

한국어의 형태소 분석

최호철(고려대)

1. 언어

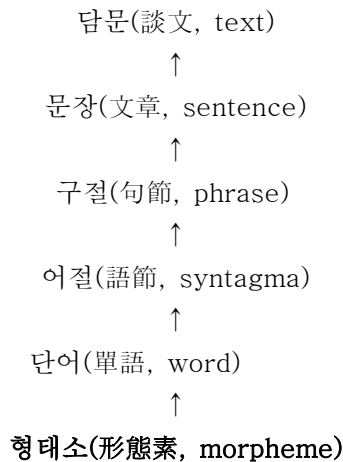
언어(言語)는 소리와 뜻이 결합된 것이므로 음성만 있는 것이라면 그것은 언어가 아닌 허언(虛言)이나 헛소리라 할 수 있고 뜻만 있는 것이라면 그것 역시 언어가 아닌 묵언(默言)이나 생각 또는 마음이라 할 수 있을 것이다.

- 1) 소리만 있다면? 허언(虛言), 헛소리.
 만구리 헬소도루 즈르바모 치누토카

- 2) 뜻만 있다면? 묵언(默言), 생각/마음.
 철이: 너, 어제 그 사람 만났지?
 석이:
 철이: 왜 말이 없어? 만났어?
 석이:
 철이: 가타부타 말을 해야지.

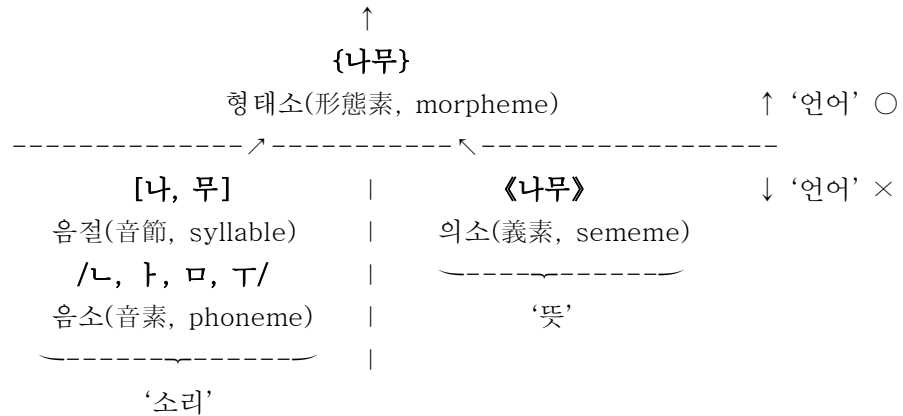
2. 언어 구성 단위

소리와 뜻이 결합된 언어 구성 단위로는 담문(談文, text), 문장(文章, sentence), 구절(句節, phrase), 어절(語節, syntagma), 단어(單語, word), 형태소(形態素, morpheme) 등이 있는데, 담문(談文, text)은 언어 구성의 최대 단위이고 형태소(形態素, morpheme)는 언어 구성의 최소 단위이다.



3. 언어 요소 단위

언어 성립에서 필수적인 소리와 뜻의 두 요소 가운데에서 음소(音素, phoneme)는 소리의 최소 단위이고 의소(義素, sememe)는 뜻의 최소 단위이므로 음소나 의소는 단독으로 언어 요소의 최소 단위가 될 수 있지만 언어 구성의 최소 단위가 될 수는 없다.



4. 형태소 분석

형태소(形態素, morpheme)는 의미를 가진 언어 구성의 최소 단위이므로 의미가 없어지는 단계 직전까지 분석할 수 있다.

- 1) ‘나무’는 <나무>라는 의미를 갖고 있으나 ‘나’ 또는 ‘무’는 어떠한 의미도 갖고 있지 않으므로 ‘나무’는 형태소가 되지만 ‘나’ 또는 ‘무’는 형태소가 되지 못한다.
- 2) ‘너무’는 <지나치게>라는 의미를 갖고 있으며, 이는 다시 <일정한 시간, 시기, 범위 따위를 벗어나 지나다>라는 의미를 갖는 ‘너ㅓ’과 <부사화>라는 의미를 갖는 ‘ㅌ’로 나눌 수 있으므로 ‘너무’는 한 형태소가 아니라 ‘너ㅓ’과 ‘ㅌ’라는 두 개의 형태소로 구성된 단어이다.
- 3) 일차적으로 언어는 음성으로 실현되고, 음성과 표기가 반드시 일대일 대응되는 것은 아니므로 형태소 분석은 표기를 대상으로 삼을 것이 아니라 발음을 대상으로 삼아야 한다.

표기: 학자 여러분!
 날개 없이 날 수 있어요?
 별에서는 붉은색이 잘 날아요.
 앞에 놓인 문체가 무엇인지 이야기해요.

발음: 학자 여러분!
 날개 업씨 날 쑤 이씨요?

벼테서는 불근새기 잘 나랴요.
아페 노힌 문제가 무어신지 이야기헤요.

4) 분석 예시

학<學> / 짜<者>

여러<許多> / 분<位>

날<飛> / 개<名詞化接辭>

업싸<沒有> / ㅣ<副詞化接辭>

나<飛> / 르<冠形詞形語尾> / 쑤<方法>

* ‘날’의 뒷말이 경음으로 발음되므로 ‘르’은 어간 발음 ‘르’이 아닌 관형사형 어미 ‘르’로 분석.

- 심을(심+을) 줄, 먹을(먹+을) 사람

- 불(보+르) 줄, 줄(주+르) 사람

* 어간 발음 ‘르’은 탈락.

- 들(드+르 ← 들+르) 줄, 벌(머+르 ← 벌+르) 사람

이싸<有> / ㄱ<平敍形語尾> / 요<相對尊待助詞>

벼트<陽光> / ㅈ서<副詞格助詞> / 는<對照助詞>

불ㄱ<紅> / ㄴ<冠形詞形語尾> / 새ㄱ<色> / ㅣ<主格助詞>

잘<容易>

나르<掉> / ㅏ<平敍形語尾> / 요<相對尊待助詞>

아프<前> / ㅈ<副詞格助詞>

노ㅎ<放> / ㅣ<被動接辭> / ㄴ<冠形詞形語尾>

문<問> / 제<題> / 가<主格助詞>

무어ㅅ<什麼> / ㅣ<敍述格助詞> / ㄴ지<疑問形語尾>

이야기<話> / 해(←하+여) / 요<相對尊待助詞>

* ‘해’는 한 형태 ‘하’<爲>와 다른 한 형태 ‘여’<平敍形語尾>의 결합으로서 형태적인 융합이어서 분석 불가.

- 해(←하여), 했(←하엿), 돼(←되어), 났(←되엿).

- 네(←너의), 내(←나의).

5. 형태소와 이형태

형태소(形態素, morpheme)는 하나의 같은 의미를 갖는 여러 형태(形態, morph)의 집합(集合, set)이며, 이형태(異形態, allomorph)는 한 형태소 안에 있는 여러 형태이다.

- 1) 형태소 표시는 두 가지 방법이 있는데, 하나는 / / 안에 모든 이형태를 나열하는 것이고, 다른 하나는 { } 안에 하나의 대표적인 이형태를 보이는 것이다.
- 2) 대표 이형태는 한 언어에서 분포가 넓고 다른 이형태의 도출이 자연스러운 이형태를 선정하는 것이 일반적이다.

/ 이형태 나열 /		{대표 이형태}
<學> /학(←학교), 향(←학문)/		= {학}
* 파열음이 비음 앞에서 비음으로 바뀐다는 설명이 더 자연스러운 ‘학’을 대표로.		
<者> /짜(←독자), 자(←청자)/		= {자}
* 평음이 파열음 뒤에서 경음으로 바뀐다는 설명이 더 자연스러운 ‘자’를 대표로.		
<許多> /여러(←여러 나라)/		= {여러}
<位> /분(←한 분), 뽀(←다섯 분)/		= {분}
* 평음이 파열음 뒤에서 경음으로 바뀐다는 설명이 더 자연스러운 ‘분’을 대표로.		
<飛> /날(←새가 날고), 나(←새가 나니)/		= {날}
* 미정의 음운 첨가보다는 기존의 음운 탈락으로 설명하는 것이 더 자연스러운 ‘날’을 대표로.		
<接辭> /개(←찌개), 깨(←덮개)/		= {개}
* 평음이 파열음 뒤에서 경음으로 바뀐다는 설명이 더 자연스러운 ‘개’를 대표로.		
<沒有> /업쓰(←없어), 업(←없고), 업(←없네)/		= {업쓰}
* 파열음이 비음 앞에서 비음으로 바뀐다는 설명이 더 자연스럽고, 발음의 중화가 일어 나기 전의 ‘업쓰’을 대표로.		
<接辭> /ㅣ(←많이)/		= {ㅣ}
<語尾> /ㄹ(←할 일), ㄴ(←먹을 것)/		= {ㄴ}
* 미정의 음운 첨가보다는 기존의 음운 탈락으로 설명하는 것이 더 자연스러운 ‘ㄴ’을 대표로.		
<方法> /쑤(←할 수), 수(←하는 수)/		= {수}
* 평음이 관형사형 어미 ‘ㄹ’ 뒤에서 경음으로 바뀐다는 설명이 더 자연스러운 ‘수’를 대표로.		
<有> /이쓰(←있으니), 인(←있지만), 인(←있네요)/		= {이쓰}
* 파열음이 비음 앞에서 비음으로 바뀐다는 설명이 더 자연스럽고, 발음의 중화가 일어 나기 전의 ‘이쓰’을 대표로.		
<語尾> /ㄱ(←먹어), ㅏ(←막아), 여(←하여)/		= {ㄱ}
* 분포가 음운적인 제약을 받는 것은 임의로 정할 수 있는데, ‘ㅏ’ 뒤에서 ‘ㄱ’로 실현되는 경우도 있어 분포가 더 넓은 ‘ㄱ’을 대표로.		
<助詞> /요(←어서요), 뇨(←하면요)/		= {요}
* ‘ㄴ’ 첨가가 일어나지 않은 경우도 있어 분포가 더 넓은 ‘요’를 대표로.		
<陽光> /벼ㅌ(←별에), 변(←별도), 변(←별만), 벼ㅌ(←별이), ㅌㅌ(←햇별에), ㅌㅌ(←햇별도), ㅌㅌ(←햇별만), ㅌㅌ(←햇별이)/		= {벼ㅌ}
* 파열음이 비음 앞에서 비음으로 바뀐다는 설명이 더 자연스럽고, 평음이 파열음 뒤에서 경음으로 바뀐다는 설명이 더 자연스럽고, 발음의 중화가 일어나기 전의 ‘벼ㅌ’을 대표로.		
<助詞> /꺀서(←집에서)/		= {꺀서}
<助詞> /ㄴ(←학교는), ㄴ(←집은), ㄴ(←난 몰라)/		= {ㄴ}
* 분포가 음운적인 제약을 받는 것은 임의로 정할 수 있는데, 미정의 음운 첨가보다는 기존의 음운 탈락으로 설명하는 것이 더 자연스럽고, 앞말의 형태가 유지되는 ‘ㄴ’을		

대표로.

<紅> /불ㄱ(←불어), 북(←불다), 불(←불고),
봉(←불네), 빨ㄱ(←길불어), 뽉(←길불다),
빨(←길불고), 뽉(←길불네)/ ={불ㄱ}

* 파열음이 비음 앞에서 비음으로 바뀐다는 설명이 더 자연스럽고, 평음이 파열음 뒤에서 경음으로 바뀐다는 설명이 더 자연스럽고, 발음의 중화가 일어나기 전의 ‘불ㄱ’을 대표로.

<語尾> /ㄴ(←맑은), ㄴ(←예쁜)/ ={ㄴ}

* 미정의 음운 첨가보다는 기존의 음운 탈락으로 설명하는 것이 더 자연스러운 ‘ㄴ’을 대표로.

<色> /새ㄱ(←색이), 생(←색만),
쌩(←혈색이), 썩(←혈색만)/ ={새ㄱ}

* 파열음이 비음 앞에서 비음으로 바뀐다는 설명이 더 자연스럽고, 평음이 파열음 뒤에서 경음으로 바뀐다는 설명이 더 자연스러운 ‘새ㄱ’을 대표로.

<助詞> /ㅣ(←밥이), 가(←나무가)/ ={가}

* 분포가 음운적인 제약을 받는 것은 임의로 정할 수 있는데, 서술격 조사 ‘이’와 구별하기 위해 ‘가’를 대표로.

<容易> /잘(←잘못), 짤(←곧잘)/ ={잘}

* 평음이 파열음 뒤에서 경음으로 바뀐다는 설명이 더 자연스러운 ‘잘’을 대표로.

<掉> /나ㄴ(←향수가 날고), 나(←향수가 나니)/ ={나ㄴ}

* 미정의 음운 첨가보다는 기존의 음운 탈락으로 설명하는 것이 더 자연스러운 ‘나ㄴ’을 대표로.

<前> /아ㅍ(←앞을), 압(←앞도), 압(←앞만)/ ={아ㅍ}

* 파열음이 비음 앞에서 비음으로 바뀐다는 설명이 더 자연스럽고, 발음의 중화가 일어나기 전의 ‘아ㅍ’을 대표로.

<助詞> /ㄱ(←집에)/ ={ㄱ}

<放> /노ㅎ(←놓아), 논(←놓네)/ ={노ㅎ}

* 마찰음이 비음 앞에서 비음으로 바뀐다는 설명이 더 자연스러운 ‘노ㅎ’을 대표로.

<接辭> /ㅣ(←깎이다), 히(←떡히다), 리(←팔리다),
기(←안기다)/ ={히}

* 분포가 형태적인 제약을 받는 것은 임의로 정할 수 있는데, 사동 접사의 ‘이’와 구별하기 위해 ‘히’를 대표로.

<問> /문(←질문, 문답)/ ={문}

<題> /제(←제목), 쟈(←숙제)/ ={제}

* 평음이 파열음 뒤에서 경음으로 바뀐다는 설명이 더 자연스러운 ‘제’를 대표로.

<什麼> /무어ㅅ(←무엇을), 무언(←무엇도),
무언(←무엇만)/ ={무어ㅅ}

* 파열음이 비음 앞에서 비음으로 바뀐다는 설명이 더 자연스럽고, 발음의 중화가 일어나기 전의 ‘무어ㅅ’을 대표로.

<助詞> /ㅣ(←밥이다)/ ={ㅣ}

<語尾> /ㄴ지(←바쁘지), ㄴ지(←맑은지),

는지(←있는지)/ ={늘지}

* 미정의 음운 첨가보다는 기존의 음운 탈락으로 설명하는 것이 더 자연스럽고, 분포가 더 넓은 ‘느지’를 대표로.

<话> /이야기(←혼사 이야기), 니야기(←뒤틀이야기)/ ={이야기}

* ‘ㄴ’ 첨가가 특별한 경우에 일어나서 분포가 더 넓은 ‘이야기’를 대표로.

6. 형태소와 한국어의 표기 원칙

한국어의 표기는 원칙적으로 형태소의 이형태 가운데에서 다른 이형태의 도출이 자연스러워서 선정한 대표 이형태를 적되, 한 이형태에서 다른 이형태의 도출이 자연스럽지 않고 발음이 확연히 달라지는 것은 해당하는 환경에서 실현되는 이형태를 그대로 적는다.

/ 이형태 /	{대표 이형태}	‘표기’
/학, 향/	={학}	‘학’
/짜, 자/	={자}	‘자’
/여러/	={여러}	‘여러’
/분, 뽀/	={분}	‘분’
/날, 나/	={날}	‘날/나’
	* ‘날’과 같은 표기로 [나는] 발음을 실현할 수 없을 때는 ‘나’로 표기.	
/개, 깨/	={개}	‘개’
/업싸, 업, 엄/	={업싸}	‘없’
	* ‘업싸’와 발음이 같고 종성을 아래에 적는 ‘없’으로 표기.	
/ㅣ/	={ㅣ}	‘이’
	* ‘ㅣ’와 발음이 같고 앞말의 형태를 밝힐 수 있는 ‘이’로 표기.	
/ㄹ, ㄴ/	={ㄴ}	‘을/ㄹ’
	* ‘ㄴ’과 발음이 같고 앞말의 형태를 밝힐 수 있는 ‘을’로 표기하되, ‘하을’과 같은 표기로 [할] 발음을 실현할 수 없을 때는 ‘ㄹ’로 표기.	
/쑤, 수/	={수}	‘수’
/이싸, 인, 인/	={이싸}	‘있’
	* 종성을 아래에 적는 ‘있’으로 표기.	
/ㄱ, ㅏ, 여/	={ㄱ}	‘어/아/여’
	* ‘ㄱ’와 발음이 같고 앞말의 형태를 밝힐 수 있는 ‘어’로 표기.	
	* ‘조허’와 같은 표기로 [조하] 발음을 실현할 수 없을 때는 ‘조하’와 발음이 같고 앞말의 형태를 밝힐 수 있는 ‘아’로 표기.	
	* ‘하어/하아’와 같은 표기로 [하여] 발음을 실현할 수 없을 때는 ‘여’로 표기.	
/요, 뇨/	={요}	‘요’
/벼ㅌ, 변, 변, 벼ㅌ,		
뼈ㅌ, 뽀, 뽀, 뼈ㅌ/	={벼ㅌ}	‘별’
	* 종성을 아래에 적는 ‘별’으로 표기.	

/케서/	= {케서}	‘에서’
* ‘케서’와 발음이 같고 앞말의 형태를 밝힐 수 있는 ‘에서’로 표기.		
/는, 은, ㄴ/	= {는}	‘는’
/불기, 북, 불, 붕, 뿔기, 뿔, 뿔, 뿔/	= {불기}	‘뿔’
* 종성을 아래에 적는 ‘뿔’으로 표기.		
/ㄴ, ㄴ/	= {ㄴ}	‘은/ㄴ’
* ‘ㄴ’과 발음이 같고 앞말의 형태를 밝힐 수 있는 ‘은’으로 표기하되, ‘예쁜’과 같은 표기로 [예쁜] 발음을 실현할 수 없을 때는 ‘ㄴ’으로 표기.		
/새기, 생, 색, 썩/	= {새기}	‘색’
* 종성을 아래에 적는 ‘색’으로 표기.		
/ㅣ, 가/	= {가}	‘가/이’
* ‘밥가’와 같은 표기로 [바비] 발음을 실현할 수 없을 때는 ‘ㅣ’와 발음이 같고 앞말의 형태를 밝힐 수 있는 ‘이’로 표기.		
/잘, 짤/	= {잘}	‘잘’
/나르, 나/	= {나르}	‘날/나’
* 종성을 아래에 적는 ‘날’로 표기하되, ‘날는’과 같은 표기로 [나는] 발음을 실현할 수 없을 때는 ‘나’로 표기.		
/아프, 압, 암/	= {아프}	‘앞’
* 종성을 아래에 적는 ‘앞’으로 표기.		
/케/	= {케}	‘에’
* ‘케’와 발음이 같고 앞말의 형태를 밝힐 수 있는 ‘에’로 표기.		
/노ㅎ, 논/	= {노ㅎ}	‘놓’
* 종성을 아래에 적는 ‘놓’으로 표기.		
/ㅣ, 히, 리, 기/	= {히}	‘히/이/리/기’
* ‘깎히다’와 같은 표기로 [까끼다] 발음을 실현할 수 없을 때는 ‘ㅣ’와 발음이 같고 앞말의 형태를 밝힐 수 있는 ‘이’로 표기.		
* ‘팔히다’와 같은 표기로 [팔리다] 발음을 실현할 수 없을 때는 ‘리’로 표기.		
* ‘안히다’와 같은 표기로 [안기다] 발음을 실현할 수 없을 때는 ‘기’로 표기.		
/문/	= {문}	‘문’
/제, 쩌/	= {제}	‘제’
/무어ㅅ, 무언, 무언/	= {무어ㅅ}	‘무엇’
* 종성을 아래에 적는 ‘무엇’으로 표기.		
/ㅣ/	= {ㅣ}	‘이’
* ‘ㅣ’와 발음이 같고 앞말의 형태를 밝힐 수 있는 ‘이’로 표기.		
/ㄴ지, 은지, 는지/	= {ㄴ지}	‘은지/ㄴ지/는지’
* ‘ㄴ지’와 발음이 같고 앞말의 형태를 밝힐 수 있는 ‘은지’로 표기하되, ‘예쁜은지’와 같은 표기로 [예쁜지] 발음을 실현할 수 없을 때는 ‘ㄴ지’로 표기.		

* ‘먹은지’와 같은 표기로 [먹는지] 발음을 실현할 수 없을 때는 ‘느지’로 표기.

/이야기, 니야기/ = {이야기} ‘이야기’

7. ‘∅’형태(영형태, 零形態, zero morph) : 의미는 있는데 형태가 없는 것.

‘날개가 없이’의 ‘가’는 <주체화> 의미의 형태 : ‘가’형태.

‘손발이 없이’의 ‘이’는 <주체화> 의미의 형태 : ‘이’형태.

‘날개 없이’의 ‘날개’ 뒤에 <주체화> 의미의 ‘∅’ 형태가 있다고 가정 : ‘∅’형태.

‘손발 없이’의 ‘손발’ 뒤에 <주체화> 의미의 ‘∅’ 형태가 있다고 가정 : ‘∅’형태.

‘날개를 펴고’의 ‘를’은 <객체화> 의미의 형태 : ‘를’형태.

‘날개^ㄴ 펴고’의 ‘ㄴ’은 <객체화> 의미의 형태 : ‘ㄴ’형태.

‘손발을 펴고’의 ‘을’은 <객체화> 의미의 형태 : ‘을’형태.

‘날개 펴고’의 ‘날개’ 뒤에 <객체화> 의미의 ‘∅’ 형태가 있다고 가정 : ‘∅’형태.

‘손발 펴고’의 ‘손발’ 뒤에 <객체화> 의미의 ‘∅’ 형태가 있다고 가정 : ‘∅’형태.

* 형태소는 같은 의미를 갖는 여러 형태의 집합이므로 <주체화> 형태소는 ‘가’형태와 ‘이’형태 및 ‘∅’형태를 아우르는 말로서 /가, ㅣ, ∅/로 표시하거나 {가}나 {이} 또는 {∅}으로 표시할 수 있고, <객체화> 형태소는 ‘를’형태와 ‘ㄴ’형태 및 ‘을’형태와 ‘∅’형태를 아우르는 말로서 /를, ㄴ, 을, ∅/으로 표시하거나 {를}이나 {ㄴ} 또는 {을}이나 {∅}로 표시할 수 있다.

* 형태(morph)는 개별적인 것을 지칭하므로 ‘가’형태, ‘이’형태, ‘를’형태, ‘ㄴ’형태, ‘을’형태, ‘∅’형태라고 할 수 있으나, 형태소(morpheme)는 하나의 같은 의미를 갖는 여러 형태의 집합이므로 ‘형태소’란 말 앞에는 구체적인 개별 형태를 나타내는 말이 올 수 없다. 예를 들어 의미를 나타내는 <주체화>나 <객체화>를 앞세워 <주체화> 형태소, <객체화> 형태소란 말은 사용할 수 있으나 개별 형태(morph)를 앞세운 ‘가’형태소, ‘이’형태소, ‘를’형태소, ‘ㄴ’형태소, ‘을’형태소, ‘∅’형태소(zero morpheme)란 말은 사용할 수 없다.

* ‘가’형태 / ‘이’형태 / ‘를’형태 / ‘ㄴ’형태 / ‘을’형태 / ‘∅’형태 : ○
 <주체화> 형태소 / <객체화> 형태소 : ○
 ‘가’형태소 / ‘이’형태소 / ‘를’형태소 / ‘ㄴ’형태소 / ‘을’형태소 / ‘∅’형태소 : ×

8. 허형태(虛形態, empty morph) : 형태는 있는데 의미가 없는 것.

좁쌀(←조ㅁ쌀)의 ‘ㅁ’은 형태는 있으나 의미가 없음 : 허형태.

송아지(←소ㅇ아지)의 ‘ㅇ’은 형태는 있으나 의미가 없음 : 허형태.

* 허형태란 말은 형태는 있으나 의미가 없는 것을 일컬으므로 위의 ‘ㅂ’이나 ‘ㅇ’과 같이 의미가 없는 형태는 모두 허형태라 할 수 있으며, 형태소는 같은 의미를 갖는 여러 형태의 집합이므로 허형태소(虛形態素, empty morpheme)는 위의 ‘ㅂ’이나 ‘ㅇ’과 같이 의미가 없는 모든 형태를 아우르는 말로서 /ㅂ, ㅇ, .../으로 나열하거나 {ㅂ}이나 {ㅇ} 등과 같이 대표 이형태로 표시할 수 있다.

* 허형태, 허형태소 : ○

9. 언어 구성과 요소의 최소 단위

언어 구성 단위		언어 요소 단위	
구성의 최소 단위		소리의 최소 단위	뜻의 최소 단위
개별적:	형태(形態, morph)	음성(音聲, phone)	의미(義味, seme)
집합적:	형태소(形態素, morpheme)	음소(音素, phoneme)	의소(義素, sememe)
개별적:	이형태(異形態, allomorph)	이음(異音, allophone)	이의(異義, alloseme)

<끝> 고맙습니다.