

An Introduction to Issues in General Linguistics

An Introduction to Issues in General Linguistics

By

Georgios P. Georgiou

**Cambridge
Scholars
Publishing**



An Introduction to Issues in General Linguistics

By Georgios P. Georgiou

This book first published 2020

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

Copyright © 2020 by Georgios P. Georgiou

All rights for this book reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the copyright owner.

ISBN (10): 1-5275-5950-5

ISBN (13): 978-1-5275-5950-9

TABLE OF CONTENTS

TABLE OF ILLUSTRATIONS	vii
PREFACE	ix
CHAPTER 1: Introduction to Linguistics	
1.1 What is language?	1
1.2 Language as a communication code.....	2
1.3 The Sign	4
1.4 Features of Language	6
1.5 The science of Linguistics	10
1.6 Levels of Linguistic Analysis.....	11
1.7 Principles of Linguistics.....	32
CHAPTER 2: Debugging the Myths about Language	
2.1 Myths about languages and language learning.....	35
2.2 Myths about Linguistics and linguists.....	38
CHAPTER 3: Speech Perception and Production	
3.1 Brief historical overview of L2 speech perception and production.....	41
3.2 Speech Perception	42
3.3 Speech Production.....	42
3.4 Is speech only a human privilege?.....	51
3.5 Phonological and phonetic differences between the sounds of different languages may impede communication.....	53
3.6 Lack of pronunciation may lead to tragedies.....	54
3.7 Factors affecting the acquisition of L2 sounds	56
CHAPTER 4: Language Learning and Education	
4.1 Why learn a second language?	67
4.2 Is it possible to learn how to pronounce a second language? ...	68
4.3 The definition of language mistakes and errors from the perspective of linguistics.....	69
4.4 Should educators correct or not the language “errors”?	71
4.5 Should a language’s grammar be taught?.....	73
4.6 Language Transfer, Error Analysis, and Inter-language	75

CHAPTER 5: Language and Society	
5.1 Even languages die.....	79
5.2 Languages that have been revived: The case of Hebrew.....	81
5.3 The phenomenon of linguistic reborrowing	83
5.4 Not All Languages are Natural: The Case of Esperanto.....	84
5.5 Do young people suffer from a lack of vocabulary?	86
5.6 The role of abbreviations in the written communication.....	87
5.7 Perceptions Regarding the Use of the Roman Alphabet for Writing Languages with a Non-Roman Alphabet: The Case of Greeklish	89
CHAPTER 6: Speech Perception and Production Models	
6.1 Speech Learning Model (SLM).....	91
6.2 Perceptual Assimilation Model (PAM).....	96
6.3 Perceptual Assimilation Model-L2 (PAM-L2).....	99
6.4 A Comparison of SLM and PAM/PAM-L2	100
REFERENCES.....	103

TABLE OF ILLUSTRATIONS

Fig. 1-1: Communication cycle model by Claude Elwood Shannon and Warren Weaver.....	3
Fig. 1-2: The signified and the signifier of the word “tree”.....	5
Fig. 1-3: The four features of language	9
Fig. 1-4: Waveform and Spectrogram in a PRAAT script for the analysis of acoustic characteristics of sounds (e.g., formant frequencies and duration).	12
Fig. 1-5: Example of the International Phonetic Alphabet (Omniglot)	13
Fig. 1-6: Main subdivisions of morphemes	16
Fig. 1-7: The lexical categories of English	20
Fig. 1-8: A brief definition of phrase, utterance, and sentence	24
Fig. 1-9: The six functions of language proposed by Jakobson.....	31
Fig. 1-10: Levels of linguistic analysis.....	32
Fig. 1-11: The principles of Linguistics	34
Fig. 3-1: Human speech organs	44
Fig. 3-2: The state of the vocal folds	45
Fig. 3-3: An overview of the vocal tract.....	46
Fig. 3-4: Manner, place, and voicing of English consonants.....	48
Fig. 3-5: The official IPA Vowel Chart.....	50
Fig. 3-6: English vowels (Received Pronunciation) (Roach, 2004).....	50
Fig. 3-7: The four processes of speech production	51
Fig. 5-1: The world’s language hierarchy (Graddol, 1997).....	81
Fig. 5-2: The letters and sounds of the Hebrew alphabet	83
Fig. 5-3: The letters and sounds of Esperanto	86
Fig. 5-4: Greeklish to Greek converter	90

PREFACE

An introduction to Issues in General Linguistics aims to show how the language system works, to cultivate a correct attitude towards language, and to familiarize readers with the science of linguistics and issues related to it. Specifically, the first chapter is a must-read by people who have little contact with linguistics in order to gain a better understanding of the way that the language system operates, as well as to familiarize them with language features. Furthermore, readers will be able to observe the main objectives and principles of linguistics, and differentiate the different levels of linguistic analysis. Chapter 2 aims to debunk myths about language by dissolving misperceptions about the language system and the science of linguistics. Chapter 3 deals with the way that humans perceive and produce language. This chapter also includes brief sections about issues surrounding pronunciation and language production. Chapter 4 discusses issues surrounding language learning, and it connects linguistics with education. Chapter 5 contains sections that discuss social phenomena related to linguistics, such as linguistic reborrowing, the language of young people, and revived or artificial languages. Chapter 6 presents two main speech models.

I tried to avoid excessive wordiness since it often creates feelings of boredom in readers. Instead, the book only focuses on essential linguistic issues, trying to explain them as briefly as possible without losing its scientific character. The first two chapters are significant as they develop an understanding of the language system mechanism and eliminate misconceptions. The other chapters contain brief sections that discuss intertemporal and timely issues in linguistics. All linguistic phenomena are accompanied by examples so that the reader can understand how they are embedded in real linguistic contexts.

The book discusses linguistic issues scientifically; that is, it neither adopts personal convictions nor uses unsubstantial arguments; instead, it employs findings that have emerged from research. The language of the book is simple so that it can be understood by a wide range of readers, even if they do not have any previous knowledge in linguistics. Therefore, *Issues in Linguistics* is aimed at linguists, philologists, language scholars,

tutors, and anyone who loves linguistics or even general readers who are curious to explore the science of linguistics.

The Author of the book

Dr George Georgiou

CHAPTER 1

INTRODUCTION TO LINGUISTICS

1.1. What is language?

In general terms, *language* is a system that describes the ability of humans to develop, acquire, maintain, and use complex communication codes. Several definitions of language have been proposed with each one linked to a different school of thought. However, a complete definition might never be expressed since language is a complex system, whereas the science that studies language, linguistics, is linked with other sciences, such as psychology and the cognitive sciences. The notable American linguist, Noam Chomsky (1957: 13), defines language as follows:

“Language is a set (finite or infinite) of sentences, each finite in length, and constructed out of a finite set of elements”.

Hadumod Bussmann gives a relatively complete and comprehensive definition of language in the *Routledge Dictionary of Language and Linguistics* (Bussmann, 1999: 627):

“Vehicle for the expression or exchanging of thoughts, concepts, knowledge, and information as well as the fixing and transmission of experience and knowledge. It is based on cognitive processes, subject to societal factors and subject to historical change and development. In this definition, language refers to a specific form of expression that is restricted to humans, and differs from all other possible languages, such as animal communication and artificial languages through creativity, the ability to make conceptual abstractions, and the possibility of metalinguistic reflection”.

As expressed above, language is used for human communication. *Communication* is the procedure in which messages of any kind are transferred. Members of a community are people who communicate with each other using one or more languages, which are based on specific rules.

1.1.1. Speech, Language, and Speaking

A distinction between *Speech*, *Language*, and *Speaking* can provide a better understanding of the human communication code.

Speech (langage) is the ability of humans to communicate orally with other humans. Speech is a universal feature that characterizes all humans in the same way.

Language (langue) is a specific system of signs that allows humans to communicate with each other. Languages (e.g., English, Chinese, and Zulu) differ with each other in terms of their grammatical rules.

The differences between speech and language are the following: (a) speech is the inherent ability for communication, while language is the result of that ability; (b) speech is created in the brain, while language is developed through social coexistence; and (c) speech is a natural ability, while language is acquired through teaching.

Speaking (parole) is the information that emerges from the use of language; e.g., words we produce, read, or listen to. This information differs from speaker to speaker. The relationship between the three terms is depicted in the equations below:

$$\begin{aligned} \textit{speaking} &= \textit{language} + \textit{speech} \\ \textit{speech} &= \textit{language} - \textit{speaking} \end{aligned}$$

1.2. Language as a communication code

Some essential elements are necessary to achieve communication. These elements constitute the *communication model* that is illustrated in the figure below:

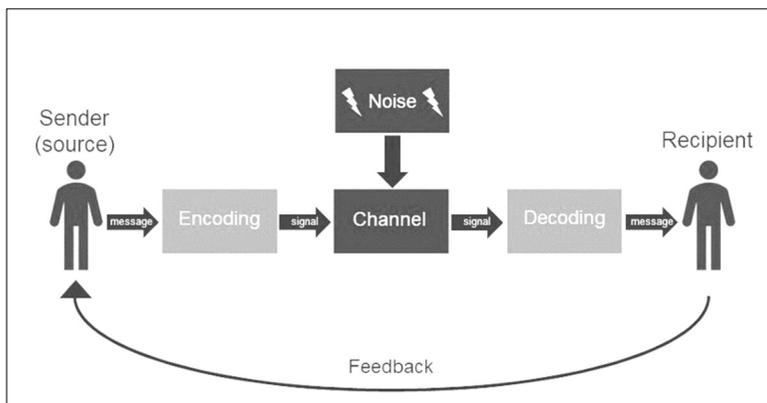


Fig. 1-1: Communication cycle model by Claude Elwood Shannon and Warren Weaver

According to the model, there is a *sender* (or a speaker in the case of language) who provides a message to a *recipient* (or a listener) through a *channel*. Nevertheless, successful communication does not only involve sending a message. The message should be in a recognizable form and have a specific meaning to be understood. Therefore, a *code* is needed for the speaker to encode the signal and the listener to decode it. The choice for the combination of form and meaning is called *encoding*, while the recognition of this combination is called *decoding*. So, we could say that code is a finite (or limited) system, which allows us to combine forms and meanings in order for the message to be understood by someone who knows the same code. Such a system consists of a sum of units and a sum of rules. In particular, the language code consists of *linguistic signs* (or just *words*) and *rules* for the use of linguistic signs in order for the listeners to be able to decode the signals.

Due to differences in each person's experiences and the knowledge, both speakers and listeners do not know precisely the same code. For example, the word "car" might recall a beautiful thought for a racing car driver who loves cars or a terrible thought for someone who has experienced a serious car crash. Therefore, as a consequence, the message of the speaker and that of the listener will be different. However, there is a constant meaning core, which is the same for all speakers; due to this core, excellent communication can be still achieved.

During communication, a *channel* is a medium for the transfer of a message. In the case of language, it can be mainly transferred through

oral or written form. Finally, as we can see from the figure above, *noise* might interfere with the reception of the signal (e.g., noise from traffic).

1.3. The Sign

To examine how signs work, we have to consider the general theory of signs, which is called *semiotics*. Semiotics is based on the theory of the famous Swiss linguist, *Ferdinand de Saussure* (1857–1913). A sign, which might be an object, a sound, or a word, etc., does not have a specific meaning unless we attribute one to it. For instance, in traffic lights, the green color does not mean in itself that “vehicles have to proceed”, but it is a meaning that was predetermined by humans to follow a common highway code. So, a sign has two components: (a) *a material form*: something that we can see, hear, touch, smell, or taste; and (b) *a concept*: a comprehensible construction which is linked with experience. The material form is called the *signifier*, while the concept is called the *signified*.

1.3.1. Types of Signs

There are three types of signs: (a) icons, (b) indices (or indexes), and (c) symbols. In *icons*, the signifier is similar to the signified; they recognizably look, sound, feel, taste, or smell the same as the signified (the signifier resembles whatever it depicts). For instance, a picture of one’s face is an icon of them. In *indices*, the signifier is directly linked in some way with the signified. For example, dark clouds are an index of impending rain. In *symbols*, the signifier is not similar to the signified. For example, the connection of a red traffic light with the instruction to stop is just arbitrary (there is not any logical relationship between the signifier and the signified; someone has decided this relationship).

1.3.2. The Linguistic Sign

As we have seen in a previous section, *words* are also called *linguistic signs*. Linguistic signs connect a sound-image (since words can be only seen or heard) with a concept. The sound-image is the *signifier*, while the concept is the *signified*. Let us take the example of the word “tree”. The word’s signified refers to the concept of “a plant that consists of a stem and branches” (this is how we define “tree” in general terms). The signifier of the word is a row of segments that gives the sound-image /t/ /r/ /i:/. These two components are illustrated in the figure below:

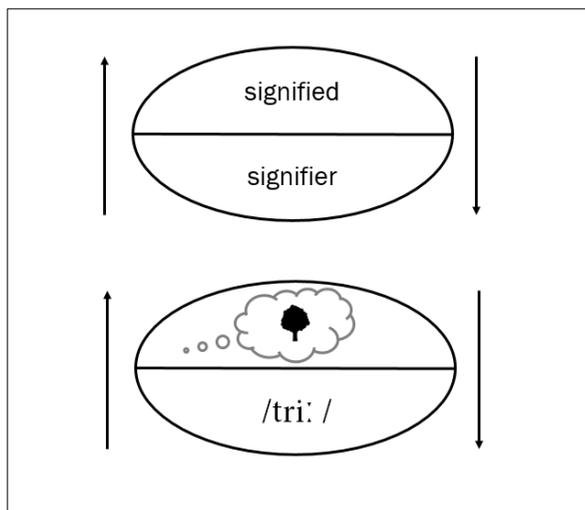


Fig. 1-2: The signified and the signifier of the word “tree”

1.3.3. Characteristics of the linguistic sign

One of the most important characteristics of the linguistic sign is its *double entity*. According to Saussure, both signified and signifier cannot be separated; that is, a sign cannot exist with the absence of one of these components. Usually, they are described as the content and the form of the sign, respectively. Another characteristic of the linguistic sign is its *conventionality*. The relationship between signified and signifier is arbitrary since there is no logical connection between a specific concept and a sound-image and, thus, this relationship cannot be explained. Such a connection exists because someone has agreed or decided to connect the two components. Furthermore, *linearity* is another feature of the linguistic sign. That is, the signified consists of segments that are pronounced consecutively and, therefore, are understood as a sequence of elements that form a speech “chain”. Finally, the combination of the linguistic sign components is *unique* as there are no words that have the same signified and signifier. However, some words share the same signifier but not the same signified. Namely, they are heard in the same way but have a different meaning: these words are called *homophones* or *homographs* [e.g., “ate-eight” or “bar (a business that sells alcoholic drinks)-bar” (a cuboid piece of any solid commodity)]. However, some words share the same signified but have a different signifier: e.g., “autumn-fall” semester

(different sound-image but identical meaning: they both mean the academic semester that usually starts in September/October).

1.4. Features of Language

People often use many communication codes, such as traffic signs, mathematical symbols, and music notes, etc. However, language code is a unique and a more sophisticated communication system compared to other ones. It consists of particular features that are common in all languages and are essential to understand how the language system works.

1.4.1. Arbitrariness

Arbitrariness is one of the most important characteristics of languages. Words and rules are not a result of a general principle or a “natural” process but, instead, someone agreed with someone else (these people are still unknown to us) to match of a particular word with a certain concept. For example, there is no explanation for why a “house” is called /haʊs/ rather than something else. Arbitrariness is also related to the rules of a language: there is not a sensible justification for why we say “how are you” instead of “how is you”. Furthermore, by looking at words in several languages, we can see that each one has a different combination of form and meaning (e.g., eye: English: /aɪ/, Spanish: /'oxo/ Greek: /'mati/) as well as different grammatical rules:

<i>Mia</i>	<i>ómorfi</i>	<i>méra</i>
Article	Adjective	Noun
<i>Um</i>	<i>dia</i>	<i>bonito</i>
Article	Noun	Adjective

The sentences above come from Greek and Portuguese, respectively. Both of them mean “a beautiful day”; however, the position of the adjective that describes the noun differs in each case. So, the rules of each language are not determined in the same manner, and there is no logical explanation for this. Another important fact is that arbitrariness is common among speakers. An individual speaker cannot change the signifier or the signified of a word because making such a change would lead to difficulties when communicating with other speakers. Moreover, even if a speaker links, for example, the word “child” with a prudent child, while

another speaker links it with an imprudent child, they will still be able to communicate with each other since the general idea of “young being” remains common between them.

Ferdinand de Saussure divided arbitrariness into two subcategories: the *absolute* and the *relative*. For example, the number “twenty” is absolutely arbitrary, whereas “twenty-one” is relatively arbitrary; the former cannot be analyzed and, hence, cannot be associated with other words, whereas the latter is derived from the word “twenty”.

Moreover, some people assume that the paradigms of *onomatopoeic* words (words that supposed to imitate sounds) and interjections prove that Saussure’s theory about the arbitrariness of the linguistic sign is incorrect. However, if we look at different languages, the same dog barks using different sound continua:

English	‘woof’
Japanese	‘wan’
Greek	‘γav’
Icelandic	‘voff’
Persian	‘haap’

Of course, the dog does not bark differently. It is the way that this barking is perceived by the speakers of each language and the historical context in which the words have been developed that form these differences. So, we cannot talk about words that were created due to the imitation of a sound. If that were true, we would have identical or similar words for the sound of barking in all languages; indeed, the English “woof” and the Persian “haap” differ significantly acoustically.

1.4.2. Variability

A linguistic sign does not remain constant over time as linguistic conventions change. If we refer to older forms of language, we can find examples in which a linguistic sign has altered its signifier. In Ancient Greek, the word water was “ὕδωρ” /iðor/ and, during the Middle Ages, the word became “νερό” /ne’ro/. Furthermore, some words may acquire additional meaning due, for instance, to technological advances: e.g., “run” = 1. to move rapidly, 2. to start a computer program; or “window” = 1. an opening in the wall, 2. a display rectangle in a computer program.

New words may emerge from the processes of derivation and composition (or compounding). In derivation, a new word is created with

the addition of an affix in an existing word: “*dis-infect*”, “*taxabil-ity*”. In composition, two words are linked to create a new word: “*inter-net*”, “*ear-phones*”. Also, the phenomenon of word borrowing can add more words to a language’s vocabulary. For example, in the 20th century, English borrowed “*paparazzi*” from Italian, “*Tamagotchi*” from Japanese, “*pogrom*” from Russian, and “*Taoiseach*” from Irish-Gaelic.

Linguistic variability is not only related to synchronies found in different periods but also to the same synchronies in different geographical locations. For instance, the Greek language spoken in Cyprus (Cypriot Greek) differs in terms of phonology, morphology, syntax, and semantics from the Greek language spoken on the mainland. The same applies to different varieties of Arabic (e.g., Levantine Arabic, and Moroccan Arabic). Even in a single country, there are other varieties (dialects or idioms) that differ from the standard language.

In sum, linguistic variability is a natural phenomenon in languages, which signifies their evolution. Linguistic signs are used in everyday communication by billions of speakers and, consequently, they do change.

1.4.3. Double articulation, language economy, and productivity

Another essential feature of language is its *double articulation*. Two central units contribute to the creation of a linguistic message. The first unit is the *first articulation* that contains meaningful units with lexical (e.g., “*door*”, “*play*”) or grammatical (e.g., “*door-s*”, “*play-ed*”) values. The second unit is the *second articulation* that contains meaningless individual units, phonemes, which, upon combination, form the first articulation units (e.g., /d/-/ɔ:/-/r/).

One of the advantages of the double articulation is that it allows us to create an infinite number of first articulation units (e.g., lexemes, morphemes) by only using a finite number of units (e.g., phonemes) that are found in the second articulation. Therefore, even though phonemes can rarely exceed the number of 50 in a language, we can create infinite words and sentences, if we combine them. This feature is called the *economy of language*. Economy of language is related to arbitrariness in which a linguistic sign has two independent and arbitrarily related components: the signifier and the signified.

A consequence of language economy is language *productivity*: the ability of humans to create and understand sentences. According to Noam Chomsky, humans can create an infinite number of sentences. Even

children, who do not receive much linguistic input in their native language, can create a great number of sentences that they have never heard before.

1.4.4. Language Universals

Every child has the ability to produce language forms from a very young age. These forms might be ungrammatical or incomplete at the early stages of language development while they are being refined as the child learns the distinct lexicon and structure of their native language. According to Chomsky, despite the external differences between the world's languages, there are more profound similarities for native speakers on a cognitive-psychological basis that begin to take shape when children listen to the input provided by their parents. Parents' speech does not help children learn the language, but it does help the biological process of language learning to begin. Thus, there are *language universals* or a universal grammar that allows children to acquire a particular language.

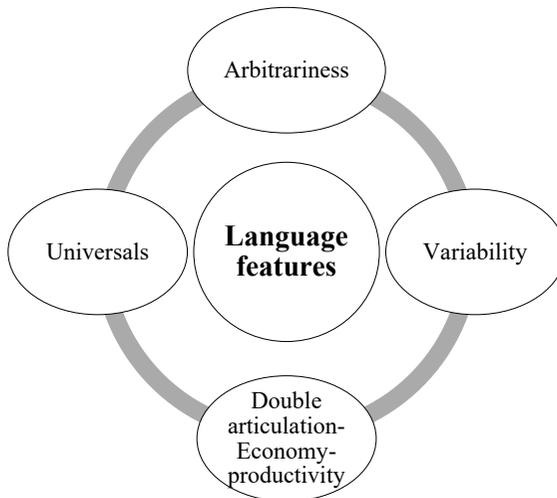


Fig. 1-3: The four features of language

There are three main categories of language universals: (a) the *substantive*, which contains categories that are necessary for the analysis of a human language (e.g., noun, verb, and number); (b) the *formal*, which includes abstract orders for linguistic analysis (e.g., the type of rules that have to be specified for linguistic analysis); and (c) the *implicational*,

which contains features that can be formulated as relationships (e.g., if there is subject-noun agreement with respect to gender, there will be also an adjective-noun agreement).

1.5. The science of Linguistics

Linguistics is the scientific investigation of language as a universal phenomenon as well as the investigation of individual languages. It aims to answer questions about the nature of language, its structure, and its relationship with the human mind and society. Also, through the examination of linguistic structures (i.e., the ways in which expression is organized and meaning is formed for communication), it aims to examine the possibility of the formulation of standard rules for all languages. In order to achieve the latter, it employs empirical data, establishes linguistic principles, and uses appropriate methodological tools. It is important to add that linguists focus on the description and interpretation of linguistic phenomena without intervening, evaluating, or having a prescriptive view; linguistics is a *descriptive* rather than a *prescriptive* science.

Today, the scientific study of language might take place in an interdisciplinary manner; namely, linguistics borrows theories and methodological tools from other sciences, such as psychology (psycholinguistics), philosophy (philosophy of language), anthropology (linguistic anthropology), philology (orthography), law (forensic linguistics), medicine (clinical linguistics), biology (cognitive linguistics), sociology (sociolinguistics), informatics (computational linguistics), and mathematics (quantitative linguistics). Also, the science of linguistics is often divided into *general/theoretical linguistics*, which studies the structure and the functions of language to formulate theories, and *applied linguistics*, which applies general/theoretical linguistic theories to provide solutions to real-life problems.

1.5.1. Linguistics and Philology

Many sciences, such as *linguistics* and *philology*, deal with language issues; nevertheless, this does not imply that they all follow the same approach. Linguistics is interested in the examination of grammatical system functions using a descriptive approach. In contrast, philology mostly follows a prescriptive approach, indicating “correct” or “wrong” forms of either an oral or written language.

Students come into contact with traditional grammar textbooks from a very young age during school courses. These grammar books

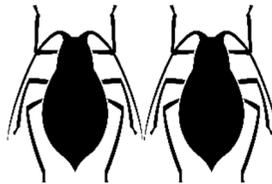
indicate which forms are “correct” or “incorrect”; “incorrect” forms are those which do not obey the rules of formal grammar. However, this approach is problematic since it only gives a minimal overview of the language system and fails to consider the evolution of language and the mechanisms that create several linguistic forms. So, misinterpretations about how the language system works are rational since all language forms spin around the “correct”/“incorrect” relationship. In conclusion, there is a clear difference between linguistics and traditional school grammar: the latter refers to *what should be said*, whereas the former analyses and records *what actually was said*.

1.6. Levels of Linguistic Analysis

Every language in the world consists of rules; meanings are formed when words are linked to each other in a predetermined way. Therefore, language does not work randomly. In the pictures below, we can see an insect called a “santle”:



Picture 1



Picture 2

-Picture 1 presents a santle.

-Picture 2 presents _____.

Ten English native speakers were asked what Picture 2 represented. All speakers said that “Picture 2 represents ‘santles’”. This response indicates that speakers know the rules of their native language. Of course, there is no insect called “santle” and this word does not even exist in the English vocabulary; however, the speakers employed a well-known rule in English: plural nouns are formed by the addition of an –s ending. This rule has never been taught to the native English speakers. In general, native speakers are often not aware of the grammatical rules in their mother tongue. For instance, it does not take any time for them to think about how to form a verb in the past tense when they want to talk about something that happened in the past.

In linguistics, grammar is a sum of rules that native speakers acquire subconsciously, which leads to the formation of particular

sentences. Also, grammar is the description and presentation of language rules by researchers. It should not be confused with traditional grammar taught at schools, which has a prescriptive character. It is important to remember that linguistics does not criticize any linguistic form; instead, it explains why one linguistic form is preferred instead of another. Language can be divided into five—or six for some researchers—main levels: *phonetics/phonology, morphology, syntax, semantics, and pragmatics*.

1.6.1. Phonetics

Phonetics [from Greek “foni” (=voice)] deals with the scientific investigation of all sounds of human speech and communication. In particular, it investigates the functions of *phones* (i.e., any speech sound/gesture regardless of whether it can change the meaning of words; see “phonemes” in Section 1.6.2. for comparison) and non-language sounds. Phonetics has three main subfields which study the following: (a) speech sound production (*articulatory phonetics*), (b) speech sound perception (*auditory phonetics*), and (c) acoustic aspects of speech sounds (*acoustic phonetics*).

Pronunciation is the ability of humans to produce sounds. However, humans can produce more sounds than those found in a specific language or languages. It is important to say that each human pronounces these sounds instinctively; nevertheless, most of the time, there is mutual intelligibility among native speakers of a particular language.

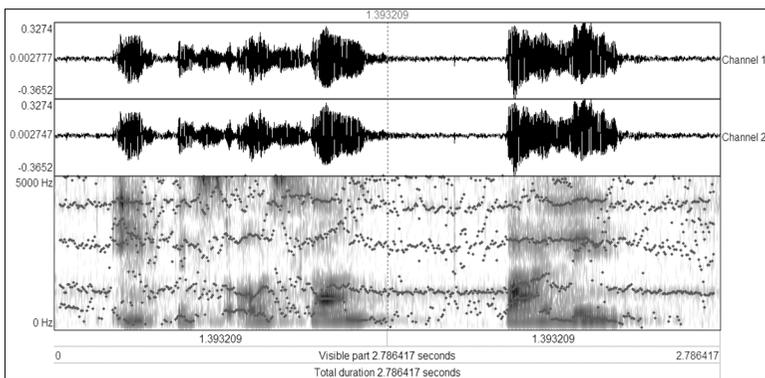


Fig. 1-4: Waveform and Spectrogram in a Praat script for the analysis of acoustic characteristics of sounds (e.g., formant frequencies and duration)

1.6.1.1. Pronunciation and Orthography

There is a complicated relationship between *pronunciation* and *orthography*. The latter rarely depicts the real pronunciation of sounds because the written word does not evolve as rapidly as the spoken word. Furthermore, the designers of the various writing systems did not consider the representation of pronunciation in the written word to a great extent. In the case of English, a grapheme (the written representation of a sound: i.e., a letter) might depict the actual pronunciation of a sound (e.g., “m” as /m/), or it might correspond to two or more different sounds [e.g., “a” as /æ/ (cat) or /ɑ:/ (star)]. To solve problems surrounding the written representation of sounds, linguists have created *phonetic writing* in which each sound is represented by a single phonetic symbol. The *International Phonetic Alphabet* (IPA) was created in 1888, and it consists of 107 letters mostly deriving from Latin (e.g., [p], [t]) and Greek (e.g., [θ], [ε]), or modifications thereof. Also, there are 52 diacritics and 4 prosodic marks. IPA is a vital tool for scientists since it helps depict the exact pronunciation of a word, and it simplifies the spelling of words which are written in alphabets that are “difficult” for many people, such as Chinese and Arabic: e.g., Chinese 树 (=tree) → /ʃu/ (shù).

International Phonetic Alphabet (IPA) ,ɪntəˈnæʃnəl fəˈnetɪk ˈælfəbet											
Consonants (pulmonic)											
	Bilabial	Labio-dental	Dental	Alveolar	Post-alveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k ɡ	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Trill	ʙ			ʀ					ʀ		
Tap or flap		ⱱ		ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateral fricative				ɬ ɮ							
Approximant		ʋ		ɹ		ɻ	j	ɰ			
Lateral approximant				l		ɭ	ʎ	ʟ			

Fig. 1-5: Example of the International Phonetic Alphabet (Omniglot)

1.6.2. Phonology

Phonology [from Greek “fóni” (=voice) and “lóγος” (=speech)] deals with the study of the *phonemes* of a particular language at levels beneath a word (e.g., syllable, onset, articulatory gestures, articulatory features, and mora) or at any level in which the structure of a sound transmits linguistic meaning. Phonemes (opposed to phones) are sounds that, if swapped with other sounds, change the meaning of the word: e.g., /θʌm/ (thumb) vs. /dʌm/ (dumb). Words that are differentiated in only one element (here, phonemes) are called *minimal pairs*. In general, phonology examines how sounds function in a particular language.

There is a big difference between phonetics and phonology. Let us take some examples from English words to understand this difference. Two speakers can pronounce the English word “pin” as [pʰɪn] and [pɪn]; in the first case, the /p/ is aspirated, whereas in the second it is plain. However, in this example, we do not have two different linguistic signs (words) but the same sign. In fact, [pʰ] and [p] are two different sounds of an abstract unit called phoneme (here, /p/). Phonemes cannot be heard; they are units created for linguistic analysis. The sounds [pʰ] and [p], which could be roughly described as “subcategories” of the phoneme /p/, are called *allophones*. Allophones cannot form minimal pairs since they depict different representations of the same phoneme. In contrast, two phonemes might form minimal pairs. For example, the phonemes /p/ and /d/ create the minimal pair /pɪn/ - /dɪn/, which consists of two different words.

In sum, phonetics takes all the characteristics of sounds (even smaller ones that are not perceivable through hearing) into account in order to understand how they are articulated, while phonology only considers characteristics that have a clear importance. The phonological transliteration of a word is included in slashes (/word/), while the phonetic transliteration is included in brackets ([word]).

1.6.3. Morphology

Morphology [from Greek “morfi” (=shape, form) + “lóγος” (= speech)] studies the different forms of words during speech. In particular, it studies conjugations and the ways that words are created, such as derivation and composition.

So, what forms are studied in morphology? As we saw before, words might have more than one meaning. Also, they might have different forms: for instance, the word “day” also has the form “days”, and the verb

“visit” also has the form “visited”. If we look for words in the dictionary, there is always a particular form for them, which is called a *lexical* form; in languages that mark grammatical cases, the lexical forms of adjectives and nouns are in the singular number and the nominative case.

Words can be analyzed in smaller units: “un-like”, “dog-s”, “receiv-ed”, and “do-ing”. These units are the smallest meaningful components of language, which are called *morphemes*. A morpheme can stand by itself and function independently as a word: e.g., “town”, “girl”, and “happy”. In this case, the morpheme is considered to be a *root*; it is a word without an embedded morpheme: e.g., “town-s” (“town”: root), “girl-ish” (“girl”: root), and “un-happy” (“happy”: root). A root might be a lexical unit (e.g., “love” as in “love-ly”) or not (e.g., “rupt” as in “corrupt”). The word “unbreakable” consists of three morphemes: “un-”, “break” (root) and “-able”. In the last example, the morphemes “un-” and “-able” cannot stand alone (they do not have a meaning by themselves) and, therefore, they have to be combined with roots in order to shape the meaning of a word. Morphemes that can stand alone are called *free* morphemes, whereas those that cannot stand alone are called *bound* morphemes.

Free morphemes can be divided into *lexical* and *functional* morphemes. Lexical morphemes are ordinary words (nouns, verbs, and adverbs) that carry information about the content of the message (e.g., “dog”, “house”, “work”, “invite”, “drink”, and “today”); these words can be easily used during communication. Functional morphemes are a language’s functional words, such as prepositions (e.g., “to”, “in”), conjunctions (e.g., “and”, “because”), articles (e.g., “the”, “a/an”), and pronouns (e.g., “it”, “theirs”) that modify the meaning of a word.

Subcategories of bound morphemes are *derivational* and *inflectional* morphemes. Derivational morphemes can change the semantic meaning of a word or the part of speech: e.g., “like” → “dis-like” (“like” ≠ “dislike”), “happy” → “happ-iness” (the adjective “happy” can become a noun, “happiness”). Inflectional morphemes modify a noun’s number, gender, and case or a verb’s number, person, mood, aspect, and tense: e.g., “boy” → “boy-s” (SING. → PLUR.); “wait” → “wait-ed” (present simple → past simple).

Root should not be confused with *stem*. As we saw before, a root is a morpheme by itself and cannot be separated into smaller meaningful parts. A stem consists minimally of a root, but it might take the form of a root plus derivational morphemes; inflectional morphemes may be added to a stem. For instance, in the word, “dislike”, “like” is a root and a stem

together, while “dislike” is a stem, which might take the forms “dislike-s” or “dislike-d”, etc.

The morphemes that can be attached to a word stem to create new words or word forms are named *affixes*. The three most common affixes are *prefixes*, *suffixes*, and *infixes*. Prefixes are placed before the stem of the word: e.g., “*dis*-appear”, “*inter*-language”, and “*down*-town”. Suffixes are placed after the word stem: e.g., “mov-*able*”, “like-*ly*”, and “help-*ful*”. Infixes are rare in several languages, including English, and they are placed inside the word stem: e.g., “cup-*s*-ful”, “narc-*o*-lepsy”, and “passer-*s*-by”.

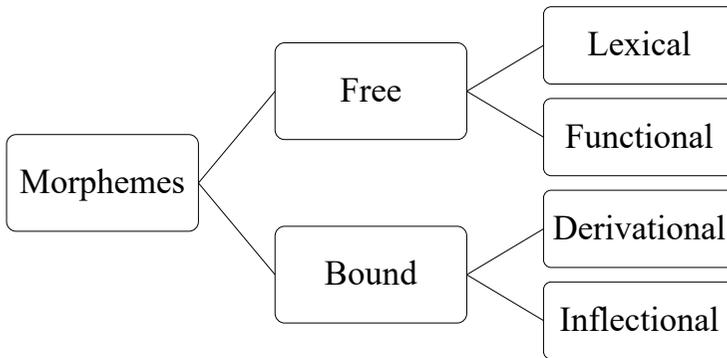


Fig. 1-6: Main subdivisions of morphemes

1.6.3.1. Allomorph

Allomorph is one of the main terms that someone might come across when studying morphology. It refers to the different forms of a morpheme that do not differentiate meaning, and they are used as supplements (i.e., the presence of one of them excludes the presence of another). Often, it is easy to predict which allomorph to use because there are particular rules. What are the plural allomorphs of the following English words: “bus” (/bʌs/), “cat” (/kæt/), and “pen” (/pen/)? The answer is “bus-*es*” (/bʌs-əz/) because “bus” ends in /s/, “cat-*s*” (/kæt-s/) because “cat” ends in a voiceless sound, and “pen-*s*” (/pen-z/) because “pen” ends in a voiced sound. Therefore, the plural in English has three allomorphs. In the above examples, the allomorphs depend on phonological processes that have particular rules. However, some allomorphs are predetermined and, therefore, do not depend on any rule. For instance, “child” becomes “children” in the plural

as a remnant of an older form of language, or the plural of some words—mostly loan-words from Latin and Greek—have kept their conjugation from their “original” language: e.g., “vertebra” → “vertebrae” and “criterion” → “criteria”.

1.6.3.2. Free variation, zero, and discontinued morpheme

It is important to note that some morphemes are found in *free variation*: that is, two or more morphemes can be used freely according to the speaker’s choice. For example, the plural of the word “curriculum” is either “curriculums” or “curricula”. Free variation exists because language systems evolve. Two variations might co-exist in a particular period, but one of them might displace the other and become dominant.

Changes in *zero morphemes* are not visible. For example, both the singular and the plural forms of the noun “sheep” are the same (“sheep” and “sheep”). Similarly, both the present simple and the past simple tenses of the verb “hit” are the same (“hit” and “hit”).

Plural	Past Tense
sheep + ∅	hit + ∅
tiger + s	visit + ed

A *discontinued morpheme* is a rare type in which a morpheme occupies two positions in a sentence. In the sentence, “*turn the light on*”, we can see that the preposition “on”, which is an essential constituent of the phrasal verb “turn on”, is located in another position, and not next to the form “turn”.

1.6.4. Syntax

While morphology deals with the structure of words, *syntax* [from Greek “sin” (= together) + “táksi” (= arrangement)] is interested in the relationship between words and the way they are combined to create larger meaningful units, such as phrases and sentences. In some theories, e.g., transformational grammar (Noam Chomsky), syntax is related to grammar. However, in other theories, it is seen as equal to other levels of linguistic analysis.

Syntaxeme refers to the minimal semantico-syntactical element of a language. We can detect syntaxemes according to their (a) categorical semantics in the world, (b) morphological form, and (c) function in a sentence.

The lecturer delivers a lecture

In the above sentence, we can replace “lecturer” with “student” or “professor” and still keep the meaning of the sentence. However, if we replace “lecturer” with “house” or “window”, the sentence becomes ungrammatical. This is because, in the second case, we have used words from different *categorical semantic classes* (specifically, we have used inanimate words). Categorical semantic class refers to words that share a semantic feature: e.g., inanimate nouns (including humans and animals, etc.) vs. animate nouns (including objects), static verbs (e.g., “know” and “like”) vs. dynamic verbs (e.g., “run” and “move”).

The mother brings her son to school.
Laziness leads the researcher to trouble.

In the above sentences, the two syntaxemes “to school” and “to trouble” have the same morphological form, but they differ in meaning. Thus, categorical semantic class is a non-morphological category because syntaxemes may have similar morphological forms but different meanings.

According to their function, syntaxemes can be (a) *free*: not depend on the context as in the case of news titles; (b) *conventional*: constitute a component of a sentence, e.g., as a grammatical subject (“Life is beautiful”), a predicate (“Life is beautiful”), or a modifier (“Life is beautiful for everyone”); and (c) *bound*: constitute components of a phrase (verb, noun, and adverb, etc.)

1.6.4.1. Lexical (syntactic) categories

In order to investigate the syntax of a sentence/phrase, we need to formulate rules for a sum of words that have similar syntactical behavior rather than for every single word.

The driver gave me a cigarette

In the above sentence, we can identify two words that have the same syntactical behavior, “the” and “a”. Specifically, they both belong to the same part of speech: they are articles. In linguistics, *lexical (or syntactic) categories* is preferred instead of the term, *part of speech*; these words do not just correspond to lexemes, but they also add grammatical value: e.g., they may indicate the number, the gender, or the case of a noun.

Therefore, the identification of lexical categories is important in syntax. It has to be added that lexical categories differ from language to language; for example, Russian and Sanskrit lack articles. Also, some languages do not make distinctions between lexical categories. For example, in Greek, nouns differ from verbs since the former have the characteristics of number, gender, and case, whereas the latter have the characteristics of number, mood, voice, aspect, and tense: e.g., “(i) ἀγάπη” (= the love); singular, feminine, and nominative vs. “ἀγαπῶ” (= I love); singular, indicative, active voice, continuous aspect, and present simple. Nevertheless, in English, the word “love” might refer to two different lexical categories (noun and verb).

The elements of lexical categories might have a *lexical* or a *grammatical* meaning. Lexical categories theoretically include infinite words that could be nouns, verbs, or adjectives, etc. Thus, these categories are called *open classes*. Lexical categories also include a number of words such as conjunctions and prepositions, etc., which accept new “members” infrequently; therefore, these categories are called *closed classes*. The words included in closed classes are not necessarily independent morphemes (e.g., “to” and “of”), especially for inflectional languages. For instance, in Spanish, gender is divided into masculine and feminine, and it appears only as a morpheme at the end of words: e.g., “un-*a* person-*a* hermos-*a*” (= a beautiful person). These grammatical features can also help us understand the relationship between the noun and the verb:

Student wita nauczyciela
Nauczyciel wita studenta

The examples above come from Polish, a highly inflectional language. The first sentence means “the student greets the teacher”. The nominative case of the subject “student” shows the subject (here, student) who acts (here, greets), while the accusative case of the word “nauczyciela” shows the object that the action goes to (here, the teacher). The second sentence means “the teacher greets the student”. So, here, the subject is the teacher who greets the student (object). To that end, grammatical cases are essential to understand “who” acts, and to “whom” or “what”.

Moreover, these categories have an impact on the semantics of words. For instance, in Greek, the category of aspect is crucial for the meaning of a word and, consequently, the meaning of a sentence. For instance, “ἔπε-*z*-*a*” (ἐπαιζα) means “I was playing”, while “ἔπε-*ks*-*a*” (ἐπαιξα) means “I played”. In this case, the differentiation in aspect

changes the tense of the verb; in the first case, it is past continuous, while in the second case, it is past simple.

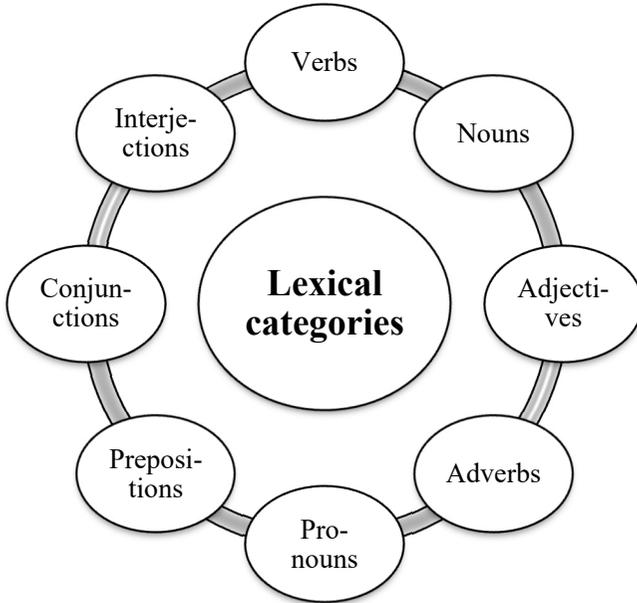


Fig. 1-7: The lexical categories of English

1.6.4.2. Phrase

A *phrase* is a group of words that works as a constituent in a sentence's syntax. A phrase is not a sentence because it is not a complete idea with a subject, a verb, and a predicate (e.g., “best friend” and “for twenty years”). Its function is to add additional meaning to the sentence. The majority of the phrases contain a keyword, which is the central constituent of the phrase. This keyword is the *head* of the phrase, and it might stand alone or accompanied by other elements: the *dependents*. In the phrase, “boiling hot water”, the head is the noun “water”. According to the distribution and their syntactical function, we can place a phrase in one of the following categories:

- *Noun phrases* (NF): their head is a noun, and they function as nouns, e.g., “*This house* is located in a small village.”