

낱말 인식 검사에 대한 어휘적 특성의 영향 분석

윤미선* 이봉원*

* 나사렛대학교 언어치료학과

Analysis of Lexical Effect on Spoken Word Recognition Test

Misun Yoon*, Bongwon Yi*

* Dept. of Communication Disorders, Korea Nazarene University

msyoon@kornu.ac.kr, cynos@kornu.ac.kr

Abstract

The aim of this paper was to analyze the lexical effects on spoken word recognition of Korean monosyllabic word. The lexical factors chosen in this paper was frequency, density and lexical familiarity of words. Result of the analysis was as follows; frequency was the significant factor to predict spoken word recognition score of monosyllabic word. The other factors were not significant. This result suggest that word frequency should be considered in speech perception test.

I. 서론

청각장애인은 말소리를 듣고 지각하여 의사소통을 하는 과정에서 어려움을 겪는다. 이러한 의사소통 상의 어려움 정도를 평가하기 위하여 말 지각 검사(speech perception test)를 실시한다. 말 지각 검사를 실시하는 목적은 청력검사도(audiogram)에 나타나는 결과만으로는 알기 어려운 정보, 즉 청각장애인이 실제 의사소통 상황에서 얼마나 상대방의 말소리를 알아 듣고 이해할 수 있는가를 평가하기 위한 것이다[1]. 청

력검사도의 경우 각 주파수별로 순음 자극을 주고 역치를 찾게 된다. 이렇게 찾아낸 역치는 청각장애 등급 판정의 준거로 사용되고 있으나 실제로 의사소통 상황에서 말소리에 대한 이해 정도를 설명하기에는 충분하지 않다. 왜냐하면 우리가 말소리를 듣고 이해하는 과정은 단순히 소리 자극에 대한 역치만으로는 설명할 수 없기 때문이다.

말 지각 검사는 대개 낱말 인식 점수(spoken word recognition score)를 측정하는 방식으로 이루어진다. 이 때 일반적으로 음소 균형을 맞춘 단음절이나 이음절의 낱말 목록을 사용한다[2] 현재 임상 현장에서 사용되고 있는 음성 균형 낱말 검사(phonetically balanced(PB) word test)는 검사 목록 안에서 음소의 균형을 확보한다는 원칙에 의해 제작되었고 그 밖의 요인은 세밀하게 고려되지 않은 것이 보통이다. 그러나 낱말의 인식은 음소 수준보다 상위의 과정을 거치게 되며, 어휘적 특성이 영향을 줄 수 있다. 즉, 낱말 중에는 더 쉽게 인식되는 것이 있으며 인식이 쉽지 않은 것도 있다.

최근 들어 단순히 음성 균형을 확보한 검사도구 외에도, 다양한 이론적 모형을 전제로 한 듣기능력평가도구의 개발이 활발히 이루어지고 있다. Kirk 등은 근접 활성화 모형(Neighborhood Activation Model; NAM)을 근거로 LNT(Lexical Neighborhood Test) 검

사를 발표하였고[3], LNT는 인공와우이식을 받은 청각장애인의 듣기능력평가에서 음성 균형을 맞춘 검사(PB)를 대체하여 사용되고 있다.

LNT는 어휘적 특성과 인식의 관계를 반영한 근접활성 모형에 따른 것으로, 이 모형은 근접 효과 또는 이웃 효과(neighborhood effect)를 설명하기 위한 것이다. 근접 효과는 특정 낱말의 인식이 어휘부에서 이 목표 낱말과 유사한 여러 다른 낱말의 영향을 받는다는 것으로, 낱말 인식에 영향을 주는 어휘적 특성으로는 유사성과 빈도가 대표적이다[4].

본 연구는 임상 현장에서 널리 쓰이고 있는 한국어 음청력검사어표[5]를 대상으로 하여 이 표에 포함되고 각각 음소 균형을 이루고 있는 것으로 보고된 4개의 목록에 대해 피험자들의 인식 결과에 차이가 있는가를 낱말 인식 실험을 통해 살펴보고, 근접 활성 모델을 근거로 대상 목록의 어휘적 특성인 빈도(frequency), 밀도(density), 어휘적 친숙도(lexical familiarity)를 측정하여 이들이 낱말 인식 점수에 주는 영향을 다중회귀분석을 통하여 살펴보았다. 본 연구의 목적은 이러한 분석을 통하여 말 지각에 미치는 어휘적 특성의 영향력을 입증하고 말 지각 검사의 제작을 위한 이론적 근거를 마련하려는 것이다.

II. 연구 방법

1. 피험자

본 연구의 피험자는 사전검사인 순음청력검사를 통해 정상청력을 가진 것으로 평가된 20대의 대학생 40명이다.

2. 낱말 듣기 검사

한국어음청력검사어표를 대상으로 낱말 듣기 검사를 실시하였다[5]. 한국어음청력검사검사는 50개의 낱말로 이루어진 4개의 목록, 총 200개의 단음절 낱말로 구성되어 있고 이비인후과에서 청각장애의 진단과 평가를 위해 가장 많이 사용하고 있는 검사이다. 피험자는 방음 부스 안에서 STARKEY AA10 청력검사기를 통해 CD에 녹음된 자극을 들었다. 자극은 순음청력검사 결과에서 측정된 피험자의 청력역치보다 0-7-14-21dB 높은 수준의 강도로, 피험자의 정면에 위치한 스피커를 통해 제공되었다. 피험자는 스피커를 통해 들리는 CD의 낱말 목록을 한 낱말마다 받아 적었다. 최대 21dB 이하의 청각자극을 제시한 것은 정상청력인은 20dB 정도의 강도에서 소리를 100% 지각할 수 있는데 [6], 본 연구에서는 정상청력을 가진 피험자에게 제한

된 청각자극을 제공하여 청각장애인의 듣기와 같은 상황을 연출하고자 했기 때문이다[7]

3. 낱말의 어휘적 특성

본 연구에서 살펴보고자 하는 어휘적 요인은 빈도(frequency), 밀도(density), 어휘적 친숙도(lexical familiarity)이다. 낱말의 빈도는 대개 일정한 규모의 말뭉치에서 출현하는 빈도를 말한다. 본고에서는 한국어 낱말의 사용 빈도 자료를 말뭉치 분석 통계 자료에서 얻었다[8]. 이 자료를 이용한 이유는 장르별 균형을 어느 정도 맞추고 있으며 실제 언어 사용 양상을 반영하고 있으며, 여러 기존 연구에서도 이미 기초 자료로 채택되었기 때문이다. 이용한 자료는 150만 어절 규모의 형태 분석 말뭉치로, 주로 문어 중심으로 구성되었다. 이 말뭉치는 형태 정보 주석이 이루어져 있으므로, 품사별로 빈도 정보를 추출할 수 있다.

낱말의 유사성은 근접어(neighbor)의 수로 측정할 수 있다. 근접어는 목표어와 음소 하나의 차이로 구별되는 낱말을 말한다. 이러한 근접어를 많이 갖고 있는 낱말은 밀도가 높다고(dense) 할 수 있고, 근접어가 드문 낱말은 밀도가 낮다고(sparse) 할 수 있다[3]. 본고에서는 빈도 정보를 얻은 자료에서 밀도 정보도 계산하는 방법을 택했다. 말뭉치에 나타난 낱말의 유형은 실제로 사용된 언어라는 점에서 사전보다도 언어 현실을 더 잘 반영할 수 있다는 장점이 있으며, 중사전 정도의 낱말 유형수가 확보되므로 사전을 이용한 분석과도 규모면에서 비교 가능하다. 단음절어의 경우 한국어의 음절 구조에 따라 네 가지 종류, 즉 초성이나 종성에서 분절음의 삭제로 만들 수 있는 근접어, 초성대치 근접어, 중성대치 근접어, 종성대치 근접어로 유형을 나누어 검색하였다.

한편, 빈도가 높은 낱말이라고 해서 일상적으로 자주 쓰이는 친숙한 것이라고는 할 수는 없다. 어휘적 친숙도는 이런 주관적인 평가를 반영한 것이다. 이 개념은 정상 청력을 가진 성인화자의 주관적 평가를 1점부터 7점까지의 척도로 매긴 것이다. 1점은 전혀 모르는 낱말, 4점은 아는 낱말이지만 그 의미를 정확하게는 모르는 낱말, 7점은 친숙도가 매우 높은 낱말을 나타낸다[9]. 본 연구에서는 29명의 대학생 화자를 피험자로 하여 이들에게 7점 척도로 대상 어휘의 친숙도를 측정하는 설문조사를 실시하였다.

4. 분석 방법

총 200개 낱말 문항에 대해 피험자 40명 중 빠르게 지각한 사람의 수를 낱말 인식 점수로 정하고, 음소균형을 맞추어 제작된 4개 낱말 목록 간의 인식 점수의 차이를 살펴보기 위하여 분산분석을 실시하였다. 또한

각 항목의 인식에 대한 어휘적 특성의 영향력을 분석하기 위하여 각 낱말의 빈도, 밀도, 어휘적 친숙도를 독립변수로 하여 다중회귀분석을 실시하였다.

III. 분석 결과

1. 목록 간의 차이

한국어음청력검사의 4개 목록에 대한 빈도, 밀도, 친숙도의 서술통계결과는 <표 1>과 같다.

<표 1> 한국어음청력검사의 어휘적 특성

	빈도		밀도		친숙도	
	평균	표준편차	평균	표준편차	평균	표준편차
목록1	216.02	303.24	13.88	5.54	6.67	37
목록2	266.20	622.43	13.02	5.49	6.63	38
목록3	268.98	413.82	13.24	5.20	6.62	42
목록4	341.76	745.47	13.08	5.92	6.68	30

평균 빈도에서 가장 높은 값을 보인 목록은 목록4이며, 밀도는 목록2가 가장 낮았고, 친숙도는 목록4가 가장 높았다. 그러나 목록 간의 빈도, 밀도, 친숙도는 분산분석 결과 통계적으로 유의한 차이를 보이지 않았다.

<표 2> 4개 목록에 대한

백분율로 환산한 낱말 인식 점수(%)의 서술통계와 분산분석 결과

	목록1	목록2	목록3	목록4
평균	53.73	52.34	64.44	54.24
표준편차	29.79	30.16	25.92	30.10
F = 7.94*				

**P<.01

검사표의 네 가지 낱말 목록에 대해 피험자들이 보인 백분율로 환산한 낱말 인식 점수의 서술통계와 분산분석 결과는 <표 2>와 같다. 백분율로 환산한 낱말 인식 점수(%)는 목록 간에 유의한 차이를 보였고 사후분석 결과 목록3의 낱말 인식 점수는 다른 목록들의 인식 점수와 모두 유의한 차이를 보였다.

2. 문항에 대한 어휘적 특성의 영향

검사표의 200개 낱말에 대해 낱말 인식 점수와 각 문항의 빈도, 밀도, 친숙도에 대한 서술통계를 보이면 다음과 같다.

<표 3> 200개 낱말에 대한 낱말 인식 점수, 빈도, 밀도, 친숙도의 서술통계

	낱말 인식 점수	빈도	밀도	친숙도
평균	22.53	274.43	13.35	6.65
표준편차	6.05	547.98	5.48	37

각 낱말에 대한 인식 점수의 평균은 22.53, 빈도의 평균은 274.43, 밀도의 평균은 13.35, 친숙도의 평균은 6.65이다.

낱말 인식 점수에 영향을 주는 요인을 분석하기 위한 회귀분석의 결과는 <표 4>와 같다.

<표 4> 회귀분석결과

	제공함	자유도	평균제곱	F
빈도	166.27	1	166.27	4.61*
간차	7101.323	198	36.04	
합계	7267.598	199		

* p<.05

단계적 투입의 분석 결과 빈도만이 낱말 인식 점수에 유의한 영향을 주는 요인이었으며 밀도와 친숙도는 통계적으로 유의한 영향을 주지 못한 것으로 나타났다.

IV. 논의 및 결론

기존 연구[5]에서 연구자는 음소 균형을 맞추어 제작한 4개의 검사목록을 피험자에게 들려주고 낱말 인식 점수를 측정하였다. 그 결과 소리의 강도를 5dB, 10dB, 15dB, 20dB, 25dB로 다르게 제시하였을 때, 5가지 상황 모두에서 목록3은 일관되게 다른 목록보다 높은 점수를 보였다[5]. 이러한 결과는 본 연구에서도 동일하게 나타나, 목록3은 자극의 강도에 관계없이 일관되게 높은 점수를 보였다. 네 가지 종류의 목록을 구성한 낱말의 선정 기준은 음소 균형을 맞추는 것이었으므로 이러한 차이는 음소 수준에서 나타난 것은 아닌 것으로 볼 수 있다. 그러나 어휘적 특성에서 4개의 목록은 서로 차이를 보였다. 본 연구에서 산출한 빈도, 밀도, 친숙도의 분석 결과 목록3의 빈도 평균은 가장 높은 빈도를 보이는 목록은 아니었으나 높은 빈도를 보인 목록4에 비해 표준편차가 작았다. 즉 목록4는 평균 빈도는 높았으나 극한값에 속하는 최고빈도 낱말이 포함되어 평균값이 상승하였고 반면 목록3은 고빈도의 낱말이 고르게 분포되어 있었다. 결과적으로 선행 연구

와 본 연구에서 나타난 검사 목록 간 낱말 인식 능력의 차이는 빈도라는 어휘적 특성에 따라 나타난 결과로 해석할 수 있다.

어휘적 특성의 영향력을 평가하기 위한 회귀분석의 결과에서도 빈도는 유의한 변인으로 측정되었다. 빈도, 밀도, 친숙도를 독립변수로 하여 분석한 결과 빈도는 낱말 인식 능력에 영향을 주는 요인으로 분석되었다. 따라서 기존 연구에서 설명할 수 없었던 목록 간의 차이는 이와 같은 어휘적 특성, 특히 빈도 효과로 설명이 가능하다.

한편, 영어의 경우에도 와우이식을 받은 아동의 음성언어 인식 능력에 어휘적 특성이 영향을 주며, 낱말의 밀도와 낱말 자체의 사용 빈도 모두 말 지각에 영향을 준다는 결과가 보고되었다[3][10] 그런데 빈도와 밀도와 같은 어휘적 특성이 모두 단음절 낱말 인식에 의미 있는 영향을 주는 영어와는 달리, 한국어 단음절 낱말에서는 빈도 효과만 의미 있는 것으로 나타났다.

이는 한국어 단음절어의 특성에서 그 원인을 찾을 수 있다. 한국어 단음절 명사 중 상당수는 자립 형식으로 쓰이기보다는 다른 단위와 결합하여 쓰이는 등의 종적인 경우가 많다. 특히 사전에서는 한자어 접사를 명사로도 등재한 경우가 많다. 즉, 우리말 단음절어는 한 낱말로 자립해서 사용되는 경우도 있지만, 적지 않은 경우 다른 낱말과 결합하는 것이 더 자연스러운 것이다. 어휘적 친숙도의 측면에서도 이런 단음절어의 특성이 드러난다. 영어의 관련 연구에서 단음절어의 친숙도는 6.9에서 7 정도로 매우 높게 나타나지만[3], 한국어 단음절 낱말의 경우 친숙도가 6.65 정도로 낱말에 따라서는 잘 사용하지 않아 친숙도가 낮은 경우가 나타나는 것이다. 잘 사용되지 않는 단음절 낱말의 경우 낱말 유사성을 측정할 때 제검토할 필요가 있다.

한국어 단음절 낱말의 인식에서의 또다른 문제는 중의성이다. 한국어의 단순한 음절 구조는 필연적으로 매우 많은 동형이의어의 존재를 가져올 수밖에 없다. 또한 음절말에서 일어나는 중화와 같은 음운론적 현상이 단음절 낱말의 인식에 영향을 줄 가능성도 있다.

따라서 향후 듣기 능력을 측정하기 위한 낱말 인식 검사를 단음절은 물론, 다음절로 확장할 필요가 있다. 실제로 기존 연구에서도 다음절 단어 인식이 단음절 단어 인식보다 유의미하게 높았으며, 따라서 청각장애 아동의 지각적 과정을 평가할 때는 다음절어 인식 검사의 효용성이 높음이 지적되고 있다[3].

지금까지 한국어음청력검사어표를 대상으로 하여 어휘적 특성이 말 지각에 주는 영향을 살펴보았다. 단음절 낱말의 인식에는 빈도(frequency)가 의미 있는 영향을 주었고, 밀도나 친숙도는 큰 영향을 주지 못했다. 빈도와 같은 어휘적 특성이 말 지각에 영향을 주

므로 청각장애인의 말 지각 능력을 평가하는 검사의 개발에는 이를 고려할 필요가 있다. 한편, 한국어 단음절 단어의 인식에 빈도 이외의 어휘적 특성이 영향을 주지 못하는 이유를 추정하였으나, 이는 다음절 낱말의 인식에 대한 어휘적 특성의 영향에 대한 고찰로 확장되어 검증되어야 하며, 추후 연구 과제로 남긴다.

참고문헌

- [1] C. Mackersie, "Test of speech perception abilities". *Current Opinion of otolaryngology Head and neck Surgery*, 10, pp392-397, 2002.
- [2] P. Jusczyk, P. Luce, "Speech perception and spoken word recognition: Past and present", *Ear & Hearing*, 23, pp.2-40, 2002
- [3] K. Kirk, D. Pisoni, & M. Osberger, "Lexical effects on spoken word recognition by pediatric cochlear implant users", *Ear & Hearing*, 16, pp.470-481, 1995.
- [4] P. Luce, D. Pisoni, & S. Goldinger, "Similarity neighborhoods of spoken words", In G. T. M. Altmann(Ed.) *Cognitive models of speech processing: Psycholinguistic and computational perspectives*. Cambridge, MA: MIT Press, 1990.
- [5] 함태영, "한국어음청력검사어표의 명료도검사성적에 관한 연구", *가톨릭의대논문집*, 제5.6호, pp.31-38, 1972.
- [6] L. Humes, Psychoacoustic considerations in clinical audiology, In J. Katz(Ed.), *Handbook of Clinical Audiology*, pp.56-72, Baltimore: Williams & Wilkins, 1994.
- [7] D. Lum, L. Braida, "perception of speech and non speech sounds by listeners with real and simulated sensorineural hearing loss", *Journal of Phonetics*, 28, pp.343-366, 2000.
- [8] 김홍규, 강범모, *한국어 형태소 및 어휘 사용 빈도의 분석*, 서울: 고려대학교 민족문화연구원, 2000.
- [9] N. Large, D. Pisoni, "Subjective familiarity of words: analysis of Hoosier mental lexicon" *Research on Spoken Language Processing* Progress report No. 22. Indiana University, pp215-231, 1998.
- [10] P. Luce, D. Pisoni, "Recognizing spoken words: The neighborhood activation model", *Ear & Hearing*, 19, pp.1-36, 1998.