

# 한국어 단어재인의 이웃(neighborhood)단위<sup>1)</sup>

조혜숙, 남기춘  
고려대학교 심리학과

## What is the neighbors of a word in Korean word recognition?

Cho Hye Suk, Nam Ki Chun  
Department of Psychology, Korea University  
E-mail : [tcoconut@hanmail.net](mailto:tcoconut@hanmail.net)  
[kichun@korea.ac.kr](mailto:kichun@korea.ac.kr)

### Abstract

The purpose of this paper is to investigate the unit of neighbor of Korean words. In English, a word's orthographic neighborhood is defined as the set of words that can be created by changing one letter of the word while preserving letter positions. For example, the words like pike, pole, and tile are all orthographic neighbors of the word 'pile'. In this study, 2 experiments were performed. In these experiments, 4 conditions of prime were included: primes sharing first letter of first syllable(1), first syllable(2), first syllable and the first letter of second syllable with target(3) and with no formal similarity with target(4). In Exp.1, RT was shortest in condition 3. In Exp.2, condition 2 had the shortest RT. We came to the conclusion that in Korean, a word's neighbor is words that share at least one syllable with the word.

### I. 서론

단어 재인에 관한 많은 여러 연구들이 어휘접근 (lexical access) 에 초점을 두어 왔고 많은 모델들이 이 과정에 영향을 주는 요인들에 대해서 조금씩 다른 설명을 해왔다. 가장 많이 논의가 되었던 요인은 빈도인데 여러 선행연구결과로부터 고빈도 단어가 저빈도 단어보다 더 빨리 처리된다는 것이 일반적으로 받아들

여지고 있다. 이러한 빈도효과(frequency effect)를 모델마다 다르게 설명하는데 serial search model(Forster, 1976)에서는 심성어휘집에 표제어가 빈도에 따라 정리되어 있다고 가정한다. 감각자극 (sensory input) 과 올바른 표제어를 매치시키는 것은 순차적으로 진행되어서 가장 가까운 고빈도 단어부터 시작한다. 따라서 고빈도 단어들은 저빈도 단어들보다 더 빨리 재인된다. [1]

Activation-based model로 대표적인 것이 McClelland and Rumelhart(1981)의 Interactive-activation model인데 이 모델에서는 고빈도 단어가 저빈도 단어보다 더 높은 활성화 수준을 가지기 때문에 빈도효과가 발생한다고 본다. 재인에 필요한 역치 수준에 이르기 위해서 고빈도 단어는 더 적은 활성화가 필요하기 때문에 저빈도 단어보다 더 빨리 재인된다는 것이다. [2]

단어재인에 영향을 주는 요인이 단어빈도외에 단어의 이웃(neighbor)인데 영어에서 이웃(orthographic neighborhood)이란 단어의 한 문자를 바꾸어서 만들 수 있는 단어들을 말한다. 예를 들어 pile 의 이웃은 pike, pine, pole, tile 이 된다. 선행연구들에 따르면 이웃크기(neighborhood) 와 이웃빈도(neighborhood frequency) 또한 단어재인에 영향을 준다. [3]

먼저 이웃크기인데 Forster의 serial search model 로 이웃크기가 단어재인에 어떤 영향을 줄지 예측해볼 수 있다. 이 모델에서는 lexicon은 빈도순서로 조직화되어 있다고 가정한다고 하였다. 즉 목표단어를 보고 그 단

어와 맞는 표제어를 심성어휘집에서 빈도순으로 찾아서 비교하여 맞으면 거기서 끝나고 아니면 다음으로 빈도가 높은 것을 체크하는 방식이다. 이 때 이웃크기는 어휘접근의 속도에 영향을 주는데 이웃크기가 클수록 탐색해야 할 목록이 많으므로 어휘접근에 걸리는 시간이 길어질 것이다.

Coltheart 등(1977)은 LDT 이용하여 이웃크기 효과를 검사하였는데 이웃크기효과가 단어에서는 없었고 비단어에서 억제효과가 있었다. 그러나 Andrews(1989)의 결과는 이와 다른데 이웃크기효과가 단어에서 있었고 단어빈도와 상호작용함을 보였다. 이웃크기가 큰 저빈도 단어는 어휘판단과제와 명명과제에서 이웃크기가 작은 저빈도 단어보다 반응시간이 더 짧았다. 반면 고빈도 단어에서는 이웃크기효과가 거의 없었다. Andrew의 결과는 serial search model 와 맞지 않는데 Andrews 는 이 결과를 Interactive-activation model 에 포함된 상호활성화기제(reciprocal activation mechanism) 로 설명하였다. 즉 이웃이 많은 단어는 sublexical constituents 로부터 더 많은 reciprocal activation을 받기 때문에 이웃이 많은 저빈도 단어가 적은 저빈도 단어에 비해서 반응시간이 짧다고 하였다. 반면 고빈도 단어는 저빈도 단어보다 높은 활성화 수준(higher resting activation level)을 가지기 때문에 reciprocal activation 의 영향을 덜 받아서 이웃크기효과가 없는 것이라고 보았다. [4],[5]

이웃크기 효과에 대한 Andrews 의 설명은 Grainger 와 동료들(1989)의 견해와 조금 다른데 그들은 단어재인에서 중요한 이웃변인은 이웃의 크기가 아니라 이웃의 빈도라고 보았다. 그들은 고빈도의 이웃을 가진 단어들이 그렇지 않은 단어들에 비해서 더 느리게 처리된다고 예측하고 이것을 serial search model과 interactive-activation model로 다음과 같이 설명하였다. serial search model 에서 고빈도 이웃은 어휘접근을 지연시키는데 왜냐하면 이 단어들이 먼저 검사되기 때문이다. 따라서 고빈도의 이웃 수가 많을수록 시간이 길어진다. Interactive-activation model 로는 고빈도 이웃을 가진 단어들이 더 느리게 처리되는 이유를 lexical nodes 간의 억제 때문이라고 설명하였다. 목표 단어에 의해서 활성화된 한 이웃 단어를 A라 하면 A는 또한 자신의 이웃들, 목표단어를 포함한, 을 억제한다. A는 고빈도 단어이므로 높은 수준의 활성화를 가지게 되고 저빈도 이웃보다 훨씬 강한 억제를 하게 된다. 따라서 이러한 억제효과 때문에 고빈도 이웃을 가진 단어들은 어휘접근이 느려진다고 하였다. [6]

위와 같이 어휘접근에 영향을 주는 이웃변인에 대한 결과들간에 논쟁이 많은데 한국어에서는 아직 이웃크기나 이웃빈도에 관해서 행해진 선행연구가 없다. 또

한 한국어에서도 어떤 단어의 철자를 바꿈으로써 생길 수 있는 단어들을 그 단어의 이웃이라고 볼 수 있는지 이웃에 대한 정의가 아직 없는 상태이다.

본 연구에서는 한국어의 이웃크기 효과, 이웃빈도효과를 알아보기 위한 선행연구로써 먼저 한국어 단어의 이웃이 어떤 단위인지를 알아보기 위해서 수행되었다. 한국어는 글자들이 옆으로 늘어서 있는 영어와는 달리 두 글자가 모여서 또는 두 글자가 모이고 그 아래에 받침 하나가 더해지는 모양을 가진다. 이는 영어 단어의 시각적 모양과는 많이 다른 형태이다.

## II. 실험 1

### 2.1 실험방법

#### (1) 실험자극

실험자극은 2음절의 받침이 없는 단어들(자음-모음-자음-모음)로 제한하였다. 하나의 목표단어에 대해서 4가지 조건을 만들었다: 1) 목표단어와 낱자 하나만 같고 다른 글자는 다른 조건, 2) 첫 번째 음절만 같고 나머지는 다른 조건, 3) 첫 번째 음절과 두 번째 음절의 초성이 같고 나머지는 다른 조건, 4) 목표 단어와 자음이나 모음을 하나도 공유하지 않는 단어(통제조건). 예를 들어 '도로' 가 목표단어인 경우 조건 1은 '두태', 조건 2는 '도해', 조건 3은 '도래', 조건 5는 '패스'이다. 점화자극과 목표단어와의 간격은 150ms 로 하였다. 한 조건이 단어 22개, 비단어 22개로 구성되었는데 단어는 고빈도 단어 11개, 저빈도 단어 11개로 구성되었다. 연대 빈도 사전을 참고하였으며 고빈도 단어의 평균은 2825, 저빈도 단어의 평균은 44 이었다. 실험자극은 피험자에게 랜덤하게 제시되었다.

#### (2) 실험절차

실험전에 연습시행을 실시하였다. 연습시행에서 제시된 자극은 실험자극에 포함되지 않은 것으로 모두 15 trial로 구성되었다. 연습시행이 끝나면 본 실험으로 들어갔는데 과제는 컴퓨터 모니터(15인치 SAMTRON 77E)에 나타나는 낱말(글자 크기 30)을 보고 단어인지 아닌지를 판단하여 키보드의 키를 누르는 것이었다.

#### (3) 실험참가자

고려대학교 재학생 104 명이 실험에 참가하였다. 한 조건당 각각 26 명이 참가하였다.

## 2.2 실험 결과

반응시간의 중앙값을 사용하였다. 2표준편차 이상의 극단값과 300ms 이하의 값은 제외하였다. 분산분석결과 빈도효과가 유의하였다( $F_{(1, 87)} = 260.755, p < .05$ ). 고빈도 단어 반응시간의 중앙값은 605ms, 저빈도 단어 반응시간의 중앙값은 717 ms 였다. 그러나 조건들간의 차이는 유의하지 않았다. 조건들간의 반응시간(중앙값)과 표준편차를 표 1 에 제시하였다.

표 1. 실험1의 반응시간(ms)과 표준편차(괄호안)

목표단어	조건	반응시간
도로	조건 1- 두태	630(151)
	조건 2- 도해	632(137)
	조건 3- 도래	622(148)
	조건 4- 패스	633(155)

## III. 실험 2

### 2.1 실험방법

점화자극과 목표단어와의 간격이 750ms 이라는 점 외에 실험자극과 실험절차는 실험 1과 동일하였다. 실험 1에 참여하지 않은 고려대학교 재학생 104명이 참여하였다.

### 2.2 실험 결과

반응시간의 중앙값을 사용하였다. 2표준편차 이상의 극단값과 300ms 이하의 값은 제외하였다. 분산분석결과 빈도효과가 유의하였다( $F_{(1, 87)} = 1118.168, p < .05$ ). 고빈도 단어 반응시간의 중앙값은 578ms, 저빈도 단어 반응시간의 중앙값은 685ms 이었다. 그러나 조건들간의 차이는 유의하지 않았다. 조건들간의 반응시간(중앙값)과 표준편차를 아래 표에 제시하였다.

표 2. 실험2의 반응시간(ms)과 표준편차(괄호안)

목표단어	조건	반응시간
도로	조건 1- 두태	615(145)
	조건 2- 도해	603(138)
	조건 3- 도래	614(152)
	조건 4- 패스	611(151)

## IV. 논의 및 결론

실험 1에서 빈도효과는 있었으나 조건들간의 차이는 통계적으로 유의하지 않았다. 그러나 반응시간을 살펴보면 조건 1(낱자 하나만 같은 조건) 과 조건 2(목표단어와 첫음절이 같은 조건), 조건 4(목표단어와 형태적으로 유사하지 않은 조건)의 반응시간은 별로 차이가 없는 반면 조건 3(목표단어와 첫음절, 두번째 음절의 낱자 하나가 같은 조건) 의 반응시간은 10ms 정도 차이가 남을 알 수 있다. 즉 목표단어와 형태적으로 가장 유사한 조건에서 반응시간이 가장 빨랐다.

실험 2는 실험 1과 조금 다른 양상을 보였다. 조건2의 반응시간이 가장 짧았고 나머지 조건의 반응시간은 비슷하였다. 조건 3의 반응시간이 가장 빠르지 않은 이유는 점화자극과 목표단어간의 시간간격 때문인 것으로 생각된다. 즉 자극간 간격이 150ms 일때는 점화자극이 빨리 사라지기 때문에 자극의 모양만 보고 목표단어를 보게 되지만 750ms 일때는 점화자극에 대한 모든 정보가 충분히 활성화될 수 있다. 따라서 점화자극의 활성화가 목표단어를 억제했을 가능성이 있다. 조건 3은 목표단어와 가장 유사한 조건이었으므로 활성화가 큰 만큼 억제효과 또한 가장 컸을 것이다.

본 연구는 한국어 단어의 이웃이 무엇인지를 알아보기 위한 실험으로 2음절 단어의 경우 목표단어와 첫음절이 같거나 첫음절과 두 번째 음절의 첫 번째 낱자가 같은 조건에서 반응시간이 가장 빠르다는 결과를 얻었다. 그러나 첫글자만 같은 조건에서는 반응시간이 빠르지 않았다. 이러한 결과는 한국어에서 단어의 이웃은 최소한 한 음절이 같아야 함을 말한다. 그리고 한국어 단어에서는 이웃이 영어에서의 정의와 다르게 내려져야 함을 의미한다. 이러한 결과를 한국어 단어에 일반화하기 위해서는 받침이 있는 2음절 단어에 대한 후속 연구도 필요하다고 본다. 또한 이러한 결과들을 바탕으로하여 한국어의 이웃크기효과와 이웃빈도효과에 대한 연구도 이루어질 수 있을 것이다.

## 참고문헌

- [1] Forster, K. I, "Accessing the mental lexicon," In R. J. Wales & E. W. Walker (Eds.), *New approaches to language mechanism*, pp. 257-287, 1976
- [2] McClelland, J. L., & Rumelhart, D. E, "An interactive-activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, Vol.88,

pp.375-407, 1981

- [3] Chris R. Sears, Yasushi Hino, and Stephen J. Lupker, " Neighborhood size and neighborhood frequency effects in word recognition," *Journal of experimental psychology*, Vol.21, No.4, pp.876-900, 1995
- [4] Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D, "Access to the internal lexicon," *Attention and performance*, vol.6, pp.535-555
- [5] Andrews, S, " Frequency and neighborhood effects on lexical access: Activation or search?," *Journal of experimental psychology: Learning, Memory, and Cognition*, vol.15, pp.802-814, 1989
- [6] Grainger, J., O'Regan, J. K., Jacobs, A. M., & Segui, J, "On the role of competing word units in visual word recognition: The neighborhood frequency effect," *Perception & Psychophysics*, vol.45, pp.189-195, 1989.