?</strong>&lt;br&gt;[W puts tray under arm, takes out pad/pencil]</td>
<td>what else drink-INF</td>
</tr>
<tr>
<td>C1:</td>
<td><strong>Tlesan, zumio.</strong>&lt;br&gt;[W points at wine image on drink menu on table]</td>
<td>beg-1stPerSgPres, wine-IO</td>
</tr>
<tr>
<td>C2:</td>
<td><strong>Tlesan, splampio koka.</strong>&lt;br&gt;[W points at Coke image on drink menu on table]</td>
<td>beg-1stPerSgPres, Coke-IO</td>
</tr>
<tr>
<td>C3:</td>
<td><strong>Tlesan, tuzio.</strong>&lt;br&gt;[W points at tea image on drink menu]</td>
<td>beg-1stPerSgPres, tea-IO</td>
</tr>
</tbody>
</table>
on table]  
[fade-out/in; clock shows time lapse; W returns]  

<table>
<thead>
<tr>
<th>Character</th>
<th>Dialogue</th>
</tr>
</thead>
</table>
| W | **Ekë splamp zuma.**  
   [W serves the wine to C1]  
   [here-is glass wine-A] |
| | **Ekë kok.**  
   [W serves the Coke to C2]  
   [here-is Coke] |
| | **Ekë splamp tuza.**  
   [W serves the tea to C3]  
   [here-is glass tea-A] |
| | **Tso jasuk?**  
   [W takes out pad and pencil]  
   [what eat-INF] |
| C1/C3 | [C1/C3 look at each other, shrug]  
   [not know-1stPerSgPres] |
| C2 | **Mwo zuan.**  
   [C2 looks W and shrugs too]  
   [not know-1stPerSgPres] |
| W | **Keh wost «Espesal Burger».**  
   [W points at "Burger Special" image on menu]  
   [this is "Burger Special"] |
| C1 | **Kap, tlesan keh.**  
   [C1 nods, points at "Burger Special" on menu]  
   [yes, beg-1stPerSgPres this] |
| W | **A tam?**  
   [W looks at C2]  
   [and you] |
| C2 | **Mwo, mwo rēdan.**  
   [C2 shakes his head]  
   [no, not want-1stPerSgPres] |
| W | **Meh, keh wost «Gogeh Pēl».**  
   [W points at "Finger Chicken" on menu]  
   [well, this is "Finger Chicken"] |
| C2 | **Ah! Meh, tlesan keh.**  
   [C2 points at "Finger Chicken" on menu]  
   [ah! well, beg-1stPerSgPres this] |
| W | **A tam?**  
   [W looks at C3]  
   [and you] |
| C3 | **Kap, tlesan trokio «Gogeh Pēl».**  
   [C2 points at "Finger Chicken" on menu]  
   [yes, want-1stPerSgPres plate-IO "Finger Chicken"-A] |
| | [fade-out/in; clock shows time lapse; W re-enters with the "Burger Special" order]  
   [here-is plate "Burger Special"-A] |
| W | **Ekē trok «Espesal Burgera».**  
   [W serves the Burger Special to C1]  
   [here-is plate "Burger Special"-A] |
| C1 | [W turns away, then returns with the other orders]  
   [here-are plates "Finger Chicken"-A] |
<table>
<thead>
<tr>
<th></th>
<th><strong>Tso josto?</strong></th>
<th>what else</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[W takes out pad and pencil]</td>
<td></td>
</tr>
<tr>
<td>C1:</td>
<td><strong>Mut, gzampēwan.</strong></td>
<td>nothing, thank-1stPerSgPres</td>
</tr>
<tr>
<td></td>
<td>[C1 shakes head]</td>
<td></td>
</tr>
<tr>
<td>W:</td>
<td>A <em>tamēt?</em></td>
<td>and you-PL</td>
</tr>
<tr>
<td></td>
<td>[W looks at C2/3]</td>
<td></td>
</tr>
<tr>
<td>C3:</td>
<td><strong>Mut, gzampēwan.</strong></td>
<td>nothing, thank-1stPerSgPres</td>
</tr>
<tr>
<td></td>
<td>[C2 shakes head, smiling]</td>
<td></td>
</tr>
<tr>
<td>C2:</td>
<td><strong>Gzampēwan.</strong></td>
<td>thank-1stPerSgPres = no</td>
</tr>
<tr>
<td></td>
<td>[C3 shakes head, makes refusal gesture]</td>
<td></td>
</tr>
</tbody>
</table>
Appendix B

A Complete List of Storyboard Images
Dalgzo tlesan.

Tso gazatsan zegio?

Kap, tlesan.

Kap, tlesan splampio zega.

Gzampëwan.

Tso josto tutsuk?

Tlesan zumio.

Tlesan splampio koka.
Tso josto?

Mut, gzampéwan.

A tamēt?

Mut, gzampéwan.

Gzampéwan.
Appendix C

Test Items

In each item, circle the most appropriate response.

1. Vukan, _____ .
   a. tamio
   b. zegio
   c. kalkēt
   d. tam

2. Tso jotso _____ ?
   a. jasuk
   b. zumio
   c. tutsan
   d. vukan

3. Ekē _____ .
   a. splamp koka
   b. splampio shkē
   c. splamp tama
   d. splamp zegio

4. «Espesal Burger»? Kap, _____ .
   a. tutsan
   b. zuuk
   c. tlesuk
   d. rēdan

5. Ekēs trokēt _____ .
   a. «Gogeh Pēla»
   b. «Espesal Burger»
   c. koka
   d. zegio

6. Keh _____ «Espesal Burger».
   a. mwo wost
   b. tutst
   c. wostan
   d. tlest

7. Ekēs splampēt _____ .
   a. «Gogeh Pēl»
   b. zuma
   c. kokio
   d. «Espesal Burgera»

8. Ekē kalk gra _____ .
   a. zuma
   b. tama
   c. kok
   d. tamēt

9. Tak, _____ .
   a. jasuk «Espesal Burger»
   b. rēdan keh
   c. tutsan shkēa
   d. gzampēwuk zega

10. Ekē kalk. _____ .
    a. Gzampēwuk
    b. Rēdan
    c. Tlesuk
    d. Gzampēwan

Your native language: ☐ English  ☐ Other (Please specify)

What foreign languages have you studied and approximately how long?

<table>
<thead>
<tr>
<th>NAME OF LANGUAGE</th>
<th>LENGTH OF TIME STUDIED</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

53
Appendix D

Coding Key to Response Type

(1/1) appropriate form and sensible
(1/0) appropriate form and nonsensical
(0/1) inappropriate form and sensible
(0/0) inappropriate form and nonsensical.

1. Vukan, _____.
   0/1 tamio
   0/0 zegio
   1/0 kalkēt
   1/1 tam

6. Keh _____ «Espesal Burger».
   1/1 mwo wost
   1/0 tutst
   0/1 wostan
   0/0 tlest

2. Tso josto _____?
   1/1 jasuk
   0/0 zumio
   0/1 tutsan
   1/0 vukan

7. Ekēs splampēt _____.
   0/0 «Gogeh Pēl»
   1/1 zumio
   0/1 kokio
   1/0 «Espesal Burgera»

3. Ekē _____.
   1/1 splamp koka
   0/0 splampio shkē
   1/0 splamp tama
   0/1 splamp zegio

8. Ekē kalk gra _____.
   1/0 zuma
   1/1 tama
   0/0 kok
   0/1 tamēt

4. «Espesal Burger»? Kap, _____.
   1/0 tutsan
   0/0 zuuk
   0/1 tlesuk
   1/1 rēdan

9. Tak, ______.
   0/1 jasuk «Espesal Burger»
   1/1 rēdan keh
   1/0 tutsan shkēa
   0/0 gzampēwuk zega

5. Ekēs trokēt ______.
   1/1 «Gogeh Pēla»
   0/1 «Espesal Burger»
   1/0 koka
   0/0 zegio

10. Ekē kalk. ______.
    0/1 Gzampēwuk
    1/0 Rēdan
    0/0 Tlesuk
    1/1 Gzampēwan