

---

# 미디어 멀티태스킹 환경에서 인터페이스의 감각양식 차이가 인지부하와 과업수행에 미치는 영향에 관한 연구

다중 자원 이론과 스레드 인지 모델을 기반으로

## The Effects of Interface Modality on Cognitive Load and Task Performance in Media Multitasking Environment

이다나, Dana Lee\*, 한광희, Kwang-Hee Han\*\*

---

**요약** 본 연구는 빠르게 발전하는 음성 기반의 디바이스가 스크린 중심의 미디어 멀티태스킹 환경에 어떤 변화를 가져올 수 있을지 확인하고자 했다. 서로 다른 자원 구조를 가진 과업을 동시에 수행할 때 정보 처리 효율이 높아진다는 이론적 근거를 토대로, 시각 주의가 필요한 과제와 음성 또는 스크린 기반의 디바이스를 활용해 정보를 검색하는 과업을 동시에 수행하는 실험이 진행되었다. 실험 결과, 과업수행 환경과 인터페이스 감각양식은 모두 인지부하에 유의미한 영향을 미쳤다. 음성 인터페이스 그룹에서 전반적으로 인지부하 수준이 높게 나타났는데, 단독으로 사용된 단일 과업 조건보다 시각 과제를 동시에 수행한 다중 과업 조건에서 시각 인터페이스 그룹과의 차이가 줄어들었다. 과업 수행도의 경우 음성 인터페이스 그룹에서 시각 과제에 대한 수행능력이 시각 인터페이스 그룹보다 더 높게 측정되었다. 이러한 결과는 멀티태스킹 환경에서 음성 인터페이스를 사용했을 때 동시적 과업을 청각 경로와 시각 경로로 나누어 처리함으로써 인지부하와 과업수행에 이점이 나타났음을 의미한다. 이는 시각 자원의 충돌이 발생하기 쉬운 스크린 중심의 미디어 멀티태스킹 환경에서 음성 기반의 디바이스가 효율적 정보 처리를 촉진시키는 잠재적 역할을 할 수 있다는 함의점을 제공한다. 본 연구는 다중 자원 이론을 통해 자원의 분산처리에 대한 이론적 증거를 제시하고, 스레드 인지 모델을 기반으로 음성 인터페이스를 활용했을 때의 이점을 더욱 구체적으로 규명하고자 했다.

**Abstract** This research examined the changes that fast-growing voice-based devices would bring in the media multitasking environment. Based on the theoretical background that information processing efficiency improves when performing multiple tasks requiring different resource structures at the same time, we conducted an experiment where participants searched for information with voice-based or screen-based devices while performing an additional visual task. Results showed that both task performance environment and interface modality had significant main effects on cognitive load. The overall cognitive load level was higher in the voice interface group, but the difference in cognitive load between the two groups decreased in a multitasking environment where the additional visual resources was required. The visual task performance was significantly higher when using the voice interface than the screen interface. Our findings suggest that voice interfaces offered advantages in the cognitive load and task performance by distributing two tasks to the auditory and visual channels. The results of this study imply that voice-based devices have the potential to facilitate efficient information processing in the screen-centric environment where visual resources collide. We provided theoretical evidence of resource distribution using multiple resource theory and tried to identify the advantages of the voice interface more specifically based on the threaded cognition model.

**핵심어:** *Voice Interface, Media Multitasking, Cognitive Load, Sensory Modality, Multiple Resource Theory*

---

\*주저자 : 연세대학교 일반대학원 인지과학협동과정 석사과정; e-mail: lulina724@gmail.com

\*\*교신저자 : 연세대학교 심리학과, 인지과학협동과정 교수; e-mail: khan@yonsei.ac.kr

■ 접수일 : 2019년 1월 14일 / 심사일 : 2019년 2월 1일 / 게재확정일 : 2019년 3월 15일

## 1. 서론

우리는 종종 TV를 틀어놓고 노트북으로 웹 서핑을 하다 스마트폰을 확인하는 등 여러 디바이스를 동시에 활용하곤 한다. 이를 ‘미디어 멀티태스킹(media multitasking)’이라고 하는데, 일상에서 접할 수 있는 디지털 디바이스가 늘어나면서, 이러한 멀티태스킹 행동은 꾸준히 증가하고 있는 추세다[1]. 딜로이트 조사에 따르면 91%의 응답자들이 TV를 보는 동안 메일을 확인하거나 SNS를 하는 등 멀티태스킹을 하고 있다고 응답했으며[2], 페이스북은 아이트래킹 실험을 통해 94%의 참가자들이 TV를 볼 때 지속적으로 스마트폰을 사용한다는 것을 확인했다[3].

이러한 미디어 멀티태스킹 행동은 일반적으로 과업수행이나 정보처리를 저해한다고 알려져 있다[4]. 여러 과업을 처리할 때 주의가 분산되고 인지 부하가 증가하기 때문이다. 특히 현재 미디어 멀티태스킹의 대상이 되는 디바이스는 대체로 스크린 기반인 경우가 많기 때문에, 시각 자원의 충돌이 발생하여 이러한 부정적인 영향은 더 심화될 수 있다.

그런데 최근 인공지능 기술의 발전을 토대로 스마트스피커와 같은 음성 기반 디바이스의 보급이 가속화되면서[5], 미디어 멀티태스킹 행동패턴이나 인지시스템에 대한 영향이 달라질 것으로 예상된다. 음성 기반 디바이스를 이용하면 시각주의 없이도 청각만으로도 정보를 얻을 수 있으므로, 다른 스크린 디바이스와와 동시적 사용이 가능해지기 때문이다. 예를 들어 TV를 보다 메일을 확인하는 경우, 스마트폰을 이용하면 메일 앱을 실행시키고 새로 온 메일을 찾는 등 시각 자원이 필요한 과업을 수행하기 위해 TV 시청을 멈추어야 하지만, 스마트 스피커를 이용하면 음성으로 메일을 읽어주기 때문에 TV에 시선을 두면서도 내용을 확인할 수 있다.

본 연구는 이렇듯 미디어 멀티태스킹 환경에서 정보처리 채널을 분산시킴으로써 과업 수행의 효율을 높이는 음성 인터페이스의 인지적 역할에 주목하였고, 이를 실증적으로 규명하기 위해 이론적 검토를 토대로 실험을 설계하였다. 실험에서 참가자들은 스크린 또는 음성 기반의 디바이스를 활용해 정보를 검색하면서, 시각주의가 필요한 추가적인 과제를 동시에 수행하였다. 멀티태스킹 환경과 인터페이스 감각양식의 영향을 확인하기 위해 참가자들의 인지부하 수준과 과업수행 정확도가 측정되었다.

본 논문의 구성은 다음과 같다. 2장에서는 멀티태스킹 행동이 인간의 인지 시스템에 미치는 영향에 대한 여러 이론을 검토하였고, 3장에서는 특히 감각양식의 간섭과 관련한 기존 멀티태스킹 문헌을 살펴보았다. 4장에서는 실험 설계와 절차 등 실험에 관련한 자세한 설명을, 5장에서는 실험 결과를 정리하였다. 6장에서는 스레드 인지 이론을 토대로 실험 결과에 대한 심층적인 해석을 진행하였고, 마지막으로 7장에서는 결론 및 연구의 향후 방향을 제시하였다.

## 2. 이론적 배경

멀티태스킹 행동이 인간의 인지시스템에 주는 영향을 설명하기 위한 다수의 이론은 ‘제한된 용량 모델(limited capacity model)’을 전제로 한다. 정보 처리를 위해 필요한 인지적 자원의 양은 한정되어 있다는 것이다[6]. 따라서 여러 과업을 동시에 수행하는 경우, 제한된 인지 자원은 각 과업의 요구에 따라 분배되어야 한다. 이 때 요구되는 자원의 총량이 실제 이용 가능한 자원보다 초과되면 ‘인지 부하(cognitive overload)’가 발생하고, 이는 곧 과업 수행능력을 약화시키는 결과로 이어지게 된다[7]. 즉, 멀티태스킹 행동이 과업 수행에 대체로 부정적인 영향을 미치는 이유는 이렇듯 인지 자원의 초과가 발생하는 경우가 많기 때문이다.

그러나 여러 과업에서 요구되는 자원의 산술적인 총량이 무조건 인지부하를 야기하는 것은 아니다. 과업의 종류에 따라 서로 다른 경로로 분산되는 경우, 인지 시스템에 과도한 부담을 초래하지 않고도 동시에 처리될 수 있기 때문이다. 본 연구는 이와 같은 과업의 분산 처리로 나타나는 인지적 이점에 대한 이론적 근거를 제공하는 ‘다중 자원 이론(multiple resource theory)’을 토대로 전반적 연구 프레임워크를 설계하고, ‘스레드 인지 모델(threaded cognition model)’을 통해 추가적인 분석을 진행하였다.

### 2.1 다중 자원 이론

다중 자원 이론(multiple resource theory)[8]에 따르면 인간은 분리된 여러 자원의 경로(resource pool)를 가지고 있으며, 각 경로는 특정한 지각적, 또는 인지적인 메커니즘을 통해 질적으로 서로 다른 구조의 과업을 처리한다. 다중 과업 조건에서 두 과업이 어떤 자원을 요구하는가에 따라 동시적 처리의 효율성이 달라지는데, 다중 자원 이론에서는 자원 구조를 4가지 차원(처리 단계, 지각적 감각양식, 시각 경로, 처리 방식)으로 구분하고 있다(그림 1).

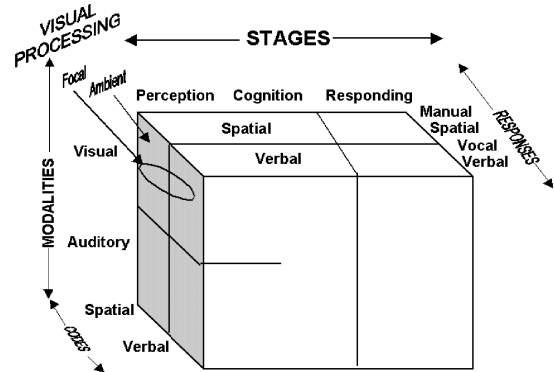


그림 1. 자원 구조에 따라 구분된 다중 자원 이론의 4차원 모형

우선, 처리 단계(processing stages)는 지각(perception)과 인지(cognition) 과정에서 요구되는 자원과 자극에 대한 반응(responding)에 필요한 자원을 구분하는 차원이다. 지각적 감각양식(perceptual modalities) 차원은 시각과 청각으로 자원 구조를 나누고, 시각 자원을 요구하는 과업은 다시 시각 경로(visual channels) 차원을 통해 중심시(focal)와 주변시(ambient)로 구분된다. 마지막으로 과업 처리방식(processing code)에 따라 공간적(spatial), 언어적(verbal)으로 구분하는 차원이 있다. 이렇듯 여러 차원으로 나뉘는 자원들이 다중 과업 환경에서 중첩적으로 요구된다면, 자원 경로의 병목(bottleneck)으로 인한 인지부하가 발생한다. 하지만 서로 다른 자원이 요구될 경우 각 경로에서 상호 간섭 없는 독립적인 과업 수행이 가능해진다.

본 논문에서는 특히 다중 자원 이론에서 제시한 지각적 감각양식 차원에서의 시각과 청각 자원의 구조적 차이에 초점을 맞추고자 하였다. 자극을 감각적으로 지각하는 단계에서는 인지나 반응 단계에서보다는 인지적 자원이 상대적으로 덜 요구되지만, 자원 경로의 충돌 여부는 처리 과정에 분명한 영향을 미친다. 즉, 동일한 지각적 자원을 요구하는 경우보다 서로 다른 자원(시각-청각)을 요구하는 다중 과업 환경에서 과업 수행도가 향상될 수 있는 것이다.

## 2.2 스레드 인지 모델

스레드 인지 이론(threaded cognition theory)은 다중 자원 이론의 변형으로, 과업 간의 상호 관련성에 초점을 맞춰 멀티태스킹 행동을 설명하고자 한다[9]. 스레드 인지 모델에서 동시적 과업은 분리된 여러 스레드(thread)로 나타난다. 각 스레드는 특정 자원을 활용해 개별 과업의 목표를 달성하며, 다중 자원 이론과 마찬가지로 동일한 자원을 요구하지 않는 한 여러 스레드는 동시에 처리될 수 있다.

다만, 스레드 인지 이론에서는 절차적(procedural) 자원의 역할이 강조된다. 절차적 자원은 과업 수행을 위해 가용한 자원을 파악하여 각 스레드에 배치하고, 개별 스레드 목표가 달성되면 이후 작업이 진행되도록 한다. 지각적(perceptual) 자원의 경우 시각, 청각 등 중첩되지 않는 하위 스레드로 나뉘어 동시에 처리될 수 있지만, 절차적 자원은 다른 자원을 관리하고 우선순위를 정하는 역할을 수행하므로 한 번에 오로지 하나의 스레드만 처리할 수 있다. 스레드 인지 모델을 이용하면, 그림 2와 같이 멀티태스킹 과업을 처리할 때 자원의 할당이 어떻게 이루어지는지 등 그 과정을 더욱 구체적으로 확인할 수 있다[10].

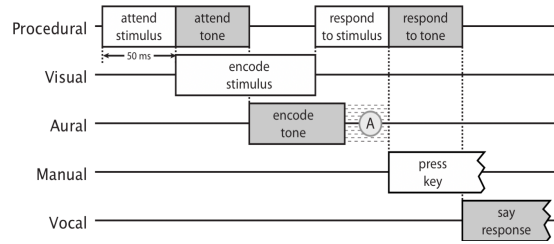


그림 2. 이중 선택 과제의 스레드 인지 모델; 시각자극에 대한 반응(흰 박스)과 청각자극에 대한 반응(회색 박스)이 거의 동시에 이루어지는 과정을 스레드로 나타내었다.

그림 2에서는 시각, 청각 자극이 동시에 주어졌을 때 이를 처리하는 과정을 보여주는데, 절차적 자원의 조정을 거쳐 각 과업이 교차적으로 진행되는 모습을 볼 수 있다. 주목할 점은 자극의 부호화(encoding) 과정에서는 시각, 청각 자원이 동시에 활용될 수 있지만, 자극에 대한 반응은 절차적 자원에서의 병목으로 인해 순차적으로 나타난다는 것이다. 즉, A 지점에서 청각 자극의 인코딩이 마무리되었음에도 바로 반응이 이어질 수 없는 이유는 절차적 자원이 시각 자극에 할당되어 있기 때문이다.

## 3. 선행연구 검토

미디어 멀티태스킹과 관련한 다수의 선행연구에서 멀티태스킹 행동이 정보처리의 효율이나 과업수행 능력을 떨어뜨린다고 보고하고 있다[11-13]. 본 논문의 연구 주제와 관련된 감각양식의 간섭(sensory interference)에 대해 직접적으로 확인한 연구는 많지 않지만, 소수의 연구에서 일관적으로 나타나는 결과는 감각양식의 간섭이 클 경우 수행능력이 더 많이 저해된다는 것이었다. Pool 등은 종이에 적힌 문제를 푸는 동안(과업 1: 시각) 음악을 들려주는 조건(과업 2: 청각)과 TV 드라마를 보여주는 조건(과업 2: 시각)을 비교했는데, 과업 2의 감각양식이 과업 1과 동일한 시각이었던 드라마 조건에서 유의미한 수행능력의 감소가 나타났다[14].

Jeong & Hwang은 특히 감각양식의 차이로 인한 구조적 간섭(structural interference, SI)에 초점을 맞춰 멀티태스킹 연구를 진행했다. 주어진 텍스트를 읽는 동안(과업 1: 시각)광고 메시지를 텍스트(과업 2: 시각) 또는 오디오 파일(과업2: 청각)로 제시했는데, Pool의 연구와 마찬가지로 감각양식의 간섭이 적었던 오디오 조건에서 과업 1의 텍스트에 대한 이해도와 기억이 높게 나타났다[15]. 이들은 최근 연구에서 감각양식의 간섭이 과업수행에 미치는 영향을 더욱 확실하게 규명하기 위해 이전보다 더 통제된 실험(시각 과제를 수행하는 동안 중 이미지를 보여주거나 종소리를 들려주거나)을 진행했는데, 마찬가지로 동일한 결과를 얻었다[16].

이렇듯 선행 연구에서 다중 자원 이론을 뒷받침하는 유의미한 결과를 도출하였지만, 최근 음성 인터페이스 등 신기술의 발전으로 달라진 미디어 멀티태스킹 환경에서도 해당 결과를 적용할 수 있을지에 대해서는 추가적인 논의가 필요하다. 텍스트 읽기, 음악 듣기 등의 과업은 디지털 디바이스를 이용한 과업과 복잡성이나 몰입도 또는 상호작용성(interactivity) 측면에서 분명한 차이가 있기 때문이다[17].

Wang 등이 거의 유일하게 디지털 디바이스 기반의 ‘온라인 채팅’을 적용해 감각양식의 차이가 미치는 영향을 살펴보았다. 참가자들은 시각적 패턴을 맞추는 과제(과업 1: 시각)와 더불어 텍스트(과업 2: 시각) 또는 보이스(과업 2: 청각)로 온라인 채팅을 하며 길을 안내하는 과업을 추가로 수행했다. 실험 결과 과업 1의 수행 능력이 보이스 채팅(30% 감소)보다 텍스트 채팅(50% 감소) 조건에서 더 크게 감소했다[18]. 디지털 기술 기반의 과업에서도 다른 연구와 일관적으로 감각양식의 간섭이 수행능력에 부정적인 영향을 미쳤던 것이다. 그런데 Wang 등의 연구에서 보이스 채팅은 사람과 대화를 나누는 방식이었다. 본 연구는 기계와의 상호작용을 전제로 하는 음성 인터페이스를 이용할 경우 인지 시스템에 미치는 영향이 달라질 수 있으므로 [19] 이에 대한 추가적인 규명이 필요하다고 판단하였다.

따라서 본 연구에서는 선행 연구에서 다루지 않았던 음성 기반의 디바이스를 포함한 최근 미디어 멀티태스킹 환경에 초점을 맞춰, 정보를 처리할 때 감각양식의 간섭이 과업수행에 어떠한 영향을 미칠지 확인하고자 했다. 이를 위해 실제 디바이스(스마트폰과 스마트스피커)를 실험에 활용하였고, 과업의 종류도 디지털 디바이스를 통해 가장 빈번하게 실행되는 키워드 기반의 정보 검색 과업으로 설정하였다.

## 4. 실험

본 연구에서는 이론적 배경과 선행연구 검토를 기반으로 멀티태스킹 상황에서는 제한된 자원의 초과로 인해 인지부하가 증가하지만, 여러 과업을 서로 다른 자원을 활용해 분산 처리할 수 있을 경우 과업수행이나 정보처리 능력의 손실이 줄어들 것임을 가정하고 실험을 진행하였다.

### 4.1 참가자와 실험 설계

정상 수준의 시력, 청력을 가졌으며 모국어가 한국어인 연세대학교 학부생 58명(남자 29명, 여자 29명)의 참가자가 실험에 참여하였다. 평균 연령은 22.87세였다. 참가자들은 실험 참여에 대한 보상으로 2 크레딧을 제공받았다.

실험은 2(Single, Multi) x 2(Visual, Audio) 참가자 간 설계(between-subject design)로 이루어졌다. 독립변인은 과업수행 환경과 과업수행 수단의 주 감각양식 이었으며, 각각 (1) 환경: 단

일(Single) 또는 다중(Multi), (2) 수단: 시각(Visual) 또는 청각(Audio)으로 나뉘어졌다. 종속변인은 각 조건에 대한 인지부하 수준과 과업 수행능력이었다. 인지부하 측정을 위해 가장 보편적으로 활용되는 주관적 설문 척도인 NASA-TLX (Task Load Index)[20]를 사용하였다. 과업 수행능력은 주관적 설문 결과를 보완할 수 있는 객관적인 척도로써 측정되었다.

## 4.2 실험 과제

### 4.2.1 검색 과제

모든 조건에서 참가자들은 주어진 제시문(예: “내년 설날은 몇 월 며칠인가요?”, “영화 배움의 개봉일은 언제인가요?”)에 대한 답을 시각 또는 청각 기반의 디바이스를 활용하여 검색하고 종이에 기록했다. 모든 제시문은 키워드 입력(스마트폰) 또는 명령어 발화(스마트스피커)를 통해 한 번에 쉽게 검색할 수 있는 것으로 선별되었다. 제시문은 참가자가 과제가 주어지는 시점을 예측하지 못하도록 불규칙적인 간격으로 알림음과 함께 주어졌다. 시각 그룹의 경우 가장 많이 쓰이는 검색 서비스인 네이버 앱[21]을 활용하도록 했고, 청각 그룹의 경우 오류 발생을 막기 위해 의인화된 캐릭터 스피커(그림 3)를 통해 실험자가 미리 녹음된 답변을 재생하는 WoZ (Wizard of Oz) 방식[22]으로 진행했다.



그림 3. 실험에서 음성 에이전트로 활용된 캐릭터 스피커

### 4.2.2 시각 비교 과제

다중 과업 조건에서 참가자들은 검색 과제와 더불어 시각적 자원이 요구되는 비교 과제를 추가로 수행하였다(그림 4). 시각 비교 과제는 멀티태스킹 관련 선행연구[18]에서 활용된 과제와 유사한 것으로, 3 x 3로 구성된 두개의 표 안의 문자열을 비교하는 과제였다. 문자열은 무작위로 추출된 2개의 영소문자로 구성되었고, 9개의 문자열은 두 표에서 서로 다른 위치에 배치되었다. 참가자들은 두 표가 동일한 문자열 세트에 구성되었는지 확인하여, 하나라도 다른 문자가 있으면 키보드를 이용해 ‘N’을, 모두 동일하면 ‘Y’를 눌렀다. 본 세트 시작 전에 10회의 연습시행이 진행되었다. 각 시행에는 시간제한이 없었으나, 참가자들은 최대한 많은 과제를 수행하도록 지시받았다.



그림 4. 다중과업 조건에서 제시된 두 가지 실험 과제; 단일과업 조건에서는 검색 과제만 제시되었다.

### 4.3 실험 절차

참가자들은 각 조건에 임의로 할당되었으며, 실험실에 도착하여 동의서를 작성하고 실험에 대한 설명을 들었다. 실험은 총 2세트(각 15분)로 구성되었고, 세트 1에는 10개, 세트 2에서는 20개의 검색 과제가 제시되었다. 각 세트 후에 인지부하를 측정하기 위한 NASA-TLX 설문을 진행하였다. 실험을 모두 마친 뒤에 디바이스 및 멀티태스킹 경험을 조사하는 사후설문을 실시하였다. 다중과업 조건 참가자의 경우 시각 비교 과제에 대한 난이도를 묻는 문항에 추가적으로 응답하였다. 모든 문항은 리커트 7점 척도로 측정되었다.

## 5. 결과

### 5.1 인지부하 수준

검색과업 수행 환경과 인터페이스의 감각양식 차이가 인지부하에 미치는 영향을 알아보기 위해 이원분산분석(two-way ANOVA)을 실시하였다. 그 결과 두 요인 모두에서 주효과가 나타났다(그림 5). 다중과업 조건 참가자들( $M = 3.63, SD = .99$ )의 인지부하 수준이 단일과업 조건( $M = 1.90, SD = 1.07$ )에 비해 유의미하게 높았다,  $F(1, 54) = 42.88, p < .000$ . 또한 청각 기반 인터페이스를 사용한 참가자들( $M = 3.14, SD = 1.36$ )이 시각 인터페이스 그룹( $M = 2.50, SD = 1.26$ )에 비해 더 높은 인지부하를 보고했다,  $F(1, 54) = 5.41, p < .05$ . 두 조건 간 상호작용 효과는 나타나지 않았다.

단일과업보다 다중과업 조건에서 인지부하 수준이 높았다는 것은, 디지털 디바이스 기반의 미디어 멀티태스킹 환경에서도 여러 과업의 동시적 처리가 인지부하를 증가시킨다는 것을 의미한다. 이론적 기반을 토대로 가정했던 것과 같이 제한된 용량 모델에 따라 대다수의 멀티태스킹 관련 선행연구를 뒷받침하는 결과가 나타난 것이다.



그림 5. 이원분산분석으로 확인한 조건 간 인지부하 수준 차이

전반적인 인지부하 수준은 청각 인터페이스 그룹에서 시각 인터페이스 그룹보다 높게 나타났는데, 세부적으로 단일, 다중 과업 조건 내에서 두 그룹 간의 인지부하 수준 차이는 각각 다르게 나타났다. 독립표본 t검정 결과, 단일과업 조건에서는 청각 기반 인터페이스를 사용한 참가자들이 시각 인터페이스 그룹에 비해 더 높은 인지부하를 보고했다,  $t(25) = 2.26, p < .05$ . 반면 다중과업 조건에서는 인지부하가 청각 그룹에서 약간 높기는 했지만 청각 그룹( $M = 3.80, SD = 1.03$ )과 시각 그룹( $M = 3.45, SD = .94$ ) 간 통계적으로 유의미한 차이는 나타나지 않았다(그림 6). 시각과제가 추가적으로 주어진 다중 과업 조건에서 두 그룹 간의 인지부하 차이가 줄어들었던 것이다.

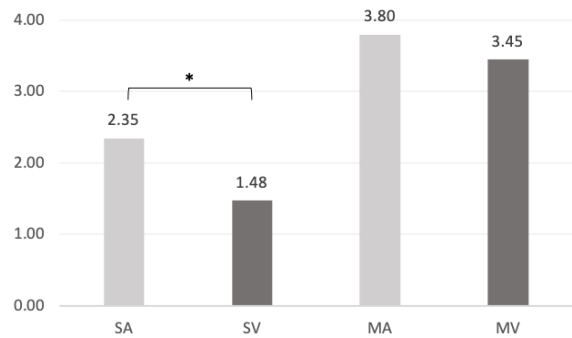


그림 6. 조건에 따른 인지부하 (평균 NASA-TLX 점수) 차이

이는 음성 인터페이스가 높은 인지부하를 야기할만 특성을 갖고 있지만, 멀티태스킹 환경에서는 스크린 기반의 인터페이스와의 차이를 줄일만한 이점이 있었다는 것을 의미한다. 즉, 다중 자원 이론에 따라 시각자원 경로에 병목이 발생했던 시각 인터페이스 그룹보다 청각 인터페이스 그룹에서 자원의 분산 처리를 통해 효율적 과업 수행이 가능했다고 해석할 수 있다.

## 5.2 과업 수행능력

### 5.2.1 검색 과제

검색과제의 수행능력은 참가자들이 종이에 기록한 답변의 정확도를 통해 측정되었다. 모든 조건에서 검색과제는 높은 수행도(평균 정답률 98.33%)를 보였으며, 조건에 따른 유의미한 차이는 나타나지 않았다.

### 5.2.2 시각 비교 과제

다만, 시각 비교 과제의 경우 청각과 시각 그룹 간의 유의미한 차이가 발견되었다. 과제의 정확도에 있어서 청각 인터페이스를 사용한 참가자들( $M=122.13$ ,  $SD=22.73$ )이 맞춘 정답 개수가 시각 인터페이스 그룹( $M=107$ ,  $SD=15.78$ )에 비해 높게 나타난 것이다,  $t(29) = 2.14$ ,  $p < .05$ (그림 7). 일반적으로 2차 과업의 수행능력은 1차 과업의 인지부하를 반영하는 지표이므로[23], 이는 인지부하 수준에서 나타난 것과 마찬가지로 감각양식의 방해가 적었던 청각 인터페이스 그룹에서 이점이 있었음을 보여준다. 특히 과업 수행능력은 주관적 척도에 비해 자원 할당의 효율성을 보다 직접적으로 나타내는 객관적인 측정치라는 점 [24]에서 더 명확한 증거라고 할 수 있다. 두 그룹 간 시각 비교 과제의 난이도에 대한 유의미한 차이는 존재하지 않았다.

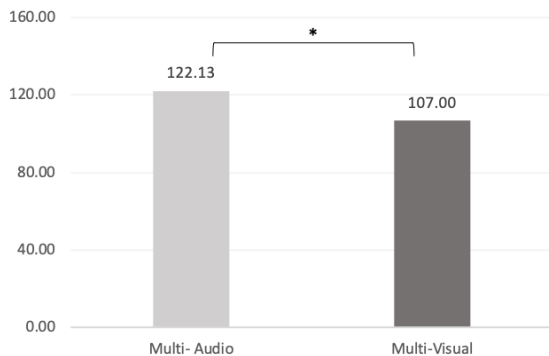


그림 7. 다중과업 조건에서 디바이스에 따른 과제 수행도 차이

### 5.3 디바이스 활용 및 멀티태스킹 경험

독립표본 t검정을 통해 디바이스 활용빈도 및 멀티태스킹 경험에 대한 참가자들의 사후설문 결과를 비교해본 결과(그림 8), 각 인터페이스에 대한 활용빈도 및 능숙도 모두 청각 인터페이스 그룹이 시각 그룹에 비해 유의미하게 낮았다( $t(56) = -25.87$ ,  $p < .000$ ,  $t(56) = -6.79$ ,  $p < .000$ ). 음성 인터페이스는 이미 보편화된 시각 기반 스크린 인터페이스에 비해 아직까지 일상적으로 쓰이지 않기 때문에, 활용도나 능숙도가 떨어지는 결과가 나타난 것으로 보인다.

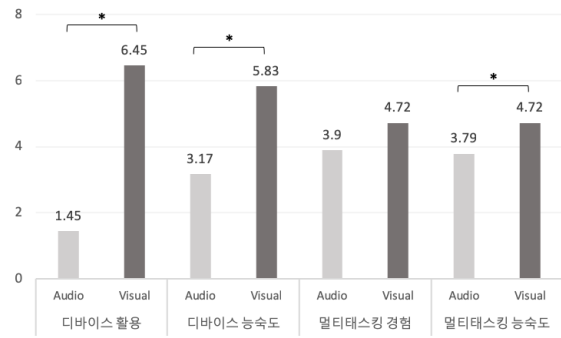


그림 8. 청각-시각 그룹 간 디바이스 및 멀티태스킹 경험 차이

이와 같은 결과는 음성 인터페이스의 주관적 인지부하 수준이 전반적으로 높게 나타난 결과에 대한 설명이 될 수 있다. 새롭고 익숙지 않은 시스템은 더 큰 인지부하를 초래하기 때문이다[25].

평소 멀티태스킹을 얼마나 자주하는지 묻는 멀티태스킹 경험의 경우 시각 그룹에서 약간 높긴 했지만, 통계적으로 유의미한 차이는 없었다. 다만, 멀티태스킹에 대한 능숙도는 시각 그룹에서 유의미하게 높다고 응답했는데( $t(56) = 2.16$ ,  $p < .05$ ), 두 그룹 간 멀티태스킹 경험에서 큰 차이가 없었기 때문에 익숙하지 않은 인터페이스를 활용해 과업을 수행했던 것이 사후 설문에 영향을 주었을 가능성이 있다.

## 6. 논의

본 연구는 음성 인터페이스가 단독으로 사용되었을 때 스크린 인터페이스에 비해 더 높은 인지부하를 발생시키지만, 시각 자원의 충돌이 발생할 수 있는 다중 과업 환경에서는 음성 인터페이스의 이점으로 인해 이러한 차이가 어느 정도 상쇄될 수 있음을 보여주었다. 과업 수행결과로 나타난 행동적 증거 (behavioral evidence)는 특정 환경에서 음성 인터페이스가 정보 처리 효율을 높이는 역할을 할 수 있음을 더 명백하게 보여주었다.

음성 인터페이스가 사용자에게 부과하는 높은 인지부하는 생소하고 능숙하지 않다는 이유 외에도 인터페이스가 가지는 특성에 기인한다고 설명하는 연구가 많다. 즉, 피드백 부족으로 인한 음성 인터페이스의 ‘보이지 않는(invisible)’ 특성이 시스템에 대한 통제감을 감소시키고[26], 이로 인해 추가적인 인지적 노력이 요구된다는 것이다[27]. 이러한 단점에도 불구하고 멀티태스킹 환경에서는 음성 인터페이스가 이점을 제공할 수 있음을 보여준 본 연구의 결과는, 다른 인터페이스와의 조합 등을 통해 음성 인터페이스의 인지적 가치를 잘 활용하는 것이 그 자체가 가진 높은 인지부하를 보완할 수 있는 하나의 방법이 될 수 있음을 의미한다.

이론적 검토에서 살펴본 스레드 인지 모델을 적용하면, 음성 인터페이스 그룹에서 나타난 이점이 구체적으로 어떤 시점에 발생했는지 추론해볼 수 있다. 그림 9a, 9b는 각 그룹에서 시각 비교 과제와 검색 과업이 자원의 교차적 분배 과정을 어떤 방식으로 거쳐 동시에 수행되는지 보여준다. 인터페이스 간의 확실한 차이는 시각 자원 경로에서 드러나는데, 우선 스크린 인터페이스를 사용한 그룹에서는 검색과제 수행 시 계속해서 시각 자원을 사용해야 하기 때문에 검색과제를 완전히 마무리해야만 시각 비교 과제를 시작할 수 있다. 반면, 음성 인터페이스 그룹에서는 검색 과업을 진행하는 동안 비교 과제의 시각 자극을 인코딩할 기회가 존재한다.

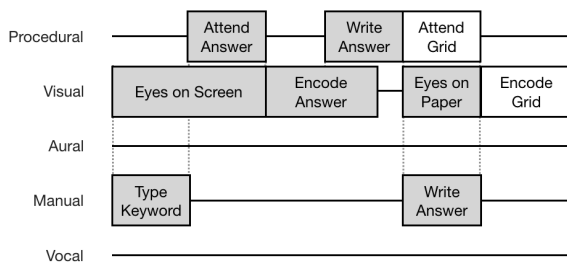


그림 9a. 스레드 인지 이론을 기반으로 추론한 스크린 인터페이스 그룹에서의 멀티태스킹 과정. 검색 과업(회색 박스)을 수행하는 동안 시각 자원이 지속적으로 활용되기 때문에 시각 비교 과제(흰 박스)의 인코딩이 뒤늦게 이루어진다.

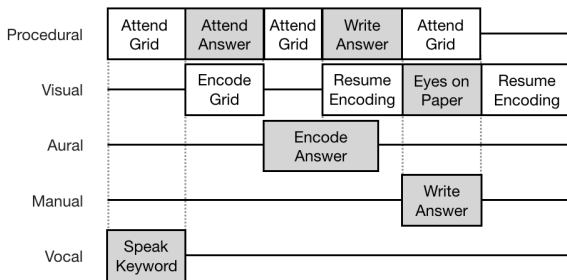


그림 9b. 스레드 인지 이론을 기반으로 추론한 음성 인터페이스 그룹에서의 멀티태스킹 과정. 검색 과업을 수행하는 동안에도 시각 자원이 추가적인 시각 비교 과제에 활용될 수 있다.

절차적 자원에서 각 스레드의 우선순위를 선행적으로 결정하기 때문에 청각 인터페이스 그룹에서도 시각 경로가 계속해서 활성화되는 것은 아니지만, 시각 그룹에 비해 상대적으로 추가적인 시각 자극을 인코딩할 기회가 더 많다. 따라서 스레드 인지 모델을 통해 확인할 수 있는 이러한 과제 전환의 기회들이 음성 인터페이스를 사용했을 때 과업수행에 긍정적인 영향을 미쳤다고 해석할 수 있다.

## 7. 결론

본 연구는 스크린 중심의 미디어 멀티태스킹 환경에서 음성 인터페이스의 잠재적 역할을 살펴보고자 했고, 실증적 실험을 통해 시각 자원의 충돌을 발생시키는 스크린 인터페이스 사이에서 음성 인터페이스가 추가적인 정보처리 경로를 제공하는 역할을 할 수 있음을 규명하였다.

본 연구는 선행 연구에서 다루지 않았던 새롭게 변화하는 미디어 멀티태스킹 환경에 초점을 맞췄고, 실제와 유사한 실험 환경을 조성해 여러 인지 이론을 기반으로 한 가설을 검증하였는데 의의가 있다. 나아가 스레드 인지 모델을 적용하여 멀티태스킹 과정을 구체적으로 추론함으로써 음성 인터페이스를 사용했을 때의 인지적 이점이 어떻게 발생되었는지 명확하게 밝히고자 했다.

지금까지 음성 인터페이스의 강점으로 주로 부각된 것은 ‘손에서 자유롭다(hands-free)’는 특성으로, 요리나 운전 등의 상황에서 손을 쓸 수 없어도 음성을 이용해 디바이스를 활용할 수 있다는 점이었다[28]. 하지만 본 연구의 결과는 보편화된 스크린 기반의 인터페이스와 비교했을 때, ‘시각에서 자유롭다(Eyes-free)’는 하나의 특성만으로도 멀티태스킹 환경에서 음성 인터페이스가 경쟁적 우위를 가질 잠재력이 존재한다는 함의점을 제공한다.

본 연구의 한계로는 참가자의 규모나 연령대가 다소 제한적이었으며, 상용화된 디바이스를 활용했지만 실험 과제가 주어진 실험실 환경이었기에 연구 결과를 일반화하기에는 보완이 필요하다. 따라서 향후 연구에서는 자율적으로 미디어 간 전환을 할 수 있고, 시각 비교 과제에 비해 인지적 노력이 적게 드는 콘텐츠 소비를 위주로 멀티태스킹을 할 때도 본 연구와 동일한 결과가 나타날지에 대한 검증을 진행하고자 한다.

또한, 스레드 인지 모델을 통해 추론한 음성 인터페이스의 인지적 이점이 실제로도 동일한 패턴을 통해 나타나는지에 대해서도 추가적인 검증이 필요하다. 향후 연구에서는 아이트래킹과 같은 연구 방법을 적용해 직접적이고 객관적인 데이터를 통해 본 연구에서 가정한 멀티태스킹 시나리오의 타당성을 확인하고자 한다. 이 외에도 각기 다른 채널에서 유입된 정보가 각각 어느 정도 깊이로 처리되는지, 또는 본 연구에서 초점을 맞춘 디바이스에서 출력되는 정보 처리 방식이 아닌 음성 명령과 같은 사용자 입력 방식이 멀티태스킹 행동에 어떤 영향을 주는지에 대해서도 추가적으로 조사해볼 예정이다.

마지막으로 본 연구 결과에서 나타났듯 단독으로 사용되었을 때보다 멀티태스킹 환경에서 유의미한 강점을 가지는 음성 인터페이스의 잠재성을 극대화하기 위한 실질적인 디자인 가이드라인을 정립할 필요가 있다. 예를 들어, 본 연구에서 확인했던 것처럼 디바이스 간 감각양식 채널의 적절한 조화는 사용자 효율을 최대한 높일 수 있는 미디어 멀티태스킹 환경을 설

계하기 위한 중요한 디자인 요소(factor) 중 하나로 고려될 수 있다.

## 참고문헌

- [1] Van Cauwenberge, A., Schaap, G. and Van Roy, R. "TV no longer commands our full attention" : Effects of second-screen viewing and task relevance on cognitive load and learning from news. *Computers in Human Behavior*. 38. Elsevier. pp. 100-109. 2014.
- [2] NewsMediaWorks. What we learned from the 2018 Deloitte Media Consumer Survey. <https://newsmediaworks.com.au/what-we-learned-from-the-2018-deloitte-media-consumer-survey> January 14, 2019.
- [3] Facebook business. Mobile and TV: Between the Screens. <https://www.facebook.com/business/news/insights/mobile-and-tv-between-the-screens#Multi-screening-has-gone-mainstream> January 14, 2019.
- [4] Wang, Z., Irwin, M., Cooper, C. and Srivastava, J. Multidimensions of media multitasking and adaptive media selection. *Human Communication Research*. 41(1). International Communication Association. pp. 102-127. 2015.
- [5] Canalys. Smart speaker installed base to hit 100 million by end of 2018. <https://www.canalys.com/newsroom/smart-speaker-installed-base-to-hit-100-million-by-end-of-2018> January 14, 2019.
- [6] Lang, A. The limited capacity model of mediated message processing. *Journal of communication*. 50(1). International Communication Association. pp. 46-70. 2000.
- [7] Mayer, R. E. and Moreno, R. Nine ways to reduce cognitive load in multimedia learning. *Educational psychologist*. 38(1). Taylor & Francis. pp. 43-52. 2003.
- [8] Wickens, C. D. Multiple resources and performance prediction. *Theoretical issues in Ergonomics Science*. 3(2). Taylor & Francis. pp. 159-177. 2002.
- [9] Salvucci, D. D. and Taatgen, N. A. Threaded cognition: An integrated theory of concurrent multitasking. *Psychological review*. 115(1). American Psychological Association. pp. 101-130. 2008.
- [10] Salvucci, D. D., Taatgen, N. A. and Borst, J. P. Toward a unified theory of the multitasking continuum: From concurrent performance to task switching, interruption, and resumption. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM. pp. 1819-1828. 2009.
- [11] Armstrong, G. B. and Chung, L. Background television and reading memory in context: Assessing TV interference and facilitative context effects on encoding versus retrieval processes. *Communication Research*. 27(3). SAGE journals. pp. 327-352. 2000.
- [12] Bolls, P. D. and Muehling, D. D. The effects of dual-task processing on consumers' responses to high- and low-imagery radio advertisements. *Journal of Advertising*. 36(4). Taylor & Francis. pp. 35-47. 2007.
- [13] Bowman, L. L., Levine, L. E., Waite, B. M. and Gendron, M. Can students really multitask? An experimental study of instant messaging while reading. *Computers & Education*. 54(4). Elsevier. pp. 927-931. 2010.
- [14] Pool, M. M., Van Der Voort, T. H., Beentjes, J. W. and Koolstra, C. M. Background television as an inhibitor of performance on easy and difficult homework assignments. *Communication Research*. 27(3). SAGE journals. pp. 293-326. 2000.
- [15] Jeong, S. H. and Hwang, Y. Multitasking and persuasion: The role of structural interference. *Media Psychology*. 18(4). Taylor & Francis. pp. 451-474. 2015.
- [16] Hwang, Y. and Jeong, S. H. Multitasking and task performance: Roles of task hierarchy, sensory interference, and behavioral response. *Computers in Human Behavior*. 81. Elsevier. pp. 161-167. 2018.
- [17] Lauricella, A. R., Blackwell, C. K. and Wartella, E. The "New" technology environment: The role of content and context on learning and development from mobile media. In *Media exposure during infancy and early childhood*. Springer. pp. 1-23. 2017.
- [18] Wang, Z., David, P., Srivastava, J., Powers, S., Brady, C., D' Angelo, J. and Moreland, J. Behavioral performance and visual attention in communication multitasking: A comparison between instant messaging and online voice chat. *Computers in Human Behavior*. 28(3). Elsevier. pp. 968-975. 2012.
- [19] Myers, C., Furqan, A., Nebolsky, J., Caro, K. and Zhu, J. Patterns for How Users Overcome Obstacles in Voice User Interfaces. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM. pp. 6. 2018.
- [20] D Hart, S. G. and Staveland, L. E. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. 52. Elsevier. pp. 139-183. 1988.



- [21] NAVER Corp. 네이버 - NAVER.  
<https://itunes.apple.com/us/app/id393499958?mt=8>  
 2019.1.14.
- [22] Maulsby, D., Greenberg, S. and Mander, R. Prototyping an intelligent agent through Wizard of Oz. In Proceedings of the INTERACT '93 and CHI'93 conference on Human factors in computing systems. ACM, pp. 277–284. 1993.
- [23] Paas, F., Tuovinen, J. E., Tabbers, H. and VanGerven, P. W. Cognitive load measurement as a means to advance cognitive load theory. *Educational psychologist*, 38(1). Taylor & Francis, pp. 63–71. 2003.
- [24] Yeh, Y. Y. and Wickens, C. D. Dissociation of performance and subjective measures of workload. *Human Factors*, 30(1). SAGE journals, pp. 111–120. 1988.
- [25] Sweller, J. Cognitive load theory, learning difficulty, and instructional design. *Learning and instruction*, 4(4). Elsevier, pp. 295–312. 1994.
- [26] Limerick, H., Moore, J. W. and Coyle, D. Empirical evidence for a diminished sense of agency in speech interfaces. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. ACM, pp. 3967–3970. 2015.
- [27] Corbett, E. and Weber, A. What can I say?: addressing user experience challenges of a mobile voice user interface for accessibility. In Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services. ACM, pp. 72–82. 2016.
- [28] Getty, D., Biondi, F., Morgan, S. D., Cooper, J. M. and Strayer, D. L. The Effects of Voice System Design Components on Driver Workload. *Transportation Research Record*. 0361198118777382. SAGE journals, pp. 1–7. 2018.