Language Acquisition and Language Learning

BBI 3206 / 3209 (Unit 1 - 7 / 7)
Program Bacelor Sastera (Bahasa Inggeris)

Wong Bee Eng, PhD
Fakulti Bahasa Moden dan Komunikasi
Universiti Putra Malaysia
43400 UPM Serdang
Selangor Darul Ehsan

Hakcipta terpelihara. Tidak dibenarkan mengeluar ulang mana-mana bahagian modul ini dalam apa jua bentuk dan dengan cara apa jua sama ada secara elektronik, salinan foto, mekanik, rakaman atau cara lain sebelum mendapat izin tertulis dari Pusat Pendidikan Luar (PPL), UPM.
MODUL PEMBELAJARAN : BBI 3206/3209 LANGUAGE ACQUISITION AND LANGUAGE LEARNING
disediakan dalam bentuk bahan pengajaran dan pembelajaran kendiri di bawah program
Pendidikan Jarak Jauh, Universiti Putra Malaysia. Sebarang pertanyaan dan cadangan untuk
memperbaiki gaya penyampaian dan isi kandungan modul ini bolehlah dikemukakan kepada penulis
dengan menggunakan alamat Pusat Pendidikan Luar.

Penulis : WONG BEE ENG, Ph.D
Fakulti Bahasa Moden dan Komunikasi
Universiti Putra Malaysia
43400 UPM Serdang
Selangor Darul Ehsan

Alamat : Unit Modul dan Bahan Kendiri
Pusat Pendidikan Luar
Universiti Putra Malaysia
43400 UPM Serdang
Selangor Darul Ehsan
Tel : 03-8946830/03-89458904
Fax : 03-89458902

Reka Bentuk Kulit dan Cetak oleh : UPM HOLDINGS SDN. BHD.
Blok F2, Bangunan MTDC-UPM
Universiti Putra Malaysia
43400 UPM Serdang
Course Contents

Unit 1: Characteristics of First Language Acquisition

Unit 2: Stages of First Language Acquisition

Unit 3: The Behaviourist Theory to Explain First Language Acquisition

Unit 4: The Role of Universal Grammar in First Language Acquisition

Unit 5: The Observable Phenomena in Second Language Learning

Unit 6: The Cognitive Approach to Second Language Learning

Unit 7: The Role of Universal Grammar in Second Language Learning
Introduction to the course/module

The course provides an introduction to the phenomena of first language (L1) acquisition and second language (L2) acquisition/learning. It will look at the characteristics and stages of first language acquisition. The course goes on to provide an explanation as to why the behaviourist theory cannot explain L1 acquisition adequately. It will also examine the role of Universal Grammar in L1 acquisition. Additionally, it will describe the role of UG and the cognitive perspective in L2 acquisition/learning. In this respect, it will also examine the observable phenomena of second language acquisition/learning.

Aims of the course

By the end of this course, you should understand and be able to discuss the fundamental concepts used in the study of language acquisition and learning.

Synopsis

The course will examine the processes of first language acquisition and second language acquisition/learning. In particular, it examines the characteristics and stages of first language acquisition. Additionally, it investigates the role of UG and the behaviourist theory in the acquisition of a first language. It also examines the role of UG and the cognitive perspective in the acquisition/learning of a second language.
Assessment

The assessment requirements for the course include:

An assignment: 25 %
Mid-semester test: 35 %
Final examination: 40 %

Assignment

The questions or topics for the assignment will be sent to you by IDEAL.

Mid-semester Test

The mid-semester test will cover units 1, 2, 3 of this module and related readings in recommended texts for first language acquisition.

Final Examination

Questions for the final examination will be based on units 4, 5, 6 and 7 of this module and chapters 1, 2, 3 and 4 of your text for second language acquisition/learning.

Text


The text provides a broad and readable overview of a wide range of topics on second language learning theories. However, only topics in chapters 1, 2, 3 and 4 of this text will be referred to and you are required to do the readings as stipulated at certain points in your reading of the guide.
Recommended Reading

L1 Acquisition


L2 Acquisition/Learning


Learning schedule

<table>
<thead>
<tr>
<th>Week</th>
<th>Topics</th>
<th>Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-2</td>
<td>Unit 1</td>
<td>7</td>
</tr>
<tr>
<td>3-4</td>
<td>Unit 2</td>
<td>7</td>
</tr>
<tr>
<td>5-6</td>
<td>Unit 3</td>
<td>7</td>
</tr>
<tr>
<td>7-8</td>
<td>Unit 4</td>
<td>7</td>
</tr>
<tr>
<td>9-10</td>
<td>Unit 5</td>
<td>7</td>
</tr>
<tr>
<td>11-12</td>
<td>Unit 6</td>
<td>7</td>
</tr>
<tr>
<td>13-14</td>
<td>Unit 7</td>
<td>7</td>
</tr>
</tbody>
</table>

ABOUT YOUR INSTRUCTOR

Wong Bee Eng, PhD

I can be contacted at the:

Department of English Language
Faculty of Modern Languages and Communication
Universiti Putra Malaysia
43400 UPM Serdang
Selangor Darul Ehsan

Telephone: 89468740 (my office)
89468778 (Department office)
(You should try ringing my office first, failing which to reach me, then you may call the department office)
My email I.D. is bee@fbm.upm.edu.my or bee@putra.upm.edu.my
You may also fax me at 03-89439951
### STUDY CHART

The course is conducted over a period of 14 weeks. The chart below is a guide to help you establish an effective study program. Good time management and careful planning are essential in achieving a successful study program.

**BBI 3206: LANGUAGE ACQUISITION AND LANGUAGE LEARNING**

<table>
<thead>
<tr>
<th>Week</th>
<th>Topic</th>
<th>Activity/Reading</th>
<th>Assessments/submission of assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Overview of the course and Unit 1</td>
<td>Read Unit 1 and complete the exercise.</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Unit 2</td>
<td>Read Unit 2, the required readings and complete the exercises.</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Unit 3</td>
<td>Read Unit 3, the required readings and complete the exercise.</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Unit 4</td>
<td>Read Unit 4, the required readings and complete the exercises.</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Unit 5</td>
<td>Read Unit 5, the required readings and complete the exercises.</td>
<td>Mid-semester examination</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Unit 6</td>
<td>Read Unit 6, the required readings and complete the exercises.</td>
<td>Submit assignment</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Unit 7</td>
<td>Read Unit 7, the required readings and complete the exercises.</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Revision</td>
<td>Self-study</td>
<td>Final examination</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
UNIT 1: Characteristics of First Language Acquisition

Introduction

Every normal human being acquires language at a very young age. The system that is acquired is a very abstract and complex one. Have you ever wondered how every single very young normal child acquires such an abstract and complex system? Before we attempt to answer this question, let us try to answer the next question.

What is language?

Language is not a concrete set of things out in the world that we can point to or measure. It is inside our brains or minds. In linguistics, we attempt to describe what we know when we know a language by formulating a grammar.

A grammar may be defined as a set of rules which characterises all and only the sentences of the language that we as speakers are able to produce and understand (Crain and Lillo-Martin, 1999: 5). To learn a language is to master the rules of the mental grammar. This mental grammar is sometimes called a generative grammar as it is a system of rules that generate or produce sentences of the language.

To answer the question in the introduction, we need to look at the characteristics of first language acquisition. The characteristics of L1 acquisition will also provide us with answers to the following questions:

- When is language learned?
- What is the course of language acquisition like?
Characteristics of Language Acquisition

Primary Linguistic Data

This has to do with the environmental experience within which children learn language. There is considerable latitude in the kind of environmental inputs that permit children to develop language. Children growing up in the same linguistic community, and ultimately learning the same language, may have a wide variety of experiences with language during their early years.

Despite the variety in experiences, however, every child in every linguistic community acquires the language of that community. Every child in the same linguistic community learns the same language, i.e. there is uniform convergence on equivalent grammars despite the fact that children all receive different input or primary linguistic data.

Now we would like to know how is it that all children manage to learn to speak and understand a language, despite being raised in environments that vary considerably? Is explicit instruction required at all for language acquisition?

Universality of language

Without special training or carefully sequenced language input, every normal child acquires a natural language. Every child in a linguistic community succeeds in converging on a grammatical system that is equivalent to everyone else’s, despite considerable latitude in linguistic experience. And children do this quite
Language Acquisition and Language Learning

rapidly and with few pitfalls. This fact is known as **universality** of language within the human species (Crain and Lillo-Martin, 1999: 6).

**Uniformity of language**

Any child can become polylingual, that is s(he) can acquire and retain all the languages s(he) is exposed to. For the child, every language is learned with equal ease. Children speak whatever language (or languages) is spoken to them. For example, if a child of English-speaking parents is taken to Russia, and raised there, the child speaks Russian. In other words, the language of a child’s parents does not determine the language of the child. This property of language acquisition is known as **uniformity** of language acquisition or language acquisition is **uniform** across languages (Crain and Lillo-Martin, 1999: 8).

**Possible Mechanisms to Explain the Acquisition Process**

Before you proceed with this topic, attempt Task 1 with your friends and tutor.

?? Task 1 (20 minutes)

Based on your observations on how infants and young children acquire language, try to answer the following questions with your friends and tutor.

1. Do children learn language through trial and error?
2. Do mothers actually correct their young children explicitly during the process of first language acquisition?
Language Acquisition and Language Learning

Trial and Error

Children actually progress through similar stages of language acquisition. This would suggest that this is not a trial-and-error process. If language acquisition is a trial-and-error process, then we would expect that children do not go through the same stages and they would not all converge on the same grammatical rules of the language being acquired. But this is not the case. There is uniform convergence on the same Final State of the grammar of the language (Crain and Lillo-Martin, 1999: 9).

There are two ways in which the trial-and-error process can be successful. First, this process will succeed if every child received the same fixed linguistic input. However, this is not the case as children acquire language in very diverse environments. Children with these varied environmental experiences all seem to form the same hypotheses about the language being learnt.

Secondly, we would have to assume that the Final State is so simple that it can be acquired universally using the trial-and-error process. But this is not the case as we will see in later units. Language is a complex abstract knowledge and to say that it is simple is a flawed assumption.

Another reason why the trial-and-error process cannot explain the acquisition process is that children make very few grammatical errors along the way.
Corrective Feedback

It has been documented that parents do not correct children explicitly for their grammatical errors (see for example Brown and Hanlon, 1970). In other words, they do not provide corrective feedback or negative evidence to their children. Parents’ feedback usually distinguish between true and false sentences, regardless of whether they are well-formed (grammatical) or ill-formed (ungrammatical). This is evidence against corrective feedback or negative evidence as contributing towards language development.

The following is a conversation to prove the above point.

Child: Nobody don’t like me.
Parent: No, say, 'Nobody likes me'.
Child: Nobody don’t like me.
(This went on for eight repetitions. Finally, the child thought she understood the correction and said: )
Child: Oh! Nobody don’t likes me.

(From Brown and Hanlon, in Crain and Lillo-Martin, 1999: 11)

In other words, to answer a question posed earlier, it seems that explicit instruction is not required at all for first language acquisition.

Imitation

It has been suggested that children acquire their first language through imitation. Evidence against this claim is that in the course of language acquisition, children
overgeneralize or overregularize certain forms. Children learning English as an L1 may produce forms such as goed (an incorrect past tense form) and foots (an incorrect plural form). These forms are not the product of imitation since these forms are not found in the input of adults. Children have internalize the rules for the past tense and plural forms but they have not applied them correctly (Crain and Lillo-Martin, 1999: 9). These overregularizations are apparent in the children’s speech for many months and they are only phased out after children have heard a number of occurrences of the irregular forms in adults’ input.

The Theory of Innate Linguistic Knowledge

A theory that is very influential has been put forth to explain the process of L1 acquisition. This is the theory of Innate Linguistic Knowledge. This theory is due to the influence of Chomsky (1965). According to this theory, children acquire language readily because it is in their genes. They are born with innate knowledge and this is known as the Language Acquisition Device (LAD). The LAD includes principles common to all human languages. And this theory further states that what is innate does not have to be learnt. Language acquisition therefore consists of learning what is peculiar to the language environment, for example, the particular words of one language or another and applying the universal principles (Crain and Lillo-Martin, 1999: 4).

The LAD is the knowledge that a child brings to the task of language acquisition. This is the means by which the child/learner analyzes the linguistic input (from parents and others). The linguistic input is called Primary Linguistic Data (PLD) (Crain and Lillo-Martin, 1999: 5). On the basis of these data, the LAD hypothesizes a series of grammars. The last grammar that is formulated is the
adult grammar, known as the **Final State**. The process can be represented schematically in the following way (Crain and Lillo-Martin, 1999: 5):

\[ \text{Input (PLD)} \rightarrow \text{LAD} \rightarrow \text{Final State} \]

### The Poverty of the Stimulus Argument

As stated earlier, the LAD is the means by which the learner analyzes the linguistic input. However, the linguistic input that the child hears is often meager or impoverished compared to what the child eventually is able to produce (Chomsky, 1987). The knowledge that is language that we have is complex and abstract but the experience of language that we receive is limited. Our minds are not able to create such complex knowledge based on such sparse information. But the knowledge must come from somewhere. Chomsky proposes that it comes from innate properties of the mind. Thus we have impoverished linguistic input on the one hand and the complex adult Final State grammar on the other hand. If the knowledge in the Final State did not come from the environment, the source must be the mind itself. This is known as the poverty-of-the-stimulus argument. The mind of the baby who knows no language is also known as the **initial zero state** or \( S_0 \) and the adult Final State, the **steady state** or \( S_s \) (Cook and Newson, 1996: 78). So a normal person is said to proceed from:

\[ S_0 \rightarrow S_s. \]
Language Acquisition and Language Learning

❓ Task 2 (30 minutes)
Explain Plato’s problem.

The Developmental Route

Children are known to follow well-defined stages in their language development, parallel to their physical development. We will examine the stages of language development in the next unit. It is said that there may be a critical period for language acquisition. Lenneberg (1967) posits that it is before the onset of puberty, after which it may be difficult for one to acquire a language completely.

Summary

The characteristics of first language acquisition are presented succinctly by Aitchinson (1989: 67) who presented Lenneberg’s four criteria (1967) as a list of six features (in Mitchell and Myles, 1998:48 – 49):

1. The behaviour emerges before it is necessary.
2. Its appearance is not the result of a conscious decision.
3. Its emergence is not triggered by external events (though the surrounding environment must be sufficiently ‘rich’ for it to develop adequately.

4. Direct teaching and intensive practice have relatively little effect.

5. There is a regular sequence of ‘milestones’ as the behaviour develops, and these can usually be correlated with age and other aspects of development.

6. There may be a ‘critical period’ for the acquisition of the behaviour.
Answers

Task 2
Suggested answer

The Greek philosopher, Plato, was among the first scholars to argue that humans had knowledge that could not have come from their experience (Crain and Lillo-Martin, 1999: 53). This parallels with the situation we have discussed with regard to L1 acquisition, that is our knowledge of language is complex and abstract but the experience of language we receive is limited. This is the conundrum called ‘Plato’s problem’ which is at the heart of Chomskyan ideas of language acquisition (Cook and Newson, 1996: 82).
UNIT 2: Stages of First Language Acquisition

Introduction

Data for determining the stages is obtained from studies of children’s natural, spontaneous production. The timing of these events varies from one child to the next. This unit is a discussion of the stages of first language acquisition of a normal child.

First Few Months

In the first few months, an infant cries, coo and babble. The babbling consists of utterances of linguistics sounds but no meaning is attached to them. Normal children have particular linguistic ways of perceiving speech sounds. According to Elmas (1985, in Crain and Lillo-Martin, 1999:26), two-week-old babies can distinguish between a voiced consonant such as b and an unvoiced consonant such as p. It is believed that voicing is one of the ways of marking categorical boundaries that distinguish between words and this ability by infants to distinguish phonemic boundaries is known as categorical perception.

Approximately 6 to 10 months

At around six months, there is more varied babbling. The utterance of the infants at this stage comprise different syllables and this happens in all language environment. Babbling is an internally driven behaviour as even deaf children babble vocally at this stage. At around ten months, the babbling changes according to linguistic environments. At this stage, deaf children stop babbling.
Language Acquisition and Language Learning

Approximately 1 year

The production of first words such as *Mama* and *Papa* occur at this stage. The child is also able to use gestures to communicate. In fact, the average child is able to combine gestures with words, often with pointing gestures, which are combined with the name for the object being pointed to.

The child continues to babble according to the sounds of her language. There is fast growing comprehension of language and most children can obey simple commands at this stage.

Approximately 1 1/2 years

At this stage, the child can put together (two) 2 successive single words. The intonation pattern of such utterances is a single word pattern for each of these words with a pause between words. These utterances are known as precursors of sentences (Crain and Lillo-Martin, 1999:27).

There is a further increase in vocabulary and primitive two-word ‘sentences’ appear. The intonation at this stage exhibits a change, that is both words are included in one intonation group, with falling intonation only at the end of the two-word sequence, without a pause between the words.

The child is able to ask for things with language now. Throughout the development of language, the child experiences computational bottlenecks (Crain and Lillo-Martin, 1999:27) which may limit the number of words a child can put together; or they may limit the use of rules to form new sentences. These
bottlenecks are attributed to some limitation in a non-linguistic cognitive capacity, for example attention span and memory span.

**Approximately 2 years**

From about two years onwards, language development should be discussed in terms of *stages* rather than age (Brown, 1973). The developmental sequence is *invariant* for all normal children. These stages correlate with the length of a child’s utterances – the average is called **Mean Length of Utterance (MLU)**. This is determined by recording a large number of speech samples from a child, writing down her utterances, and finding the average number of words that appear in the utterances.

The MLU is a measure of a child’s computational capacity. The different rates of language development among children is due to differing processing capacities. At a given processing capacity, the kinds of utterances produced are constant. However, the MLU is not an accurate measure of a child’s grammatical competence as it provides us with information about a child’s stage of language development relative to other children.
Task 1  (20 minutes)

Calculate the Mean Length of Utterance (MLU) for the following utterances produced by a two-year-old-child.

- Sam milk?
- See toy?
- No eat?
- What Mama doing?
- Where doggie?
- Where cat go?

Stage I

This occurs at approximately the age of two years. The child has a vocabulary of approximately 400 words and an MLU of 1.75. The child produces many single word utterances and two- and three-word ‘sentences’ which express semantic concepts contained in a single clause. The word order of the child’s utterances follows adult word order but without grammatical words.

Stage II

This occurs at approximately 2½ to 3 years of age. At this stage, the child has an MLU of approximately 2.25 and a vocabulary of approximately 900 words. The child now has within its grasp some grammatical devices. For a child acquiring English as its first language will have some determiners, pronouns, progressive
ending -ing, and past tense marker -ed. The child can talk about things around
them as well as absent objects and past events. They also form generalizations
based on input, for example, the past tense marker -ed and the plural marker -s.
These are called overregularizations or overgeneralizations, that is the
deduction of a rule based on the linguistic input that the child hears in its
environment.

Approximately 3 to 3 ½ years

Stage III

This stage occurs at approximately 3 to 3 ½ years of age. The MLU of the child
is 2.75 and it has a vocabulary of approximately 1200 words. These include
auxiliary verbs, prepositions and other grammatical morphemes. The child also
begins to use syntactic transformations, that is rules which create new sentence
types out of basic declarative sentences.

An example of this is the operation of Subject-Auxiliary Inversion to form
Yes/No questions:

Subject       Auxiliary
1a.  ‘John       is       clever.’

Auxiliary       Subject
1b.  ‘Is       John       clever?’
Another syntactic transformation that is available to the child at this stage is *Wh-Movement* to form *Wh*-questions:

```
2a. He punched Ali.
2b. Who did he punch it?
```

In both of the examples above, the symbol *i* indicates the original position of the auxiliary (in (1)) or the object (in (2)).

**Approximately 3 ½ to 4 years**

**Stage IV**

This stage occurs when the child is approximately 3 ½ to 4 years old. The MLU of the child is 3.50 and its vocabulary is approximately 1500 words. Now the child is able to form multi-clause sentences. Relative clauses, complement clauses and conjoined sentences appear in the child’s utterances.

Example of a **relative clause**:

```
The boy who stole my book is Ali.
```

Example of a **complement clause**:

```
She said that Ali stole my book.
```
Example of a conjoined sentence:

Ali stole my book and my pen.

At the same time, the child still overregularizes many irregular forms of verbs.

Approximately 4 to 5 years

Stage V

The child experiences this stage at approximately 4 to 5 years of age. The MLU of the child is 4.0. The child now has within its control a vocabulary of about 1500 words. He or she is able to produce conjunctions and subordinate clauses with temporal terms before and after. The child is able to engage in social conversations with peers. Some metalinguistic abilities such defining words and correcting its own grammatical errors are observed. In other words, the children at this stage have some conscious awareness of the properties of language.

After five years

After Stage V

After Stage V, that is after five years of age, it is difficult to explain language acquisition in terms of stages or MLU. The child produces more complex sentences although they may not be necessarily longer. Now, the learning of the basics of grammar and rules is almost complete.
Language Acquisition and Language Learning

From the age of five to ten years, there is further increase in vocabulary. By now, the child has learned most of the exceptions to rules which they have overregularized. The learning of these exceptional forms is not completed until after ten years of age.

After puberty, there is very little change in syntax or pronunciation. However, the child’s vocabulary continues to grow from the teenage years to adult life because concepts associated with many words are not accessible to children. Moreover, vocabulary learning is fundamentally memorization. But children can cope with general rules that underlie the productivity of human languages. There is an increase in ability to use language stylistically. Conceptual development also continues, for example, the child is able to give logical explanations.

Task 2 (20 minutes)

Complete this task with your friends and tutor.
State whether each of the following structures involves Subject-Auxiliary Inversion transformation or Wh-Movement transformation? Identify the position from where the movement or extraction takes place.

a. Is John going to the shop?
b. Who did he kick?
c. Were the boys naughty?
d. Who hugged John?
e. Are the students clever?
f. What was eaten by John?
g. Is John’s favourite colour blue?
h. Was the boy who kicked John punished by the principle? Will the students that went to France be present today?

Summary

This unit has been an overview of the stages of first language acquisition. These stages show that the development of language is not only fast but consistent across children. Children from different language backgrounds show similar developmental stages at similar age ranges. The kind of errors that children make indicate that they are able to detect the regularities in the linguistic input that they hear and they produce novel and rule-governed forms. This is apparent in the overregularization or overgeneralization errors that they make. Additionally, children are not known to make errors in word order, complex clauses, and transformations. In the earlier stages, they may not apply a transformation at all but once they use it, they seldom make errors in the form.
Language Acquisition and Language Learning

Answers

Task 1
Answer
The MLU of this child at this stage is 2.5.

Task 2
Answers

a. **Subject-Auxiliary Inversion**: Is John *t* going to the shop?
b. **Wh-Movement**: Who did he kick *t*?
c. **Subject-Auxiliary Inversion**: Were the boys *t* naughty?
d. **Wh-Movement**: Who *t* hugged John?
e. **Subject-Auxiliary Inversion**: Are the students *t* clever?
f. **Wh-Movement**: What *t* was eaten by John?
g. **Subject-Auxiliary Inversion**: Is John’s favourite colour *t* blue?
h. **Subject-Auxiliary Inversion**:
   Was the boy who kicked John *t* punished by the principle?
i. **Subject-Auxiliary Inversion**:
   Will the students that went to France *t* be present today?
UNIT 3: The Behaviourist Theory to Explain First Language Acquisition

Introduction

As mentioned in Unit 1, some possible mechanisms have been proposed to explain the L1 acquisition process. Such mechanisms include trial and error, corrective feedback, and imitation. Such mechanisms are actually subsumed under the theory of Behaviourism to explain the process of learning, including the learning of a language.

Behaviourism

Behaviourism, a theory of learning (Bloomfield, 1933; Skinner, 1957; Thorndike, 1932; Watson, 1924, in Mitchell and Myles, 1998: 23), sees language learning as any other kind of learning, that is as a formation of habits. This notion stems from work in psychology. With regard to language learning, this theory is anchored in structural linguistics.

The most fundamental tenet of this theory is that the mind is not a proper object of study because this theory studies only observable behaviour, that is actions, and relates these actions to other observable actions (Crain and Lillo-Martin, 1999: 34). Mental processes are unobservable. Thus it claims that there are physically mediated correlations between environmental stimuli and behavioural responses. In other words, one can apparently condition an animal’s (or a human’s) natural response by changing the environment in specific ways.
In this theory, the role of the environment is thus crucial in shaping the child’s learning and behaviour. Skinner’s book *Verbal Behaviour* (1957) outlined in detail his behaviourist view of learning as applied to language. This view would see language learning as following a similar course as the learning of mathematics or music. In doing this, children would use general principles of learning which apply across domains. They would make errors that differ from one child to another, depending on the child’s environment. And the level of ability achieved by different individuals would be dissimilar, that is they would differ widely in linguistic abilities. In other words, just as people differ widely in mathematical knowledge and musical ability, they would differ widely in linguistic abilities (Crain and Lillo-Martin, 1999: 33).

However, we have seen in Units 1 and 2 that language development does not occur in this manner. Every child achieves a similar Final or Steady State grammar and they go through similar stages in attaining this state, regardless of the different linguistic input that they receive.

Before we proceed further, let us examine two methodologies of the Behaviourist theory: classical conditioning which was discovered by Pavlov (1927) and operant conditioning, a technique associated with B. F. Skinner (1938, in Crain and Lillo-Martin, 1999: 35 – 40).

1. Classical Conditioning (Ivan Pavlov, 1927)

A prototypical experiment Pavlov used to perform involved dogs. He knew that dogs will salivate naturally when they come face-to-face with food. This naturally occurring behaviour is termed an unconditioned response. The stimulus for an unconditioned response is called an unconditioned stimulus.
Pavlov then experimentally conditioned dogs to respond in the same way (to salivate in this case) even when food was not present. In order to produce the response he wanted, he sounded a bell each time the dog was given some food. After many trials, the dog associated the sound of bell with the given food. Eventually, the dog would salivate when he heard the bell, whether or not it was given food. Th sound of bell is then called the conditioned stimulus, and the salivation is termed a conditioned response.

This experiment indicated that one can condition an animal or a human’s response by changing the environment in certain ways. This is summarised in the figure below.

Classical Condition (Pavlov, 1927; adapted from Crain and Lillo-Martin, 1999: 35)
Task 1  (15 minutes)

Complete this task with your friends and tutor.

When applied to a linguistic phenomenon such as learning the meaning of a word (in this case, milk), the process can be represented in the following figure. Fill in the blanks of the figure.

Unconditioned Stimulus = milk
Unconditioned Response = drink?

Conditioning =
Say the word milk when milk is present

Conditioned Stimulus =
Conditioned Response =

What can you say about the conditioned response and the unconditioned response to the word?

The goal in this case would be to get the learner to respond to the word milk in the same way as s/he would respond to the object milk. In order to do this, the
Language Acquisition and Language Learning

parent simply has to utter the word *milk* whenever the child is in the presence of milk.

In this case, this theory assumes that to understand a verbal stimulus is to produce a behavioural response appropriate to the associated object. Of course, this cannot be true. One problem is that it is difficult to identify the unconditioned response that the child will utter in the presence of milk since other responses are possible:
- The child might grab the bottle or cup containing the milk, or
- The child might drink it and spit it out.

Then the question remains as to whether these behaviours are associated with the word *milk*?

There are occasions when a person who understands perfectly the meaning of a word and yet might not respond to it at all. However, according to the behaviourist theory, a child who does not respond to a word in the same way on each occasion does not understand the meaning of the word.

2. Operant Conditioning

Techniques of this form of conditioning, associated with B. F. Skinner, are extended to behaviours which are normally not associated with any particular stimulus. In this form of conditioning, test animals are usually deprived of food throughout the experiment. In this case, food is used as positive reinforcement for a behaviour to be learned.
A bird in a cage, deprived of food might start to move around randomly. After a while, it might peck on a disk by chance and food (positive reinforcement) drops into a tray in the cage. The random behaviour is continued and after a time, the bird pecks on the disk again. After a series of such pairings, the pigeon associates pecking the disk with the receipt of food.

In this type of experiment, the initial pecking of the disk, a random behaviour, is called a **free operant**. After many trials, this behaviour becomes a **conditioned operant** (Crain and Lillo-Martin, 1999: 37). Skinner argued that even the most complex behaviours may be conditioned in the same way. For example, verbal behaviour may be acquired through operant conditioning with **shaping** or progressive approximation of utterances (this technique is used to teach behaviours that are not free operants) (Crain and Lillo-Martin, 1999: 37 – 38).

In most cases of language learning, this technique of progressive approximation does not explain how children learn to speak. In fact, Chomsky (1959) reviewed Skinner’s book *Verbal Behaviour* and accorded it his (Chomsky’s) strongest criticisms. Three reasons are offered by Crain and Lillo-Martin (1999: 39 – 40) to explain why the behaviourist theory cannot explain first language acquisition adequately.

Firstly, children develop vocabulary too rapidly to explain what happens in Behaviourist terms. They learn as many as nine words per day and this is more than what they should be capable of if they were relying on reinforcement. Sometimes children learn a word after only one exposure to it.

Secondly, children use words in various forms although they have never heard them used in the these forms before. For example, if a child is shown this picture [●] and is told that it is called a wug, the child would be able to produce the
form wags when shown two of them \[\text{[\text{dog} \text{cat}]}\], even though they have not heard the form before.

Children also speak and understand completely novel (new) sentences. The behaviourist theory states that children’s understanding and production of language would be limited to sentences they have been exposed to, that is those sentences which they have been conditioned to understand.

**Summary**

This has been a simple, common sense account of language development. B. F. Skinner tried to account for verbal behaviour using the tenets of Behaviourism, but this theory has been too simple to account for the complexities of linguistic knowledge.
Answers

Task 1
Suggested answer

Unconditioned Stimulus = milk
Unconditioned Response = drink?

Conditioning = Say the word milk when milk is present

Conditioned Stimulus = word milk
Conditioned Response = drink

What can you say about the conditioned response to the word and the unconditioned response to the item?

The unconditioned response to the word is the same as the unconditioned response to the item.
UNIT 4: The Role of Universal Grammar in First Language Acquisition

Introduction

We have seen that the knowledge that a child brings to the task of language acquisition is known as the Language Acquisition Device (LAD). And the linguistic input that the child gets from its caretakers is the Primary Linguistic Data (PLD). Based on the data, the child hypothesizes a series of grammar and it eventually reaches the Final State or the adult grammar. This process is universal, uniform and rapid. This is the theory of Universal Grammar, a linguistic theory which aims to describe the mental representations of language which are stored in the human mind.

Universal Grammar

This theory postulates that there are universal linguistic principles that are specific to grammar formation (Crain and Lillo-Martin, 1999: 55). These universals form Universal Grammar (UG), which is a ‘theory of the internal organisation of the brain of the learner’. The UG approach, proposed by Chomsky (1981, 1986a, 1986b) claims that all human beings inherit this set of principles. Thus UG is part of our biological blueprint, and the principles of UG are observed in every natural language. These universal principles emerge early in the course of language development and they are invariant and apply to all natural languages.
Although children do not speak at birth, they will master the universal linguistic principles on the basis of minimal linguistic input. First they learn the words and when these are mastered, they put the words together to form simple sentences. Once these sentences emerge, then it is obvious that the UG principles are operative in them. In other words, these principles are not learned. Instead they are acquired with very little linguistic experience.

Although UG apply uniformly to all languages, there is variation between languages, so UG cannot contain everything there is to know about any one language (Crain and Lillo-Martin, 1999: 55). Languages differ in systematic and limited ways which are captured by parameters. These parameters possess a number of open values or choices/options which characterize differences between languages.

There are at least three reasons for assuming that UG is innately given as part of human genetic endowment. The first is that in L1 acquisition the final state knowledge attained is underdetermined by input (that is what they acquire ultimately is more than what they had been exposed to). The second is that a child can construct a mental grammar on the basis of positive evidence (what they hear) alone in a surprisingly short space of time; given the complexity of language it is difficult to believe that this would be possible on the basis of hypothesis testing starting from a blank slate. The third is that even where children are apparently provided with negative evidence (corrective feedback), they do not or cannot make use of it.

UG sees languages as having no specific rules for forming grammatical constructions. Instead there are universal innate principles and a finite number of options/choices or parameters as to how they apply. To master a language that one is exposed to is to fix the values of those parametrized principles of the said
language. The parameter values have to be fixed on the basis of direct and positive experience. A set of principles with associated parametric values will result in a particular grammatical system known as an internal-language. Thus the end state reached in any natural language is a system with an infinite number of syntactic structures but one which is generated by a finite number of universal principles and the parameters associated with them.

Principles

We have seen that in this theory (UG), the child’s initial state for language learning is comprised of a set of invariant universal principles which are wired to the language faculty and which forms a child’s genetic blueprint for a grammar. The principles specify limited possibilities of variations. These are the parameters that need to be fixed in one of a few possible ways (Saleemi, 1992: 58, in Mitchell and Myles, 1998:49). This makes the task of language learning easier and more constrained for the child.

An example of a principle is the structure dependence principle. This principle is stated as:

All grammatical operations are structure-dependent (Radford, 1997: 12).

The principle specifies that ‘all grammatical operations are sensitive to the grammatical structure of the sentences they apply to’ (Radford, 1997: 14 – 15) and it holds for all grammars of all natural languages. This means that ‘words are regrouped into higher-level structures which are the units which form the basis of language’ (Mitchell and Myles, 1998: 49). Study the following examples.
1. *He* bought a book.
3. *My friend who is from Portugal* bought a book.

In the above sentences *he, my friend* and *my friend who is from Portugal* are the same kind of groupings and they perform the same role, that of the Subject. Such groupings are known as phrases and the examples above are **Noun Phrases**. The main element or the **head** in Noun Phrases is a pronoun (*he*) or noun (*friend*). The second part of each of the above sentences (*bought a book*) is a **Verb Phrase** as the head is a verb (*bought*). All languages in the world have sentences which consist of at least a Noun Phrase (NP) and a Verb Phrase (VP) as shown in the example below.

2'. 

\[
\text{[n: My friend] [vp bought a book]}
\]

These phrases may contain other phrases or whole sentences.

This principle of Structure Dependency explains many of the operations we perform on languages. For example, how do we form a **yes-no** question in English? The canonical (typical) word order of a declarative sentence in English is Subject-Verb-Object.

4. *She is* tall.

\[ \]

5. *Is she tall?*
To form a yes-no question from the above sentence, we invert the subject (*she*) and the auxiliary (*is*).

In the following sentences, we do the same.

6.  *His friend is tall.*

7.  *Is his friend tall?*

8.  *His friend who is from Portugal is tall.*

9.  *Is his friend who is from Portugal tall?*

The yes-no questions formed from the sentences above are not based on the linear order of the sentence but are structure-dependent. We did not move every first, second, third or n-th word in the sentences. If we did that, we would generate ungrammatical sentences. For example, if the rule for forming yes-no questions is to move every second word in the declarative sentence, then we would have *Friend his is tall?* (for (6)) and *Friend his who is from Portugal is tall?*\(^1\) (for (8)).

Actually what we did was to move the **first auxiliary in the main clause** to the beginning of the sentence, thus inverting the subject and the auxiliary. The auxiliary in example (8) is the second *is*. The subject in example (6) is *his friend* and in (8), it is *his friend who is from Portugal.*

---

\(^1\) The asterisk (*) indicates that the sentences are ill-formed or ungrammatical.
Although the examples are taken from English, all languages are organized hierarchically in terms of phrases, that is the structure dependence principle is common to all languages. However, there are other rules that differ between languages. These are the parameters that have been mentioned earlier.

**Parameters**

We will now turn our attention to an example of a parameter. This is the **head parameter** and it is stated as:

The parameter that determines the relative positioning of heads with respect to their complements (Radford, 1997; 20).

The following are examples of phrases in English. Each phrase has a head and a complement.

<table>
<thead>
<tr>
<th>English</th>
<th>Phrasal category</th>
<th>Head</th>
<th>Complement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noun Phrase (NP)</td>
<td>the boy (N)</td>
<td>with a book</td>
<td></td>
</tr>
<tr>
<td>Verb Phrase (VP)</td>
<td>kick (V)</td>
<td>the boy</td>
<td></td>
</tr>
<tr>
<td>Prepositional Phrase (PP)</td>
<td>under (P)</td>
<td>the table</td>
<td></td>
</tr>
</tbody>
</table>

Since the head in each of the phrases above comes before its complement, English is a **head-first** language. The setting for the head parameter is head-first.

Now look at the following examples of Japanese phrases.
Japanese Phrasal category

Noun Phrase (NP)
- hon o motteiru otokonoko (N)
- book with boy
- 'the boy with a book'

Verb Phrase (VP)
- otokonoko o keru (V)
- boy kick
- 'kick the boy'

Prepositional Phrase (PP)
- tsukue no shita (P)
- table under
- 'under the table'

In the Japanese phrases above, we see that the complement comes before the head. Japanese is thus a head-last language. The setting of the head parameter in Japanese is head-last.

In sum, we can say that parameters are language-specific knowledge.

L1 Acquisition

We have seen in the previous section that the core element of a phrase is the head. Complements optionally modify the head. Another type of modifier is the Specifier. This modifier is also optional. Study the following English examples.
**Language Acquisition and Language Learning**

<table>
<thead>
<tr>
<th>Phrasal Category (XP)</th>
<th>Specifier</th>
<th>Head ($X^0$)</th>
<th>Complement</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP</td>
<td>John's</td>
<td>book</td>
<td>under the table</td>
</tr>
<tr>
<td>VP</td>
<td>always</td>
<td>reads</td>
<td>the same book</td>
</tr>
<tr>
<td>AP</td>
<td>amazingly</td>
<td>good</td>
<td>in mathematics</td>
</tr>
<tr>
<td>PP</td>
<td>always</td>
<td>with</td>
<td>an excuse</td>
</tr>
</tbody>
</table>

XP is the phrasal level, where X is either N, V, A or P and the head is represented as $X^0$ ($V^0$, $A^0$, $P^0$).

All phrases are arranged in a hierarchical manner, with an optional Specifier modifying an $X'$ (pronounced as X-bar), which consists of an $X^0$ (the head). This head is modified by an optional Complement (Mitchell and Myles, 1998:57). Thus the phrase in English can be represented as:

![Phrase Diagram](image-url)
This is a labeled tree diagram. In first language acquisition, children would know that sentences are made of phrases which comprise the Specifier-Head-Complement structure. They don’t have to learn this since this is part of the child’s innate knowledge. But they would not know the exact ordering of these elements in their language. They need linguistic input in order to set the head parameter. The number of possibilities with regard to the ordering of these three elements is constrained. The following are the possibilities:

1. Specifier-Head-Complement (like the English language)
2. Specifier-Complement-Head
3. Head-Complement-Specifier

According to Radford (1997: 22), children acquiring English as their first language appear to set the head parameter at its appropriate head-first setting from around 18 months of age, that is from the two-word stage. And they seem to know how to project productively the head (V\(^0\)) categories into the X' categories, and the X' categories into the XP categories (Towell and Hawkins, 1994: 65). The following are examples of utterances of a 20-month-old boy (examples from Radford, 1997: 22):

<table>
<thead>
<tr>
<th>Head (V(^0))</th>
<th>Complement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Touch</td>
<td>heads</td>
</tr>
<tr>
<td>Cuddle</td>
<td>book</td>
</tr>
<tr>
<td>Want</td>
<td>crayons</td>
</tr>
</tbody>
</table>
Children do not use different orders for different words of the same category. This model of language acquisition enables us to account for the rapidity and relatively error-free manner in which children acquire their native language. Acquiring this aspect of word order involves the simple task of setting a binary (two-way) either head-first or head-last parameter at its appropriate value. In other words, UG would tell the child that the only possible choices are for languages to be head-first or head-last (Radford, 1997: 22).

Now complete the following task.
Language Acquisition and Language Learning

Task 1 (60 minutes)

1. Define the terms **principles** and **parameters**.

2. Provide an example of a **principle**.

3. Provide an example of a **parameter**.

4. Name the two (2) settings of the head parameter and provide an example in each case.

5. When is this parameter set in L1 English children?

6. What are the three (3) elements in a phrase?

7. Draw a labeled tree diagram of the following phrase: *the girls in the building*

8. What does it mean when we say language acquisition (L1 acquisition in this case) is constrained? Illustrate this by using the arrangements of the three elements in (6) to form a phrase.

9. How do L1 learners acquire the phrase structure?
Answers
Task 1
Suggested answers

1. Define the terms principles and parameters.
   Principles are invariant and they are wired to the language faculty. Together they form Universal Grammar (UG) and this is part of a child's genetic blueprint for a grammar. These principles of UG are observed in every natural language. The principles specify limited possibilities of variations. These are the parameters that need to be fixed in one of a few possible ways.

2. Provide an example of a principle.
   The Structure Dependence Principle

3. Provide an example of a parameter.
   The Head Parameter

4. Name the two (2) settings of the head parameter and provide an example in each case.
   Head first (English) and Head last (Japanese)

5. When is this parameter set in L1 English children?
   At around 18 months of age

6. What are the three (3) elements in a phrase?
   Specifier, Head, Complement
7. Draw a labeled tree diagram of the following phrase: *the girls in the building*

```
NP
  
Specifier
  the

N'
  
N₀
  girls

Complement
  in the building
```

8. What does it mean when we say language acquisition (L1 acquisition in this case) is constrained? Illustrate by using the arrangements of the three elements in (5) to form a phrase.

The number of possibilities with regard to the ordering of elements in a structure is limited. The number of possibilities with regard to the ordering of elements of the three elements (specifier, head, complement) in a phrase is constrained. In other words, language acquisition is constrained. The following are the possibilities:

specifier-head-complement
specifier-complement-head
head-complement-specifier

9. How do L1 learners acquire the phrase structure?

They project productively the head (X₀) categories into the X' categories, and the X' categories into the XP categories.
UNIT 5: The Observable Phenomena in Second Language Learning

Introduction

According to Towell and Hawkins (1994: 7), any approach to the formulation of a theory of second language acquisition or learning would have to offer an account for the observable phenomena of the process involved. There are altogether five observable phenomena in second language acquisition (Towell and Hawkins, 1994: 7 – 16). They are transfer of properties of the first language (L1) into the second language, staged development, systematocity in the attainment of knowledge of the second language (L2), variability at some stages of the development of the second language and incompleteness in second language acquisition.

The Observable Phenomena Of Second Language Acquisition

Transfer of Properties of the L1 into the L2

In second language acquisition or learning (SLA/SLL), there is transfer of linguistic properties from the learner’s L1 into the target language or L2. This is a pervasive feature of the process (Towell and Hawkins, 1994: 7). In fact, this was the main focus of attention for researchers involved in the study of the nature of second languages in the 1960s. This phenomenon affects not only the phonology (pronunciation) of the L2 learners, it is also observed in the morphology (the
internal structure of words), lexicon (vocabulary) and discourse (how sentences are put to use in the course of communication).

It is a well-documented fact (and most of us would have observed this) that when there is an obvious difference in the L1 and L2 on a particular grammatical property, L2 learners seem to transfer the L1 property into the L2. In this case, the L2 learner would produce patterns of speech not found in the L2. As stated earlier, transfer can affect all linguistic levels. We will discuss transfer in the three of these levels – phonological, morphological and syntactic.

Transfer Of Phonological Properties

This can be observed among some Malaysian L2 learners of English. For example, L1 Malaysian Chinese speakers of L2 English would generally pronounce words such as thirsty (pronounced as [θəˈstiː]) and mother (pronounced as [ˈmʌðər]) as [θəˈstiː] and [ˈmʌðər] respectively. The voiceless and voiced interdental fricatives /θ/ and /ð/ are not present in Chinese and thus the L1 Chinese speakers tend to substitute them with the voiceless and voiced alveolar stops /t/ and /d/ respectively.

Transfer Of Morphological Properties

In the case of morphological transfer, it has been observed that L1 Malaysian Chinese speakers of L2 English tend to add the adverb already after some verbs as in:
1. *I eat already (to indicate I have (already) eaten)

This is probably due to the fact that in Chinese, there is a verbal suffix *le* (for perfect aspect) (Norman, 1988: 157). The learners transfer this property into their L2. The Chinese equivalents of (1) is:

1'. *I eat already
   wǒ chī le

Transfer Of Syntactic Properties

As for syntactic transfer, it is not uncommon to hear the following being uttered by a Malaysian L1 Tamil speaker of L2 English, especially in the initial stages of SLA.

2. *I rice eat (the correct version is I eat rice)

In the above example, the Tamil L1 speaker has reversed the verb and object. In English, the word order is Subject-Verb-Object while in Tamil, it is Subject-Object-Verb. Thus the production of the above utterance by the L2 learner.

So far, we have looked at examples where the L1 and L2 properties are structurally different. It has also been observed that transfer can occur when the L1 and L2 properties are structurally identical (Towell and Hawkins, 1994: 9). For example, Zobl (1984, in Towell and Hawkins, 1994:9) noted that in the
Language Acquisition and Language Learning

acquisition of the indefinite and definite determiners a and the in English, L1 French and Spanish speakers seem to acquire these properties faster than L1 Chinese and Russian speakers. The explanation offered is that French and Spanish also make a distinction between indefinite and definite determiners while Chinese and Russian do not.

Staged Development in SLA/SLL

The L2 learner generally progress from initial-state grammars towards the target language. In the initial stages, the learner is usually influenced by transfer. In this respect (staged development), L2 learners resemble L1 learners. However, the initial-state grammar of the L2 learner is different from that of the L1 learner.

The initial state of the L1 learner is a zero state (S₀, that is the learner has not acquired any language yet) while that of the L2 learner is an initial state (S₀, where the learner has complete control of at least one language)¹ (Cook and Newson, 1996: 126).

Clahsen and Muysken (1986, in Towell and Hawkins, 1994: 10 – 11) compared the development of word order patterns in German among L1 learners and L2 learners whose first languages include Italian, Spanish, Portuguese and Turkish. They found that L2 go through a number of stages in the acquisition of word order in German. Initially, they produce word order patterns that were ungrammatical in German. Towell and Hawkins suggest that these word order patterns are probably a result of transfer from the L2 learners’ first languages. The following were the stages (taken from Towell and Hawkins, 1994: 10):

¹ See also Units 1 and 7.
**Verb separation**

3. Er hat heute ein Buch gekauft
   He has today a book bought

**V2 (Verb two)**

3'. Heute hat er ein Buch gekauft
   Today has he a book bought

**Verb final**

3'''. Ich glaube dass er heute ein Buch gekauft hat
   I think that he today a book bought has

Subsequently, the L2 learners acquired the appropriate word order patterns just like L1 learners do. However, the word order patterns were different for both sets of learners.

In other words, both L2 and L1 learners approximate the target language in stages (staged development) for this property but the stages they go through are different.

**Systematicity in SLA/SLL Across Learners**

Studies have shown that learners from different L1 backgrounds develop L2 linguistic knowledge in a way that is not directly attributable either to their L1, or to the L2 input (Towell and Hawkins, 1994: 11). Patterns in the development of accuracy on grammatical morphemes in English across a range of L2 learners of different ages, from different language backgrounds and learning English under
different conditions have displayed similar patterns in the development of accuracy on grammatical morphemes in English (Dulay and Burt, 1973; Dulay and Burt, 1974; Bailey, Madden & Krashen, 1974). The studies by Dulay and Burt (1973 and 1974) showed that child learners whose first languages were Spanish and Cantonese learning L2 English produced some grammatical morphemes more accurately than others. In other words, whether they spoke Spanish or Cantonese, they acquired these morphemes in a particular order. The following is the sequence of the acquisition of some of the morphemes (from Dulay, Burt and Krashen, 1982: 206; examples from the present author).

4. Progressive: The girl is playing.
   Plural s: The girls are dirty.
   Past regular: The girl jumped.
   Past irregular: The girl drank the milk.
   Possessive ‘s: The girl’s toy
   Third person singular present tense: The girl jumps.

From the findings of these studies, it was concluded that learners from different L1 backgrounds, acquiring an L2 either in a naturalistic or classroom environment, go through the same stages of development, that is there is systematic development in the SLA process.

Variability in SLA/SLL

The grammars of L2 learners at particular stages of development appear to allow more than one structural form for a given construction where the TL has only one form (Towell and Hawkins, 1994: 13). In other words, there is variability in the
learner’s intuitions about and production of certain L2 properties. For example, an 11-year-old Portuguese learner of L2 English produced the following forms with the same meaning (no and don’t to perform the task of negation) in a card game (Ellis, 1992; in Towell and Hawkins, 1994: 13):

5.  
No look my card.

Don’t look my card.

Variation of this nature is usually temporary although in certain cases, they may persist.

Incompleteness in SLA/SLL

Generally, very few L2 learners appear to be fully successful in the way that native speakers are (Towell and Hawkins, 1994: 14). In a study conducted by Johnson and Newport (1989), it was found that L1 Chinese and L1 Korean speakers learning English as an L2 in the United States of America acquired the target language to varying degrees of proficiency depending on when they first arrived in the US. All the learners had had at least five years of exposure to English in their adopted country (US). In the study the learners had to complete a grammaticality judgement task, that is they had to judge whether or not the given items (which included determiners a/the, plural, past, third person singular present tense, pronouns, yes/no questions, word order, and so on) were grammatical. The main factor determining a subject’s degree of success in attaining native-like proficiency was the age at which she or he was first consistently exposed to the L2. Learners who attained native-like judgements were those who were first consistently exposed to the L2 before the age of seven. The older the learner is
Language Acquisition and Language Learning

when he or she was first exposed to the target language, the less accurate s/he was in the task.

Now, complete the following task.

🤔 Task 1 (20 minutes)

Answer the following questions.

1. What are the observable phenomena of SLA/SLL?
2. Provide an example of phonological transfer.
3. Provide an example morphological transfer.
4. Provide an example of syntactic transfer.
5. Provide an example of the phenomenon of staged development in SLA/SLL.
6. What can we deduce from Dulay and Burt’s findings (1973 and 1974)?
7. Explain briefly the phenomenon of variability in SLA/SLL. Quote an example.
8. Describe briefly the findings of Johnson and Newport (1989).
Answers
Task 1
Suggested answers

1. What are the observable phenomena of SLA/SLL?

   Transfer of properties of the first language (L1) into the second language;
   Staged development in the acquisition of properties of the L2;
   Systematicity in the attainment of knowledge of the second language (L2);
   Variability at certain stages of the development of the second language;
   Incompleteness in the acquisition of properties in the L2.

2. Provide an example of phonological transfer.
   Any example that you have observed.

3. Provide an example morphological transfer.
   Any example that you have observed.

4. Provide an example of syntactic transfer.
   Any example that you have observed.

5. Provide an example of the phenomenon of staged development in SLA/SLL.
Language Acquisition and Language Learning

a. **Verb separation**
   Er hat heute ein Buch gekauft
   He has today a book bought

b. **V2 (Verb two)**
   Heute hat er ein Buch gekauft
   Today has he a book bought

c. **Verb final**
   Ich glaube dass er heute ein Buch gekauft hat
   I think that he today a book bought has

6. What can we deduce from Dulen and Burt’s findings (1973 and 1974)?
   Learners from different L1 backgrounds, acquiring an L2 either in a
   naturalistic or classroom environment, displayed similar patterns in the
   development of accuracy on grammatical morphemes in English. In
   other words, they go through the same stages of development, that is
   there is systematic development in the process of SLA/SLL.

7. Explain briefly the phenomenon of variability in SLA/SLL. Quote an
   example.
   Variability refers to the phenomenon where L2 learners at particular
   stages of development appear to allow more than one structural form
   for a given construction where the TL has only one form (Towell and
   Quote any example you have observed.
8. Explain briefly the findings of Johnson and Newport (1989). The main factor determining an L2 learner’s degree of success in attaining native-like proficiency in US English was the age at which the learner was first consistently exposed to the L2, that is the age at which the learner first arrived in the US. Learners who attained native-like judgements were those who were first consistently exposed to the L2 before the age of seven. The older the learner is when he or she was first exposed to the target language, the less accurate s/he was in the grammaticality judgement task.
UNIT 6: The Cognitive Approach to Second Language Learning

Introduction

The researchers who are interested in cognitive approaches are those who are interested in the learning component of second language acquisition or learning (SLA/SLL) (Mitchell and Myles, 1998: 72). These are also known as psychological approaches. The focus is mainly on staged development and systematicity. They do not necessarily recognize that there is an encapsulated module of linguistic ability in SLA/SLL (Towell and Hawkins, 1994: 45). In this unit, we will look at a perceptual saliency approach and two information processing models. In the former, we will discuss the learnability/teachability hypothesis while in the latter, we will discuss McLaughlin’s information processing model and Anderson’s ACT* model.

The Perceptual Saliency Approach

According to Slotin (1970s and 1980s, in Mitchell and Myles, 1998: 74), humans are programmed to perceive their environment, make sense of it and organise this information in certain ways. In other words, it is perceptual saliency which drives the learning process. Slotin came up with principles and universals for L1 acquisition. These universals are cognitive in nature and are different from linguistic universals. Some researchers have appealed to similar psychological principles in order to explain SLA/SLL phenomena.
Learnability/Teachability Hypothesis


Pienemann (based on work done by Meisel, Claibsen and Pienemann, 1981, in Towell and Hawkins, 1994:46) made the assumptions that L2 learners:

- will attend to and acquire ways of expressing actors, actions and people or things acted upon (because these are more salient) before they will attend to and acquire adverbials dealing with place, time and manner of event.

- will attend to and acquire the beginnings and ends of sentences before the middles of sentences as the former are more salient.

Wolfe Quintero (1992, in Towell and Hawkins, 1994: 46) made the assumption that the continuous phenomena are more salient than the discontinuous, so that learners will attend to she picked up the book (continuous phenomenon) before they will attend to she picked the book up (discontinuous phenomenon).

The notion of learnability came about from the observation that L2 learners follow a fairly rigid route in the learning of certain grammatical structures (Pienemann, in Mitchell and Myles, 1994: 77). The structures become learnable when the previous steps in the acquisitional or developmental path have been acquired. This is the learnability hypothesis. In other words, the learners cannot acquire a complex structure straight away.
Pienemann proposes that effective teaching will take place only if learners have successfully passed through the preceding stages on the developmental route (in Mitchell and Myles, 1998: 79). This is the teachability hypothesis. A pedagogical implication that can be drawn from this hypothesis is that by looking at the productions for a given structure of an L2 learner, and by placing them on the developmental route, a teacher can assess what this learner can be expected to learn next.

Now complete the following task.

atra? Task 1 (20 minutes)

State whether the following are true (T) or false (F)

1. L2 learners follow a rather rigid route in learning certain structures.
2. By looking at the productions of an L2 learner, a teacher can assess what this learner can be expected to learn next.
3. L2 learners can be taught a complex structure quite easily in a formal learning environment.
4. Learners will initially attend to and acquire ways of expressing ‘actors’, ‘actions’ and people or things acted upon’.
5. Learners will see the beginnings and ends of sentences as more perceptually salient than the middles of sentences.
6. Learners will find ‘discontinuous’ phenomena more salient than ‘continuous’ phenomena in sentences.
Information Processing Models

Such models have been developed by cognitive psychologists and have been adapted to L2 language processing. These models see complex behaviours as building on simple processes. We will discuss McLaughlin’s Information Processing Model (1987, 1990) and Anderson’s ACT* model (1983, 1985) (in Mitchell and Myles, 1998: 85 – 86).

The main characteristics of such an approach include the following (Mitchell and 1998: 85):

- Humans are seen autonomous and active beings.
- The mind is a general-purpose, symbol-processing system.
- Complex behaviours comprise modular simpler processes.
- These processes can be isolated and studied independently of other processes.
- Processes take time and we can make predictions about reaction time.
- The mind is a limited-capacity processor.

McLaughlin’s Information Processing Model

In this model, SLA/SLL is seen as the acquisition of a complex cognitive skill. To learn a second language is to learn a skill as the different aspects of the task have to be practised and integrated into fluent performance (McLaughlin, 1987:
133 - 4). Applied to SLA/SLL, the model works in the following way (Mitchell and Myles, 1998: 86):

a. Learners first resort to controlled processing in the L2 which involves the temporary activation of a selection of information nodes in the memory. Such processing needs a lot of attentional control and is constrained by the limitations of the Short-Term Memory (STM).

b. Through repeated activation or practice, structures (utterances/sentences) first produced by controlled processing become automatic. Automatized structures are stored as units in the Long-Term Memory (LTM). These can be made available quickly with minimal attentional control. These automatized skills are not easy to erase or modify.

c. When a sequence becomes automatic (through automatic processing), controlled processes are free to deal with higher levels of processing. This explains the incremental or step by step nature of learning.

d. This continuous movement from controlled to automatic processing results in a constant restructuring of the linguistic system of the L2 learner.

This model can explain variability and fossilization (incompleteness) to a certain extent. Restructuring destabilizes some structures (already acquired) in the interlanguage (IL) and leads to temporary reappearance of L2 errors. This leads to variability. Fossilization (incompleteness) would arise as a result of a controlled process becoming automatic before it is native-like. Since automatic processes are difficult to modify, they are likely to remain in the learner’s IL resulting in a stable but an inappropriate and erroneous construction.
Andersen’s ACT* Model

This is another information processing model from cognitive psychology which has been applied to aspects of SLA/SLL. ACT stands for Adaptive Control of Thought and the asterisk refers to a revised and updated version of the initial ACT model. According Mitchell and Myles (1998: 87), this model is different from McLaughlin’s information processing model in the sense that it is more wide-ranging. The terminology used is also different. It is similar to McLaughlin’s model in that practice leading to automatization plays a central role in SLA/SLL.

Two types of knowledge are posited by this model. They are declarative knowledge (knowledge that which is similar to controlled processes) and procedural knowledge (knowledge how which is similar to automatic knowledge). In terms of learning, practice leading to automatization enables declarative knowledge to become procedural knowledge. Three kinds of memories are proposed by this model. They are a working memory and two kinds of long-term memories (LTM), a declarative LTM and a procedural LTM. These different kinds of knowledge are stored differently.

This model can be applied to certain aspects of SLA/SLL, especially those which require proceduralization and automatization as it is general cognitive model of skill acquisition. Learning a rule consciously, for example the third person singular present tense –s marker, would constitute declarative knowledge as the L2 learner might not be able to consistently produce the marker in a conversation in real time. After much practice, this rule (knowledge) will probably become proceduralized and the marker will be produced each time it is required. The
move from declarative knowledge to procedural knowledge takes place in three stages (Anderson, 1985, in Towell and Hawkins, 1994: 203):

a. The **cognitive stage** where a description of the procedure is learned.
b. The **associative stage** where a method for performing the skill is worked out.
c. The **autonomous stage** where the skill becomes more and more rapid and automatic.

Now complete the following task.

---

❓ Task 2 (60 minutes)

1. What does ACT* stand for? What does the star stand for?

2. How is this model different when compared to McLaughlin's information processing model? How is it similar?

3. What are the two types of knowledge posited by this model? How do they work?

4. What are the three kinds of memory posited by this model?

5. The early stages of learning how to drive is likened to one type of knowledge. What is it?
6. With practice, this knowledge in (5) becomes another type of knowledge. Name this type of knowledge.

7. Why can this model be applied to certain aspects (those which require *proceduralization* and *automatization*) of SLA/SLL?

8. According to Anderson (1985), the move from *declarative knowledge* to *procedural knowledge* takes place in three stages. Name them and provide a brief explanation for each of these stages.

9. Complete the following:

<table>
<thead>
<tr>
<th>McLaughlin's Model</th>
<th>Anderson's Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short-Term Memory</td>
<td></td>
</tr>
<tr>
<td>Long-Term Memory</td>
<td></td>
</tr>
<tr>
<td>Controlled Processes</td>
<td></td>
</tr>
<tr>
<td>Automatic Processes</td>
<td></td>
</tr>
</tbody>
</table>
Strengths and Weakness of Cognitive Approaches

Most of the studies based on such approaches are carried out in the laboratory, thus enabling the researchers to control the variables under study (Mitchell and Myles, 1998: 98). This very advantage is also a weakness as it assumes that particular aspects of language can be studied in isolation without taking into consideration the interaction between the different language modules (phonology, syntax, semantics, and so on).

Such approaches have also revealed the processes involved in the speeding up of the acquisition process, thus enabling us to draw pedagogical implications from the findings. However, these approaches cannot explain adequately what the mental grammar of the L2 learner consists of and what constrains the learner’s hypotheses about the language.
Answers

Task 1

Suggested answers

State whether the following are true (T) or false (F)

1. L2 learners follow a rather rigid route in learning certain structures. (T)
2. By looking at the productions of an L2 learner, a teacher can assess what this learner can be expected to learn next. (T)
3. L2 learners can be taught a complex structure quite easily in a formal learning environment. (F)
4. Learners will initially attend to and acquire ways of expressing ‘actors’, ‘actions’ and people or things acted upon’. (T)
5. Learners will see the beginnings and ends of sentences as more perceptually salient than the middles of sentences. (T)
6. Learners will find ‘discontinuous’ phenomena as more salient than ‘continuous’ phenomena in sentences. (F)

Task 2

Suggested answers

1. What does ACT* stand for? What does the star stand for?

   Adaptive Control of Thought. The star refers to a revised and updated version of Anderson’s initial ACT model.
2. How is this model different when compared to McLaughlin’s information processing model? How is it similar?
   It is a more wide-ranging model and the terminology is different. The similarity is that practice leading to automatization also plays a central role.

3. What are the two types of knowledge posited by this model? How do they work?
   Declarative knowledge (knowledge that, similar to controlled processes) and procedural knowledge (knowledge how, similar to automatic knowledge). Practice leading to automatization enables declarative knowledge to become procedural knowledge.

4. What are the three kinds of memory posited by this model?
   A working memory; two kinds of long-term memories – a declarative LTM and a procedural LTM.

5. The early stages of learning how to drive is likened to one type of knowledge. What is it?
   Declarative knowledge (knowing that)

6. With practice, this knowledge in (5) becomes another type of knowledge. Name this type of knowledge.
   Procedural knowledge (knowing how)

7. Why can this model be applied to certain aspects (those which require proceduralization and automatization) of SLA/SLL?
   Because it is a general cognitive model of skill acquisition.
8. According to Anderson (1985), the move from declarative knowledge to procedural knowledge takes place in three stages. Name them and provide a brief explanation for each of these stages.
   a. The cognitive stage where a description of the procedure is learned.
   b. The associative stage where a method for performing the skill is worked out.
   c. The autonomous stage where the skill becomes more and more rapid and automatic.

9. Complete the following:

   McLaughlin’s Model   Anderson’s Model

   Short-Term Memory    Working Memory
   Long-Term Memory     Declarative LTM and Procedural LTM
   Controlled Processes Declarative knowledge
   Automatic Processes  Procedural knowledge
UNIT 7: The Role of Universal Grammar in Second Language Learning

Introduction

In Unit 5 we saw that second language acquisition or second language learning was very similar to first language (L1) acquisition – there is staged development and systematicity in the process. Second language (L2) learners are also faced with the same logical problem of constructing a grammar of the L2 on the basis of fragmentary input and of having to construct abstract representations on the basis of the samples of language they actually encounter.

However, unlike a first language learner, the second language learner is someone who has already acquired a developed steady-state grammar for a first language with the parameters set to the values of the L1. Thus the L2 learner is armed with mature competence in an L1 which may or may not have similar syntactic properties to the target language. In fact, it is claimed that the initial state of older L2 learners also includes more developed cognitive abilities as well as problem-solving skills. Thus the initial state of the L2 learner is not a zero state (S₀) but a complete grammar of the L1. This is the fundamental difference between L1 acquisition and second language acquisition (SLA) or second language learning (SLL). The S₀ (zero state)¹ comprises the invariant principles of UG with options or parameters attached to them. This fact affects the assumptions we make with regard to SLA/SLL in relation to UG. With development the L2 initial state restructures towards a terminal state. This state of affairs can be summarized as follows:

¹ See also Units 1 and 4.
$S_i \rightarrow S_e$

where $S_i$ is the initial state in SLA and $S_e$ is the terminal state (Cook and Newson, 1996: 126). Not only is the initial state in SLA different from the initial state in L1 acquisition, but the terminal state may also differ in that L2 learners do not all achieve the same competence. Some acquire near native competence while others fossilize at a very elementary level.

Thus L2 learners come to the task of SLA/SLL with a steady L1 grammar state. In addition, the L2 learner is more mature cognitively and thus is able to solve problems and deal with abstract concepts. The L2 learner may also have different motivations or needs for learning an L2 (L1 acquisition takes place because there is the basic human need to communicate) (Mitchell and Myles, 1998: 61). If L2 learners come to the task of SLA/SLL with a steady L1 grammar state, does this mean that a potential L2 learner has no further access to the principles and parameters of Universal Grammar (UG)? If this is the case, would they only be able to learn another language by falling back on the properties of L1? This state of affairs has given rise to four logical possibilities with regard to the role of UG in L2 acquisition. They are the full access, no access, indirect access and partial access hypotheses (Mitchell and Myles, 1998: 61 – 62).

**Access to UG in L2 Acquisition**

**The Full Access to UG View**

This view is also known as the direct access view. It states that L2 learners may employ the principles of UG and set the parameters without any reference their
L1 values. In other words, UG can be accessed directly in L2 acquisition and L1 and L2 acquisition are basically similar processes and the differences that are observed are due to cognitive maturity and in the needs of the learner (Mitchell and Myles, 1998: 61). In this view, learners are therefore said to have parallel competences in L1 and L2, that is there are two instantiations of UG. The poverty-of-the-stimulus argument also applies to L2 learners in so far as anything that the L2 speaker could not have learnt from the environment must be an innate property of the mind (Cook and Newson, 1996:292). One way of showing that L2 learners have full access to UG in SLA is to show that they indeed know principles and parameters they could not have learnt (see for e.g. Flynn, 1996; Thomas, 1991; and Travis and MacEachlan, 1992).

The No Access to UG View

This hypothesis or view states that UG is not involved in L2 acquisition. It is said that UG atrophies with age and L2 learners have to resort to more general problem-solving skills (Mitchell and Myles, 1998: 61). Arguments in favour of the no access to UG view include the following:
1. the L2 knowledge is usually not so complete (see for e.g. Bley-Vroman, 1989; Johnson and Newport, 1989);
2. the L2 gets fossilized, and
3. L2 learners vary in ways that L1 learners do not.

In Unit 5, we have seen that L2 learners do indeed exhibit these phenomena.
The Indirect Access to UG View

This view sees UG as not being directly involved in L2 acquisition but it is indirectly accessed via the L1 (Mitchell and Myles, 1998: 61). Therefore, there will be only one instantiation (that is one working example of UG) made available to the L2 learner. In this case, the parameters are already fixed to the settings for the L1. There are two sub-views for this hypothesis.

The first sub-view is proposed by Schachter (1996). She combines this view with the notion of a critical period for L2 acquisition. She claims that the UG principles that are accessible to an L2 learner are those that are already instantiated in the learner's L1. However, she also claims that child L2 learners may be able to access UG for the task. But she added that there is a critical period or several critical periods for the successful acquisition of L2 principles and/or parameter settings if these are not instantiated in the child learner’s L1. She refers to the critical period as a Window of Opportunity and further claims that a child L2 learner passes through different Windows for the different components of the target language (Schachter, 1990: 188).

The second sub-view of the indirect access view is proposed by Schwartz and Sprouse (1994). They claim that learners transfer all the parameter-settings from their L1 in the initial stage of L2 acquisition. Then they revise their hypotheses when the L2 cannot be 'mapped' onto the L1 settings. The L2 learners then form new hypotheses but these are constrained by UG. In this variant of the indirect access view, learners initially access UG via the L1 and then directly at the more advanced stage of the SLA/SLL process.
The Partial Access to UG View

This view came to the fore only recently. The partial access to UG view or hypothesis proposes that some principles and parameters seem to be unproblematic to reset (for e.g. the head parameter) while others are more difficult or impossible to reset (for e.g. Subjacency). The result is that L2 learners produce grammars that are neither like their L1 nor the target language (Mitchell and Myles, 1998: 68). Recent work on SLA within this framework suggests that learners are unable to reset parameters of UG but they are still able to model syntactic properties of the L2 which are different from the first language (L1) through the universal principles made available by UG (see for e.g. Hawkins and Chan, 1997; Wong and Hawkins, 2000).

Evaluation of the UG-Based Approach

Criticism of the UG Theory

The UG theory is a linguistic theory which aims to describe and explain human language. It is relevant to SLA research indirectly in that it helps us to understand the acquisition process and what it is that learners have to acquire. However, this theory has been criticized for focusing on some aspects of language but not others. In the past, it focused mainly on syntax. Recently, more interest has been shown in phonology, morphology and the lexicon. However, semantics, pragmatics and discourse are still not considered. The object of study is still the sentence and its internal structure, rather than any larger unit of language. This includes the study of smaller units (words, morphemes and phonemes) and how different elements relate to one another.
Language Acquisition and Language Learning

The major criticism is that it studies language clinically, in a vacuum, as a mental object rather than a social or psychological one. This approach views the speaker/learner not as an individual with varied characteristics, nor as a social being but as an idealized receptacle for the UG blueprint (innate knowledge). The emphasis is not on the speaker/learner (the person) but on language as the object of study. In spite of these criticisms, it is highly influential as a linguistic theory and is the most sophisticated tool for analyzing language today (Mitchell and Myles, 1998:69).

UG and SLA

Weaknesses

The UG-based approach has been criticized for the same reasons as the theory itself. In the past, SLA research has concerned itself with syntax mainly. Now, the areas of phonology, morphology and the lexicon are being investigated. Semantics, pragmatics and discourse are still ignored. The approach is concerned mainly with the developmental linguistic route followed by L2 learners. It ignores the social and psychological variables which affect the rate of the acquisition process. It is interested in the modelling of linguistic competence and not the aspect of naturalistic performance (Mitchell and Myles, 1998:70).

Strengths

The UG approach to SLA research has been highly influential and has yielded much knowledge about the L2 acquisition process. In addition, it is a
sophisticated descriptive tool for linguistic analysis. It has been used to describe the interlanguage, the L2 and the L1 of the second language learners. Researchers are able to formulate well-defined and focused hypotheses in empirical work using this framework. As a result, SLA/SLL research has increased our knowledge of human language.

It has also been useful in explaining some facts about SLA/SLL. For example, it has informed our understanding of the phenomena of staged development and systematicity in SLA, that is if learners are constrained by UG, their development should be staged and systematic (just like L1 development is). This theory can also explain transfer/cross-linguistic influence in terms of principles and parameters, that is whether or not parameters can be reset (Mitchell and Myles, 1998: 70 - 71).

In sum, we can say that this is a good theory as it makes explicit statements of the grounds it aims to cover and the claims it makes, by having systematic procedures for theory evaluation, by attempting to explain and describe some of the L2 phenomena.

Task 1 (60 minutes)

1. What are the logical possibilities/hypotheses concerning the role of UG in L2 acquisition/learning?
Language Acquisition and Language Learning

2. Hypothesis 1:
   Argument/studies:

3. Hypothesis 2:
   Argument/studies:

4. Hypothesis 3:
   Argument/studies:

5. Hypothesis 4:
   Argument/studies:

6. In the UG approach, how is the learner or speaker viewed as?

7. What is the object of study in this approach?

8. What is a major criticism of this approach?

9. List the weaknesses of this approach.

10. List the strengths of this approach.
Answers

Task 1

Suggested answers

1. What are the logical possibilities/hypotheses concerning the role of UG in L2 acquisition/learning?
   Full Access/Direct Access to UG
   No Access to UG
   Indirect Access to UG
   Partial Access to UG

2. Hypothesis 1:
   Argument/studies:
   The Full Access (Direct Access) to UG View (see pages 66 – 67)

3. Hypothesis 2:
   Argument/studies:
   The No Access to UG View (see page 67)

4. Hypothesis 3:
   Argument/studies:
   The Indirect Access to UG View (see page 68)

5. Hypothesis 4:
   Argument/studies:
   The Partial Access to UG View (see page 69)
6. In the UG approach, how is the learner or speaker viewed as?
   This approach views the learner or speaker not as an individual with varied
   characteristics, nor as a social being but as an idealized receptacle for the UG
   blueprint (innate knowledge).

7. What is the object of study in this approach?
   Language is the object of study in this approach.

8. What is a major criticism of this approach?
   The major criticism is of this approach is that it studies language clinically, in
   a vacuum, as a mental object rather than a social or psychological one.

9. List the weaknesses of this approach.
   See page 70.

10. List the strengths of this approach.
    See pages 70 and 71.