

Day 3-In Search of the Orthographic Word

Lecture 1

Orthography Seminar given at Bibles International, Grand Rapids, MI, July 6-9, 2010

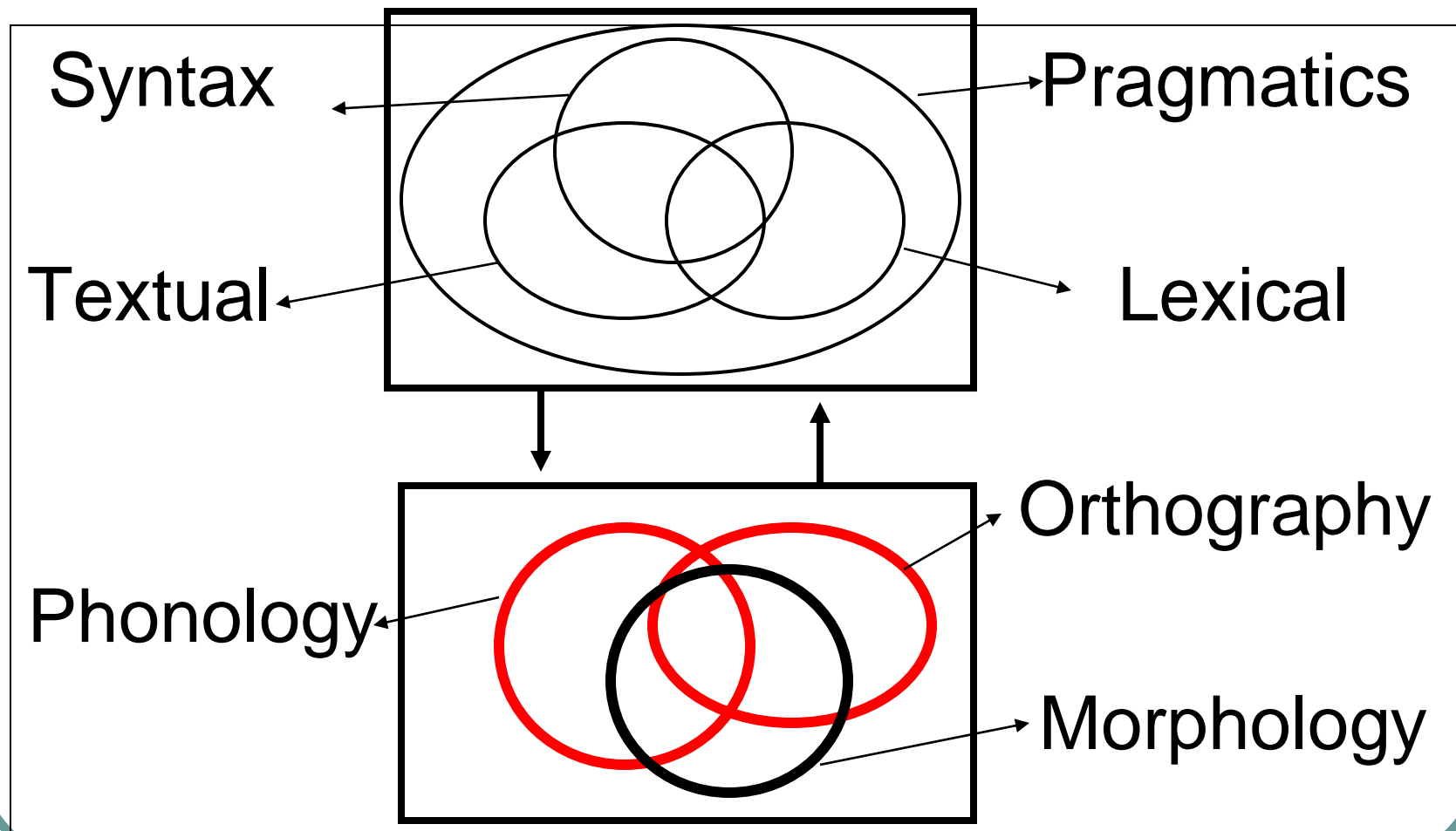
Ettien N. Koffi, Ph.D.
Professor Linguistics
Saint Cloud State University,
Saint Cloud, MN, USA

Lecture Outline

1. A Review
2. Morphological Typologies
3. Phonetic « Word »
4. Phonological «Word »
5. Morphological « Word »
6. Word Formation Processes
7. Orthographic «Word »

A Review of the Bottom-up Processor

Interactive Model of Reading Comprehension



Word Superiority Effect-WSE-1

- **The Maximal Projection Unit of Reading:** The minimal unit of reading is the grapheme. They organize into a larger unit called the syllable. Syllables organize yet in a larger unit called the “orthographic word”. This is the unit that all readers rely on to read. This has been called the “**Word Superiority Effect**” or the “**Word Advantage**”.

Word Superiority Effect-WSE-2

- **“Minimalist Hypothesis”** i.e., WSE, claims that the whole orthographic word has to be seen before reading can proceed. The claim that the whole word must be seen is also referred to as the **“Holistic Word Recognition”**: **Exterior letters** (the first last letters) of words have been found to be particularly useful in reading.

Morphological Typology of World Languages

Source: *Language Universals and Linguistic Typology* by Bernard Comrie. 1981, pp. 39-44

Isolating Languages

- **Definition:** “An Isolating language is one which has **no morphology**, i.e., at least ideally, a language where **there is one to one correspondence between words and morphemes**. Each word is invariable, there being no morphological variation, for instance no tense.”
- **Example:** Vietnamese

Agglutinating Languages

- **Definition:** “A word may consist of more than one morpheme,; moreover, a given morpheme h**but the boundaries between morphemes in the words are always clear-cut** as at least a reasonably invariant shape, so that the identification of morphemes in terms of their phonetic shape is straightforward.”
- **Example:** Turkish


Fusional Languages

- **Definition:** “In a fusional language, however, **there is no such clear-cut boundary between morphemes**, the characteristic of a fusional language being that the expression of different categories within the same word is fused together to give a single, unsegmentable morph.”

Polysynthetic Languages: A Little of All

- **Fuzzy Boundaries:** This classification is more theoretical than practical. All three typologies are found in English. Most pronouns give the impression that English is an isolating language. The word **<unforgettable>** gives the impression that English is an agglutinating language while the ...

Typology and the Orthographic Word

- ...suppletive words such as **<were>** denote a fusional characteristic.
-  **It is all a matter of degrees.** These typologies have implications for decisions about what constitutes an orthographic word in a given language.

Phonetic Word

- “The possibility of pausing thus supplies a certain kind of support for the division of a sentence into words,”

Anderson (1987:151)


- **Pause Criterion**

“Any segment of a sentence bound by two successive points at which pausing is possible is a word,” Koffi (1990:153)


Phonological Word-1

- In some African languages, the phonology of the language gives some strong clues as to what constitutes “**a phonological word**,” i.e., a stretch of utterance in which [+ATR] (tense vowel) or [-ATR] (lax vowel) vowel harmony applies is a phonological word.

Phonological Word-2

-  This is not a strong criterion because vowel harmony can spread across word boundaries.


Orthographic Word

- “In general, the layman looks to writing and classes as a word whatever he finds written between two successive spaces,”
Hockett (1975:166)
-  This is useful only for languages that already have a writing system.

Definitions and Classifications of Morphology

Structural Morphology vs.
Functional Morphology

Definition of Morphology

- **Definition of “morphology” and “morpheme”:** Morphology is the branch of linguistics that studies the **structure of words** and the various processes that a language uses to **create new words**.
-  The latter is very important in terminology management and lexical modernization..

Definition of the Morpheme


- **Definition of “morpheme”:** The morpheme is the smallest linguistic unit that has:
 1. a meaning
 2. and/or a function.

Focus on Structural Morphology

Structural Morphology

1. It studies the structure of words to see if they are free morphemes or if they consist roots and bound morphemes
2. It examines words to see the kinds of bound morphemes (affixes) that they have.

Free Morphemes

- **Free Morphemes:** They are sometimes called “lexemes.” They have the following characteristics:
 1. They can occur by themselves
 2. They have independent dictionary meanings
-  They are often represented in writing as independent orthographic words

Bound Morphemes

1. **Bound morphemes/Affixes:** Bound morphemes are morphemes that cannot stand by themselves. For instance, we cannot have the morphemes <-s>, <ed>, <-ing>, <-ly>, <un->, <sub->, <con->, etc. by themselves. They must always be attached to a root. Such morphemes are called “bound” morphemes.

Bound Morphemes-1

- **Bound Morphemes:** Bound morphemes are morphemes that cannot occur by themselves. There are **5 types** of bound morphemes in world languages:
 1. **Prefixes:** They attach themselves to the left of the root as in <**im**possible>

Bound Morphemes-2

2. **Suffixes:** They attaches themselves to the end of the root or other suffixes as in <helpfulness>
3. **Infixes:** They insert themselves inside of words. There is no polite infix in English!

Classification and Definition of Affixes

4. **Interfixes:** usually a vowel inserted between two roots to make the pronunciation easier as in <infant-**i**-cide>, <dorm-**it**-ory>, etc.
5. **Circumfixes/discontinuous morphemes:** a group of morphemes that attach themselves to different parts of a word or a syntactic structure as in : <Have-En>, <Be-EN>, <Be-ING> according Kaplan (1995)

Focus on Functional Morphology

Inflectional Morphology vs.
Derivational Morphology

Functional Morphology

- Functional morphology studies the function(s) and/meaning(s) of morphemes. It is divided into two categories:
 1. Inflectional morphology
 2. Derivational morphology

Inflectional Morphology

- **Definition:** Inflectional morphemes are morphemes that add **grammatical information** to the root. There are only 10 of them in English:

N0.	Morpheme	Meaning	Examples
1.	<-s>	Present tense	eats
2.	<-s>	Plural	girls
3.	<-’s>	Genitive	The boy’s

Table of Inflectional Suffixes

NO.	Morpheme	Meaning	Examples
4.	<-ed>	Past tense	learned
5.	<-en>	Irregular participle	eaten
6.	<-ing>	Present participle	reading
7.	<-er>	Comparative Adj.	bigger
8.	<-est>	Superlative Adj.	biggest
9.	<-er>	Comparative Adv.	faster
10.	<-est>	Superlative Adv.	fastest

Inflectional Processes-1




1. **Regular Inflections:** these are morphemes that can be added to words in a predictable and systematic fashion: <bible**s**>
2. **Irregular inflections:** The irregularity may be due to affixation or to internal vowel changes as <child**ren**>, <ox**en**> or <ma**n**> vs. <me**n**>

Inflectional Processes-2

1. **Zero forms:** Irregular inflection can be indicated by a zero morpheme added as in **<one deer>** vs. **<two deer>**.
- **Suppletion:** Extreme cases of irregularity so that the inflected forms have no phonetic similarity with the original root as in **<Be>** and its various conjugated forms: **<am, are, is, was...>**

Focus on Derivational Morphology

Derivational Morphology

- **Definition:** A derivational morpheme is a bound morpheme (affix) that helps create a new word out of an existing root by changing the part of speech of the original root. The derivation can be done through
 1. **Suffixation** as in <create>  <creation>
 2. **Prefixation** as <little>  <belittle>
-  English has over 60 derivational suffixes but only a handful of derivational prefixes.

Complex Derivational Processes

- Languages may have very complex derivational processes consisting of multiple prefixation, suffixation, infixation, interfixation, or circumfixation:
- **Derived Words = (Prefix) + Root + (Suffix)**

 <en+light+en+ment+s>

Semi-Affixes

Definition

- “...Those elements that stand midway between words and suffixes. Some of them are only used as suffixes though their word character is still recognizable.”

Marchand (1969:356) quoted in Koffi (1990:164) Examples:

1. <-hood> as in <boyhood>
2. <-wise> as in <clockwise>

Focus on Morphophonology

Definition

- **Morphophonology:** Several phenomena happen at the intersection of phonology and morphology. It is the area of linguistics that studies:
 1. Changes that happen to bound morphemes as a result of affixation
 2. Changes that happen to the root as a result of affixation.

Suffixation-1

- **Data Set 1**

<He/she talk**s**>

<He/she finish**es**>

<He/she bug**s**>

- **Data Set 2**

<He/she talk**ed**>

<He/she treat**ed**>

<He/she bug**ged**>

-  It has been claimed that one cannot read English accurately without knowing this morphophonological rule.


Suffixation-2

- **Spirantalization Data Set 1**

<electric**c**> vs. <electric**ic**ity>

<president**t**> vs. <president**ial**>

<dent**t**> vs. <dent**ure**>


-  It has been claimed that one cannot read English accurately without knowing this morphophonological rule.

Suffixation-3

- **Data Set on the Trisyllabic Laxing Rule:**

<se**e**cret> vs. <se**e**cretary>

<bi**i**ble> vs. <bi**i**blical>

 It has been claimed that one cannot read English accurately without knowing this morphophonological rule.

Suffixation-4

- **Spirantalization Data Set 1**

<offen**d**> vs. <offens**ive**>

<comprehend**d**> vs. <comprehens**ive**>


<attend**d**> vs. <attent**ive**>

<affir**m**> vs. <affirmat**ive**>

<dismiss**ss**> vs. <dismiss**ive**>

<decla**re**> vs. <declarat**ive**>

Morphophonological Principles in English Orthography-1

-  **Suffixation Principle:** It seems that the designers of the orthography expected readers to use their knowledge of predictable allomorphic rules when reading **suffixes**. When the rules were **not sufficiently predictable**, they helped the reader by representing the morphophonological rules explicitly in the orthography.

Prefixation-1

- **Data Set 1**

<possible> vs. <impossible>

<legal> vs. <illegal>

<responsible> vs. <irresponsible>


<tolerant> vs. <intolerant>

<conceivable> vs. <inconceivable>

This same process apply so the prefix

<con->

Morphophonological Principles in English Orthography-1

-  **Prefixation Principle:** It seems that the designers of the orthography systematically reproduced morphophonological changes orthographically if the prefixation involved changes in place of articulation and/or manner of articulation.

Focus on Word Formation Processes


Table Summary of WFP-1

NO	Name	Examples
1.	Derivational Process	create +ion
2.	Compounding	high + way
3.	Blending	information + commercial = infomercial
4.	Clipping	gymnastics = gym
5.	Coining	Kleenex, xerox, google

Table Summary of WFP-2

NO	Name	Examples
6.	Back-formation / conversion	a party = to party
7.	Borrowing	déjà-vu
8.	Abbreviation	CIA
9.	Reduplication	willy-nilly
10.	Proverbial Words	a johnny-come lately

Note about WFPs-1

- **Exceptions:** The word formation processes of blending, clipping and abbreviation do not seem to exist in Anyi and the dozen or so West African languages that I have worked on. It seems that these WFPs come as a result of a long literacy tradition.
-  Do you have examples to the contrary?

Reduplication

Reduplication and the Orthographic Word-1

- **Reduplication:** It is by far the most common word formation process in many African languages. How should it be represented in the orthography?
- **Anyi example:** <cinman>: to walk
<cinmancinman+lɛ>: loitering
<cinman> <ciman+lɛ>
<cinman-cinman+lɛ>

Reduplication and the Orthographic Word-2

- **Orthographic Decision-Making:** Should a reduplicated word be written as a single orthographic word, a hyphenated orthographic word, or two separate orthographic words?
- **Anyi Solution:** Reduplicated words are written as a single orthographic word. What solution have you adopted?

The Universal Word Formation Apparatus

Word Formation, Spell-Out Rules and the Lexicon

List of Morphemes	Word Formation Rules	Readjustment Rules	Spell-Out Rules	Dictionary
1. friend				
2. Boy hood				
3. recite al			[+idiosyncrasies]	
4. ignore ation			X [-L1]	
5. mountain al	X			

Universality of the Derivational Apparatus

- Presumably Scalise's derivational apparatus is universal. All languages have a similar mechanism that allows well-formed words into the lexicon or blocks ill-formed words from entering the lexicon. In written languages, these rules are stated explicitly in the form of orthographic (spelling) rules.
- The operation of the morphological component has been slightly revised to include Readjustment Rules, Scalise (1986:31, 63)

Explanation of the Apparatus-1

- **Free Morphemes and the Lexicon/ Dictionary:**
 1. The word **<friend>** is a free morpheme, that is, a word that is indivisible into parts. It is not formed by any word formation rule. Therefore, it goes directly into the lexicon.
 2. The word **<boyhood>**, on the other hand, is a morphologically derived word. It is composed of the free morpheme **<boy>** and the suffix **<-hood>**. The combination goes directly into the lexicon without modification.

Explanation of the Apparatus-2

- **Affixation and the Lexicon/ Dictionary:**
 1. The word **<recital>** which is formed by adding the suffix **<-al>** to the root **<recite>** as the combination **<*reciteal>** is not a real word.
 2. To make its way into the dictionary, it has to transit through the **Spell-Out Rules**. These are spelling rules to insure that the word enters into the lexicon only if it conforms to the spelling conventions of the language.

Explanation of the Apparatus-3

- **Lexical Gap and the Lexicon/ Dictionary:**
 1. The word <*ignoration> is blocked from entering the lexicon because <ignore> cannot accept the suffix {-ation}.
 2. This suffix is very productive in English. It is therefore hard to explain why <adoration> which is derived from <adore> is a well-formed word but <*ignoration> cannot be derived from <ignore>.
 3. This is an example of **lexical gap**: a word that could be but has not yet been created.

Explanation of the Apparatus-4


- **Illicit Word Formation:**
 1. The word <*mountainal> is blocked right away because no word formation process allows its formation.
 2. For some reason, the noun <mountain> cannot take the suffix <-al>.

Proverbial Words

Proverbial Words-1

- **Definition:** “Proverbial words are generally complete sentences which behave like singles. Proverbial words, as their name indicates, are words which constitute a proverbial statement,”
Koffi (1990:197)

Proverbial Words-2

-  **Orthographic implications:**
 1. Should be written as a single orthographic word as in *Johnnycome latelies?*
 2. Should the word inside of the statement should be written according to their dictionary forms as in *Johnny come latelies?*

Proverbial Words-3

3. Should there be dashes at word boundaries as in *Johnny-come-latelies*?

In Search of the Orthographic Word

The Relevance of Morphology in Orthography

- “Samuel Johnson, who standardized the spelling of thousands of modern words, used people’s perception of the anatomy of words as a rationale in his decisions, and that is one of the reasons that the spellings of English words notoriously do not always reflect their sounds; often they reflect the morphological structure instead,” Pinker (1999:45). *Words and Rules*.

The Orthographic Word-1

- **Free Morpheme Principle:**

Free morphemes are written as independent orthographic words.

- **Bound Morpheme Principle:**

Bound morphemes and the roots to which they are affixed are written together as a single orthographic word.

The Orthographic Word-2

- **The Uniqueness of Lexical Representation Principle**

Each lexical item must have a unique spelling. Chomsky (1970:281)

The Orthographic Word-3

- **Lexical Spelling Principle**

Subject to rare exceptions, each word should be written in one way only, and its orthography form should in most cases be based on the pronunciation it has in isolation. Jones (1967:227)

The Orthographic Word-4

- **The Allomorphic Principle**

Allomorphic alternations should have separate graphemic representation in the orthography.” Koffi (1990:137)

- Failure to do so makes the orthography opaque and calls for a lot of cognitive computation while reading.

The Orthographic Word-5

- The formation of the future (and the aorist) in Greek and its representation in the orthography:
 1. **Guttural:** <k, γ, χ> or <σσ> + <σ> is written in the orthography as <ξ>
<διωκ+σ+ω> becomes <διωξω> (I will persecute) in the orthography.

The Orthographic Word-6

- **Labial:** <π, β, φ> or <ππ> + <σ> is written in the orthography as <ψ> as in <πμεπ+σ+ω> becomes <πμεψω> (I will send) in the orthography.

The Orthographic Word-7

- **Dental:** <τ, δ, θ> or <ζ> + <σ> is written in the orthography as <σ> as in <πειθ+σ+ω> becomes <πεισω> (I will believe/persuade) in the orthography.
- Source: Voelz (1993:54-64)

Binding, Clipping and Abbreviation

The Orthographic Word-8

- **Lexical Reduction Principle**

Words created as a result of blending, clipping, or abbreviation should be represented as a single orthographic word.

Homework

Exercise 1

- How would you classify your language(s) typologically:
 1. Isolating language?
 2. Agglutinating language?
 3. Fusional language?
 4. Polysynthetic (all of the above)?

Exercise 2

- What is the most common word formation process in your language(s)?
 1. Derivation?
 2. Compounding?
 3. Reduplication?
 4. Others. Which ones?

Exercise 3

- How does the orthography of your language(s) deal with allomorphic variations?
 1. Are the morphemes written the same everywhere?
 2. Are the allomorphs represented in the orthography?
 3. Others? What is the rationale?

Exercise 4

- How has the orthography dealt with the problem of long words? Are they written as:
 1. A single orthographic word?
 2. A dash between words?
 3. Others. Which ones?
 4. What is the rationale behind your decisions?

Selected References

Selected References-1

1. Anderson, Stephen. 1985. *Phonology in the Twentieth Century*. Chicago: The University of Chicago Press.
2. Hockett, Charles F. *A course in Modern Linguistics*. New York: MacMillan Publishing Co., Inc.
3. Comrie, Bernard. *Language Universals and Linguistic Typology*. Chicago: The University of Chicago Press.

Selected References-2

4. Jones, Daniel. 1967. *The Phoneme: Its Nature and Use*. New York: Cambridge University Press.
5. Kaplan, Jeffery P. 1995. *English Grammar: Principles and Facts*, 2nd ed. Englewood Cliffs, New Jersey: Prentice Hall.

Selected References-3

6. Koffi, Ettien N. 1990. *The Interface between Phonology, Morpho(phono)logy in the Standardization of Anyi Orthography*. Indiana University Ph.D. dissertation: Bloomington, Indiana. Also available online at: www.orthographyclearinghouse.org

Selected References-4

7. Scalise, Sergio. 1986. *Generative Morphology*. *Studies in Generative Grammar* 18. Riverton, USA: Fortis Publications.
8. Voel, James W. 1993. *Fundamental of Greek Grammar*, 2nd ed. St. Louis, MO: Concordia Publishing House.