인지과학, 제34권 제4호
Korean Journal of Cognitive Science
2023, Vol. 34, No. 4, 277～297
http://dx.doi.org/10.19066/cogsci.2023.34.4.002

한글 두 글자 단어와 비단어의 어휘판단에 글자 빈도, 글자 유형, 받침이 미치는 영향: KLP 자료의 분석

신 명 석　　박 창 호†
전북대학교 심리학과

본 연구는 한국어 심성어휘 데이터베이스(KLP-DB)의 분석을 통해 글자 빈도, 글자의 모음 유형, 받침 유무 등 글자 수준 정보가 두 글자로 된 단어와 비단어의 어휘판단에 어떤 영향을 주는지를 알아보고자 하였다. 반응시간과 오반응률에 대한 위계적 회귀분석을 실시한 결과 단어의 어휘판단에는 단어빈도가 중대한 영향을 미치지만, 첫째 글자의 빈도, 첫째 글자와 둘째 글자의 모음 유형과 받침 유무와 같은 글자 속성이 영향을 미쳤고, 두 글자의 모음 유형의 조합 및 둘째 글자의 빈도와 받침 유무의 조합도 영향을 주었다. 비단어의 어휘판단에는 첫째 글자와 둘째 글자의 빈도, 첫째 글자의 모음 유형, 첫째 글자와 둘째 글자의 받침 유무와 같은 글자 속성이 영향을 미쳤고, 두 글자의 사용빈도의 조합, 모음 유형의 조합, 및 첫째 글자의 빈도와 받침의 조합도 영향을 주었다. 단어빈도는 단어의 어휘판단에서 강력한 영향을 미쳤으며, 글자 속성은 단어보다 비단어의 판단에서 더 일관적인 영향을 미쳤다. 본 연구의 결과는 어휘판단과제에서 단어와 비단어 목록의 구성 및 반응시간의 해석에 글자 속성의 문제를 충분히 고려해야 함을 가리킨다. 글자 속성의 효과에 대한 이해는 단어 재인 과정의 이해에도 기여할 것이다.

주제어 : 한글, 비단어, 글자 빈도, 글자 유형, KLP

† 교신저자: 박창호, 전북대학교 심리학과, 전북 전주시 덕진구 백제대로 567, 인문사회관 720호
연구분야: 인지심리학
E-mail: finnegan@jbnu.ac.kr

단어는 문장의 기초 단위이므로, 단어 재인(word recognition)에 대한 많은 연구가 수행되어 왔다. 단어의 처리과정은 자소와 음소의 대응 정도, 다의어 여부, 인접 단어와의 유사성 등 여러 요인에 따라 영향을 받으며, 이런 요인들은 사용되는 언어에 따라 달라진다. 특히 한글 단어는 몇 개의 시각적으로 조직된 글자의 조합으로 만들어지며, 각 글자가 단독 단어 혹은 단어 내의 유의미한 성분으로 사용되기도 하는 특징이 있다(Simpson & Kang, 2004). 이 때문에 같은 알파벳 언어로 간주되지만, 낱자가 일렬로 배치되는 영어 단어에 대한 연구 결과가 한글 단어 처리에 곧이곧대로 적용되지 않을 가능성이 있다(Lee & Taft, 2011). 본 연구는 한글 글자의 속성이 단어 재인에 영향을 미칠 가능성을 탐구하기 위해, 종종 고려되는 글자 빈도 외에도 글자(모음) 유형과 받침 유무와 같은 시각적 속성이 어휘판단에 미치는 영향을 분석하고자 한다. 이 연구를 위해 한글 단어와 비단어에 대한 어휘판단시간과 오반응률을 수집해 놓은 한국어 심성어휘 데이터베이스(KLP-DB) (이광오 외, 2017)를 사용할 것이다.

한글의 음절은 글자 단위와 상응하므로(표음 심도가 얕음), 글자 단위의 속성이 한글 단어나 글자의 처리에 영향을 줄 수 있다(이광오, 1995). 글자의 빈도에 따라서 단어우월효과가 달라졌으며(고은태, 배성봉, 이윤형, 2022), 사용 빈도가 높은 한 글자 단어의 어휘판단시간이 더 빨랐으며(배성봉, 박권생, 이혜원, 이광오, 2016), 명명 시간도 더 빨랐다(이광오, 1993). 그러나 단어의 첫째 글자의 빈도가 높을 때 어휘판단시간이 증가한 결과도 있는데(권유안, 2012), 이 효과는 다른 언어에서도 관찰되었다(Carreiras, Alvarez, & DeVega, 1993). 이처럼 한 글자와 단어의 첫째 글자에서 음절 빈도 효과가 상이하게 관찰된 결과는 글자 수준 정보와 단어 수준 정보. 그리고 둘 간의 경합을 고려함으로써 설명될 수 있다. 예를 들어 친숙한 단어에서 낱자 재인이 힘들었는데(Healy, 1980), 이것은 단어에 대한 전역 처리로 인해 글자에 대한 처리가 억제되거나 약화될 수 있음을 가리킨다. 그렇다면 단어 여부가 불분명한 한 글자에서의 글자 빈도 효과와 두 글자 이상 단어에서의 글자 빈도 효과가 상이할 것이라고 예측할 수 있을 것이다. 두 음절 자극에 대해서는 첫째 글자의 빈도 효과와 함께 둘째 글자의 빈도 효과를 고려해볼 수 있는데, 스페인어 연구에서 마지막 음절의 빈도가 높은 단어가 더 빠르게 명명되는 촉진적인 글자 빈도 효과가 관찰되기도 하였다(Levelt & Wheeldon, 1994). 한글에서의 음절빈도효과는 불분명한데(권유안, 2020), 글자가 단어 속에서 사용되면서 단어 빈도와 글자 빈도의 상호작용이 발생하기 때문일 가능성도 있다. 단어 수준의 정보(예: 단어 빈도)와 글자 수준의 정보(예: 글자 빈도)의 상호작용을 보기 위해서는 어휘판단과제에서 두 변인을 함께 조작해야 하는데, KLP-DB에서 두 변인이 무선적으로 조작된 자료를 얻을 수 있다.

글자 빈도 효과는 또한 비단어에서도 관찰된다. Carreiras, Alvarez와 DeVega(1993)는 스페인어 비단어를 소리내어 읽을 때 음절 빈도가 높을수록 명명 시간이 짧아지는 촉진적인 빈도 효과를 관찰하였으나(Carreiras & Perera, 2004에서 반복 검증됨), 어휘판단과제에서는 그 반대로 어휘판단시간이 길어지는 억제적인 빈도 효과를 관찰하였다. 이는 과제에 따라서 글자 빈도 효과가 다르

게 나타날 수 있음을 보여준다. 어휘판단과제에서의 글자 빈도 효과는 글자의 처리가 단어의 처리에 영향을 줄 가능성을 보여준다. 그렇다면 비단어가 가진 속성도 단어의 어휘판단에 영향을 줄 수 있을 것이다. Stone과 Van Orden(1993)은 어휘판단과제에서 비단어의 유형(발음 불가능, 발음 가능, 단어와 동음 비단어)에 따라 단어의 어휘판단시간과 오반응률이 달라지는 결과를 얻었다. 배성봉(2018)은 비단어가 한자로 이루어지면 단어의 어휘판단 시간과 오반응률이 증가하는 결과를 관찰하였다. 이처럼 비단어가 가진 속성도 어휘판단에 영향을 준다면, 비단어 내의 글자 속성에 대한 더 자세한 검토가 필요할 것이다. 본 연구는 어떤 글자 속성이 비단어의 판단에 영향을 미치는지를 살펴보고자 한다.

한글 글자의 시각적 속성도 단어나 글자의 처리에 영향을 줄 수 있다(박권생, 1995). 한글 글자는 초성, 중성, 종성이 네모 형태로 조직되므로, 대체로 외형이 서로 비슷하다. 그러나 글자 속 모음의 종류나 받침의 유무에 따라 글자의 형상이 조금씩 다르다. 예컨대 모음에 긴 수평선이 있는 횡모음 글자(예: '고')와 긴 수직선이 있는 종모음 글자(예: '가')는 서로 다른 형상을 갖는 것으로 보인다. 흔히 글자 유형은 글자 속 모음의 종류, 그리고 받침의 유무에 따라 구별되는 글자의 시각적 유형을 말한다. 글자 유형은 글자 내 낱자의 시각적 집단화가 다른 방식으로 이뤄지게 하여 자모나 받침의 처리에 영향을 줄 수 있다. 종모음 유형에서는 첫 자음(초성) 중심의 집단화가 일어나고, 횡모음 유형에서는 모음(중성) 중심의 집단화가 일어나는 것으로 보인다(김정오, 1994; 김정오 김재갑, 1992). 또한, 한 글자의 판단에서 종모음 글자는 횡모음 글자보다 더 잘 판단되었다(박창호, 2006). 그리고 어휘판단과제에서 종모음 글자로만 이루어진 자극(단어 혹은 비단어)의 처리가 횡모음 글자로만 이루어진 자극(단어 혹은 비단어)보다 더 빨랐다(민수영, 이창환, 2018). 이상의 연구에서 글자는 단독으로 사용되거나, 혹은 종모음 혹은 횡모음만으로 이루어진 단어만을 사용되었는데, 대체적으로 종모음 유형의 처리가 횡모음 유형의 처리보다 더 빨랐다. KLP-DB에서 사용된 단어들은 종모음, 횡모음 유형의 글자들이 모두 조합되어 있으므로, 첫째 및 둘째 글자 모두에서 글자의 모음 유형이 어휘판단에 미치는 영향을 세부적으로 확인할 수 있다. 기존 연구를 바탕으로 예상해 보면, 두 글자가 모두 종모음이거나 한 글자가 종모음인 조건에서 다른 조건보다 어휘판단 시간이 더 빠를 것이다.

한글 글자는 받침이 있거나(받친글자), 받침이 없다(민글자). 민글자(예: 가, 고)는 받침이 있는 글자(예: 갑, 곳)보다 더 간단하게 보이고, 더 잘 식별될 수 있다. 한 개의 민글자 혹은 받친글자 내 표적 낱자를 탐지하는 시간이 낱자 수가 증가할수록 더 길어졌다(이준석, 김경린, 1989). 김민식과 정찬섭(1989)은 순간 노출된 글자에서 초성보다 받침의 식별률이 떨어진다는 결과를 얻었다. 김정오와 김재갑(1992)는 단어우월효과 연구에서 받친글자에서 음절이 낱자보다 열등하게 처리된다는 것을 발견하였으며, 박창호(2006)는 한글 글자의 사용여부를 판단하는 과제를 사용한 연구에서 한글 글자의 초성과 중성이 한 단위로 집단화되어 받침보다 먼저 처리된다고 주장하

였다. 이런 연구들을 볼 때, 글자의 받침 유무는 글자의 처리에 영향을 주며, 나아가 단어 혹은 비단어의 판단에도 영향을 주는 것으로 보인다. 그러나 단어에서 받침의 유무는 어휘판단과제와 명명과제의 수행에는 영향을 주지 못하였다(이광오, 1993; 최양규, 1986). 이 결과는 한글 단어가 글자 단위로 처리되기 때문에 각 글자 속의 자모의 수나 받침 유무가 어휘판단이나 명명에 영향을 주지 못하는 것에 기인할 가능성이 있다. 혹은 두 글자 단어의 처리는 단어 여부가 불분명한 한 글자의 처리와 다를 가능성이 있다(이준석, 김경린, 1989). 이상의 연구를 고려할 때 본 연구에서도 단어의 받침 유무는 어휘판단에 영향을 주지 못할 것으로 예상해볼 수 있다. 하지만 비단어에는 어휘속성의 영향이 제한적이므로 글자의 받침 유무가 어휘판단에 영향을 줄 수 있을 것이다. 두 글자 단어와 비단어에서 받침 유무의 효과를 검토하는 것은 단어 내의 받침이 어휘판단에 미치는 영향을 자세히 판단하는 데 도움이 될 것이다.

그동안 어휘판단 연구에서 비단어는 과제 구성에 필요한 채움자극(filler) 정도로 사용되어 왔고, 비단어에 대한 관심은 크지 않았다. 그러나 앞에서 언급하였듯이, Stone과 Van Orden(1993)은 비단어의 유형에 따라서 실험 참가자들의 수행 전략이 변화할 수 있음을 보여주었고, Luker와 Pexman(2010)은 비단어가 단어와 유사해질수록 어휘판단시간이 증가하며 단어빈도효과도 커진다는 것을 보여주었고 이를 단어다움(wordlikeness)으로 설명하였다. 한글에서도 비단어의 유형과 비율에 따라서 단어 자극의 수행이 달라지는 결과가 관찰되었다(배성봉, 2018). 한글 단어와 비단어는 한두 획의 차이로 구별될 정도로 외형적인 차이가 작다. 그래서 단어와 비단어의 판단에 어휘 속성 외에도 글자(모음) 유형이나 받침 유무와 같은 글자 속성이 영향을 미칠 가능성이 높다고 생각된다.

이상에서 논의한 점을 살펴보기 위해 본 연구는 KLP-DB를 분석하여, 두 글자 단어 및 비단어에 대한 어휘판단에 글자 빈도, 글자 유형, 받침 여부와 같은 글자 수준 정보. 그리고 단어 빈도가 어휘판단시간에 미치는 효과를 검토하고자 한다. 선행 연구를 바탕으로 한 예측은 다음과 같다. 글자 속성의 효과는 글자가 단어의 일부인지 아닌지에 따라, 그리고 글자의 위치에 따라 달라질 것이다. 단어의 경우 첫째 글자의 빈도가 증가할수록 어휘판단시간이 증가할 것이다(권유안, 2012). 또한, 종모음 글자(들)이 빨리 처리될 것이다(민수영, 이창환, 2018). 글자(들)의 받침 유무는 단어의 어휘판단시간에 영향을 미치지 못할 것이다(최양규, 1986). 그러나 선행 연구는 주로 단어를 대상으로 하였으므로, 이런 예측이 비단어에는 적용되지 않을 가능성이 높다. 비단어 판단에는 본 연구에서 검토하는 여러 글자 속성이 모두 어느 정도 영향을 미칠 수 있을 것이다. 본 연구는 단어와 비단어의 판단에 어떤 글자 속성(들)이 차별적인 영향을 미치는지를 검토하여, 단어와 비단어의 재인 과정의 특성을 알아보고자 한다.

방 법

재료

한국어 심성어휘 데이터베이스(KLP-DB; 이광오 등, 2017)는 한국어 심성어휘집의 구조와 검색 과정에 대해서 알아보기 위해 연구한 결과로서, 여기에 3만 여개의 단어와 3만 여개의 비단어를 대상으로 수집한 어휘판단 데이터가 있다. 데이터에는 단어의 경우 단어 빈도, 단어 길이, 품사와 어휘판단시간의 평균과 표준편차, 오반응률의 평균과 표준편차가 정리되어 있으며 비단어의 경우 비단어의 길이, 품사와 어휘판단시간의 평균과 표준편차, 오반응률의 평균과 표준편차가 정리되어 있다.

본 연구는 KLP-DB 중 두 글자 단어와 비단어 자료를 이용하고자 하는데, 두 글자 단어가 가장 흔히 사용되고 있고, 단어-비단어의 판단이 비교적 명백하기 때문이다. KLP-DB에서 두 글자 단어와 비단어 자료는 각각 12,881개였다. 총 25,762개의 자극 중에서 “과”와 “광”과 같이 이중모음이 포함된 자극을 분석에서 제외하였으며 평균 오반응률이 20%를 초과하는 자극들도 분석에서 제외하였다. 분석에 사용된 자극은 총 18,456개(단어 : 8,906개, 비단어 : 9,550개)였다.

예측변인

본 연구에서 선택한 글자 속성 변인은 글자 빈도, 모음의 유형, 받침 유무이다. 글자 속성은 두 글자에서 첫째 위치와 둘째 위치를 구별하여 처리하였다. 단어 수준의 변인은 단어 빈도이다. 표 1에 글자 속성별로 단어와 비단어의 빈도 평균이 제시되어 있다. 단어 빈도는 KLP-DB에 기재된 단어 빈도 자료를 사용하였다. 글자 빈도는 국립국어원의 세종 말뭉치 자료를 기반으로 계산되었다. 글자 빈도는 단어에서의 (첫째 혹은 둘째) 위치별로 계산되었는데, 예컨대 둘째 글자의 빈도는 해당 글자가 단어 내에 둘째 글자로 사용된 모든 단어들의 빈도를 합산한 값이다. 모음 유형은 글자의 모음이 종모음(예: ‘가’)인지, 횡모음(예: ‘고’)인지를 가리킨다. 받침 유무는 글자에 받침이 없는 민글자인지, 받침이 있는 받친글자인지를 구별한다.

표 1의 단어 및 글자 빈도 자료는 자연로그 값인데, 원래 값의 범위(단어 빈도: 2 ~ 9,908, 첫째 글자 빈도: 1 ~ 312,077, 둘째 글자 빈도: 1 ~ 365,099)가 너무 크기 때문이다. 단어의 첫째 글자의 빈도 평균은 10.03이었으며 둘째 글자의 빈도 평균은 9.96이었다. 비단어의 첫째 글자의 빈도 평균은 9.10이었으며 둘째 글자의 빈도 평균은 9.30이었다. 글자 유형은 단어에서 첫 글자가 종모음 유형인 자극의 수는 5,659개이었으며 이 중에서 둘째 글자가 종모음 유형인 자극은 3,884개, 횡모음 유형인 자극은 1,775개이었다. 첫 글자가 횡모음 유형의 글자인 자극은 3,247개이었으며 이 중에서 둘째 글자가 종모음 유형인 자극은 2,212개, 횡모음 유형인 자극은 1,035개

〈표 1〉 단어와 비단어별로 본 단어와 글자 빈도의 평균(자연로그 변환), 그리고 모음 유형과 받침 유형의 수 (괄호 안은 표준편차)

			단어	비단어
빈도 (Ln)	단어빈도		4.97 (1.42)	-
	글자빈도	첫째 위치	10.03 (1.56)	9.10 (2.21)
		둘째 위치	9.96 (1.46)	9.30 (1.79)
모음 유형 (개수)	종-종		3,884	4,029
	종-횡		1,775	2,238
	횡-종		2,212	2,112
	횡-횡		1,035	1,171
받침 유무 (개수)	민-민		1,257	850
	민-받친		2,176	2,037
	받친-민		1,977	2,332
	받친-받친		3,496	4,331
	계		8,906	9,550

주1. 종: 종모음 글자; 횡: 횡모음 글자; 민: 민글자; 받친: 받친글자.
2. 빈도는 조사한 수치를 자연로그(Ln)로 변환한 값이다.

이었다. 비단어에서는 첫째 글자가 종모음 유형인 자극은 6,267개이었으며 이 중에서 둘째 글자가 종모음 유형의 자극은 4,029개, 횡모음 유형의 자극은 2,238개이었다. 첫째 글자가 횡모음인 글자의 자극은 3,283개이었으며 이 중에서 둘째 글자가 종모음 유형의 자극은 2,112개, 횡모음 유형의 자극은 1,171개이었다. 글자 받침은 단어에서 첫째 글자에 받침이 없는 자극은 3,433개이었으며 이 중에서 둘째 글자에도 받침이 없는 자극은 1,257개, 받침이 있는 자극은 2,176개이었다. 첫째 글자에 받침이 있는 자극은 5,473개이었으며 이 중에서 둘째 글자에 받침이 없는 자극은 1,977개, 받침이 있는 자극은 3,496개이었다. 비단어에서 첫째 글자에 받침이 없는 자극은 2,887개이었으며 이 중에서 둘째 글자에도 받침이 없는 자극은 850개, 받침이 있는 자극은 2,037개이었다. 첫째 글자에 받침이 있는 자극은 6,663개이었으며, 둘째 글자에 받침이 없는 자극은 2,332개, 받침이 있는 자극은 4,331개이었다.

자료 분석

KLP-DB에서 단어 및 비단어에 대한 반응시간(RT)는 어휘판단 과제에서 실험 참가자가 정반응

을 하였을 때의 반응시간이다. 오반응의 반응시간은 입력되지 않았다. 오반응률은 단어를 비단어로, 비단어를 단어로 판단한 비율이다. 오반응률이 매우 높은 일부 단어/비단어 자료는 분석에 부적합하므로, 본 연구는 정확률이 80% 이상인 자료만을 사용하여 분석하였다.

한글 두 글자 단어와 비단어의 글자 속성이 어휘판단시간과 오반응률을 유의하게 설명하는지 알아보기 위하여 KLP-DB의 어휘판단시간과 오반응률에 대한 위계적 다중회귀분석(hierarchical multiple regression analysis)을 실시하였다. 예측(독립)변인은 자극을 구성하는 첫째 및 둘째 위치에 있는 각 글자의 빈도, 각 글자의 모음 유형(종, 횡), 각 글자의 받침 유무로서 총 6개였다. 단어 자극의 경우에는 단어 빈도도 예측변인에 포함되었다. 글자 및 단어 빈도는 양적 변인이었으며, 글자의 유형과 받침 유무는 범주적 변인이었다. 변인 간 상호작용의 효과를 알아보기 위하여 다음 변인 간 상호작용항을 예측변인에 포함하였다. 즉 첫째와 둘째 글자의 빈도, 첫째와 둘째 글자의 모음 유형, 첫째와 둘째 글자의 받침 유무, 그리고 첫째 글자의 빈도와 모음 유형, 첫째 글자의 빈도와 받침 유무, 둘째 글자의 빈도와 모음 유형, 둘째 글자의 빈도와 받침 유무 등이 포함되었다. 상호작용항은 개별 예측변인들을 투입한 다음 단계에 투입하였다.

자료에 대한 상관 분석과 위계적 회귀분석에 SPSS 20을 이용하였다. 범주형 변인은 더미 코딩을 하여 입력하였다.

결 과

단어와 비단어를 나누어 글자 속성의 영향을 분석하였다. 어휘판단시간의 평균은, 단어의 경우 611ms(SD = 56), 비단어의 경우 673ms(SD = 66)이었으며, 전체 평균은 643ms(SD = 69)이었다. 단어에 대한 오반응률은 4.2%(SD = 4.9)였고, 비단어에 대한 오반응률은 5.3%(SD = 5.5)이었으며, 전체 평균은 4.8%(SD = 5.3)이었다. 반응속도와 정확률의 교환은 발생하지 않은 것으로 보이는데, 이에 대해서는 뒤에서 상술할 것이다. 단어와 비단어별 예측변인 간 상관관계가 표 2에 제시되어 있다. 단어와 비단어가 서로 다른 상관관계를 보인 결과를 살펴보면, 첫째 글자의 빈도와 둘째 글자의 빈도는 단어(r = .156)와 비단어(r = -.188)에서 상반된 유의미한 상관을 보였다. 또한 단어의 첫째 글자 빈도와 첫째 글자의 모음 유형은 유의한 상관을 보였지만(r = -.045), 비단어에서는 유의하지 않았다. 첫째 글자 빈도와 둘째 글자 모음 유형은 단어에서 유의하지 않았지만, 비단어에서는 유의한 상관을 보였다(r = .040). 둘째 글자 빈도는 첫째 글자의 받침 유무의 상관은 단어의 경우 유의하지 않았지만, 비단어의 경우 유의하였다(r = .082). 또한, 첫째 글자와 둘째 글자의 받침 유무 간의 상관은 단어에서 유의하지 않았지만 비단어에서는 유의한 부적 상관이 나타났다(ϕ = -.054). 그리고 첫째 글자 받침 유무와 둘째 글자 모음 유형은 단어에서 유의한 부적 상관이 나타났지만(ϕ = -.035), 비단어에서는 유의하지 않았다.

〈표 2〉 단어(좌하 삼각형 영역)와 비단어(우상 삼각형 영역)의 예측변인 간 상관계수

	1	2	3	4	5	6
1. 첫째 글자 빈도	-	-.188***	-.007	.040***	-.331***	.123***
2. 둘째 글자 빈도	.156***	-	.014	-.079***	.082***	-.346***
3. 첫째 글자 모음 유형	-.045***	-.015	-	.000	-.055***	-.001
4. 둘째 글자 모음 유형	.011	-.101***	.005	-	-.003	-.026*
5. 첫째 글자 받침 유무	-.443***	.002	-.092***	-.035**	-	-.054***
6. 둘째 글자 받침 유무	.037***	-.365***	.006	-.074***	.005	-
7. 단어 빈도	.201***	.212***	-.025*	-.040***	-.072***	-.115***

주 1. 표의 아래쪽은 단어, 위쪽은 비단어의 상관계수이다.

2. 글자 빈도와 단어 빈도를 제외한 모든 변인은 범주형 변인이었다. 빈도 자료에 대해서는 Pearson 상관계수, 범주형 자료에 대해서는 phi 상관계수, 빈도 자료와 범주형 자료 간에는 Point Biserial 상관계수가 계산되었다.

3. * $p < .05$, ** $p < .01$, *** $p < .001$

분석에 사용된 두 글자 자극에 대한 어휘판단시간이 단어 빈도와 글자 빈도, 종과 횡 모음 유형, 받침 유무에 따라 다른지를 독립표본 t-검증으로 살펴보았다(표 3). 단어 빈도와 글자 빈도에서 상위 50%에 속하는 자극은 고빈도 집단으로, 하위 50%에 속하는 자극은 저빈도 집단으로 분류하여 집단별 어휘판단시간과 오반응률 평균을 계산하였다. 단어의 경우, 단어 빈도가 낮은 단어에 비해 높은 단어의 어휘판단시간이 빨랐으며($t(8844.1) = 43.430$, $p < .001$), 오반응률도 더 낮았다($t(7792.9) = 34.456$, $p < .001$). 단어에서 첫째 글자가 고빈도일 때보다 저빈도일 때 어휘판단시간이 더 느렸으며($t(8415.958) = 3.593$, $p < .001$) 오반응률도 더 높았다($t(8904) = 2.326$, $p < .05$). 비단어에서는 반대로 첫째 글자가 고빈도일 때보다 저빈도일 때 어휘판단시간이 빨랐으며($t(8890.4) = -24.852$, $p < .001$), 오반응률도 더 낮았다($t(8093.3) = -17.149$, $p < .001$). 둘째 글자의 빈도는 단어에서 첫째 글자와 마찬가지로 고빈도일 때보다 저빈도일 때 어휘판단시간이 더 느렸지만($t(8234.869) = 2.636$, $p < .01$), 오반응률에서는 그 차이가 유의하지 않았다. 비단어에서는 둘째 글자가 고빈도일 때보다 저빈도일 때 어휘판단시간이 더 빨랐으며($t(8376.228) = -9.955$, $p < .001$), 오반응률도 더 낮았다($t(8167.733) = -11.184$, $p < .001$). 단어에서 첫째 글자와 둘째 글자 모두에서 모음 유형은 어휘판단시간과 오반응률에 차이를 보이지 않았다. 반면에 비단어에서는 첫째 위치에서 종모음 글자보다 횡모음 글자가 어휘판단시간이 느렸으며($t(9548) = -2.597$, $p < .01$), 오반응률도 더 높았다($t(6525.0) = -2.162$, $p < .05$). 둘째 위치에서 이와 반대로 종모음 글자보다 횡모음 글자가 어휘판단시간이 더 빨랐으며($t(9548) = 2.434$, $p < .05$), 오반응률이 더 낮았다($t(7209.6) = 3.926$, $p < .001$). 받침 유무와 관련해서 보면, 단어의 첫째 글자에 받침이 없

〈표 3〉 단어 빈도 집단, 글자 위치별 글자 빈도 집단, 모음 유형, 받침 유무에 따른 판단시간(RT, ms)과 오반응률(ERR, %)의 평균과 표준오차(괄호 안)

		단어(RT)		단어(ERR)		비단어(RT)		비단어(ERR)	
		평균	*t*	평균	*t*	평균	*t*	평균	*t*
단어 빈도	고	588 (.7)	43.430***	2.5 (.08)	34.456***	-	-	-	-
	저	635 (.8)		5.9 (.05)		-		-	
첫째 글자 빈도	고	609 (.8)	3.593***	4.1 (.07)	2.326*	692 (1.0)	-24.934***	6.5 (.07)	-17.149***
	저	614 (.9)		4.3 (.08)		659 (.9)		4.5 (.09)	
둘째 글자 빈도	고	610 (.8)	2.636**	4.2 (.08)	.278	681 (1.1)	-9.955***	6.1 (.09)	-11.184***
	저	613 (.9)		4.2 (.07)		667 (.9)		4.8 (.07)	
첫째 글자 모음 유형	종	611 (.8)	-.048	4.2 (.07)	.817	671 (.8)	-2.597**	5.2 (.07)	-2.162*
	횡	611 (1.0)		4.1 (.08)		675 (1.2)		5.5 (.10)	
둘째 글자 모음 유형	종	611 (.7)	.108	4.2 (.06)	.548	674 (.9)	2.434*	5.5 (.07)	3.926***
	횡	611 (1.1)		4.1 (.09)		670 (1.1)		5.0 (.09)	
첫째 글자 받침	무	607 (1.0)	-5.310***	4.1 (.08)	1.462	669 (1.2)	-3.664***	5.1 (.10)	-2.838**
	유	614 (.8)		4.2 (.07)		674 (.8)		5.4 (.07)	
둘째 글자 받침	무	610 (1.0)	-1.205	4.3 (.09)	2.529*	666 (1.1)	-7.407***	5.0 (.09)	-4.529***
	유	612 (.7)		4.1 (.07)		676 (.8)		5.5 (.07)	

주 1. *t*-검증의 자유도는 모두 1,000 이상의 큰 수이므로(표 1의 글자 수 참조) 표에 제시하지 않았다.
2. 단어 빈도와 글자 빈도에서 고빈도 집단은 상위 50%에 속하고, 저빈도 집단은 하위 50%에 속하는 자극들이다.
3. * $p < .05$, ** $p < .01$, *** $p < .001$

을 때보다 있을 때 어휘판단시간이 더 느렸지만($t(8904) = -5.310$, $p < .001$), 오반응률은 차이가 없었다. 단어의 둘째 글자의 받침 유무는 어휘판단시간에는 유의한 차이를 보이지 않았지만($p > .1$), 오반응률은 받침이 없을 때보다 있을 때 감소하였다($t(8904) = 2.529$, $p < .01$). 비단어의 경우 첫째 글자에 받침이 없을 때보다 있을 때 어휘판단시간이 길었고($t(9548) = -3.664$, $p < .001$), 오반응률도 더 높았다($t(5672.3) = -2.838$, $p < .01$). 비단어의 둘째 글자도 받침이 없을 때보다 있을 때 어휘판단시간이 길었으며($t(9548) = -7.407$, $p < .001$), 오반응률도 높았다($t(6690.4) = -4.529$, $p < .001$).

각 예측 변인들이 어휘판단시간에 어떤 영향을 미치는지를 알아보기 위해 위계적 회귀분석을 실시하였다. 단어의 분석에는 단어 빈도의 변인을 가장 처음에 투입하여 단어 빈도를 통제한 다음 그다음 단계에 개별 예측변인 집합을 투입하였고, 세 번째 단계에 상호작용항을 투입하였다.

비단어의 분석에는 첫 단계에 개별 예측변인을, 그다음 단계에 상호작용항을 투입하였다.

단어의 어휘판단시간에 대한 회귀분석 결과가 표 4에 제시되어 있다. 첫 단계에서 단어 빈도는 유의하였다(R^2 = .231, F(1, 8904) = 2682.887, p < .001). 단어 빈도가 증가할수록 어휘판단시간이 감소하는(β = -.481, t = -51.797, p < .001) 패턴은 이후 분석에서도 유지되므로, 이에 대한 추가 언급은 하지 않는다. 두 번째 단계에서 개별 예측변인들 집합을 투입하였을 때 설명량은 유의하게 증가하였고(ΔR^2 = .006, F(6, 8898) = 12.538, p < .001), 단어 빈도를 포함한 전체 예측변인들의 모형은 유의하였다(R^2 = .238, F(7, 8898) = 396.997, p < .001). 세부적으로 보면, 단어의 첫째 글자 빈도와 둘째 글자 빈도가 증가할수록 각각 어휘판단시간이 증가하였다(첫째 글자: β = .058, t = 5.410, p < .001, 둘째 글자: β = .025, t = 2.444, p < .05). 그리고 단어의 둘째 글자가 종모음일 때보다 횡모음일 때 어휘판단시간이 감소하였다(β = -.021, t = -2.185, p < .05). 또한, 단어의 첫째 글자는 글자에 받침이 없을 때보다 있을 때 어휘판단시간이 증가하였지만(β = .045, t = 4.282, p < .001), 둘째 글자에 받침은 없을 때보다 있을 때 어휘판단시간이 감소하였다(β = -.039, t = -3.899, p < .001).

셋째 단계에서 상호작용항을 투입하였는데, 전체 설명량, R^2 = .241가 되었고, 이 증가분은 통계적으로 유의하였다(ΔR^2 = .003, F(7, 8891) = 5.737, p < .001). 상호작용항의 투입으로 둘째 글자 빈도는 더 이상 유의하지 않게 되었고, 첫째 글자 모음의 유형이 유의한 것으로 변화하였는데, 첫째 글자가 종모음일 때보다 횡모음일 때 어휘판단시간이 감소하였다(β = -.030, t = -2.565, p < .05). 유의한 상호작용항을 살펴보면 첫째 글자 모음 유형과 둘째 글자 모음 유형 간의 상호작용이 통계적으로 유의하였다(β = .064, t = 4.825, p < .001). 상호작용 효과를 분석하기 위해 독립변인의 수준별로 자료를 분류하여 비교를 하였다. 표 5에서 보듯이 둘째 글자가 종모음 유형일 때는 첫째 글자가 종모음일 때보다 횡모음일 때 어휘판단시간이 빨랐고(t(6094) = 2.293, p < .05), 둘째 글자가 횡모음 유형일 때는 첫째 글자가 횡모음일 때보다 종모음일 때 어휘판단시간이 빨랐다(t(2808) = -3.520, p < .001).

또한, 둘째 글자의 빈도와 둘째 글자 모음 유형 사이의 상호작용도 통계적으로 유의하였다(β = -.037, t = -3.254, p < .01). 이를 살펴보기 위해 빈도 평균으로부터 1 *SD* 이상 떨어진 집단(상위 16%)과 아래로 1 *SD* 이상 떨어진 집단(하위 16%)을 비교하였다. 표 6을 보면 둘째 글자의 빈도 효과는 둘째 글자가 종모음일 때(7ms; t(1051.062) = 2.329, p < .05)와 횡모음일 때(12ms; t(248.361) = 2.124, p < .05) 모두 유의하였는데, 특히 저빈도-횡모음 조건(622ms)의 어휘판단시간은 나머지 세 조건의 평균(610ms)보다 유의하게 느렸다(t(504.793) = -3.588, p < .001).

비단어의 어휘판단시간에 대한 회귀분석 결과가 표 7에 제시되어 있다. 글자 속성들은 전반적으로 비단어의 어휘판단시간을 유의하게 설명하였다(R^2 = .296, F(6, 9543) = 668.331, p < .001). 비단어의 첫째 글자와 둘째 글자의 빈도는 증가할수록 어휘판단시간이 증가하였다(첫째 글자: β = .493, t = 53.341, p < .001, 둘째 글자: β = .365, t = 39.276, p < .001). 그리고, 첫째 글자가

〈표 4〉 단어의 어휘판단시간에 대한 단어 빈도와 글자 속성 및 상호작용항의 위계적 회귀분석 결과

단계	예측 변인	β	*SE*	*t*	*p*	*VIF*	R^2	ΔR^2
1	(상수)		1.913	369.154	<.001		.231	
	단어 빈도	-.481	.370	-51.797	<.001	1.00	$F(1, 8904)=2682.9$***	
2	(상수)		1.406	434.245	<.001			
	단어 빈도	-.500	.384	-51.921	<.001	1.09		
	첫째 글자 빈도	.058	.390	5.410	<.001	1.35	.238	.006
	둘째 글자 빈도	.025	.400	2.444	<.05	1.26		
	첫째 글자 모음	-.004	1.094	-.464	.642	1.02	$F(7, 8898)=397.0$***	$F(6,8898)=12.5$***
	둘째 글자 모음	-.021	1.139	-2.185	<.05	1.03		
	첫째 글자 받침	.045	1.212	4.282	<.001	1.28		
	둘째 글자 받침	-.039	1.186	-3.899	<.001	1.19		
3	(상수)		1.892	308.459	<.001			
	단어 빈도	-.502	.384	-52.072	<.001	1.09		
	첫째 글자 빈도	.087	.769	4.088	<.001	5.27		
	둘째 글자 빈도	.031	.822	1.453	.146	5.32		
	첫째 글자 모음	-.030	1.351	-2.565	<.05	1.56		
	둘째 글자 모음	-.050	1.434	-4.273	<.001	1.64		
	첫째 글자 받침	.045	2.028	2.587	<.05	3.59		
	둘째 글자 받침	-.050	1.933	-3.050	<.01	3.18		
	두 글자의 빈도 간 상호작용	.002	.187	.232	.816	1.07	.241	.003
	글자 모음 유형 간 상호작용	.064	2.329	4.825	<.001	2.05	$F(14, 8891)=202.1$***	$F(7,8891)=5.7$***
	글자 받침 유무 간 상호작용	.008	2.261	.430	.667	4.49		
	첫째 글자 빈도, 모음 상호작용	-.022	.721	-1.894	.058	1.53		
	첫째 글자 빈도, 받침 상호작용	-.017	.855	-.931	.352	3.86		
	둘째 글자 빈도, 모음 상호작용	-.037	.776	-3.254	<.01	1.50		
	둘째 글자 빈도, 받침 상호작용	.017	.922	.860	.390	4.45		

주 1. * $p < .05$, ** $p < .01$, *** $p < .001$

〈표 5〉 단어의 첫째 글자와 둘째 글자의 모음 유형 조합에 따른 어휘판단시간(ms)과 표준오차

		둘째 글자 모음 유형	
		종모음	횡모음
첫째 글자 모음 유형	종모음	612 (.9)	608 (1.3)
	횡모음	609 (1.2)	616 (1.8)

〈표 6〉 단어의 둘째 글자에서 빈도와 모음 유형 조합에 따른 어휘판단시간(ms)과 표준오차

		둘째 글자 모음 유형	
		종모음	횡모음
둘째 글자 빈도	고빈도	609 (1.7)	609 (5.0)
	저빈도	615 (2.3)	622 (2.8)

주. 고빈도 집단의 빈도는 평균보다 1 *SD* 이상 더 높으며, 저빈도 집단의 빈도는 평균보다 1 *SD* 더 낮다.

종모음일 때보다 횡모음일 때 어휘판단시간이 증가하였다(β = .035, t = 4.062, p < .001). 또한, 첫째 글자와 둘째 글자 모두 글자에 받침이 없을 때보다 있을 때 어휘판단시간이 증가하였다(첫째 글자: β = .180, t = 19.705, p < .001, 둘째 글자: β = .152, t = 16.565, p < .001).

상호작용항을 투입한 모형의 회귀분석 결과도 유의하였으며(R^2 = .307, F(13, 9536) = 324.338, p < .001), 설명량의 증가분은 유의하였다(ΔR^2 = .011, F(7, 9536) = 21.059, p < .001). 글자 속성 변인의 개별 효과는 상호작용항을 투입하기 전 단계의 결과와 거의 같으므로, 이에 대한 서술은 생략한다. 통계적으로 유의한 상호작용 결과에 대해 서술하면 다음과 같다. 첫째 글자 빈도와 둘째 글자 빈도 사이의 상호작용이 통계적으로 유의하였다(β = 078, t = 8.900, p < .001). 표 8을 보면 첫째 글자와 둘째 글자가 모두 고빈도일 때(720ms) 어휘판단시간이 나머지 세 조건의 평균(631ms)보다 더 길었다(t(672) = -6.901, p < .001).

첫째 글자 모음 유형과 둘째 모음 유형 간의 상호작용도 통계적으로 유의하였다(β = -.058, t = -4.750, p < .001). 표 9를 보면 첫째 글자가 종모음일 때는 둘째 글자의 모음 유형 간 차이는 거의 없으나(p > .1), 첫째 글자가 횡모음일 때는 그 차이가 약 12ms로 증가하였다(t(3281) = 4.922, p < .001).

비단어의 첫째 글자의 빈도와 첫째 글자의 받침 유무의 상호작용도 통계적으로 유의하였으므로(β = -.160, t = -7.882, p < .001), 표 6의 경우처럼 고빈도 상위 16% 집단과 저빈도 하위 16% 집단을 비교하였다. 그 결과 표 10에서 보듯이 첫째 글자가 고빈도 상위 16% 집단에 속할 때 민글자와 받친글자의 차이는 통계적으로 유의하지 않은 반면(p > .1), 둘째 글자가 저빈도 하

〈표 7〉 비단어의 어휘판단시간에 대한 글자 속성 및 상호작용항의 위계적 회귀분석 결과

단계	예측 변인	β	*SE*	*t*	*p*	*VIF*	R^2	ΔR^2
1	(상수)		1.498	434.014	<.001			
	첫째 글자 빈도	.493	.279	53.341	<.001	1.16		
	둘째 글자 빈도	.365	.345	39.276	<.001	1.17	.296	
	첫째 글자 모음	.035	1.198	4.062	<.001	1.00		
	둘째 글자 모음	-.011	1.192	-1.222	.222	1.01	$F(6, 9543)=$ 668.3***	
	첫째 글자 받침	.180	1.312	19.705	<.001	1.13		
	둘째 글자 받침	.152	1.289	16.565	<.001	1.14		
2	(상수)		2.066	313.004	<.001			
	첫째 글자 빈도	.650	.673	29.156	<.001	6.84		
	둘째 글자 빈도	.375	.806	17.239	<.001	6.52		
	첫째 글자 모음	.063	1.510	5.822	<.001	1.62		
	둘째 글자 모음	.020	1.488	1.828	.068	1.60		
	첫째 글자 받침	.196	2.262	12.492	<.001	3.40		
	둘째 글자 받침	.157	2.314	9.523	<.001	3.74		
	두 글자의 빈도 간 상호작용	.078	.151	8.900	<.001	1.07	.307	.011
	글자 모음 유형 간 상호작용	-.058	2.480	-4.750	<.001	2.08	$F(13, 9536)=$ 324.3***	$F(7, 9536)=$ 21.1***
	글자 받침 유무 간 상호작용	-.012	2.695	-.613	.540	5.67		
	첫째 글자 빈도, 모음 상호작용	-.008	.566	-.739	.460	1.51		
	첫째 글자 빈도, 받침 상호작용	-.160	.704	-7.882	<.001	5.64		
	둘째 글자 빈도, 모음 상호작용	.009	.672	.825	.409	1.60		
	둘째 글자 빈도, 받침 상호작용	-.006	.861	-.292	.770	5.58		

주. * $p < .05$, ** $p < .01$, *** $p < .001$.

위 16% 집단에 속할 때는 그 차이(35ms)는 통계적으로 유의하였다($t(211.880) = -8.703$, $p < .001$).

KLP-DB에는 어휘판단시간과 오반응률이 모두 보고되어 있으나 본 연구는 어휘판단시간을 위주로 자료를 분석하였다. KLP-DB에 대한 선행 연구(이광오 등, 2017)는 어휘판단시간과 오반응

〈표 8〉 비단어의 첫째 글자 빈도와 둘째 글자 빈도의 조합에 따른 어휘판단시간(ms)과 표준오차

		둘째 글자 빈도	
		고빈도	저빈도
첫째 글자 빈도	고빈도	720 (17.0)	658 (3.7)
	저빈도	630 (3.7)	602 (3.7)

주. 고빈도 집단의 ~~글자~~ 빈도는 평균보다 1 *SD* 이상 더 높으며, 저빈도 집단의 빈도는 평균보다 1 *SD* 더 낮다.

〈표 9〉 비단어의 첫째 글자와 둘째 글자의 모음 유형 조합에 따른 어휘판단시간(ms)과 표준오차

		둘째 글자 모음 유형	
		종모음	횡모음
첫째 글자 모음 유형	종모음	671 (1.0)	672 (1.4)
	횡모음	679 (1.4)	667 (1.9)

〈표 10〉 비단어의 첫째 글자 빈도와 첫째 글자 받침 유무의 조합에 따른 어휘판단시간(ms)과 표준오차

		첫째 글자 받침 유무	
		민글자	받친글자
첫째 글자 빈도	고빈도	690 (2.6)	699 (6.6)
	저빈도	593 (3.8)	628 (1.4)

주. 고빈도 집단의 ~~글자~~ 빈도는 평균보다 1 *SD* 이상 더 높으며, 저빈도 집단의 빈도는 평균보다 1 *SD* 더 낮다.

률은 정적 상관을 보였으므로, 속도-정확도 교환은 발생하지 않았다고 판단하였다. 본 연구에서도 오반응률의 패턴은 대체로 어휘판단시간의 패턴을 따랐는데, 단어의 경우 판단시간과 오반응률의 상관계수, r = .711이었고, 비단어의 경우 그 상관계수, r = .698로서 모두 통계적으로 유의하였다. 즉 판단시간이 길어질 때 오반응률도 증가하는 경향이었으므로, 전반적으로 반응속도-정확도의 교환은 없었다고 볼 수 있다. 표 3을 보면, 단어 자극의 첫째 글자의 받침 유무는 판단시간에 유의한 차이를 보였으나 오반응률의 차이는 없었으며, 둘째 글자의 받침 유무는 오반응률에 차이를 보였으나 판단시간의 차이를 보이지는 않았다.

오반응률에 대한 위계적 회귀분석 결과도 어휘판단시간에 대한 회귀분석 결과와 거의 같았다. 구체적으로 살펴보기 위해, 단어의 오반응률에 대한 분석에서 유의하게 관찰된 변인이 판단시간 분석 결과와 다른 방향의 회귀계수를 보이는지를 살펴보았다. 단어의 둘째 글자 빈도와 첫째 글

자의 받침 유무는 어휘판단시간에 유의한 예측 변인이었지만 오반응률에는 유의하지 않았다(둘째 글자 빈도 : β = .036, t = 1.579, p > .1; 첫째 글자 받침 유무 : β = .015, t = .801, p > .1). 첫째 글자의 빈도와 둘째 글자의 빈도 사이의 상호작용은 어휘판단시간에 대해서 유의하지 않았지만 오반응률에 대해서는 유의하였다(β = .021, t = 2.124, p < .05). 하지만 첫째 글자가 고빈도일 때 둘째 글자가 고빈도인 경우(4.1%,)와 저빈도인 경우(4.2%)의 오류율의 차이는 유의하지 않았으며, 첫째 글자가 저빈도일 때에도 고빈도 둘째 글자(4.8%)와 저빈도 둘째 글자(5.3%) 간의 오류율의 차이는 유의하지 않았다. 비단어의 오반응률 자료에서 유의하게 관찰된 변인은 대부분 판단시간의 분석과 일치하는 방향의 회귀계수를 보였다. 예외적으로, 둘째 글자의 모음 유형에서 둘째 글자의 모음이 종모음일 때보다 횡모음일 때 오반응률이 감소하는 결과를 보였다(β = -.025, t = -2.643, p < .01). 또한, 첫째 글자와 둘째 글자의 받침 유무 사이의 상호작용도 오반응률 분석에서 나타났다(β = -.058, t = -2.631, p < .01). 전체적으로 봐서, 위계적 회귀분석에서 관찰된 총 41개(단어 22개, 비단어 19개)의 효과 중 3개(7.3%)에서만 속도-정확도 교환의 가능성이 관찰되었다. 전반적으로는 속도-정확도 교환이 거의 발생하지 않은 것으로 보인다. 그러므로 본 연구의 분석은 어휘판단시간에 대한 것만으로 충분하다고 생각된다.

논 의

본 연구는 KLP-DB의 어휘판단 자료를 분석한 결과 글자 속성이 단어와 비단어의 처리에 미치는 영향이 비슷하지만 차이가 있음을 확인하였다. 상호작용효과까지 고려하였을 때, 단어의 어휘판단에 영향을 미치는 글자 속성은 첫째 글자의 빈도와 첫째와 둘째 글자의 모음 유형과 받침 유무였다. 또한, 첫째 글자와 둘째 글자의 모음 유형 조합도 영향을 미쳤고, 둘째 글자의 빈도와 모음 유형의 조합도 영향을 미쳤다. 반면에 비단어의 어휘판단에 영향을 미치는 속성은 첫째와 둘째 글자의 빈도와 첫째 글자의 모음 유형, 첫째 글자와 둘째 글자의 받침 유무였다. 그리고 첫째 글자와 둘째 글자의 빈도의 조합도 영향을 미쳤으며 단어와 마찬가지로 첫째 글자와 둘째 글자의 모음 유형 조합도 영향을 미쳤다. 또한, 첫째 글자의 빈도와 받침 유무의 조합도 영향을 미쳤다.

단어 빈도는 위계적 분석의 모든 단계에서 강력한 변인으로서 회귀계수 절대값이 가장 컸다(표 4). 즉 단어 빈도가 높을수록 판단시간이 짧아졌다. 단어빈도효과는 고빈도 단어가 저빈도 단어에 비해 더 효율적으로 처리되기 때문으로 보이는데(Monsell, Doyle, & Haggard, 1989) 한글 단어(이광오, 1993)를 포함한 다양한 언어에서 관찰된다. 단어 빈도는 네덜란드어에서 어휘판단시간의 변량을 약 33% 설명하였는데(Brysbaert, Stevens, Mandera, & Keuleers, 2016), 본 연구에서는 어휘판단시간의 약 23%를 설명하였다.

단어에서 첫째 혹은 둘째 글자의 빈도가 증가할수록 어휘판단시간이 증가하였다. 그런데 표 3의 빈도 집단 비교에서 고빈도 글자 집단은 저빈도 글자 집단보다 어휘판단시간이 더 빨랐는데, 여기에서는 단어빈도가 통제되지 않았다. 이와 달리 표 4의 회귀분석에서 단어빈도를 통제하였을 때에는 글자 빈도가 높을수록 어휘판단시간이 더 길어졌다. 이것은 단어빈도와 글자 빈도의 상관이 높아서, 단어빈도를 통제하였을 때 글자 빈도의 효과가 역전된 결과로 보인다. 단어의 첫째 글자에서 빈도가 증가할수록 어휘판단시간이 증가한 결과는 예상한 바였다. 권유안(2020)은 한글 음절(글자)의 빈도 효과가 일관적이지 않다고 보았는데, KLP-DB의 분석 결과 단어와 비단어 모두에서 글자 빈도가 증가할수록 어휘판단시간이 증가하는 억제적인 글자빈도효과가 관찰되었다. 억제적인 글자빈도효과는 심성어휘집에서 단어의 첫째 글자를 공유하는 다른 단어의 경쟁적인 활성화로 인해 나타나는 지연된 효과로 설명된다(Carrieras, Alvarez, & DeVega, 1993). 글자빈도효과는 특히 저빈도 단어의 첫째 글자를 공유하는 단어 중에서 고빈도 단어가 있을 때 나타나는 것으로 보이는데(권유안, 2020), 많은 수의 단어를 분석한 본 연구에서는 전반적으로 억제적인 글자빈도효과가 나타났다.

이 같은 억제적 현상은 단어의 둘째 글자에 대해서도 관찰되었다. 둘째 글자의 빈도 효과는 스페인어에서 촉진적으로 보고되기도 하였지만(Levelt & Wheeldon, 1994), 그 효과를 관찰하지 못한 보고도 있었다(Carreiras & Perea, 2004). 본 연구에서 두 글자의 조합에 따라 단어와 비단어가 구별되는데, 각 글자는 단독으로 혹은 다른 단어의 일부로 쓰일 수 있었다. 그래서 어휘판단에서 첫째 글자 못지않게 둘째 글자도 중요한 역할을 할 수 있고, 둘째 글자가 활성화하는 심성어휘집의 단어들에 대한 탐색과 경쟁도 활발하였을 가능성이 있다. 이로 인해 둘째 글자의 사용빈도가 높을수록 판단시간이 지체될 가능성이 있다. 하지만 단어에서 둘째 글자의 빈도 효과는 상호작용항을 투입하였을 때 사라졌는데, 이는 단어에서 둘째 글자의 빈도가 안정적인 효과가 아닐 수 있음을 시사한다.

단어뿐만 아니라 비단어에서도 글자의 빈도 효과가 나타났다. 이런 결과는 스페인어 비단어에서도 억제적인 글자(음절) 빈도 효과가 보고된 결과와 일치한다(Carreiras, Alvarez, & DeVega, 1993; Carreiras & Perera, 2004). 비단어의 각 글자 역시 단어와 마찬가지로 단어의 일부로 쓰일 수 있어 어휘판단에서 비단어의 첫째 글자와 둘째 글자 모두 중요한 역할을 할 수 있다. 특히 첫째 글자의 빈도가 높을수록 그 글자로 시작하는 단어들이 심성어휘집 내에서 빨리 활성화될 수 있고, 이런 혼동 요인에 의해 어휘판단시간이 느려졌을 가능성이 있다.

본 연구의 결과에서 단어의 첫째 글자와 둘째 글자 사이의 모음 유형의 상호작용 효과는 유의하였는데(표 5), 횡모음으로만 이루어진 단어 자극에서 어휘판단시간이 가장 느린 것은 선행연구와 비슷한 결과였다(민수영과 이창환, 2018). 반면에 비단어에서는 횡모음으로만 이루어진 자극에서 어휘판단시간이 가장 짧았다. 횡모음-횡모음 조합의 단어 혹은 비단어의 수는 거의 같았으며, 여러 조합 중 해당 조합의 비율(약 12%)이 비교적 낮았다. 이로 인해 횡모음-횡모음 단

어가 어휘접근에 더 불리하였을 가능성이 있다. 그러나 사례 수가 가장 많은 종모음-종모음 조합의 단어가 특히 빨리 판단되지는 않으므로 이런 가능성은 추후에 더 분석되어야 할 것이다.

단어에서 첫째 글자에 받침이 있을 때 어휘판단시간이 증가하는 결과가 관찰되었는데, 이 결과는 받침의 효과를 관찰하지 못한 선행 연구(조규영, 진영선, 1991; 최양규, 1986)로부터 예상되지 않은 것이다. 본 연구에서 받침 유무의 효과를 조금 더 살펴보면, 두 글자 모두에 받침이 없을 때의 평균 어휘판단시간은 607ms였으며 두 글자 모두에 받침이 있을 때의 평균 어휘판단시간은 614ms였다. 특히 첫째 글자에 받침이 없을 때보다 있을 때 평균 어휘판단시간이 유의하게 증가하였고, 이는 단어에서 받침의 유무가 한글 시각단어재인에 영향을 미칠 가능성이 있음을 보여준다. 선행 연구와 같은 실험 설계에서는 비교적 적은 수의 단어 자극을 사용하므로 받침의 효과가 잘 드러나지 않았을 가능성이 있다. 본 연구에서 단어의 첫째 글자의 받침 효과는 상호작용 항이 투입된 이후에는 관찰되지 않았으므로, 강력한 효과로 생각되지는 않는다.

비단어의 경우 첫째 글자와 둘째 글자 모두에서 받침이 있을 때 어휘판단시간이 증가하는 결과가 나타났다. 즉 받침으로 인한 정보처리 부하의 증가로 판단시간이 길어지는 통상적인 결과가 관찰되었다. 받침의 효과는 단어 수준 정보의 영향이 제한되었기 때문에 발생한 것으로 볼 수 있다. 비단어 처리에서 받침의 효과는 한글이 글자 단위로 처리된다는 연구(예: Simpson & Kang, 2004)에 의문을 제기한다. 한글 글자가 자음과 모음으로 이루어진 글자핵(민글자)과 받침 정보, 그리고 출현속성 등을 단위로 처리된다는 연구(박창호, 2006)가 비단어 속 글자에 대한 본 연구의 결과를 어느 정도 설명할 수 있는 것으로 보인다.

본 연구의 분석에 따르면, 단어의 판단시간은 단어 빈도가 높고, 첫째 및 둘째 글자 빈도가 낮은 조건에서 가장 짧았다. 이런 조건에서 단어의 처리가 신속하게 일어나는 반면, 글자 수준의 처리는 느리게 일어남으로써 어휘판단에 경쟁적인 후보들을 빨리 만들어내기 어려울 것이다. 그런데, 단어 빈도가 높고, 첫째 글자 빈도가 낮아도, 둘째 글자 빈도가 높은 조건에서 어휘판단시간이 길어졌다. 여기에 해당하는 예로서 “핥다”에서 첫째 글자의 빈도는 낮아도, 접미사인 둘째 글자 “-다”의 빈도와 단어의 빈도는 높다. 그러나 이런 종류의 단어는 5% 미만이며, 이를 제거하여도 전체 결과는 별로 달라지지 않았다. 비단어의 경우에는 글자 빈도가 낮고, 종모음을 가지고, 받침이 없는 글자가 빨리 판단되었다. 선행 연구로부터 종모음 글자에 대한 판단시간이 횡모음 글자보다 빠르며, 민글자의 정보처리 부하가 받친글자보다 더 낮으리라고 예상할 수 있다. 그래서 비단어의 처리에는 글자 빈도라는 하향적 요인과 글자 모음 유형 및 받침이라는 상향적 요인이 모두 작용한다고 해석할 수 있다. 글자 속성이 어휘판단에 일정한 영향을 미친다는 것을 보여준 본 연구의 결과는 어휘판단과제에서 단어와 비단어의 목록을 구성하고 어휘판단시간을 해석할 때 첫째 글자의 빈도 외에도 글자 유형, 받침 여부 등 형태적인 측면을 포함한 글자 속성의 효과를 충분히 고려해야 함을 시사한다.

글자 속성의 영향은 단어와 비단어의 어휘판단과정에서 상이한 것으로 보인다. 단어는 심성어

휘집에서 탐색될 때 비로소 단어로 판단된다. 친숙한 혹은 고빈도 단어는 심성어휘집에 더 빨리 접근함으로써 글자 속성이 영향을 미칠 겨를을 주지 않을 것이다. 예컨대 친숙한 단어에서 낱자 재인이 힘들었다(Healy, 1980). 이와 달리 비단어 혹은 저빈도 단어는 비교적 늦게 처리되면서 나중에 글자 속성의 영향이 개입할 여지가 생길 것이다. 사실 비단어와 단어는 한두 획이나 낱자로 구별될(예: '밉다'와 '맙다') 뿐이지만, 단어가 심성어휘집에 재빨리 성공적으로 접근하게 되면 이런 차이는 영향을 미치지 못하는 것 같다. 이러한 생각은 Coltheart, Rastle, Perry, Langdon, 및 Ziegler(2001)의 이중경로모형(dual route cascaded model)과 일치한다. 이중경로 모형은 단어가 철자 입력 어휘집을 거쳐 음운 출력 어휘집으로 가는 어휘적 비의미 경로(lexical nonsemantic route)와 비단어가 자소-음운 변환 시스템을 거쳐 음운 시스템으로 가는 경로(grapheme-phoneme rule system)를 구분하는데, 후자에서는 낱자와 자소의 처리가 순차적으로 진행된다. 이런 틀에서 본 연구의 결과를 해석해 보면, 비단어 판단시간(약 673ms)이 단어 판단시간(약 611ms)보다 약 62ms 느린 것은 순차적 변환으로 인한 것일 수 있다. 그러나 단어와 비단어 여부는 사전에 구별되는 것이 아니므로, 심성어휘집에의 접근과 글자 속성의 처리가 동시적으로 진행된 다음, 비단어의 경우 글자의 속성에 대한 추가 분석이 벌어진다는 가설을 생각해 볼 수 있다.

본 연구는 KLP-DB 중에 두 글자 단어 혹은 비단어를 사용하였다. 비록 두 글자 단어가 가장 널리 쓰이는 글자 수이지만, 본 연구의 결과를 세 글자 이상의 단어에 적용하는 데에는 한계가 있다. 세 글자 이상의 단어/비단어의 처리에 글자 속성이 어떤 영향을 미치는가도 추후에 검토될 필요가 있다. 그리고 글자 유형 중, 사례 수가 낮은 "과", "광"과 같은 복모음 자극은 제외되었다. 이로 인해 더 복잡한 글자 속성을 충분히 다루지 못하였다. 비록 본 연구가 낱자 수준의 글자 속성을 검토하였지만, 낱자를 구성하는 종말, 수평선, 사선, 교차점, 곡선 등의 더 기초적인 시각 속성의 영향(Fiset et al., 2008)을 고려하지는 못하였다. 또한 본 연구는 3원 상호작용 이상을 분석하지 않았으며, 글자 타입 빈도(권유안, 2012)도 고려하지 않았는데, 그럴 경우 본 연구에서 검토한 예측변인들과 관련해서 과도한 공선성이 예상되기 때문이었다. 이와 같은 글자의 고차적인 특성은 추후에 체계적으로 연구할 필요가 있을 것이다.

KLP-DB에 있는 방대한 단어와 비단어 자료의 덕분으로, 본 연구는 어휘판단 과제에서 여러 글자 속성의 영향과 상호작용을 분석할 수 있었다. 소규모 자료를 이용하는 실험 설계 연구(예: 이은희, 이은주, 2019)는 많은 변인들의 효과와 그 상호작용을 체계적으로 분석하는 데 한계가 있다. 본 연구의 결과는 비교적 소수의 글자 혹은 단어/비단어를 사용한 연구에서 포착하기 어려운 것이라고 생각된다. KLP-DB와 같은 메가데이터를 이용하여 여러 관련 변인들을 통계학적으로 통제하고 그 기여도를 분석함으로써(예: Brysbaert, et al., 2016), 단어재인의 복합적인 특성을 탐구할 수 있을 것이다.

참고문헌

강범모, 김흥규 (2009). **한국어 사용 빈도.** 서울: 한국문화사.

고은태, 배성봉, 이윤형 (2022). 한글 글자 지각시 단어우월 효과: 어휘 속성의 역할. **한국심리학회지: 인지 및 생물, 34**(1), 11-24.

권유안 (2012). 첫 음절 토큰 빈도와 타입 빈도가 단어 및 유사 단어 어휘 판단 시간에 미치는 영향. **한국심리학회지: 인지 및 생물, 24**(4), 315-328.

권유안 (2020). 한국어 시각적 단어 재인에서 음절 빈도 효과의 고찰. **한국심리학회지: 인지 및 생물, 32**(4), 291-303.

김민식, 정찬섭 (1989). 한글 자모구성 형태에 따른 자모 및 글자 인식. **인지과학, 1**(1), 27-75.

김정오 (1994). 한글 글자인식에서의 받침 정보. **성곡논총, 25**(1), 693-727.

김정오, 김재갑 (1992). 한글 단어재인에 있어서 글자처리와 낱자의 지각. **한국심리학회지: 인지 및 생물, 4,** 36-51.

민수영, 이창환 (2018). 한글 단어 재인에 영향을 미치는 변인: 음절 형태를 중심으로. **인지과학, 29**(4), 193-220.

박권생 (1995). 한글 단어 재인을 위한 시각 처리 과정. **한국심리학회지: 인지 및 생물, 7**(2), 61-78.

박창호 (2006). 낱자의 집단화가 글자 지각에 미치는 영향: 글자 사용 여부 판단 과제에서. **한국심리학회지: 인지 및 생물, 18**(3), 173-185.

배성봉 (2018). 비단어의 유형과 비율이 어휘판단에 미치는 영향. **한국심리학회지: 인지 및 생물, 30**(3), 301-307.

배성봉, 박권생, 이혜원, 이광오 (2016). 한국어 단어재인에서 한 글자 단어 열등 효과. **언어과학연구, 77,** 109-125.

이광오 (1993). 한글 단어인지과정에서 표기법이 심성어휘집의 구조와 탐색에 미치는 영향. **한국심리학회지: 인지 및 생물, 5,** 26-39.

이광오, 구민모, 남기춘, 박기남, 박태진, 배성봉, 이창환, 이혜원, 조증열 (2017). 한국어 심성어휘집 연구를 위한 어휘판단 데이터베이스. **한국심리학회지: 인지 및 생물, 29**(4), 395-410.

이은희, 이은주 (2019). 단어 빈도와 길이가 20대 성인과 노인의 어휘판단에 미치는 효과. **특수교육논총, 35**(3), 93-110.

이준석, 김경린 (1989). 한글 낱말의 처리단위. **인지과학, 1**(2), 221-239.

조규영, 진영선 (1991). 회전된 단어 읽기에서 음절의 수 및 시각의 효과. **한국심리학회지: 실험 및 인지, 3,** 63-75.

최양규 (1986). 음절수가 한글단어 재인 반응시간에 미치는 영향. 석사학위 청구논문, 부산대

학교.

Brysbaert, M., Stevens, M., Mandera, P., & Keuleers, E. (2016). The impact of word prevalence on lexical decision times: Evidence from the Dutch Lexicon Project 2. *Journal of Experimental Psychology: Human Perception and Performance, 42*(3), 441-458.

Carreiras, M., Alvarez, C, J., & DeVega, M. (1993). Syllable frequency and visual word recognition in Spanish. *Journal of Memory and Language, 32*(6), 766-780.

Carreiras, M., & Perera, M. (2004). Naming pseudowords in Spanish: Effects of syllable frequency. *Brain and Language, 90*(1-3), 393-400.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*(1), 204-256.

Fiset, D., Blais, C., Éthier-Majcher, C., Arguin, M., Bub, D., & Gosselin, F. (2008). Features for identification of uppercase and lowercase letters. *Psychological Science, 19*(11), 1161-1168.

Lee, C. H., & Taft, M. (2011). Subsyllabic structure reflected in letter confusability effects in Korean word recognition. *Psychonomic Bulletin & Review, 18*, 129-134.

Levelt, W. J., & Wheeldon, L. (1994). Do speakers have access to a mental syllabary?. *Cognition, 50*, 239-269.

Lupker, S. J., & Pexman, P. M. (2010). Making things difficult in lexical decision: the impact of pseudohomophones and transposed-letter nonwords on frequency and semantic priming effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(5), 1267-1289.

Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). Effects of frequency on visual word recognition tasks: Where are they?. *Journal of Experimental Psychology: General, 118*(1), 43-71.

Simpson, G. B., & Kang, H. (2004). Syllable processing in alphabetic Korean. *Reading and Writing, 17,* 137-151.

Stone, G. O., & Van Orden, G. C. (1993). Strategic control of processing in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 19*(4), 744-774.

1차 원고 접수: 2023. 10. 17
1차 심사 완료: 2023. 11. 10
2차 원고 접수: 2023. 11. 26
2차 심사 완료: 2023. 12. 15
3차 원고 접수: 2023. 12. 22
최종 게재 확정: 2023. 12. 23

(*Abstract*)

The Effect of Syllable Frequency, Syllable Type and Final Consonant on Hangeul Word and Pseudo-word Lexical Decision: An Analysis of the Korean Lexicon Project Database

Myong Seok Shin ChangHo Park

Department of Psychology, Jeonbuk National University

This study attempted to find out how lexical decision of two-syllable words or pseudo-words is affected by syllabic information, such as syllable frequency, syllable (i.e. vowel) type, and presence of final consonant (i.e. batchim), through the analysis of the Korean Lexicon Project Database (KLP-DB). Hierarchical regression of RT data showed that lexical decision of words was influenced by the frequency of the first syllable, the syllable type of the first and second syllables, batchim for the first and second syllables, and also by the interaction of the two syllable types and the interaction of syllable frequency and batchim of the second syllable. For pseudo-words lexical decision was influenced by the frequency of the first and second syllables, syllable type of the first syllable, and batchim for the first and second syllables, and also by the interaction of the two syllable frequencies, the interaction of the two syllable types, and the interaction of syllable frequency and batchim of the first syllable. Word frequency had a strong effect on lexical decision of words, while syllabic information had a stable effect on the lexical decision of pseudo-words. These results indicate that syllabic information should be seriously considered in constructing word and pseudo-word lists and interpreting lexical decision time. Understanding the effect of syllabic information will also contribute to the understanding of word recognition process.

Key words : Hangul, pseudo-word, syllable frequency, syllable types, Korean Lexicon Project