

현 대 국 어 음 변 화 규 칙 의 기 능 부 담 량

이 성익

서울대학교 인문대학 국어국문학과

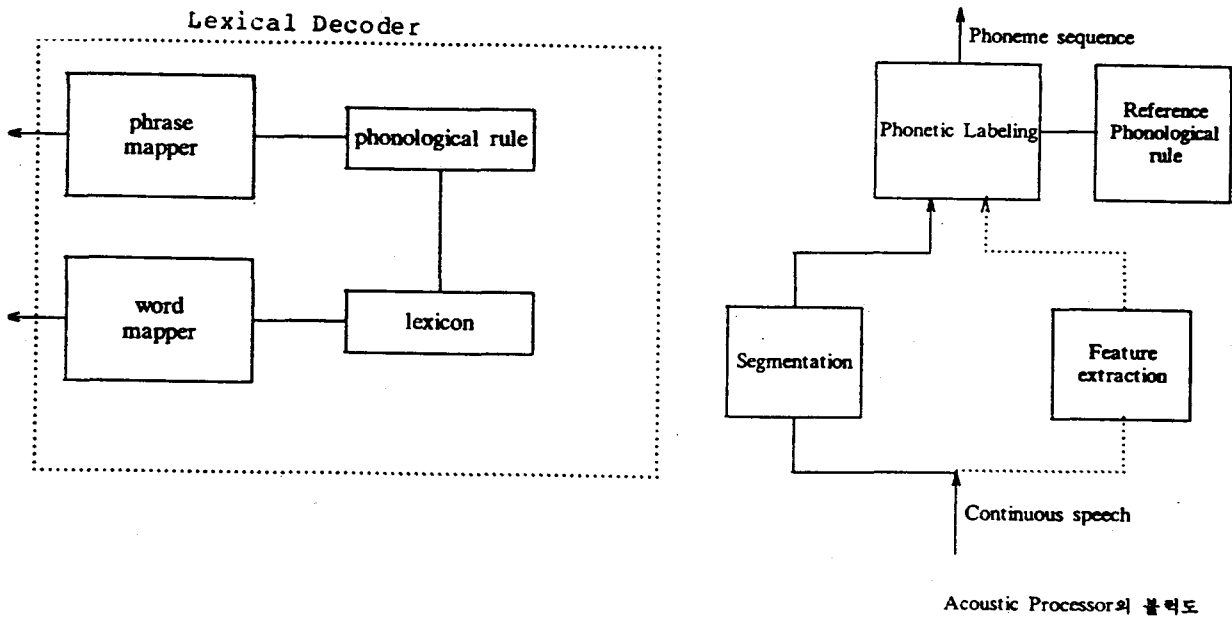
서울 관악구 신림동 산56-1 (우 151-742)

요약

이 연구는 현대국어에 나타나는 음성규칙, 음운규칙, 형태음운규칙, 기타 등등의 발생 빈도를 조사하여, 각 규칙이 차지하는 기능 부담량(functional load)을 알아 보려는 것이다.

1. 서론

이런 규칙이 그 언어의 음변화 과정에 동원되는 빈도수, 즉 규칙의 기능 부담량을 체계적으로 파악할 수 있으면 좋겠다는 착상을 이미 이상익(1979)부터 한 바 있으나, 아직까지도 그 연구를 착수하지 못하여 왔다. 이제 그러한 연구를 더 이상 늦출 수 없다는 생각을 갖게 된 까닭은 음성 인식, 음성합성, 언어이해 등등의 분야에서 음성/음운규칙의 개입을 모두 설정하고 있기 때문이다. 과학기술원(1988)에서 한 예를 보이면 다음과 같다.



언어학자는 음성/음운규칙의 발견 및 기술뿐 아니라, 그것이 어떤 언어 체계에서 차지하는 비중에 대해 상대적 평가를 내릴 수 있는 방법도 강구하여야 할 것이다. 이것은 발인 음운 현상의 기술에서도, 가령 대구개음화규칙이 발인 방언에서 차지하는 기능 부담량은 0 이라 할 수 있는데 비해, 다른 방언들에선 어느 정도라고 한다는가 하는 편리한 척도를 마련해 줄 것이다. 한 개인의 개인 방언에 대해서도 마찬가지로 적용할 수 있을 것이다.

이제 본 연구를 위하여, 이희승 편 '국어대사전' (개정판)에서 각 페이지의 왼쪽 맨 밑 어휘를 하나의 뽑아 총 4,370 개의 어휘를 dBASE III 플러스 프로그램을 이용하여 디스켓에 입력시켰다. 입력 방식으로는 (2)에서와 같이 철자 표기, 음운 전사, 음성 전사의 3 방식으로 각 어휘를 기록하되; 철자 표기에는 자간 경계를 -로 표시하고; 음운 전사에는 단어 경계를 =로, 형태소 경계를 +로 기록하고; 음성 전사에는 음절 경계를 /로 기입하도록 하였다(이상익 1989 참조).

이러한 각 어휘에 각각 품사(word class), 형태론적 자격(morphological status), 그리고 차후의 의미론적 연구를 위하여 용법(usage), 완전 동음이의어(complete homonym), 불완전 동음이의어(incomplete homonym), 다의어(polysemy), 어원(etymology)을 밝혀 부호나 숫자로 기록하였다(이상익 1990 참조).

여기서 기록하지 못한 정보로는 이상익(1986)에 이미 소개하였듯이 빈도수, 유의어(또는 근사 유개념 genera proxima), 피면화성(affectivity), 연상성(associativity) 등이 있겠으나, 기초 연구가 진행되면 후일 추가할 수 있을 것이다.

이상의 설명을 전제로 실제 입력된 기록의 앞부분만을 보이면 (2)의 요외 같다.

(2) 철자 표기	음운 전사	음성 전사	WC	MS	US	HC	HI	PS	ET
#1 ka	ka	ka		ex		5	5	3	k
#2 ka-ky9k-ca-yu-hwa	kaky9k=cayuhwa	ka/gy9k/	n	cd					c
		1 21							
		Ja/yu/hwa							
		26							
#3 ka-koq-mu-y9k	kakoq+mUy9k	ka/goq/mu/y9k	n	cd					
		33							
		1 2 32 21							

#4	ka-k7-U4-ki	kAkwi+U4+ki	kA/gwi/U4/gi	n	d	d	k
		33 95 26	1 1				
#5	ka-nan	kanan	ka/nan	n			k
			2				
#6	ka-n4n-k7-m9k-ta	ka+n4q+kwi=	ka/n4q/gwi/	vi	cd		k
		23 95	2 1				
		m9k+Da	m9k/Da				
etc.		26	21				

위의 기록상 각음 전사 체제에 쓰인 기호는 다음과 같다.

모음

단순모음

短모음

장모음

	i	—	4	ㅏ	u	:	I	—:	6	ㅓ:	U
	e		9	ㅜ	o	:	E	:	8	ㅛ:	O
	3		a			:	5	:	A		

ㅑ 7/wi

ㅑ: Y/wi

ㅓ ∅

ㅓ: 2

이중모음

短모음

장모음

	ㅓ	4y				ㅓ:	6y				
			ㅓ	yu				ㅓ:	yU		
	ye		y9	ㅓ	yo	:	yE	:	y8	ㅓ:	yO
	y3		ya			:	y5	:	yA		

ㅑ we ㅑ w9

ㅑ: wE ㅑ: w8

ㅓ w3 ㅓ wa

ㅓ: w5 ㅓ: wA

자음

	b		d		j		g		l (← 음성전사에만 이용)	
ㅍ	p	ㅌ	t	ㅈ	c	ㅊ	k	ㅅ	s ㅎ	h
ㅃ	B	ㅆ	D	ㅉ	J	ㅊ	G	ㅅ	S	
ㅍ	P	ㅌ	T	ㅈ	C	ㅊ	K			
ㅍ	m	ㅌ	n			ㅇ	q			
		ㄹ	r							

2. 음변화 규칙의 목록과 그 적용 사례의 검색

본 연구를 위하여 우선 국어 음변화 규칙의 목록을 작성하면 다음과 같다.

A. 음성규칙

- 1) 유성음화
- 2) 음절말 유성 자음의 불파화: 인[an̚], '안드'처럼 하면 안 됨.
- 3) ㄱ 구개음화
- 4) ㅋ 구개음화
- 5) ㆁ 구개음화
- 6) 설측음화

B. 음운규칙

- 21) 음절말 장애음 중화 현상
- 22) ㄷ 구개음화
- 23) 비음화
- 24) 유음화
- 25) 격음화
- 26) 경음화
- 27) 자음군 단순화
- 28) 邊子音化: 낮까지(→ 날까지)→ 닉까지[nak̚kkaʝi]
- 29) y - 월음화
- 30) w - 활음화
- 31) ㄴ 심입
- 32) 단보음화
- 33) 징모음화
- 34) ㅎ 탈락현상
- 35) 두음법칙
- 36) ㄹ 탈락
- 37) 으 탈락
- 38) 구개음 뒤의 y - 탈락
- 39) ㄹ → ㄴ

1. 기타 (발음 해독 규칙)

- 91) 음절 초성 규칙: 높이 → 노피
- 92) 외 → 웨 / 예
- 93) 외 → 으 / 이
- 94) < → t (자이시옷)
- 95) 외 → wi
- 96) 외 → 에
- 97) ye → e

위의 규칙들 외에도 ㅁ, ㄴ, ㅅ, ㅌ, 리, ㅎ 불규칙 변화 등의 형태음운 규칙이 추가될 수 있겠으나, 본 연구에서는 개입시키지 않겠다. 그 까닭은 앞서의 4370 어항 데이터가 사전에서 무작위로 뽑은 자료이기 때문에 활용어미 부분이 나타나지 않은 특색이 있기 때문이다.

다음 단계로 (2)에서 보인 3 가지 표기 중, 특히 음운 전사와 음강 전사 밑에 이미 기록했듯이 규칙의 번호를 붙이는 일이 필요하다. 이 작업은 규칙들이 적용된 사례를 빠짐없이 검색하도록 주의해서 해야 한다.

3. 발음 빈도 조사 통계

각 규칙들이 적용된 위치는 보통 형태소 경계가 표시된 음절 전사와 음절 경계가 표시된 음운 전사에 나타나고 있다. 이를 음절 별, 형태소 별 (예: 제1 음절에 규칙 1, 즉 유성음화가 1304 번 적용되어 최고 빈도이므로 밑줄을 긋고, 제 6 형태소에는 1번 적용됨을 보임) 도 집계하면 아래와 같다.

A. 음절규칙

	제1	2	3	4	5	6	7	8	9	10	11	합계
규칙 1: 음절	<u>1304</u>	816	541	213	81	34	9	6	4	1		3008
형태소						1						<u>1</u>
												3009
규칙 2: 음절	<u>1636</u>	1596	878	507	166	88	46	17	9	2		4945
형태소	4	6			1							<u>11</u>
												4956
규칙 3: 음절	107	<u>190</u>	160	68	28	14	5	2	1			575
형태소			1									<u>1</u>
												576

규칙 5: 음절	1	<u>44</u>	30	17	5				1			98
규칙 6: 음절	371	<u>396</u>	212	117	37	26	14	5	1			1519
형태소		1										<u>1</u>
												<u>1520</u>
총계												10159

B. 음운규칙

	제1	2	3	4	5	6	7	8	9	10	11	합계
규칙 21: 음절	184	<u>237</u>	142	93	40	14	7	2	1			720
형태소	30	28	15	3	3	2						<u>81</u>
												801
규칙 22: 음절		1		1								2
형태소	1	<u>2</u>	2				1					<u>6</u>
												8
규칙 23: 음절	4	17	3		1	1						26
형태소	<u>47</u>	21	24	3	1	3	1					<u>100</u>
												126
규칙 24: 음절			2	1								3
형태소	<u>23</u>	4	20	3	1		1					<u>52</u>
												55
규칙 25: 음절		<u>25</u>	35	10	4	2						76
형태소	1	9	2									<u>12</u>
												88
규칙 26: 음절		7	58	19	6	4						94
형태소	19	<u>389</u>	260	127	64	29	10	6	1			<u>905</u>
												999
규칙 27: 음절	1											1
형태소	<u>4</u>	2	2	1	1							<u>10</u>
												11
규칙 28: 형태소	<u>2</u>		1		1							4
규칙 29: 음절		<u>1</u>										1
규칙 31: 음절		1	<u>18</u>	3		1	1					24
형태소		3	11		1							<u>15</u>
												39

규칙 32: 음절	30	<u>152</u>	30	21	4	7	1			1	246
형태소	1			1			1				<u>3</u>
											249
규칙 33: 형태소	<u>1072</u>	81	275	62	36	13	8	4	1	1	1553
규칙 34: 음절		2	2	4	2	1	1			1	13
형태소	1	<u>23</u>	8	11	4	2	2			2	<u>66</u>
											79
규칙 35: 음절	<u>106</u>										106
규칙 38: 음절	11	<u>52</u>	25	8	4	1	1	1			103
형태소		1									<u>1</u>
											104
규칙 39: 음절		1	1	3							5
형태소		<u>61</u>	17	18	4	1	1	1			<u>103</u>
											<u>108</u>
총계											4331

다. 기타

	개1	2	3	4	5	6	7	8	9	10	11	합계
규칙 91: 음절	13	<u>218</u>	133	59	19	11	2		2	1	2	460
형태소		14	3	3	1							<u>21</u>
												481
규칙 92: 음절			1	1								2
형태소	<u>58</u>	35	12	19	10	2	3	1	1		2	<u>142</u>
												145
규칙 93: 형태소	7	10	1	<u>17</u>	1	7	1	1				45
규칙 94: 형태소	<u>18</u>	16	6	4	1							45
규칙 95: 형태소	<u>45</u>	30	11	9	4		1	1	2			103
규칙 96: 형태소		1	<u>7</u>	3		1						12
규칙 97: 형태소		<u>1</u>		1								<u>2</u>
총계												833

3. 결론

2에서 보인 바와 같이, A.음성규칙 발생 총계 (10,159) 가 B.음운규칙 발생 총계 (4,331)보다 2 배 이상 더 많다.

가장 발생 빈도가 높은 규칙부터 10개만 순서대로 정리하면 다음과 같다.

규칙 번호	명칭	발생 빈도
A 2	음절말 유성자음의 불파화	4956
A 1	유성음화	3009
B 33	강모음화	1553
A 6	설측음화	1520
B 26	경음화	999
B 21	음절말 장애음 중화 현상	801
A 3	시 구개음화	576
C 91	음절 조정 규칙	481
B 32	단모음화	249
C 92	외 → 웨	145

음성 인식-합성 등의 절차에서 특히 위와 같은 규칙들은 고빈도로 적용되는 것들이니 우선적으로 처리하는 기제를 설정해야 할 것이다.

한편 2장에서 예상 목록에 올렸던 규칙들 중(예 30) w-활음화(36) r 탈락(37) 등 탈락 규칙이 4370 어항 이내에서는 전혀 발생되지 않았다는 점도 흥미롭다. 29) y-활음화(97) ye→e 규칙은 1,2번의 발생 빈도를 보이는 점도 주목된다.

3 장의 28 개 규칙중 밑줄 친 최고 빈도가 제1 음절/형태소에 11, 제2에 13, 제3에 3, 제4에 1번 집계된다. 또 28개 규칙중 18 개 규칙은 음절/형태소 양쪽에 의존되고(sensitive), 7 개 규칙은 형태소만, 3 개 규칙은 음절에만 의존하여 집계된다.

참고 논저

- 과학기술원 통신연구실 1988 한국어 음성인식 시스템 개발 연구. 최종 보고서.
 이상익 1979 국어 음운론에 있어서의 공모성에 대한 재론. 한글 165호, 3-30.
 _____ 1986 계량언어학(어휘론 부분): 어휘부의 음운-형태-의미론적 구조.
 언어 11 권 2호, 323-339.
 _____ 1989 국어 어휘목록의 형태-음운론적 구조 연구 -- 계량언어학적 표준
 조사 --. 어학연구 25 권 1호, 111-128.
 _____ 1970 국어 어휘부의 계량언어학적 연구. 기록 강신항 선생 화갑기념
 논문집(정글권대). 태학사, 331-346.