

#### 저작자표시-비영리-변경금지 2.0 대한민국

#### 이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

• 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

#### 다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건 을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer 🖃





#### 언론정보학석사학위논문

# 온라인 문서에서 하이라이트의 효과 및 생성 방법에 관한 연구

2014년 8월

서울대학교 대학원 언론정보학과 오 종 환

#### 국문초록

온라인에서 받아들이는 정보의 양이 급격히 늘면서 그에 따른 인지적 부하가 생가난다. 그 결과 온라인 문서의 모든 내용을 정독하기 보다는 부분적으로 훑어 읽는 '스킴 리딩(skim reading)'이 일반화되고 있다. 이러한 읽기 방식의 변화는 글의 내용을 효율적으로 파악할 수 있도록 도와주는 보조 도구의 필요성을 제기하였다. 이 연구에서는 하이라이트(highlight)가 문장의 일부분을 눈에 띄게 밝은 색으로 표시함으로써 온라인 문서를 보다 효율적으로 읽는데 도움을 줄 것이라고 가정하고 몇 가지 실험을 수행하였다. 첫 번째 실험에서는 참여자들이 스킴 리딩 그룹과 정독 그룹으로 나뉘어 각각 세 가지 다른 조건의 하이라이트가 표시된 글을 읽게 하였다. 이와 동시에 아이트래커(eyetracker)를 이용하여 참여자의 시선이 어느 부분에 오래 머무는지 측정하였으며, 글을 읽고 나서는 이해도를 테스트하였다. 그 결과 하이라이트는 글의 이해도 향상에 도움을 주며, 실제 시선도 더 오래 머무는 것을 관찰하였으며, 이러한 효과는 스킴 리딩 그룹에서 두드러지게 나타남이 밝혀졌다. 두 번째 실험에서는 커뮤니케이션 소스의 개념을 기존의 미디어 연구가 아닌 하이라이트 생성 방법에 적용함으로써. 심리적인 차이가 불러오는 인지(perception)의 차이를 탐구하고자 하였다. 사람들은 같은 하이라이트 문서를 보고도 컴퓨터 알고리즘(algorithm)에 의한 하이라이트 생성 방법 보다는, 다수 사람들의 직접 참여하여 작성한 크라우드소싱(crowdsourcing) 방법에 더 높은 신뢰도를 보였다. 더 나아가 프로토타이핑(prototyping)의 단계로, 모바일 기기에서 사람들의 읽기행동 패턴을 수집하여 자동으로 하이라이트를 생성하는 시도를 하였다. 예비 실험을 통해 사람들이 모바일 기기의 화면에서 손가락을 터치하는 위치와 각각의 문장이 화면에서 표시되는 위치, 그리고 문장의 노출시간을 자동으로 계산하는 어플리케이션을 제작하였다. 이 읽기 데이터를 바탕으로 하이라이트를 자동으로 생성하는 방법을 시험해 보고 결과를 기록하였다. 추가적으로 하이라이트 생성이 실용화되기 위한 앞으로의 연구방향을 논의하였다.

주요어: 하이라이트, 읽기 행동, 소스(source), 스킴 리딩(skim reading), 온라인 문서

학 번: 2012-22955

## 목차

1.	서론	<u>-</u>	1
2.	관련	면 연구 및 연구문제	6
	2.1.	스킴 리딩의 특징	6
	2.2.	온라인 문서에서의 하이라이트의 표시	7
	2.3.	하이라이트 생성 방법	11
	2.3	.1. 컴퓨터 알고리즘을 이용한 방법	11
	2.3	.2. 크라우드소싱을 이용한 방법	14
	2.4.	하이라이트를 생성하는 소스에 따른 차이	17
3.	실현	] 1: 하이라이트 조건에 따른 내용 이해도와 시선의 차이	25
	3.1.	개요	25
	3.2.	측정	26
	3.3.	과정	31
	3.4.	결과 및 분석	31
4.	실현	] 2: 소스에 대한 심리적 차이	43
	4.1.	개요	43
	4.2.	조건과 처리	44
	4.3.	측정	47
	4.4.	과정	51

4	4.5.	결고	ł 및 분석	53			
5.	하이	]라이	트의 생성	55			
4	5.1.	개요	2	55			
4	5.2.	영성	) 촬영을 통한 예비 실험	58			
	5.2	.1.	측정	59			
	5.2	.2.	과정	60			
	5.2	.3.	결과	62			
4	5.3.	프로	르토타입 제작	63			
	5.3	.1.	측정	63			
	5.3	.2.	결과	65			
6.	논의	]		67			
7.	결론	<u>!</u>		74			
8.	참고	2 문 한	원	76			
ΑP	A RSTR A CT 82						

## 표 목차

표 1. 참여자에게 부여된 6개의 실험 세트30
표 2. 평균중심화 후 각 선택지를 고른 횟수의 평균32
표 3. 읽기 방법/하이라이트 조건에 따라 선택한 단어 횟수의
이원분산분석
표 4. 읽기 방법/하이라이트 조건에 따른 선택한 문장의 횟수와 고난도
문제 점수의 이원분산분석39
표 5. 각 AOI 에 단위 면적당 시선의 평균 누적시간40
표 6. 읽기 방법/하이라이트 조건에 따라 선택한 단어 횟수의
이원분산분석40
표 7. 신뢰도와 대표성의 측정 항목49
표 8. 평균중심화 이후 소스에 따른 각 인지요소 점수의 평균55
표 9. 각 소스에 따른 인지요소 점수의 일원분산분석55
표 10. 글에서 직접 하이라이트를 표시한 문장과 노출 시간 간의
피어슨 상관관계61

## 그림 목차

그림 1. 읽기 방법 및 하이라이트 조건에 따라 선택한 단어의 횟수37
그림 2. 읽기 방법 및 하이라이트 조건에 따라 선택한 문장의 횟수와
고난도 문제 점수
그림 3. 하이라이트 조건에 따라 각 AOI 간의 시선이 머문 시간 비교
40
그림 4. 소스의 종류에 따른 심리적 차이를 알아보기 위한 실험용 웹
사이트 화면49
그림 5. 모바일 화면에서 글 읽는 화면을 관찰하기 위한 비디오 실험57
그림 6. 문장의 노출 시간을 측정하기 위해 고안한 4가지 방법61
그림 7. 각 문장 별 모바일 기기 화면 노출시간을 세 가지 다른
방법으로 계산한 결과와 사용자의 직접 하이라이트 표시 횟수61
그림 8. 프로토타입으로 수집한 각 문장의 평균 노출시간과 문장
중요도66

#### 1. 서론

인터넷 등의 온라인 서비스의 확장으로 인해 전자적인 형태로 공급되는 정보가 급증함에 따라 개인용 컴퓨터의 모니터나 스마트폰 등의 전자 기기에서 글을 읽는 경우가 증가하였다. 이는 정보의 과잉공급(information overload)를 초래하게 되었으며, 이에 따라 한정된 시간에 주어지는 많은 정보를 읽기 위해 필연적으로 인지적 부하(cognitive overload)를 수반하게 된다. 인지적 부하는 사람들이 글을 읽는 습관에도 변화를 가져오게 되어 스킴 리딩(skim reading)이 일반화되었다. 스킴 리딩에서 특징적으로 나타나는 행동들이 관찰되었는데, 글을 집중하여 순차적으로 정독하기 보다는, 전반적으로 훑어가며 중요하다고 생각하는 정보만 선택적으로 읽거나 비선형적(non-linear) 순서로 읽는 것이다(Liu, 2005), 특히 스킴 리딩은 온라인 뉴스 등 정보를 얻기 위해 글을 읽는 상황에서 두드러진다. 이는 사람들이 온라인에서 읽는 글들은 문학 작품이기 보다는 뉴스 등 정보를 전달하는 글인 경우가 상대적으로 많기 때문이다. 이런 종류의 글에서는 내재된 문학적 가치나 수사학적인 의미를 파악하기보다는 중심 내용을 빠른 시간 내에 내용을 정확하게 파악하는 것이 중요하다(Shaikh & Chaparro, 2004).

스킴 리딩은 많은 양의 문서를 빠르게 읽을 수 있다는 장점이 있지만, 짧은 시간 안에 글을 읽기 때문에 필연적으로 동반하는 문제점이 있다.

예컨대 내용을 적게 기억하거나 이해도가 떨어지는 등의 문제가 지적되었다(Muter & Maurutto, 1991; Duggan & Payne, 2009; Dyson & Haselgrove, 2001). 스킴 리딩이 가진 단점을 보완하고자 하는 방법이 오랫동안 사용되어 왔는데, 하이라이트가 대표적이다. 여기서 하이라이트는 중요한 부분에 밝은 색으로 줄을 그어 강조하는 방법을 의미한다. 이 방법은 종이 문서에서도 읽기 과정에서 정보 습득 과정을 보조하고자 사용되어온 방법으로, 문서를 나중에 다시 읽거나 해당 문서를 처음 보는 사람이 읽을 때, 효율적으로 이해할 수 있도록 돕는 역할을 한다. 하이라이트가 시선을 끄는 원리는 인지적인 면에서 설명하자면 폰 레스트로프 효과(von Restroff effct)로 설명될 수 있는데, 여러 항목 가운데 지각적으로 구별되는 항목이 있다면, 이를 다른 항목에 비해 더 잘 기억하게 되는 현상을 말한다(Hunt, 1995). 이러한 하이라이트의 효과는 스크린 기반의 장치 위에서 글을 읽는 상황에도 점차 널리 사용되고 있다. 온라인 환경에서는 정보 처리 기술의 도움으로 원본의 손상 없이 하이라이트를 생성하거나 수정한 문서를 다수의 사람에게 쉽고 빠르게 전달하는 것이 가능하다. 이를 이용하여 제공하는 다양한 알고리즘과 어플리케이션들이 개발되고 있다(Mani, 2001).

정보를 얻기 위한 글에서 문서에 존재하는 중요 내용을 찾아내고 획득하는 것이 목적이라면, 단위 시간당 획득한 정보량이 효율성을 나타내는 지표라고 볼 수 있을 것이다. 글을 정독했을 때는 많은 시간을 들였기 때문에 글의 중요 내용을 파악할 수 있는 가능성이 높겠지만 정독을 하지 않아도 스킴 리딩 만으로도 중요한 정보를 파악할 수 있다면 효율적인 읽기를 했다고 볼 수 있을 것이다. 반면에 정독을 했음에도 원하는 정보를 찾지 못했을 경우에는 많은 시간을 낭비하게 된다. 따라서 스킴 리딩 상황에서 글에 포함된 중요한 정보를 빠르게 파악할 수 있도록 도와주는 보조 도구의 존재는 사람들의 읽기 행위에 이득을 가져다 줄 수 있을 것이라고 기대할 수 있다.

적절한 하이라이트가 온라인 미디어 환경에서 스킴 리딩 환경에서 글의 내용을 파악과 이해에 도움을 줄 수 있다면, 문서에 하이라이트를 생성하기 위한 방법을 고려해보아야 할 것이며, 가능한 방법을 세 가지로 나누어볼 수 있다. 첫 번째는 전문가에 의한 방법으로, 문서의 원작자 또는 문서의 내용과 관련된 분야의 전문가가 중요한 부분에 하이라이트를 표시하는 방법이다. 가장 정확하고 신뢰도 높은 하이라이트를 생성해낼 수 있는 방법이기는 하지만, 많은 문서에 전문가가 일일이 투입되어 하이라이트를 표시하는 것은 비용이 많이들뿐만 아니라 현실성이 부족하다. 두 번째로는 컴퓨터의 알고리즘을 활용한 방법이다. 컴퓨터가 단어 사용의 반복이나 문장 구조 분석을 통한 일정 수준의 의미를 파악하여 텍스트의 주요 문장을 추출하고 하이라이트를 생성하는 방법이다. 컴퓨터는 인간의 언어를 근본적으로 이해할 수 없기 때문에 여러 한계가 지적되어 왔으나(Morris, Kasper, &

Adams, 1992), 관련 기술의 발달로 정확도가 점점 높아지고 있으며 특히 뉴스 미디어 등 설명적인 글을 파악하는 데 비교적 높은 정확도를 보이고 있다(Chi et al., 2005). 셋째는 크라우드소싱(crowdsourcing)의 방법으로, 대중의 집단지성(collective intelligence)을 활용하는 방법이다. 전문가가 아닌 다수의 비전문가가 참여하여 원하는 정보, 아이디어를 얻거나 서비스를 제공하는 방법으로, 다수의 참여가 용이한 온라인 커뮤니티를 배경으로 등장하였다. 앞서 언급한 전문가가 직접 작업할때의 비용 문제, 컴퓨터 알고리즘 사용시의 정확도 혹은 독자의 신뢰도문제를 해결하기 위해 가장 최근에 제시된 방법이다. 참여 인원이 많을 수록 정확한 하이라이트를 찾아낼 수 있는 구조이지만, 참여 인원이 적은 문서에서는 정확도를 보장하기가 어렵고, 전문가가 직접 표시하는 방법에 비해 다소 오류가 발생할 수 있을 것이다.

본 연구에서는 온라인에서 스킴 리딩의 보조 도구로서 사용되는 하이라이트의 효과를 온라인 문서에서 검증하고자 한다. 각기 다른 하이라이트 조건을 주고 읽기를 수행한 후에 글의 내용을 얼마나 파악했는지를 문제 풀이를 통해 파악하고자 한다. 이와 더불어 아이트래커를 이용하여 실제로 시선이 하이라이트에 오래 머무는 지를 측정해 본다. 아울러 하이라이트를 생성하는 방법으로 컴퓨터 알고리즘과 비교적 최근에 새롭게 제안된 크라우드소싱이라는 두 가지하이라이트 생성 방법을 비교한다. 컴퓨터 알고리즘을 이용한 방법과

함께 새롭게 연구되고 있는 크라우드소싱이 하이라이트 생성에 도움을 줄 수 있는지를 탐구하고자 한다. 아울러 사람들이 하이라이트를 보고느끼는 심리적인 차이에도 주목하였다. 실제 하이라이트의 수준 외에도다른 부수적인 요소가 하이라이트가 생성된 문서에 대한 평가에 영향을미치는 지를 살펴본다. 하이라이트를 생성 과정과 정보를 발생하는커뮤니케이션의 과정의 유사성에 주목하여, 하이라이트 생성 주체를소스(source)로 보고, 수용자들이 이를 어떻게 인지하는지를 실험을통해 비교해 볼 것이다. 그리고 최근에 널리 사용되는 모바일 기기의특성을 활용하여 사람들이 모바일에서 수행하는 읽기 행동 패턴을 통해하이라이트를 자동으로 추출할 수 있는지를 살펴보고자 한다. 이를 위해사용자의 행동 패턴을 수집하는 어플리케이션의 프로토타입을 제작해보았다.

#### 2. 관련 연구 및 연구문제

#### 2.1. 스킴 리딩의 특징

Liu(2005)는 사람들의 읽기 습관이 종이 문서 환경에서 디지털 문서 환경으로 옮겨가면서 스킴리딩이 일반적인 행태가 되어가고 있음을 설문 조사를 통해 밝힌 바 있다. 스킴 리딩을 할 때에는 핵심 단어만을 골라 찾아 읽거나(keyword spotting), 비선형적으로 읽어 나가는(nonlinear reading) 특징을 보였다. 그리고 글의 전체를 모두 읽지 않고 선택적으로 일부만을 읽으며, 하나의 글을 반복해서 읽지 않고 한 번만 읽는 행동이 더 두드러지게 나타난다는 것을 발견했다.

스킴 리딩할 때에는 빠른 시간 안에 중심 내용을 파악할 수 있다는 장점이 있지만, 세세한 정보나 추론이 필요한 부분을 기억하지 못하는 경우가 많았다(Duggan & Payne, 2009). Dyson과 Haselgrove(2001)은 웹 페이지와 같이 스크롤과 멈춤을 반복하여 글을 읽을 수 있는 환경에서 읽는 속도를 달리하여 나눈 두 집단 간의 실험적 비교를 통해 빠르게 읽기를 할 경우 전반적인 글의 내용에 대한 이해도가 떨어짐을 밝혔다. Muter와 Maurutto(1991)는 컴퓨터 화면과 종이 문서에서 글을 스킴 리딩을 하거나 정독을 하는 과제를 부여한 후 단어 기억 수준을 비교했다. 정독했을 경우에는 두 조건 간 기억하는 단어의 수는 큰 차이가 없었지만,

스킴 리딩 조건에서는 정독 조건과 비교하여 걸린 시간에 비해 기억한 단어의 수가 크게 적었다. 이를 근거로 스킴 리딩의 보조 도구가 필요함을 주장하였다.

안구의 움직임을 측정하여 스킴 리딩을 할 때 사람들이 행동 패턴을 읽으려는 시도도 있었다. Duggan과 Payne(2009)은 아이트래커(eye-tracker)를 사용하여 스킴 리딩을 할 때 독자의 시선이 움직이는 패턴을 관찰함으로써 사람들이 글을 읽을 때 중요한 부분을 찾아내기 위해 다양한 전략을 쓰고 있음이 밝혀졌다. 독자들은 각 문단의 첫 문장을살피는데 시간을 들이는 등의 전략을 사용하는 것이 관찰되었으나, 문단의 앞부분을 주로 읽은 참여자와 뒷부분을 더 많이 본 참여자 간의 전체적인 내용 기억에는 크게 차이가 없었다. 이로부터 중요한 부분을 빠른 시간 내에 직접적으로 도와 줄 수 있는 보조적인 장치의 필요성을 고려해볼 수 있다.

#### 2.2. 온라인 문서에서의 하이라이트의 표시

이러한 필요성에 따라 연구자들은 문서 내에서 중요한 부분을 강조하는 방법에 대한 연구를 진행했다. 그 중에 한 가지는 특정 부분의 크기를 달리하는 것이다. Hornback과 Frokjaer(2001)는 페이지에서 썸네일(thumbnail) 형식으로 문서를 보여주면서 중요하다고 판단되는 부분을 다른 부분보다 더 크게 보여주는 어안(fisheve) 인터페이스를

제시하였다. 크기를 통해 일부가 강조가 됨으로써 주목을 할 수 있는 효과가 있었지만, 이는 일부의 크기가 변함으로써 화면에 문서를 보여줄 때 문서 전체의 배치가 뒤틀리는 등의 문제가 발생하여 일반적인 온라인 문서에 모두 적용하기에는 힘들다는 한계점이 있었다. 이러한 문제를 개선하고자, 본문에 목차를 먼저 제시하는 전통에 착안하여, 화면을 나누어 개요와 자세한 내용을 따로 제시하는 방법이 제안되었다. 이와 함께 사용자 개개인의 프로파일에 기반하여 중요한 부분이 사용자 별로 다르게 제시되는 인터페이스가 만들어졌다(Harama et al., 2003). 비교적 소형의 모바일 화면에서 글을 읽는 경우가 많아지면서, 글의 내용을 보여주기에 앞서 요약을 제공하는 방법도 자세한 제안되었다(Alam et al., 2003). 그러나 이러한 방법은 본문에 앞서 추가적인 정보가 제공되면서 오히려 정보의 양이 늘어난다는 단점이 있다. 또한 화면을 나누어 제공할 수 있거나 모바일 화면에 특화되는 등. 특정 인터페이스에서만 유용하다는 점이 지적된다.

이에 반해 하이라이트는 중요한 방법을 강조하는 방법으로 종이 문서 환경에서부터 읽기의 보조 도구로 널리 사용되어 왔으며 이에 대한 연구도 활발히 이루어져 왔다(Liu, 2005). 과거에는 교육학 분야에서 하이라이트 연구가 활발했다. Fowler 등(1974)은 학습 자료에 하이라이트를 표시하는 실험을 통해, 하이라이트를 통한 학습이 이를 통해 학습하는 학생의 기억을 유지시키는 데 도움이 된다는 사실을 밝혔다. 이러한 종이 문서에서의 하이라이트의 효과를 디지털 환경에서도 검증하는 연구가 이어졌다. Kol과 Schcolnik(2000)은 디지털 문서를 활용한 교육을 위해 학생들을 대상으로 한 실험을 수행하면서 스킴 리딩을 할 때 전체적인 개요 제공과 하이라이트를 포함하는 보조 도구를 제작하여 제공함으로써 학습 효율을 향상시킬 수 있음을 증명하였다. Chi 등(2005)은 문서에 하이라이트를 어떻게 표시해주는 것이 효율적인 방법인지를 연구하였다. 사람의 뇌가 정보를 인지하는 원리로부터 핵심 단어에 하이라이트를 표시할 뿐만 아니라 몇몇 그 단어를 둘러싼 중요한 문장에도 다른 색상으로 함께 하이라이트를 표시하는 방법을 제안했다.

하이라이트를 보조 도구로 사용하는 것은 인지적으로 특이한 부분이 시선을 끌고 더 오랜 기억을 남긴다는 폰 레스트로프 효과에 근거한다. Chi 등(2007)은 이러한 효과가 하이라이트가 포함된 온라인 문서에서도 적용됨을 증명하고자, 서로 다른 방법으로 하이라이트가 포함된 문서를 제시하면서 사람들의 시선이 하이라이트에 더 오래 머무는 것을 실험적으로 확인하였다. 이 실험에서는 참여자에게 실험에 앞서 내용에 관련된 문제를 제시하고 하이라이트 유무에 따라 답을 찾아내는 데 걸린 시간의 차이를 측정하였다.

그러나 일반적으로 사람들이 글을 읽을 때 문제의 답을 찾는 과정을 거치는 것이 아니기 때문에, 사전에 제시된 문제에 답을 찾는 구조는 현실과 거리가 있다. 또한 이러한 하이라이트의 역할이 빠른 시간 내에 스킴 리딩을 하는 데에만 효과가 있을 것인지, 혹은 시간을 두고 꼼꼼하게 정독하는 조건에서도 같은 효과가 있는 것인지를 비교하지 않았다. 하이라이트가 스킴 리딩에서 읽기의 효율을 높이는 도구임을 주장하기 위해서는 서로 다른 조건에서 비교해야 한다. 앞서 언급한 Duggan과 Payne(2009)의 실험에서는 스킴 리딩과 정독의 조건을 나누어 결과를 비교하였지만, 스킴 리딩의 행태를 관찰했을 뿐하이라이트와 같은 보조 도구에 대한 제안은 부족했다.

이러한 기존의 연구들을 참조하고 부족했던 점을 보완하고자 본 연구에서는 스킴 리딩 상황과 전체를 정독하는 상황을 비교하여 각각의 경우에 하이라이트의 효과를 확인한다. 이와 더불어 내용에 대한 하이라이트 유무에 따라 글을 읽은 이해도를 측정하여 하이라이트 조건에 따라 내용 이해도에 차이가 있는지를 관찰한다. 이와 함께 아이트래커를 사용하여 실제로 하이라이트가 없는 부분에 비해 하이라이트에 시선이 오래 머무는 지를 측정해본다.

연구문제1. 스킴 리딩 상황에서 글의 중요 부분을 하이라이트를 통해 강조하여 보여 주는 것이 효과적인가?

가설1-1. 중요한 부분에 하이라이트가 적절히 표시되어 있으면, 글의 중요한 내용에 대해 더 잘 기억하며 글에 대한 이해도를 높일 수 있을 것이다.

가설1-2. 하이라이트의 효과는 전체를 정독할 때 보다 스킴 리딩을 할 때 더 뚜렷하게 나타날 것이다.

가설1-3. 하이라이트로 표시된 부분에 그렇지 않은 부분에 비해 시선이 오래 머물게 될 것이다.

#### 2.3. 하이라이트 생성 방법

온라인 문서의 중요한 내용의 추출은 당연히 글의 원작자 혹은 해당 분야에 대한 이해가 깊은 전문가가 직접 수행하는 것이 정확할 것이다. 그러나 온라인에서 읽게 되는 방대한 양의 정보를 일일히 전문가가 개입하여 중요 부분을 표시하는 것은 비현실적이며, 비용과 시간의 한계성의 문제가 수반된다. 따라서 이에 대한 대안으로 컴퓨터의 알고리즘을 이용한 자동 요약(automatic summarization) 방법과 다수의 비전문가의 집단 지성을 활용한 크라우드소싱(crowdsourcing)을 이용한 방법이 제안되었다.

#### 2.3.1. 컴퓨터 알고리즘을 이용한 방법

컴퓨터 알고리즘에 의한 자동 요약의 분야에 대한 연구는 주로 컴퓨터 과학이나 공학 분야에서 활발하게 진행되어 왔다. 하이라이트를 생성하는 것은 다시 말해 컴퓨터의 연산 능력을 이용하여 문서의 중요한 부분을 추출하는 것을 의미한다. 이는 컴퓨터 과학에서 "문서 요약"이라고 불리는 분야로 오랜 시간 활발히 연구되어오고 있다.

문서 요약에 관한 연구는 크게 추출 요약(extract)과 생성 요약(abstract)로 나뉜다(Mani, 2001). 추출 요약은 문서에 존재하는 단어나 구, 문장을 그대로 골라내는 방법이고, 생성 요약(abstract)은 문서의 내용을 새로운 문장으로 재구성하는 기술로 자연어 이해 및 생성기술이 필요한 방법이다. 하이라이트의 생성은 문장에서 중요한 부분을 추출하여 표시를 하는 것이기 때문에 두 가지 중 추출 요약의 방법과 직접적으로 연관되어 있다.

컴퓨터 알고리즘 원리는 인간의 사고과정과는 달리 글의 구조나 언어학적인 특징을 활용한다. Moens와 Dumortier(2000)는 글의 문법적 구조와 언어적 단서들을 이용하여 온라인 잡지의 글들에서 자동으로 중요한 부분을 추출하는 방식을 제시하였다. 이러한 방식은 글의 주제와 상관없이 다양한 내용의 글들에 보편적으로 적용할 수 있다는 장점을 가진다. Chi 등(2005)은 단어나 문장의 반복 빈도와 위치를 분석하는 알고리즘을 통해 하이라이트를 자동으로 생성하는 방법을 제시하였다. 또한 알고리즘에 의해 생성된 하이라이트를 표시한 글을 읽었을 때,

하이라이트가 없는 글을 읽었을 때보다 사람들이 내용을 더 정확히 파악하였고, 이를 통해 알고리즘으로 생성된 하이라이트가 어느 정도 유용함 보여주었다.

웹 페이지 등 온라인 문서가 지닌 독특한 특성을 활용한 연구도 등장하였다. Buyukkokten 등(2001)은 웹 페이지의 글을 문단, 목록, 이미지를 설명글 등 의미를 가지는 단위로 구분한 후, 이 의미 단위들 안의 단어들의 빈도 분석을 통해 문장들의 순위를 매겼다. 문장 내 단어들은 글 전체 내 등장 빈도가 높을 수록 더 중요한 문장으로 여겨졌고, 높은 순위를 기록한 문장들을 추출할 수 있었다. Alam 등(2003)은 HTML(HyperText Markup Language)을 통해 생성된 페이지가 구조화 하기 쉽다는 특성을 가지고 있으며, 이를 이용해 만들어 낸 페이지 요약(summarization)은 웹 브라우징과 페이지 간 네비게이션을 쉽고 빠르게 할 수 있다는 장점이 있음을 주목하였다. 이러한 중요한 문장을 선별하는데 빈도, 문장의 위치 등 계량화 하기 쉬우 요소들을 주로 사용하는 것을 알 수 있다.

컴퓨터 알고리즘에 의한 생성은 일반적으로 연산 시간이 오래 걸리지 않기 때문에 문서가 작성됨과 동시에 거의 실시간으로 하이라이트의 생성이 가능하다는 장점이 있다. 그러나 컴퓨터가 인간과 같은 사고과정을 거치지 않기 때문에 내용을 이해하는 것처럼 보이나 사실은 인간과 완전히 같은 방식으로는 내용을 분석할 수는 없다는 치명적인

단점이 공존한다. 이에 따라 문서 요약 분야에서는 지속적으로 정확도나신뢰도에 관한 의문이 제기되어왔다. Morris 등(1992)은 컴퓨터알고리즘을 통해 추출하는 방법의 한계를 지적하였는데, 같은 시간에요약된 내용을 읽은 후의 이해도가 전체 문장을 읽었을 때와 크게 다르지않다고 설명했다. Hahn과 Mani(2000)는 알고리즘 추출법이 가지고있는 문제 중의 하나는 이러한 요약 알고리즘을 평가하는 방법이부족하다는 점을 지적하였다. 이 연구에서는 이용자들이 컴퓨터알고리즘을 신뢰할 수 있어야 함을 전제하였고, 그러기 위해서는 추출된결과물의 수준을 평가하는 기준과 방법이 더욱 고도화되어야 한다고주장하였다.

#### 2.3.2. 크라우드소싱을 이용한 방법

크라우드소싱은 전문가가 아닌 다수의 비전문가가 참여하여 사회적, 기술적 과제를 포함한 다양한 문제를 해결하는 방법이다. 온라인에서 다양한 사용자가 참여할 수 있다는 장점을 가지고 있지만, 광범위한 네트워크가 기반이 되어야 하기 때문에, 기술적 발달에 따라 비교적 최근에 개발되기 시작한 분야이다. 크라우드소싱은 적은 비용으로 다양한 분야의 데이터를 수집할 수 있다는 장점을 가지고 있으며, 이를 이용한 다양한 연구들도 발표되고 있다. 과거의 대부분의 연구에서 크라우드소싱을 통해 수집된 데이터들은 컴퓨터 알고리즘 방법의 연습

데이터로 쓰이거나, 사람들의 인지능력을 통해 컴퓨터 알고리즘을 보완할 목적으로 사용되어 왔다(Sabour, Bontcheva, & Scharl, 2012).

하이라이트 생성 연구에 크라우드소싱을 이용한 연구는 아직 많지 않다. 그러나 이와 비슷한 방법을 사용한 연구들이 꾸준히 있어 왔다. 하이라이트 이전에 읽기의 보조 도구로 쓰이는 주석(annotation)을 공동으로 작성하는 방법들이 교육학 분야에서 연구되었다. 오프라인 환경에서도 Davis와 Huttenlocher(1995)는 협동 학습을 위해 주석을 공유하는 시스템을 제안했다. 이 연구에서는 글의 특정 부분에 적절한 주석이 삽입됨으로 독자들이 해당 내용에 더 집중하고 높은 참여도를 보인다는 것을 밝혀냈다. 네트워크 기술의 발달로 다수가 온라인에서 동시에 통신하는 작업이 가능해짐에 따라 Cadiz 등(2000)은 450명의 참여자가 1,250개의 문서에서 9천여 개의 주석을 공유하는 실험을 수행하였고 이로써 온라인 문서에서 크라우드소싱의 가능성을 암시하였다.

중요 부분을 추출하는 목적과는 다르지만 온라인 상의 공동작업에서 하이라이트를 사용하는 연구도 발표되었다. Shen과 Sun (2004)는 서로 다른 색으로 하이라이트를 표시함으로써 다수의 이용자가 하이라이트를 이용한 공동작업이 가능함을 보였으며, 이를 확장함으로써 하이라이트를 이용한 크라우드소싱이 가능할 것임을 예측할 수 있다. 전자책 분야에서는 이러한 방법이 상용화되어 활용되고 있다.

아마존(Amazon)의 전자책 리더인 킨들(Kindle)에서는 사람들이 책을 읽으면서 중요하다고 생각하는 부분이나, 나중에 다시 보고 싶은 중요한 정보에 하이라이트를 표시한다는 점에 착안하여, 그들이 판매한 전자책에 표시된 하이라이트의 위치에 대한 데이터를 수집하였다. 그리고 사용자가 원할 때, 이러한 데이터를 처리하여 많은 사람들이 하이라이트로 표시한 부분을 추려서 독자에게 보여주는 방법을 사용하고 있다.

최근의 연구에서 특히 문서 요약 분야에 크라우드소싱을 도입하려는 노력이 계속되고 있다. Lloret 등(2013)은 크라우드소싱을 이용한 요약문 생성 연구에서 사람들이 어떤 대상을 설명할 때 특정한 패턴을 보인다는 것을 관찰하였다. 특정 장소를 찍은 사진을 글로 설명하라는 과제에서 참여자들은 장면을 설명하기 위해 공통적 요소를 택한다는 것을 보여주었다. 이 실험에서는 단순한 크라우드소싱만을 이용한 제대로 된 요약문 생성은 쉽지 않았지만, 중재자와 같이 데이터를 걸러주기 위한 요소가 추가된다면 수집된 데이터의 일관성을 좀 더확보할 수 있을 것이라고 주장했다. Walter와 Back(2013)은 크라우드소싱을 통해 수집된 데이터를 텍스트 마이닝(text mining) 기법을 통해 분류한 것이 전문가들의 의사결정에 도움이 될 수 있음을 보여주었다.

Gillick과 Liu(2010)의 연구와 같이 비전문가들에 의해 구성된 크라우드소싱을 통한 데이터의 한계를 지적하는 연구들도 있다. 하지만 정교한 절차를 거쳐 수집된 크라우드소싱 데이터와 잘 가공된 결과물이 사람들의 의사결정에 미칠 긍정적인 가능성을 제시한 연구들도 존재한다. 크라우드소싱에 대해 전문가 자문단, 상호 평가 등(Gilick & Liu, 2010, Sabour, Bontcheva, & Scharl, 2012, Walter & Back, 2013)과 같이 수집된 데이터의 신뢰도 문제를 해결할 수 있는 장치들이 제안되었다. 크라우드소싱은 많은 숫자의 이용자만 확보가 된다면, 복잡한 알고리즘에 의한 연산 없이도 의미 있는 결과물을 낼 수 있는 가능성이 있다. 크라우드소싱은 근본적으로 사람이 직접 참여하기 때문에 글의 의미에 대한 인지과정을 바탕으로 한 방법이다. 따라서 구조 등의 요소만을 사용하는 컴퓨터 알고리즘과는 다른 결과를 기대할 수 있을 것이다.

#### 2.4. 하이라이트를 생성하는 소스에 따른 차이

앞선 가설 1-1과 같이 더 정확히 표시된 하이라이트가 읽기 수행에 도움이 된다면 보다 양질의 하이라이트를 생성하는 방법을 개발하는 것은 문서 요약 및 하이라이트 연구 분야에서 중요한 요소임을 증명하는 일일 것이다. 현재에도 컴퓨터 알고리즘에는 다양한 종류가 존재하며 적용하는 글의 종류에 따라서도 그 결과가 달라진다. 크라우드소싱의

방법 역시 데이터의 수집 방법이 얼마든지 다양할 수 있으며 시행에 따라다른 수준의 결과물을 도출할 것이다. 그러나 실제로 컴퓨터 알고리즘에의한 추출 방법이나 크라우드소싱의 방법 중에서 어떤 방법을 이용하는것이 더 좋은 결과를 나타내는지 비교하는 것은 본 연구의 범위를 벗어난다. 더 고도화된 방법으로 정확한 하이라이트가 생성이 된다고해도 사람들이 이를 실제 읽기 생활에 적용하게 하기 위해서는 실제하이라이트의 정확성 외에도 다른 여러 요인들이 작용한다. 이연구에서는 위에서 언급한 두 가지 하이라이트 생성 방법에 대한하이라이트 수용자의 심리적 반응에 의한 차이를 알아보고자 한다.

소스(source)의 개념은 커뮤니케이션 분야에서 다양한 각도로 연구되어왔다. 소스는 말 그대로 어떤 정보의 출처를 의미하는 것이다. 그러나 반드시 최초에 정보를 생성한 사람을 의미하는 것은 아니며 겉으로 드러나는 소스(visible source)외에도 정보의 전달자(channel), 혹은 다른 수용자(other receiver)도 소스가 될 수 있다. Chaffee(1982)는 수용자(receiver)의 입장에서 어떤 뉴스의 소스를 인식할 때에 실제 정보가 만들어진 정보의 소스(visible source)와 그것을 전달하는 대상(channel)을 명확하게 구분하지 않으며, 동일하게 소스의 역할을하고 있음을 지적했다. 온라인 뉴스의 예를 들면, 겉으로 드러나는 소스인 뉴스 편집자, 기술적인 소스인 컴퓨터, 온라인 뉴스를 선별하여 전달해 준 다른 수용자 모두가 소스의 기능을 할 수 있다는 것이다.

본 연구에서는 커뮤니케이션 연구에서 다루는 소스의 개념을 두 가지하이라이트 생성 방법에 적용하였다. 컴퓨터 알고리즘에 의한하이라이트의 생성은 문서를 전달하는 컴퓨터라는 기술적인 요소가소스의 역할을 하는 것이다. 크라우드소싱에 의한하이라이트 생성은수용자와 같은 입장에서 주어진 정보를 받는 다른 수용자가 소스의역할을하게 된다. 두 가지 소스는하이라이트의 생성에 있어 수용자가다른 입장을 갖게한다. 컴퓨터 알고리즘에 의한 방법은 사용자가기본적으로 컴퓨터가 글을 내용을 인지하는 원리가 인간의 그것과다름을 알고 있기 때문에 사람들이 결과물을 신뢰하고 사용하지않는다는 문제가지적된 바 있다(Hahn & Mani, 2000).

한편 크라우드소싱 방법이 컴퓨터 알고리즘이 가진 이런 문제를 해결할 수 있는 가능성이 있음이 다른 분야의 연구를 통해 간접적으로 확인할 수 있다. Sundar와 Nass(2001)의 온라인 뉴스 연구에 대한 실험은 본 실험의 기본 틀이 되었다. 이 실험에서 참여자는 선별된 온라인 뉴스를 제공받으며, 제공된 뉴스의 소스의 종류에 따른 수용자의 '인지(perception)'의 차이를 평가하였다. 여기서 설명하는 인지에는 하이라이트 결과물에 대한 신뢰도(credibility), 선호도(liking), 질(quality), 대표성(representativeness)이 포함된다. 참여자를 4개의 그룹으로 나누어 소스의 종류를 뉴스 편집자, 컴퓨터, 다른 이용자, 자기 자신으로 구분하였으며, 참여자에게는 각각의 소스에 의해 기사가

주어졌다고 알려주었으나, 실제로는 모든 참여자에게 같은 기사가 주어졌다. 그리고 참여자에게 전달된 기사의 출처가 어디라고 믿는지에 따라 나타나는 인지의 차이를 살폈다. 그 결과, 전체적으로 다른 사용자들에 의해 전달된 기사가 가장 많은 인지 항목에서 높은 점수를 받았으며, 특히 뉴스 편집자나 자기 자신에 의해 선별된 것보다 기사의 질이 높다고 평가했다.

소스의 속성에 대한 인지로 인해 나타나는 결과는 커뮤니케이션의 다양한 분야에서 연구되었다. Flanagin과 Metzger(2007)은 웹 사이트 제공자에 대한 친밀도와 웹 사이트의 속성에 따른 상대적인 인지된 신뢰도(perceived credibility)에 영향을 미치는 정도를 측정하였다. Hu와 Sundar(2010)는 건강 정보에 관한 연구에서 정보의 생성자뿐만 아니라, 웹 사이트, 게시판, 블로그(blog), 개인 홈페이지 등 정보를 선택하는 채널에 따라 수용자에게 인지되는 신뢰도의 차이가 있음을 보였다.

본 연구의 대상인 온라인 문서에서 하이라이트의 추출은 다량의 정보로부터 소량의 중요 정보를 추출하는 행위이다. 규모와 형식에 차이가 있지만, 이러한 구조에서 온라인 뉴스에서 중요한 뉴스만을 추출하여 이용자에게 제공하는 것과 하이라이트 추출과의 유사성을 관찰할 수 있다. 크라우드소싱은 '다른 사용자들'이라는 소스에 의해하이라이트가 결정되는 것이며, 컴퓨터 알고리즘은 '기술적 요소에 의한

소스(technological source)'역할을 한다. 하이라이트를 추출해서 제공하는 소스가 컴퓨터 알고리즘(기술적 소스)인지, 아니면 크라우드소싱(소스로서 작용하는 다른 사용자들)인지의 여부에 따라 인지의 차이가 나타나는지 관찰할 수 있을 것이다.

그 동안의 소스의 구분을 다루는 연구들은 뉴스 등의 미디어를 대상으로 이루어진 것이 대부분이고, 하이라이트의 생성과 같은 새로운 분야에 적용한 사례를 찾기 어렵다. 그래서 아직 하이라이트 생성물의 결과물을 평가하는 기준도 마련되어 있지 않았다. 이 연구에서는 기존의 비슷한 연구에서 사용한 척도를 참조하여 하이라이트 추출이라는 속성에 맞는 인지 요소를 선별하였다. Sundar와 Nass(2001)가 수행한 연구에 사용한 네 가지 인지 요소 중 본 연구에서 다루는 하이라이트의 추출과 다소 거리가 있는 질(quality)과 선호도(liking)를 제외하였다. 질이나 선호도를 측정하는 설문에는 소스에 의해 선택된 온라인 뉴스의 내용을 포함하는데, 하이라이트의 추출이라는 특성 상 전 체 글에서 중요한 부분만을 선별하는 것이기 때문에 전체 글의 내용에는 변함이 없고 새로운 내용을 포함하지는 않는다. 따라서 이 가운데 신뢰도(credibility)와 대표성(representativeness)을 측정하기로 했다. 또한 하이라이트 추출 연구는 이를 이용한 어플리케이션을 제작하여 사람들이 일상 생활에서 사용하게 하는 데 궁극적인 목적이 있다. 따라서 이를 실제로 이용할 때 도움이 되는지를 측정하는 척도가 필요할 것이기 때문에, 하이라이트 결과물에 대한 참여자의 인지를 측정할 때 효용성(usefulness)의 척도를 포함하였다. 신뢰도는 여러 연구에서 가장 일반적으로 측정되는 요소이며, 대표성은 하이라이트가 가진 특성상 전체 문서를 대표해야 하는 속성과 일치한다. 효용성은 이 하이라이트의 존재로 인해, 하이라이트가 없을 때에 비해 글을 읽을 때 유용하게 사용될 수 있는지에 대한 척도를 의미한다

연구문제2. 하이라이트를 추출하는 두 가지 소스(source)에 따른 심리적 차이가 하이라이트에 대한 평가에 영향을 미치겠는가?

가설2-1. 독자는 컴퓨터 알고리즘에 의해 생성된 하이라이트와 크라우드소싱 방법을 통해 추출된 하이라이트에 대한 신뢰도를 다르게 인지할 것이다.

가설2-2. 독자는 컴퓨터 알고리즘에 의해 생성된 하이라이트와 크라우드소싱 방법을 통해 추출된 하이라이트에 대한 대표성을 다르게 인지할 것이다.

가설2-3. 독자는 컴퓨터 알고리즘에 의해 생성된 하이라이트와 크라우드소싱 방법을 통해 추출된 하이라이트에 대한 효용성을 다르게 인지할 것이다.

앞서 지적했듯이 컴퓨터 알고리즘의 하이라이트 생성원리는 인간의 인지 과정과는 다르다는 한계를 가지고 있다. 반면에 크라우드소싱을 이용한 방법은 다른 수용자가 인간의 인지 과정을 통해 선별하는 것이기 때문에 컴퓨터 알고리즘의 단점을 극복할 수 있는 대안 혹은 보조적용도로도 사용될 수 있을 것이다. 아마존의 킨들에서 크라우드소싱을실용화한 서비스를 제공하는 사례를 볼 때, 다수의 사용자가 확보가된다면 상업적용도로 활용이 가능할 것이다.

그러나 크라우드소싱을 이용한 방법에도 단점이 있다. 하나의 문서에 민을만한 수준의 하이라이트가 생성되기 위해서는 다수의 참여자가 확보되어야 한다. 따라서 이미 아주 많은 사용자를 확보한 소수의 사업자가 아니라면 실용성을 확신하기 어렵다. 왜냐하면 하이라이트의 생성과 공유 과정이 추가적인 작업을 요하는 과정을 거치기 때문에 광범위한 자발적 참여를 기대하기 힘들다.

본 연구에서는 이러한 크라우드소싱의 한계를 극복하기 위한 아이디어를 제안한다. 온라인 읽기 환경에 최근 스마트폰과 같은 모바일기가 널리 활용되고 있다. 모바일 화면은 기존의 데스크톱 컴퓨터스크린에 비해 화면 크기에 제한이 있기 때문에, 전체 문서의 일부만이화면에 노출이 되며, 독자는 이를 스크롤(scroll)하면서 읽게 된다. 그러면 화면에 많이 노출되는 부분과 그렇지 않은 부분이 구분될 것이고, 노출 시간이 긴 부분은 사용자가 더 주의 깊게 읽었다고 볼 수 있다. 따라서 이 읽기 패턴 데이터를 수집하여 여러 사람들이 공통적으로 오래주의를 기울인 부분에 하이라이트를 표시하는 방법이 가능하다. 모바일

독자들의 화면 스크롤을 통한 자연스러운 읽기 행위를 통해 수집된 패턴 데이터를 크라우드소싱에 응용하여 하이라이트를 생성하는 방법을 제안하며, 실험을 통해 가능성을 검증해 보고자 한다.

연구문제3. 사용자의 읽기 패턴 데이터를 이용하여 온라인 문서에서 하이라이트를 추출할 수 있는가?

# 3. 실험 1: 하이라이트 조건에 따른 내용이해도와 시선의 차이

#### 3.1. 개요

온라인 문서에서 스킴 리딩을 할 때, 중요한 부분에 하이라이트를 통해 강조하여 보여 주는 것이 효과적인지를 확인하기 위한 실험을 진행하였다. 실험 조건은 2(읽기 방법) × 3(하이라이트 조건)으로 구성하였다. 독립변인으로 제시된 하이라이트 조건은 한 명의 참여자가 세 가지 조건 모두에서 이루어지는 피험자 내(within-subject) 실험으로 설계하였으며, 측정이 각각 중요한 부분에 하이라이트가 되어 있는 조건(GH조건, good-highlighted)과 중요하지 않은 부분에 하이라이트가 되어 있는 조건(BH조건, bad-highlighted), 그리고 하이라이트가 없는 조건(NH조건, non-highlighted) 등 세 가지 조건이 주어졌다. 그리고 이러한 하이라이트의 효과가 스킴 리딩 조건에서 더욱 뚜렷하게 나타날 것이라는 가정에 따라 실험 그룹을 둘로 나누어 피험자 간(between-subject) 실험으로 설계하였으며, 한 그룹에서는 스킴 리딩을 하게 하고(스킴 리딩 그룹), 다른 그룹은 전체 문장을 정독하게 하였다(정독 그룹). 이와 동시에 하이라이트로 표시된 부분에 시선이 더 오래 머물 것이라는 가설을 검증하기 위하여 아이트래커를 사용하여

시선이 머무는 시간을 측정하였다. 그리고 한 참여자가 여러 조건에서 측정 하였으므로, 반복측정 이원분산분석(two-way ANOVA with repeated measures)으로 결과를 분석하였다.

#### 3.2. 측정

실험은 39명의 참여자를 모집해서 실험실에서 진행되었다. 참여자들은 남자 17명, 여자 22명으로 구성되었으며 평균 나이는 24.1세이다. 모두 대학생 또는 대학원생이며 다양한 전공을 배경으로 가지고 있었다. 참여자 가운데 17명은 스킴 리딩 그룹에, 22명은 정독그룹에 속했다. 두 그룹 간 인원의 차이는 아이트래커 실험 때문에 발생하였는데, 장비가 시선을 추적하는 데 실패하는 경우 양쪽의 그룹 인원을 조정하였다. 스킴 리딩 그룹에 속한 참여자 가운데 시선 추적에 실패한 경우가 많아서, 스킴 리딩 그룹에 참여자을 더 배정하였다. 시선 추적에 실패한 참여자는 10명으로, 시선 데이터 분석에 사용된 데이터는 29명(스킴 리딩 그룹 15명, 정독 그룹의 14명)의 데이터가 사용되었다. 아이트래커는 높은 도수의 안경을 쓰거나, 실험 도중 참여자의 눈의 위치가 아이트래커를 벗어나는 경우 추적에 실패하는 것으로 보인다.

자극으로 사용된 문서는 1,000자 이내의 한글 문서로, TOEFL 읽기 문제집으로부터 내용을 발췌하여 국문 번역을 통해 재구성하였다. 스킴 리딩 그룹에게는 문서 당 45초의 제한 시간을 두어 스킴 리딩을 해야 하는 환경을 강제하였으며, 정독 그룹에는 제한 시간을 두지 않고 글을 충분히 읽게 하였다. 제한 시간을 통해 읽기 조건을 구성하였기 때문에 스킴 리딩 그룹의 참여자들은 연습 시행을 통해 주어진 시간을 어떻게 분배하여 읽어야 할지 가늠하는 기회를 갖도록 하였다. 각 참여자들은 연습 시행을 1회 실시한 후 3회의 본 실험의 시행을 진행하였다. 실험 전체에 걸리는 시간은 30-40분 정도였다.

글의 하이라이트 조건은 세 가지로 나뉘는데, 첫 번째 조건은 GH조건(good-highlighted)으로, 글 전체를 구성하는 문장 중에서 핵심 내용에 해당하는 문장 5개에 하이라이트가 표시된 문장이다. 두 번째 조건은 BH조건(bad-highlighted)으로, 앞의 조건과 마찬가지로 5개 문장에 하이라이트가 되어 있지만 여기서는 글의 내용을 이해하는 데 비교적 중요하지 않은 문장에 하이라이트가 표시되어 있다. 세 번째 조건은 NH조건(non-highlighted)으로, 하이라이트가 표시되지 않은 조건이다. 글은 세로로 세워진 23인치 모니터 한 화면에 모두 표시되며 실험 도중 스크롤은 하지 않는다. 이렇게 구성한 이유는 불필요한 조작으로 주의를 분산시키지 않게 하기 위함이며, 시선 추적 데이터의 오차가 발생하는 확률을 낮추기 위해서이다.

참여자마다 받게 되는 문서의 내용과 순서는 같지만, 하이라이트 조건은 순서의 영향을 받지 않기 위해 무작위로 섞어서 실시하였다. 각 실험 참여자가 읽게 되는 3개의 문서는 각각 다른 세 가지 하이라이트 조건(GH조건, BH조건, NH조건)으로 주어졌으며, 하이라이트는 문장 위에 노란색이 덧칠된 형태로 제시되었다.

내용 이해도는 총 9개의 문제를 풀면서 측정된다. 5개의 단어 기억 문제, 3개의 문장 기억 문제, 그리고 1개의 고난도 문제로 구성되었다. 기억 문제는 "다음 [단어/문장] 중 글을 이해하는 데 가장 중요한 [단어/문장]을 고르시오."라고 주어진다. 단어 기억 문제의 선택지는 4지선다형으로 구성되어있는데, 선택지의 각 항목은 중요도에 차이가 있는 서로 다른 단어 4개로 구성되었다. 각 선택지에서 단어는 GH단어(good-highlighted, GH조건에서 하이라이트로 표시된 부분에서 선택), BH단어(bad-highlighted, BH 조건에서 하이라이트로 표시된 부분에서 선택), NH단어(non-highlighted, 어느 조건에서도 하이라이트 표시되지 않은 부분에서 하나씩 선택), 마지막으로 EX단어(extraneous, 글에 존재하지 않는 단어)로 구성되었다. 문장 기억 문제는 중요도가 다른 세 개의 문장을 고르는 형태로 3지 선다형으로 구성이 되었으며, 각 선택지는 단어 문제 선택지에서 EX단어를 제외한 나머지 세 종류의 단어와 비슷하게 주어졌다. 세 가지 문장은 각각 단어는 GH문장(good-highlighted, GH조건에서 하이라이트로 표시된 문장 가운데 선택), BH문장(bad-highlighted, BH 조건에서 하이라이트로 표시된 문장 가운데 선택), NH문장(non-highlighted, 어느 조건에서도 하이라이트 표시되지 않은 문장 가운데 선택)으로 이루어졌다. 그리고 고난도 문제는 좀더 깊은 이해를 요하는 문제로 만들어졌으며, 이 문제에서 참여자들은 6개의 문장들 중 3개를 골라 읽은 글의 요약을 완성하는 과제를 수행하였다. 고난도 문제는 3개를 맞추면 2점, 2개를 맞추면 1점을 부여하였다. 즉, 단어 선택 문제의 경우각 조건에 따라 4개의 항목(GH단어, BH단어, NH단어, EX단어)을 고른 횟수를 서로 비교해서 측정하며, 문장 선택 문제의 경우 각 조건에따라 3개 항목(GH문장, BH문장, NH문장)을 고른 횟수를 비교하여이루어진다. 고난도 문제의 경우는 각 조건에서 획득한 점수를 서로비교한다.

시선의 추적을 위해 글이 표시된 부분을 영역으로 나누어 분석하였다. 이 영역을 판독 영역으로 잡고 본 논문에서는 AOI(Area of Interest)라고 칭한다. GH조건에서 하이라이트로 표시된 부분(gAOI, good-AOI), BH 조건에서 하이라이트로 표시된 부분(bAOI, bad-AOI), 어떤 조건에서도 하이라이트로 표시되지 않은 부분이나 글이 아니라 화면 테두리 등 그 외 모든 부분(nAOI, none-AOI)으로 나누어 분석하였다. 각 AOI에 시선이 머문 시간은 밀리 초(ms) 단위로 총 누적시간을 측정하였다. 시선 추적에 실패하지 않았을 때, 시선은 추적에 실패하지 않았을 때 4개의 AOI(gAOI, bAOI, nAOI) 가운데 한 곳에 머물러 있다고 보았다. 여기서 독립변인은 읽기 방법(스킴 리딩, 정독)과

하이라이트 조건(GH조건, BH조건, NH조건)이며, 시선이 각 AOI에 머문 누적 시간을 종속변인으로 측정하였다.

자극으로 주어진 3개의 글은 각각 생물학, 역사, 사회학 분야의 정보 전달 형식의 글이다. 하이라이트 조건의 순서에 의한 효과를 배제하기 위해 참여자들에게 제시된 글 내용의 순서는 고정하고 하이라이트 조건의 순서를 고르게 분배하였다. 참여자들에게는 라틴스퀘어 방식으로 짜인 6개의 실험 세트 중 한가지를 받게 된다. 실험 세트의 구성은 표 1과 같다.

표 1. 참여자에게 부여된 6개의 실험 세트

	글 1: 생물학	글 2: 역사	글 3: 사회학
세트 1	비하이라이트(NH)	중요 하이라이트(GH)	비중요 하이라이트(BH)
세트 2	비하이라이트(NH)	비중요 하이라이트(BH)	중요 하이라이트(GH)
세트 3	중요 하이라이트(GH)	비하이라이트(NH)	비중요 하이라이트(BH)
세트 4	중요 하이라이트(GH)	비중요 하이라이트(BH)	비하이라이트(NH)
세트 5	비중요 하이라이트(BH)	비하이라이트(NH)	중요 하이라이트(GH)
세트 6	비중요 하이라이트(BH)	중요 하이라이트(GH)	비하이라이트(NH)

### 3.3. 과정

실험은 아이트래커가 설치된 컴퓨터가 있는 실험실 환경에서 진행되었다. 참여자는 아이트래커가 설치된 컴퓨터 앞에 앉게 되며 시선 추적을 위한 캘리브레이션(calibration) 과정을 거친다. 참여자의 인구통계정보를 먼저 입력 받으며, 실험은 연습 시행 1회을 시행한 후 글 1, 글 2, 글 3 순서로 본 시행 3회를 수행하게 된다. 연습 시행과 본 시행은 똑 같은 과정을 거친다. 시행 1회는 글 읽기와 설문 응답으로 구성되며, 각 참여자가 속한 그룹에 따라 글 읽는 시간이 다르게 주어진다. 실험이 진행될 때, 하이라이트가 표시된 조건(GH조건 및 BH조건)에서 참여자들은 자신이 보고 있는 하이라이트가 두 조건 중에 어떤 것인지 알 수가 없게 하였다.

스킴 리딩 그룹 참여자는 45초가 지나면 자동으로 설문 응답으로 넘어가게 되며, 정독 그룹은 원하는 시간에 버튼을 눌러 다음 화면으로 넘어갈 수 있다. 글 읽기 단계에서는 주어진 조건에 맞춰 하이라이트가 표시되거나 표시되지 않은 글을 읽는다. 설문 응답은 위에서 설명된 총 9개의 문제를 풀게 된다. 문제를 다 풀면 다음 시행으로 넘어갈 수 있다. 그리고 글을 읽는 동안 참여자의 시선은 설치된 아이트래커에 의해 녹화된다.

### 3.4. 결과 및 분석

글의 내용을 얼마나 이해했는지의 차이를 분석하기 위해, 평가 항목에서 각 선택지를 고른 횟수를 측정하였다. 각 그룹을 나누는 것은 피험자 간(between-subject) 방식이지만, 그룹 내에서 참여자는 피험자 내(within-subject) 방식으로 세 가지 하이라이트 조건을 모두 한 번씩 되며 이에 따라 세 차례 글을 한 차례씩 읽도록 했다. 이 때 각 글이 포함하는 내용의 난이도나 분야 등 내재적 속성까지 완전히 통제할 수는 없었다. 이 같은 효과를 상쇄시키기 위해 각 글이 받은 항목당 점수를 사용하여 모든 데이터를 평균 중심화(mean centering) 했다. 이러한 방법을 사용한 것은 각 문서가 가진 속성이 반영되는 것을 막고, 그룹 혹은 하이라이트 조건의 차이로 인한 각 항목 점수를 종속변수화할 수 있게 하기 위해서이다(표 2). 분석에 포함된 참여자 수는 N=39(스킴리딩 그룹 17명, 정독 그룹 22명)이다.

표 2. 평균중심화 후 각 선택지를 고른 횟수의 평균(괄호 안은 표준편차)

	스킴 리딩 그룹			정독 그룹		
	GH조건	BH조건	NH조건	GH조건	BH조건	NH조건
GH단어	.00 (.96)	83 (1.12)	36 (1.01)	.96 (.70)	.07 (.89)	.48 (.68)
BH단어	06 (.63)	.62 (1.06)	19 (.57)	33 (.33)	.23 (.77)	34 (.62)
NH단어	.07 (.60)	.25 (.82)	.01 (.59)	38 (.5)	17 (.49)	.13 (.50)
EX단어	01 (.48)	04 (.50)	.53 (.61)	26 (.28)	14 (.53)	26 (.24)
GH문장	.02 (.85)	36 (.85)	24 (.8)	.25 (.73)	.17 (.83)	.33 (.64)
BH문장	24 (.70)	.34 (.76)	.20 (.73)	13 (.63)	.01 (.81)	25 (.29)
NH문장	.23 (.68)	.03 (.72)	.04 (.43)	12 (.56)	18 (.36)	08 (.53)
고난도	09 (.54)	22 (.63)	14 (.44)	.18 (.44)	.19 (.53)	.20 (.35)

각 읽기 그룹 및 하이라이트 조건에 따라 참여자가 각 항목의 이해도 점수에 차이가 있는지 확인하기 위하여, 2(읽기 방법) × 3(하이라이트 조건)의 이원분산분석(two-way ANOVA with repeated measures)을 시행한 결과를 표 3에 나타내었다. 참여자 집단을 스킴 리딩 그룹과 정독 그룹으로 나누었기 때문에 읽기 방법은 피험자 간 분석이며, 참여자가 세가지 하이라이트 조건에 모두 참여하였기 때문에 하이라이트 조건은 피험자 내 분석으로 진행되었다. 추가로 하이라이트 조건일 때와하이라이트 조건과 읽기 방법의 상호작용이 가져오는 유의한 차이를 가져오는 구체적 요인을 알아보고자 Tukey's HSD 사후검증을 실시하였다.

먼저 "내용을 이해하는 데 가장 중요한" 단어 고르는 문제에서 각 단어의 선택 횟수를 비교해보았다. GH단어를 고르는 횟수에 있어서 읽기 방법과 하이라이트 조건이 모두 유의미한 영향을 주었고, BH단어를 고르는 횟수에는 하이라이트 조건이 NH단어를 고르는 횟수에는 읽기 방법이 영향을 주었다. 글에 없는 EX단어를 고르는 횟수에는 읽기 방법과 하이라이트 조건 모두 영향을 주었고 두 독립변인 간의 상호작용도 관찰되었다.

실제 내용 상 중요한 단어인 GH단어를 고른 횟수는 곧 정답을 맞춘 횟수라고 볼 수 있다. 두 가지 변인에 따른 결과의 차이를 비교해본 결과 읽기 방법에 따른 차이가 유의미했다[F(1, 36)=34.995, p < .001]. 똑같이 중요한 부분에 하이라이트가 표시되어 있지만 정독 그룹은 더 많은 시간이 주어졌기 때문에 중요 부분을 잘 파악한 것을 암시한다. 이는스킴 리딩의 일반적인 단점이 드러난 결과로, Duggan & Payne(2009)이실험을 통해 확인한 단점과 상통하는 결과를 보여주었다. 한편하이라이트 조건 간의 차이도 드러났는데 사후검증 결과는 BH조건일 때보다 GH조건일 때 GH단어를 더 많이 선택한 다는 결과가 유의미하게나타났다[F(2, 73)=7.515, p<.01]. 하이라이트가 중요한 부분에 있을경우, 중요한 부분을 파악하는 정도가 높아지고, 반대로 하이라이트가 제대로 표시되어 있지 않을 경우에는 하이라이트가 없는 것보다 혼란을일으켜 실제로 중요한 단어를 고르는 횟수가 줄어든 다는 것을 의미한다.

BH단어를 고른 비율을 보면 하이라이트 조건에 따른 차이가 유의미하게 나타났다[F(2, 73)=12.122, p<.001]. 사후 검증 결과 BH조건-GH조건, BH조건-NH조건 간의 차이가 모두 유의미하게 나타났는데, BH단어는 BH조건에서 하이라이트로 표시된 단어 이기 때문에, 하이라이트의 영향을 받아 다른 두 조건보다 더 많이 BH단어를 골랐음을 의미한다. 즉 실제로 중요하지 않은 단어임에도 불구하고, 사람들은 잘못된 하이라이트가 표시되어 있을지라도 하이라이트가 표시되어 있다는 이유로 이 단어를 중요한 단어라고 착각하게 된다. 읽기 방법에 따른 그룹 간 차이는 통계적 유의미성이 없는 것으로

나타났다[F(1, 36)=3.449, p=n.s.]. 잘못된 하이라이트에 의한 영향은 특정 읽기 방법에서만 나타난다고 확신할 수 없음을 의미한다.

하이라이트와 관계없는 부분에서 무작위로 선택한 단어인 NH단어는 임기 방법에 따른 그룹 간 차이만 유의미하게 드러났다[F(1, 36)=6.333, p < 0.05]. 그림 1의 그래프에서도 확인할 수 있듯이 보면 하이라이트가 존재할 경우(GH조건, BH조건)에 정독 그룹보다 스킴 리딩 그룹에서 이 단어를 선택하는 횟수가 많았음을 알 수 있다. 스킴 리딩을 할 때에는 글의 모든 단어를 다 읽을 시간이 부족하기 때문에 정독 집단에 비해 중요한 단어가 아닌 NH단어를 중요한 단어를 착각하여 고르는 횟수가 많음을 의미한다. 이 역시 스킴 리딩의 단점이 드러난 결과이다.

글에 존재하지도 않고 내용과 상관없는 단어인 EX단어를 선택하는 결과도 변인에 따른 유의미한 차이를 나타났다. EX단어를 선택하는 횟수는 하이라이트 조건에 따른 차이를 보였고, 사후 검증을 실시한 결과, NH조건-GH조건, NH조건-BH조건 양쪽에서 유의미하게 드러났다[F(2, 73)=4.640, p<.05]. 중요한 부분에 정확히 표시된 하이라이트(GH조건)이건, 부정확하게 표시된 하이라이트(BH조건)이건 상관없이 하이라이트만 존재한다면, 글과 상관없는 단어를 고르는 횟수가 줄어든다는 것이다. 흥미롭게도, 이러한 현상은 정독을 하는 조건에서는 나타나지 않으며, 스킴 리딩 시에만 유의미하게 나타나는 것으로 드러났다[F(1, 36)=18.488, p<.001]. 특히 EX단어를 고른

횟수는 읽기 방법과 하이라이트 조건 두 종류의 독립변인 사이에 상호작용이 관찰되었다[F(2, 73)=5.469, p<.01]. 상호작용은 그림 1(좌하)의 그래프에서도 확인할 수 있는데, 정독을 할 때는 하이라이트와 관계없이 글과 상관없는 EX단어를 고르는 횟수는 상당히 낮다. 스킴 리딩 상황에서도 GH조건과 BH조건 간에는 큰 차이가 없지만, 하이라이트가 주어지지 않는다면 EX단어를 고르는 횟수가 상당히 증가하는 것을 알 수 있다.

이어서 표 4와 그림 2에는 좀더 내용을 깊이 이해해야만 풀 수 있는 고난도 문제를 풀게 하고 그 점수를 서로 비교한 결과를 함께 나타내었다. 문장을 고를 때에는 단어를 고르는 것에 비해 대체적으로 유의미한 차이가 많이 나타나지 않았지만, 읽기 방법에 따라 GH문장을 고르는 횟수에 차이가 있었고, 하이라이트 조건에 따라 BH문장을 고르는 횟수에 차이가 있었다. 또한 고난도 문제의 점수는 읽기 방법에 따라 나는 그룹 간에 유의미한 차이를 보여주었다.

표 3. 읽기 방법/하이라이트 조건에 따라 선택한 단어 횟수의 이원분산분석

		Sum of		Mean			
		Squares	df	Square	F	Sig.	
	읽기 방법	22.990	1	22.990	34.995	.000	***
GH단어	하이라이트 조건	14.520	2	7.258	7.515	.001	* *
	상호작용	.130	2	.067	.069	.933	
	읽기 방법	1.896	1	1.896	3.449	.072	
BH단어	하이라이트 조건	11.690	2	5.843	12.122	.000	***
	상호작용	.300	2	.152	.315	.731	
	읽기 방법	1.779	1	1.779	6.333	.016	*
NH단어	하이라이트 조건	.886	2	.443	1.069	.349	
	상호작용	1.999	2	1.000	2.413	.097	
	읽기 방법	4.344	1	4.344	18.488	.000	***
EX단어	하이라이트 조건	2.050	2	1.025	4.640	.013	*
	상호작용	2.415	2	1.208	5.469	.006	**

(\*\*\*: p<.001, \*\*: p<.05, \*: p<.01)

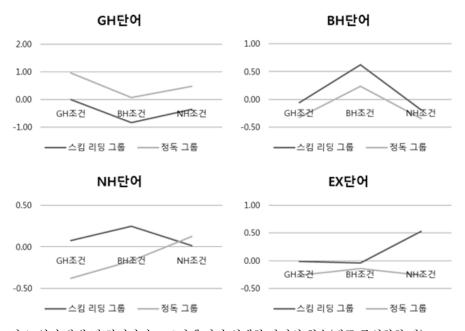


그림 1. 읽기 방법 및 하이라이트 조건에 따라 선택한 단어의 횟수(평균 중심화한 값)

내용 이해도 측정 문제의 정답이라고 할 수 있는 GH문장을 고르는 횟수는 그룹에 따라 나뉜 읽기 방법에 따라 영향을 받았다[F(1, 36)=6.570, p<.05]. 정독 그룹이 스킴 리딩 그룹에 비해 모든 하이라이트 조건 하에서 대체로 높은 점수를 받았음을 알 수 있다. 반면에 BH문장을 고를 때에는 읽기 방법 보다는 하이라이트 조건에 영향을 받음을 확인하였다[F(2, 73)=3.537, p<.05]. 사후 검증 결과 GH조건과 BH조건에서 차이를 보였다.

문장 및 단어를 고르는 결과에서 대체로 비슷한 패턴을 발견할 수 있다. GH문장은 대체로 읽기 방법에서는 정독 그룹에서 고르는 횟수가 높으며, BH조건에서 선택한 횟수가 가장 낮고, GH조건과 NH조건에서 비교적 높다. BH문장에 대해서는 스킴 리딩 그룹에서 고르는 횟수가 대체로 높으며, 하이라이트 조건은 BH조건에서 높고 GH조건과 NH조건에서 낮다. 이러한 패턴은 문장과 단어에서 공통적으로 발견할수 있다. 그러나 문장을 묻는 문제에서는 단어 문제에 비해 대체로 전반적인 유의도가 감소하는 것을 확인할수 있었다. 기억해서 대답해야하는 선택지의 단위가 크기가 영향을 준다는 것을 예상할 수 있다. 고난도 문제에서는 읽기 방법에 따라서만 유의한 결과를 나타냈다. 글을 정독할 때 스킴 리딩할 때에 비해 확연히 높은 점수를 받았다[F(1, 36)=13.142, p<.001]. 반면에 하이라이트 조건의 차이는 결과에 거의 영향을 미치지 못하였다[F(2, 73)=.170, p=n.s.].

표 4. 읽기 방법/하이라이트 조건에 따른 선택한 문장의 횟수와 고난도 문제 점수의 이원분산분석

		Sum of		Mean			
		Squares	df	Square	F	Sig.	
	읽기 방법	5.759	1	5.759	6.570	.015	*
GH문장	하이라이트 조건	1.320	2	.660	1.301	.278	
	상호작용	.640	2	.318	.627	.537	
	읽기 방법	1.417	1	1.417	2.361	.133	
BH문장	하이라이트 조건	2.907	2	1.453	3.537	.034	*
	상호작용	1.709	2	.855	2.080	.132	
	읽기 방법	1.463	1	1.463	3.318	.077	
NH문장	하이라이트 조건	.354	2	.177	.679	.510	
	상호작용	.329	2	.165	.631	.535	
711	읽기 방법	3.265	1	3.265	13.142	.001	***
고난도	하이라이트 조건	.087	2	.043	.170	.844	
문제	상호작용	.077	2	.039	.152	.859	

(\*\*\*: p<.001, \*\*: p<.05, \*: p<.01)

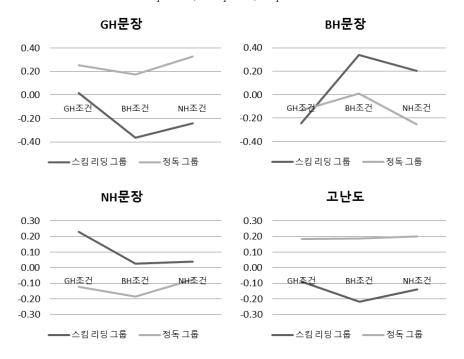


그림 2. 읽기 방법/하이라이트 조건에 따라 선택한 문장의 횟수와 고난도 문제 점수

표 5. 각 AOI(판독 영역)에 단위 면적당 시선의 평균 누적시간(괄호 안은 SD)

		스킴 리딩 그룹			정독 그룹	
	GH조건	BH조건	NH조건	GH조건	BH조건	NH조건
gAOI	24.16 (12.08)	17.97 (8.91)	18.99 (7.46)	57.50 (18.33)	60.27 (29.95)	60.02 (18.63)
bAOI	13.44 (2.97)	19.02 (5.98)	15.28 (4.36)	46.29 (15.14)	57.08 (21.49)	49.90 (12.26)
nAOI	13.40 (3.83)	12.49 (3.61)	14.15 (3.80)	46.92 (18.71)	47.79 (14.36)	46.78 (7.50)

표 6. 읽기 방법/하이라이트 조건에 따라 선택한 단어 횟수의 이원분산분석

		Sum of		Mean			
		Squares	df	Square	F	Sig.	
	읽기 방법	32857.0	1	32857.0	93.530	.000	***
gAOI	하이라이트 조건	56.0	2	27.8	.100	.905	
	상호작용	341.0	2	170.4	.614	.545	
	읽기 방법	26886.0	1	26886.0	108.100	.000	***
bAOI	하이라이트 조건	986.0	2	493.0	5.202	.009	* *
	상호작용	101.0	2	50.7	.534	.589	
	읽기 방법	24840.0	1	24840.0	104.200	.000	***
nAOI	하이라이트 조건	2.4	2	1.2	.031	.970	
	상호작용	26.6	2	13.3	.340	.713	

(\*\*\*: p<.001, \*\*: p<.05, \*: p<.01)

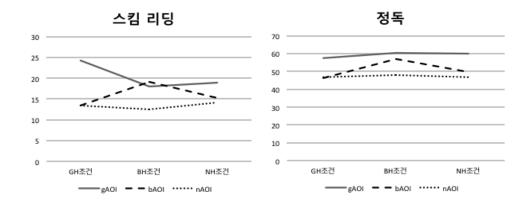


그림 3. 하이라이트 조건에 따라 각 AOI간의 시선이 머문 시간 비교

전체 실험 과정에서 동시에 진행한 아이트래커를 이용한 시선 추적결과도 같은 방법으로 이원분산분석(two-way ANOVA)을 실시하였다. 시선 추적 실험은 글의 내용을 이해하는 것과 상관없이 단순히 시선이 하이라이트에 더 오래 고정되는지를 확인하고자 하는 실험이다. 스킴 리딩 그룹에게는 주어진 시간이 적기 때문에, 정독 그룹에 비해 모든 경우에 누적 시간이 적을 것임이 자명해 보이지만. 일원분산분석(one-way ANOVA)를 사용하지 않은 이유는 하이라이트와의 상호작용을 관찰하기 위함이다. 이 때 독립변인은 그대로 2(읽기 방법) × 3(하이라이트 조건) 이며, 종속변인은 각 AOI에 단위 면적당 시선이 머무르는 누적 시간이다. 이해도 측정과 마찬가지로 읽기 방법은 피험자 간(between-subject), 하이라이트 조건은 피험자 내(within-subject) 분석을 하였다. 각 AOI 마다 표시된 면적이 동일하지 않기 때문에 머무른 절대 시간을 사용하지 않고 단위 면적으로 나누어 사용했다. 단위 면적은 화면에 한글 한 글자가 표시되는 면적으로 계산 했기 때문에 한 글자당 평균 어느 정도의 시선이 누적되었는지 짐작할 수 있다. 측정의 단위는 밀리 세컨드(ms, 1/1000초)이다. 모든 참여자는 아이트래커를 통해 시선 추적을 시도하였지만, 시선 추적이 실패하여 충분한 시간이 기록되지 않은 참여자는 제외하였다. 그 결과 분석에 포함된 참여자 수는 N=29(스킴 리딩 그룹 15명, 정독 그룹 14명)이다.

모든 AOI에 유의미하게 정독 그룹과 스킴 리딩 그룹의 시선 누적시간에 대략 관찰하여도 2배 이상 큰 차이를 보였는데, 시간에 제한을두지 않았기 때문에 당연한 결과이다(표 5). 그림 3에는 그림 1과 그림 2와는 달리 각 그룹별로 그래프를 나누어 그렸다. 이번 실험은 시선의누적 시간을 측정하는 것인데, 두 그룹 간 차이를 비교하는 것은 의미가없으므로 그 대신에 각 하이라이트 조건에서 각 AOI에 머무는 시선의길이를 관찰할 수 있게 표현하였다.

AOI 가운데 gAOI는 GH조건에서 하이라이트가 표시된 부분이기 때문에 글에서 내용 상 가장 중요한 부분이라고 볼 수 있다. 이 부분은 하이라이트 조건에 관계없이 모든 조건에서 가장 시선 고정이 길게 되는 것을 관찰할 수 있었다. 글의 내용 상으로도 중요한 부분이기 때문에 하이라이트의 조건과 상관없이 gAOI에는 대체로 시선이 많이 고정되는 것을 확인할 수 있다. 그러나 여섯 가지 조건 가운데 유일하게 스킴 리딩 그룹의 BH조건에서 gAOI[M=17.97, SD=8.91]보다 bAOI[M=19.02, SD=5.98]에 시선이 더 오래 고정되었음을 확인할 수 있다. 글을 상세히 읽을 시간이 부족한 스킴 리딩 상황에서는 하이라이트의 영향을 받아 중요하지 않은 부분에 하이라이트가 존재할 지라도 시선을 더 오래 머물게 함을 확인할 수 있다. 사후 검정결과에서도 bAOI에 대해 GH조건과 BH조건 사이에 유의미한 시선 길이의 차이를 보였다[F(2, 54)=5.202, p<.01](표 6).

# 4. 실험 2: 소스에 대한 심리적 차이

### 4.1. 개요

이 실험은 서로 다른 커뮤니케이션 소스로 분류된 컴퓨터 알고리즘(기술적 소스)과 크라우드소싱 방법(다른 다수의 수용자들이 소스)이 수용자로부터 서로 다른 심리적인 반응을 불러일으키는지 확인하는 것을 목적으로 한다. 서로 다른 소스의 종류가 심리적 요인으로 작용하면서, 수용자가 하이라이트에 대해 인지 요소(신뢰도, 대표성, 효용성)에 차이를 가져오는지를 확인한다. 실험은 두 개의 참여자 그룹에 하이라이트가 똑같이 표시된 글을 제시하면서, 각각 컴퓨터 알고리즘과 크라우드소싱을 이용하여 생성된 하이라이트라고 달리 설명한다. 글을 읽은 후에는 참여자로부터 받는 하이라이트에 대한 평가에 차이가 있는 지를 알아본다.

실험 참여자는 24명을 모집하였으며 19세에서 33세 사이의 대학생 및 대학원생을 대상으로 하였다. 실험에서 모든 참여자는 똑 같은 하이라이트가 표시된 4개의 문서를 받지만 2개는 컴퓨터 알고리즘으로 작성된 것이라고 안내 받으며, 다른 2개는 크라우드소싱으로 작성된 것이라고 안내를 받는다. 실험은 모두 하이라이트가 표시된 문서를 주고 읽는 시간의 제한을 두어 스킴 리딩을 유도하는 조건에서 이루어진다. 참여자에게는 4개의 문서가 주어지며 정보를 전달하는 목적의

설명문이다. 문서 하나 당 1분의 제한 시간이 주어지며 중간에 다음 화면으로 넘어갈 수 없다. 각각의 문서를 읽은 후에는 하이라이트에 대한 평가 항목이 주어지며 방금 읽은 문서에 표시된 하이라이트에 대한 신뢰도, 대표성, 효용성을 측정하게 된다. 이와 함께 문서 내용에 집중하게 하기 위해 삼지선다형과 정오를 맞추는 문제로 구성된 퀴즈 4문항을 함께 풀게 된다.

### 4.2. 조건과 처리

모든 참여자는 실험에서 4개의 문서를 읽는다. 각각의 문서는 예술, 기술, 과학, 역사 분야의 설명하는 글이다. 각각의 글은 유명 포털 사이트에 게재된 기사와 대학수학능력시험 모의고사에서 내용을 발췌하여 분량과 내용을 재구성하였다. 문서의 길이는 각각 1,000자 내외로 하였으며, 부정적이거나 긍정적인 반응을 불러일으킬 만한 소재를 담는 글을 가급적 배제하고 가치중립적이고 단순 지식을 전달하는 내용인 글을 위주로 선별하였다. 이러한 글의 종류를 선택한이유는 실제로 온라인에서 문학 작품 등 보다 단순 정보 전달의 글을 많이 접하기 때문이다(Shaikh & Chaparro, 2004). 글에 실제로 표시된하이라이트는 실험자가 임의로 중요한 문장을 선별하여 표시하여참여자가 하이라이트의 도움을 받으면서 읽을 수 있도록 한다.하이라이트는 문장 단위 혹은 문장이 2개 이상의 문장이 길게 이어진

복문인 경우에는, 쉼표를 기준으로 나누어서 표시하기도 하였다. 한 글에 4-5개의 문장을 선정하여 표시하였으며, 한 단락(paragraph) 당 2개이하로 제한하였다.

실험은 실험 환경에 맞는 웹 페이지를 제작하여 이루어졌다. 실제참여자가 실험을 수행하는 공간은 따로 마련한 원격지에서이루어졌으나 통제된 환경에 가깝게 설정하였다. 평소에 참여자가편하게 다루는 컴퓨터가 있는 장소에서 실험 전과 실험 후에 전화를 통해지속적으로 통제를 하였으며 온라인을 통해 참여자가웹 사이트에서입력하는 내용을 실시간으로 실험 진행자가 통제하는 방식으로이루어졌다. 참여자가 하이라이트를 표시한 소스가 어느 것인지를참여자가확실하게 알도록하기위해실험 내내끊임없이 반복하여말로설명함과 동시에 화면에 표시된 내용 등으로 참여자를 지속적으로상기시켰다. 이러한 방법은 웹 사이트의 인터페이스와 표시된 글, 그리고전화를 통한육성을 이용한 방법으로 조작이 이루어졌다.

먼저 온라인 인터페이스에서는 각 문서를 읽기 전에 하이라이트의 추출 방법에 대한 설명이 보조하는 이미지와 함께 제공이 된다. 사용자는 이 화면의 글을 꼼꼼하게 읽고 넘어가도록 요구를 받으며, 실험 진행자에 의해 진행시간이 실시간으로 점검 받기 때문에 충분한 시간을 가지고 넘어가지 않은 참여자에 대해서는 다시 자세히 읽을 것을 요구 받는다. 화면에 글을 읽는 내내 화면의 하단에 "본문에 표시된 하이라이트는

[컴퓨터 프로그램의 자동추출/다수의 참여자에 의한 직접 표시]에 의해 작성된 것입니다"라는 메시지를 표시하였다. 참여자는 글을 읽으면서 지속적으로 표시되는 남은 시간을 중간중간 확인하게 되므로, 남은 시간을 표시하는 글을 이 메시지와 함께 노출하였다. 또한 어떤 소스에 의해 하이라이트가 생성된 것인지는 다른 글들에 비해 눈에 띄게 표시되도록 굵게 표시하였으며 글꼴의 크기도 크게 표시하였다. 문서를 읽은 다음에는 하이라이트에 대한 평가가 이어지는데 역시 이 화면에도 어떤 방법으로 작성된 것인지에 대한 설명이 포함되어 있다. 하이라이트의 인지를 측정하는 문항에도 매 문항에 하이라이트를 생성하는 방법에 대한 설명이 포함되어 있다.

실험 참여자들은 원격지에서 화면을 보고 실험에 대한 통제를 받게 되므로 참여자가 통제에서 벗어나거나 다른 변인의 영향을 받을 위험이 있기 때문에 모든 참여자는 실험 내내 전화로 지속적인 통제를 받는다. 문서를 읽기 전에는 실험에 대한 설명과 함께 각 하이라이트 생성 방법에 대한 내용을 실험 진행자가 직접 설명해주며, 같은 말을 반복할 경우통제에 대한 집중도가 떨어질 것을 우려하여 화면에 표시된 내용과는 다른 문장으로 재구성하여 설명하였다. 특히 화면에서 하이라이트 생성 방법이 표시된 부분을 지속적으로 재확인 시키면서 참여자가하이라이트 생성방법을 계속해서 상기시키도록 하였다. 어떤 소스로생성된 문서를 읽었는지 실험 내내 정확히 인지하고 있는 것이

중요하므로 다 읽은 후에는 조작적 검증으로 어떤 방법으로 하이라이트가 생성된 문서를 읽었는지 참여자가 직접 확인하도록 하였다.

## 4.3. 측정

소스의 종류는 독립변인에 해당하며, 인지 요소(perception)의 측정은 종속변인에 해당한다. 이를 평가하는 방법은 참여자가 직접 점수를 주는 것에 의해 측정이 된다. 관련 연구에서 사용한 척도를 함께 참고하였다. 여기에는 신뢰도, 대표성, 효용성을 포함한다. 각 질문 항목은 표 7과 같다. 응답자들로 하여금 그들이 받은 추출된 하이라이트에 대한 전반적인 신뢰도와 대표성을 평가하도록 하며, 응답은 1점(전혀 동의하지 않는다)부터 5점(매우 동의한다)까지의 5점 리커트 척도(Likert scale) 상에서 이루어졌다.

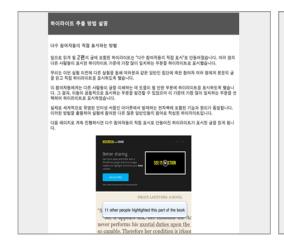
신뢰도(credibility)의 측정은 Yang과 Oliver(2004)는 온라인 뉴스기사 인지를 측정하면서 중요성, 신뢰도, 오락성 등 세 가지 차원, 12개의 항목을 활용한 바 있다. 본 연구에서는 그 중에서 신뢰도 인지를 측정한 4개 항목 가운데 3개의 항목을 하이라이트 추출이라는 상황에 맞게 수정하였다. 각각의 항목은 믿을 만한지(believable), 확신이가는지(convincing), 정확성이 있는지(accurate)를 묻는다.

대표성(representativeness)의 측정은 Sundar와 Nass(2001)의 온라인 뉴스 선택 연구에서 사용된 대표성 항목 중에 5 가지 가운데 하이라이트 추출이라는 상황에 맞는 3가지 요소를 추렸으며 각각의 항목은 중요한 부분인지(important), 유용한 정보가 포함되었는지(informative), 적절한지(relevant)를 기준으로 측정하도록 했다.

효용성(usefulness)을 측정하는 척도로는 Davis(1989)가 사용한 인지된 효용성(perceived usefulness)을 측정하는 척도에 관한 연구에서 도출된 결과를 참조하였다. 이 연구에서 요인 분석을 통해 밝혀낸 6가지 척도 가운데 시간이 한정된 실험 연구에서 평가하기 어려운 요소를 제외한 3가지 척도를 사용하였다. 하이라이트가 내용 파악에 도움이되는지(job performance), 하이라이트의 도움으로 수월하게 읽을 수 있었는지 (makes job easier), 하이라이트가 작업을 효율적으로 만드는지(effectiveness)를 묻는 질문을 포함하였다.

표 7. 신뢰도와 대표성의 측정 항목

	측정 항목
신뢰도	이 [컴퓨터 알고리즘/다른 참여자들]에 의해 추출된 하이라이트는 믿을만해 보인다 이 [컴퓨터 알고리즘/다른 참여자들]에 의해 추출된 하이라이트에 대해 확신이 간다. 이 [컴퓨터 알고리즘/다른 참여자들]은 정확성이 있게 하이라이트를 추출하는 것으로 보인다
대표성	이 [컴퓨터 알고리즘/다른 참여자들]에 의해 추출된 하이라이트는 중요한 부분이다 이 [컴퓨터 알고리즘/다른 참여자들]에 의해 추출된 하이라이트는 적절성 있게 표기되었다. 유용한 정보가 이 [컴퓨터 알고리즘/다른 참여자들]에 의해 추출된 하이라이트에 포함된다.
효용성	이 [컴퓨터 알고리즘/다른 참여자들] 인해 이 글의 내용을 더 잘 파악할 수 있었다. 이 [컴퓨터 알고리즘/다른 참여자들]의 도움으로 글을 수월하게 읽어 나갈 수 있었다. 이 [컴퓨터 알고리즘/다른 참여자들]이 이 글의 읽기 과정을 효율적으로 만들어주는 것 같다.



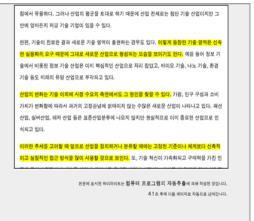


그림 4. 소스의 종류에 따른 심리적 차이를 알아보기 위한 실험용 웹 사이트 화면

하이라이트 추출 방법 설명 화면(좌)과 제시된 글을 읽는 화면(우)

글의 내용을 적극적으로 파악하려는 노력이 없으면 하이라이트의 도움을 받거나 글의 내용에 집중력이 떨어질 가능성이 있다. 따라서 인지 요소에 대한 평가와 함께 글의 내용 이해도를 측정하는 퀴즈를 하이라이트 평가와 함께 제시하였다. 이전의 유사한 연구를 참조하였는데, Dyson과 Haselgrove(2001)는 문장의 길이와 읽는 속도가 이해도에 주는 영향을 평가할 때, 글의 내용에 대해 물어보는 퀴즈를 삼지선다형과 정보가 맞았는지 여부를 묻는 확인하는 문제 형태를 동시에 출제하여 평가하였다. 이들이 사용한 방법을 참고하여 내용에 대한 이해도는 삼지선다형과 글의 내용에 관한 진술에 대한 정오(그렇다/아니다)를 판단하는 문제를 각각 2문제씩 출제한다. 퀴즈는 글의 주제를 묻는 쉬운 문제부터 깊이 이해해야 풀 수 있는 문제까지 난이도를 다양하게 하였다. 독자가 글을 읽으면서 하이라이트에 자연스럽게 도움을 받는 것이 아니라 하이라이트에만 의식적으로 집중하여 문제를 푸는 것을 피하기 위해. 내용 이해도의 측정 퀴즈는 하이라이트가 있는 부분을 포함하여 글의 다양한 부분에서 고르게 출제하였다. 퀴즈에는 각 보기 외에도 "모름" 항목을 추가하여 참여자가 아무렇게나 찍어서 맞추는 것을 방지하였다.

#### 4.4. 과정

실험은 원격지에서 연구를 위해 제작한 웹 사이트를 통해 이루어지지만 앞서 설명한 바와 같이 참여자가 진행하는 과정을 온라인을 통해 실험진행자가 일대 일로 통제하며, 전화를 통해 음성으로도 계속해서 의사 소통하는 방식으로 진행되었다. 실험을 시작하면서 전화를 걸어 참여자가 반드시 실험이 진행되는 20여 분 동안 다른 방해 요소 없이 집중할 수 있는 환경에 있는지 확인한다. 그렇지 않을 경우에는 실험을 중단하고 충분한 환경이 마련되었을 때 실험을 진행하였다. 실험의 내용에 대해서는 하이라이트가 표시된 글을 제한 시간 동안 읽게 되며, 하이라이트가 잘 표시되었는지 평가하는 문항과 퀴즈를 풀게 된다고 설명한다. 그리고 온라인 메신저 혹은 문자를 통해 실험 진행자로부터 페이지 접속 주소와 참여자 코드를 부여 받는다. 참여자는 이 주소를 직접 입력하여 접속하게 되며, 참여자마다 자기가 속한 그룹에 따라 다른 주소를 받게 된다. 참여자 코드는 웹 사이트에 참여자가 직접 입력함으로써 데이터베이스에 저장할 때에 참여자를 구분하는 용도로 사용되다.

설명에 따라 실험 사이트에 접속하면, 성별과 나이 등 인구통계정보를 입력한다. 다음 화면으로 넘어가면 앞으로 읽게 될 문서에 생성되는 하이라이트를 어떻게 추출했고 표시했는지를 상세하게 설명한 페이지로 넘어간다. 사용자는 이 페이지를 대충 읽고 넘어갈 수 없도록 실험 진행자에 의해 통제 받는다. 설명의 내용에는 인터넷 상의 모든 문서를 전문가가 선택할 수 없으므로. 새로운 방법을 통해 선택하는 과정이라는 설명이 포함되며, 구두로 다시 한 번 반복하여 조건을 인식시킨다. 이 화면을 넘어가면 첫 번째 문서인 예술 주제의 글을 읽게 된다. 1.000자 내외로 이루어진 글에는 4-5개의 하이라이트가 표시되어 있으며, 참여자는 마우스로 스크롤 하며, 글을 읽는다. 사이트 하단에는 지속적으로 하이라이트 생성방법이 표시되며, 남은 시간이 초 단위로 표시되어 참여자가 시간을 분배하여 읽을 수 있도록 하였다. 이 시간은 파일럿 테스트를 통해 결정하였으며, 제한 시간 동안 성실하게 글을 읽고 문제를 풀면 5점 만점에 평균적으로 3점 정도가 나오는 선에서 결정하였다. 시간이 지나면 자동으로 설문 페이지로 넘어가며, 여기에서 방금 읽은 하이라이트에 대한 평가를 하며, 이와 함께 퀴즈를 풀게 된다. 이와 같은 과정은 기술, 과학, 역사에 대한 글에서도 반복된다. 그리고 2개의 문서를 읽게 되면 소스의 조건이 바뀌게 된다. 여기서 다시 새로운 소스에 대한 설명을 읽고. 구두로도 소스의 종류에 대해 상세히 듣는다. 순서에 따른 편향을 없애기 위해 24명의 참여자는 무작위로 절반의 참여자는 컴퓨터 알고리즘 소스 조건을 먼저 실험하고. 나머지 절반의 참여자는 크라우드소싱 조건을 먼저 실험한다. 어떤 조건을 먼저 실험하게 되든 간에 4개의 자극을 모두 실험하게 되며, 각 소스에 따라 2개씩의 문서를 읽게 되는 것은 같다.

### 4.5. 결과 및 분석

인지 요소에 따라서 어떤 점수가 나왔는지를 분석하였다. 각 인지 요소인 신뢰도, 대표성, 효용성은 각각 3개의 항목으로 측정되었기 때문에, 각 요소는 3개 항목의 합으로 계산되었다. 각 인지 요소를 구성하는 각 3 가지 항목들의 상관관계를 확인하고 내적 합치도(internal consistency)를 계산하였다. 제시된 각 4개의 글(예술, 기술, 과학, 역사)에 대한 참여자들의 기본 지식이나 각 글이 가지고 있는 내적인 난이도, 실제로 표시된 하이라이트의 정확성 등이 인지 요소 항목 측정에 영향을 미칠 여지가 다분하다. 따라서 이러한 차이를 피하기 위해 Sundar와 Nass(2001)의 유사한 실험에서 분석 방법을 참조하여 각 점수에 대한 평균 중심화(mean centering)을 실시하여 점수를 변환하였다. 이를 테면, 신뢰도의 평균 점수가 x이고, 예술, 기술, 과학, 역사 지문에서 신뢰도의 점수가 각각 a, b, c, d로 나타났다면, (a-x), (b-x), (c-x), (d-x)를 변환된 점수로 사용한다. 연구에서 확인하고자 하는 독립변인은 소스의 종류 한가지 이지만, 한 명의 참여자가 2개의 소스 조건에서 각각 2번씩 차례로 측정하도록 하기 위해. 한 명의 참여자에 대하여 2(소스의 종류) × 2(글의 순서) = 4회의 관찰을 하게 되었다. 관심이 있는 독립변인은 소스의 조건 한 가지이고, 한 피험자가 서로 다른 조건을 반복하여 측정하였으므로. 피험자 내 비교를 진행하기

위해 반복측정 일원분산분석(one-way ANOVA with repeated measured)을 실시하였다.

각 인지 요소에 대한 측정 항목의 내적 합치도는 Cronbach's  $\alpha$ 를 통해 계산하였는데, 신뢰도( $\alpha$ =.87), 대표성( $\alpha$ =.85), 효용성( $\alpha$ =.73)등 세 항목 모두에 대해 내적 합치도가 우수함을 확인하였다. 모든 참여자에 대한 조작적 검증의 결과, 24명 중 24명 모두가 읽었던 문서의 하이라이트 조건을 모두 정확히 인식을 하고 있던 것으로 나타났다.

표 8에 나타난 수치는 평균 중심화를 하였기 때문에 각 인지요소의 평균을 0으로 둔 값이다. 결과를 보면 세 가지 인지 항목 간에 유의미한 차이가 있었던 인지 요소는 신뢰도이다. 크라우드소싱으로 추출한 것으로 생각된 하이라이트(M=.55, SD=2.3)가 컴퓨터 알고리즘으로 추출한 것으로 생각된 하이라이트(M=-.55, SD=2.13)보다 신뢰도가 높은 것으로 나타났다[F(1, 23)=4.610, p<.05]. 대표성과 효용성 등 다른 인지 요소들도 평균적으로 크라우드소싱이 높은 점수를 나타냈지만, 유의미한 값은 아니었다(표 9).

표 8. 평균중심화 이후 소스에 따른 각 인지요소 점수의 평균(괄호 안은 표준편차)

	소스					
	컴퓨터 알고리즘	크라우드소싱				
신뢰도	55 (2.13)	.55 (2.3)				
대표성	33 (2.19)	.33 (2.05)				
효용성	38 (2.37)	.38 (2.19)				

표 9. 각 소스에 따른 인지요소 점수의 일원분산분석

	Sum of	df	Mean Square	F	Sig.	
	Squares					
신뢰도	29.260	1	29.260	4.610	.043	*
대표성	10.670	1	10.670	1.718	.203	
효용성	13.500	1	13.500	2.070	.164	

(\*\*\*: p < .001, \*\*: p < .05, \*: p < .01)

# 5. 하이라이트의 생성

# 5.1. 개요

컴퓨터 알고리즘은 본질적으로 인간의 인지 과정과는 다른 메커니즘을 갖기 때문에 사람이 직접 수행한 결과물보다는 신뢰감을 적게 준다는 것을 실험 2를 통해 보였다. 그렇다면 이에 대한 대안으로 크라우드소싱이 하이라이트 생성에 유용한 방법으로 응용될 수 있을 것이고, 이를 활용한 프로토타입(prototype)을 실제로 만들어보는

실험은 추후에 실제 어플리케이션을 만들 때 도움이 될 수 있을 것으로 보인다. 크라우드소싱은 또한 많은 참여자가 확보되어야 정확한 결과를 내는 구조이기 때문에, 비교적 사람들이 적게 읽는 문서에서는 그 효과를 보기가 힘들다. 또한 독자에게 직접 하이라이트를 표시하라고 요구하는 것은 쉬운 일이 아니고 실험 2에서 보았듯이 부정확한 하이라이트는 잘못된 내용 이해를 가져올 수도 있다.

본 연구에서는 이러한 어려움을 극복하고자 자연스러운 읽기행동만으로 하이라이트 생성에 참여 할 수 있는 새로운 아이디어를 제안한다. 모바일 기기의 화면 크기는 전체 글의 일부 만을 화면에 표시하며, 어떤 부분을 오래 읽을수록 그 부분은 화면에 오래 표시될 것이라는 점에 착안하였다. 글을 읽는 어플리케이션을 제작하여 독자가하나의 글을 자연스럽게 읽으면 화면에 많이 노출 되는 시간을 누적하여 저장한다. 가장 많이 노출 된 부분이 사람들에게 많이 읽힌 부분이라고 가정한다면, 적당한 양의 데이터가 쌓였을 때, 이 부분에 하이라이트를 표시하여 다음 독자에게 제공할 수 있을 것이다.

이를 구현하기에 앞서 모바일 기기에서 사람들이 실제로 화면의 어떤 부분을 읽는지 등의 패턴을 알아내는 것이 필요하다. 모바일 기기에서의 읽기 패턴은 전통적인 종이에서의 그것과 상당한 차이점이 있다. Biedert 등(2012)에 따르면, 독자들이 모바일 기기에서 글을 읽을 때 화면의 최상단과 최하단은 잘 바라보지 않는다. 또한 한 화면에 표시된 내용전체를 읽은 후 다음 부분으로 이동하는 방법보다는 더욱 상세한 수준에서 단락마다 혹은 행을 옮길 때 마다 자주 스크롤을 사용하여화면을 이동해가며 글을 읽는다는 것을 알아냈다. 이를 참조하여 실제로 영상으로 사람들이 모바일 기기에서 글을 읽는 행동을 관찰하고 이를

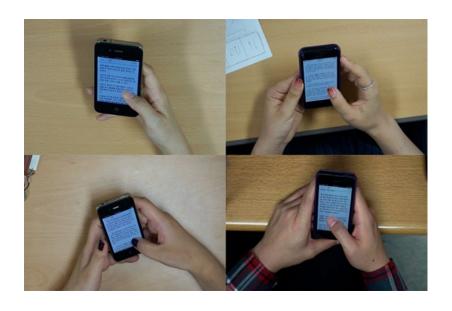


그림 5. 모바일 화면에서 글 읽는 화면을 관찰하기 위한 비디오 실험

분석하여 어떤 패턴을 보이는가 탐색하였다. 녹화된 영상을 통해 화면을 터치하는 손가락을 움직이는 방식이나 모바일 기기를 쥐는 방법 등을 관찰하였다. 영상의 내용을 분석하여 어플리케이션을 제작하기에 앞서 화면에 표시되는 문장의 노출 시간을 직접 영상을 보며 측정하여 하이라이트를 표시해 보았다. 또 이렇게 표시된 하이라이트와 사람들이 직접 중요하다고 생각하여 표시한 하이라이트를 비교하여 상관관계를 계산하였다.

한편 정확도를 향상시키기 위해 화면 상에 노출되는 모든 문장을 독자가 읽고 있는 것은 아니라는 점을 고려하였다. 예컨대 엄지 손가락으로 스크롤을 한 상태로 화면의 아랫부분을 가리고 있다면, 손으로 가린 부분을 제외한 화면 상단에 있는 문장의 노출 시간만을 측정하는 것이 더욱 합리적일 것이다. 만약 사람들이 화면의 맨 윗줄만 주로 읽는 패턴이 관찰된다면 굳이 화면에 표시된 모든 문장의 노출 시간을 측정할 필요는 없을 것이다. 이러한 여러 가지 가정을 바탕으로 노출 시간을 계산하는 방법을 달리하여 여러 방법으로 측정하여 보았다. 그리고 이러한 방법으로 실제로 모바일 어플리케이션의 프로토타입을 제작하여 각 문장의 노출 시간을 측정하여 데이터를 수집하고, 가장 많이 노출된 문장들을 하이라이트로 표시하여 제공할 수 있도록 하였다.

# 5.2. 영상 촬영을 통한 예비 실험

## 5.2.1. 측정

영상 촬영 실험은 연구 아이디어가 가능성이 있는지를 검증하고 모바일 독자의 읽기 습관을 관찰하는 것을 목표로 하였다. 실험은 14명의 참여자를 모집하여 두 그룹으로 나누었다. 각 그룹의 참여자는 2개의 설명하는 글을 각각 모바일 기기와 종이에서 읽게 하였다. 제시된 글은 각각 약 1,000자(약 300단어) 내외로 이루어져 있고 각각 13개, 23개의 문장으로 이루어져 있었다. 사용자 마다 같은 환경에서 화면을 보고 글을 읽도록 하기 위해 같은 기기를 사용하였다. 사용된 모바일 기기는 Apple사의 아이폰4가 사용되었고 한 화면에 17줄이 표시되도록 하였다.

모바일 화면에서 주어진 글을 읽는 모습을 비디오로 촬영하고 관찰하였다. 촬영할 때에는 모바일은 조작하는 손과 화면의 내용이 화면에 보이도록 하였다(그림 5). 이 때, 화면에 보이는 각각의 문장이얼마나 노출이 되는지 비디오의 타임 스탬프(time stamp)를 이용하여측정하였다. 촬영된 화면을 보며 각각의 화면이얼마나 노출되는지를 직접 기록하여 누적하였다. 노출 시간을 계산하는 방법도 3가지 서로다른 방법을 사용하였다. 화면의 맨 위의 한 문장의 노출 시간만을계산하는 방법(first\_one), 맨 위와 그 바로 아래 두 문장에 대해 노출시간을 재는 방법(first\_two), 그리고 화면에 노출된 모든 문장의 노출시간을 재는 법(whole on screen)을 사용하였다(그림 6).

### 5.2.2. 과정

영상 촬영 실험에 참여한 사람은 두 그룹으로 나뉘며, 한 참여자 당 모바일에서 글 읽기와 하이라이트 직접 표시 등 두 가지 작업을 수행하였다. 각 참여자가 모바일 기기에서 글을 읽을 때에는 촬영을 통해 기록하고, 종이에서 읽을 때는 따로 기록하지 않고 다 읽은 후에 "하이라이트가 있으면 글의 중요한 내용을 파악하는 데 도움이 될 것 같은 부분에 직접 표시를 해보라"는 지시를 통해 종이 문서에 직접 하이라이트를 작성하도록 했다. 첫 번째 그룹에게는 글 1을 모바일 기기에서 읽게 하였고 이 작업이 끝나면, 글 2가 인쇄된 종이를 제시하고 이를 읽고 난 후에 하이라이트를 형광펜을 이용하여 표시하게 하였다. 두 번째 그룹은 반대로 글 2를 모바일 기기에서 읽고. 글 1를 종이에서 읽고 하이라이트를 직접 표시하게 하였다. 그리고 글 1에 대하여 첫 번째 그룹으로부터는 비디오 촬영 결과를 얻고 두 번째 그룹으로부터는 직접 표시된 하이라이트를 모을 수 있었다. 글 2에 대해서는 방법을 반대로 하여 같은 방법으로 결과물을 얻어 노출 누적 시간과 직접 하이라이트를 서로 교차 비교하였다.

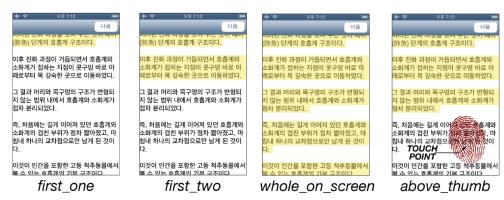


그림 6. 문장의 노출 시간을 측정하기 위해 고안한 4가지 방법

표 10. 글에서 직접 하이라이트를 표시한 문장과 노출 시간 간의 피어슨 상관관계

		글1	글:	2
	r	p	r	р
first_one	.68	.03*	.33	.39
first_two	.54	.12	.22	.64
whole_on_screen	.07	.83	.06	.8

 $<sup>(*:</sup> p \le .05, **: p \le .01, ***: p \le .001)$ 

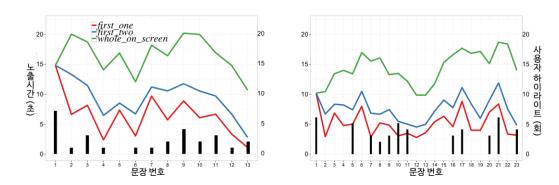


그림 7. 각 문장 별 모바일 기기 화면 노출시간을 세 가지 다른 방법으로 계산한 결과(선 그래프)와 사용자의 직접 하이라이트 표시 횟수(막대 그래프)

## 5.2.3. 결과

각각의 글에서 비디오 촬영을 통한 노출 시간과 사용자가 직접 표시한 하이라이트 간의 상관관계를 계산 하였다. 화면의 가장 상단에 노출되는 문장의 노출 시간(first\_one)과 사용자가 직접 작성한 하이라이트 사이에 상관관계가 유의미하게 관찰되었다 첫 번째 글에서 상관관계가 있음이 드러났으며[r(11)=.68, p<.05], 두 번째 글에서는 뚜렷하게 관찰되지는 않았다[r(21)=.33, p=n.s.]. 그림 7에는 각각 문장의 노출 시간을 3가지 다른 방법으로 측정한 결과를 선 그래프로 나타내었고, 종이로 주어진 각 문장에 표시된 하이라이트의 누적된 개수를 막대 그래프로 나타내었다.

한편 촬영된 영상을 통해 관찰한 읽기 습관에서 글을 읽을 때, 14명의 참여자 중 14명 모두가 한 손으로 모바일 기기를 쥐고 엄지손가락을 이용하여 화면을 스크롤을 하는 것을 관찰하였다. 그리고 이 때 항상화면 하단이 엄지 손가락으로 가려지는 것을 발견하였다. 이러한 관찰을 통해 화면의 하단에 손가락 등으로 가려지는 지점은 읽히지 않는 경우가 많았음이 밝혀졌는데 이를 프로토타입 제작에 반영하도록 했다.

## 5.3. 프로토타입 제작

### 5.3.1. 측정

비디오 실험의 결과를 바탕으로 실험의 두 번째 단계에서는 화면에 어떤 문장이 노출되고 얼마나 오랜 시간 표시되는지를 측정할 수 있는 프로토타입 어플리케이션을 제작하였다. 크라우드소싱을 이용한 하이라이트 생성의 가능성을 실제 사용자가 쓸 수 있는 어플리케이션의 실용성 확인을 위한 예비 실험의 형태로 실시하였다. 참여자는 51명을 모집하였으며 모두 대학생이거나 대학원생으로 다양한 전공을 가지고 있었다. 자극으로 사용된 문서는 총 3개로 각각 312, 232, 324글자로 이루어져 있으며 각각 20, 13, 23문장으로 이루어져 있었다. 온라인 신문기사에서 발췌한 글을 사용하되 큰 긍정이나 부정적인 감정을 불러일으킬만한 가치 판단의 문제를 담지 않은 정보를 전달하는 글을 사용했다.

어플리케이션은 Apple사의 아이폰4에서 구동되도록 제작되었다. 이는 서로 다른 참여자 간에도 동일한 노출 정보량을 보장하기 위함이다. 이 실험은 참여자들이 읽기에 편한 공간에서 이루어졌다. 참여자가 모바일 기기에서 글을 읽으면 화면에 자연스럽게 독자의 읽기 패턴 데이터가 기록되어 서버로 전송되어 데이터베이스에 저장되도록 설계하였다. 읽기 패턴 데이터에는 각 문장의 노출시간, 화면에서의 좌표,

그리고 화면에 스크롤 일어날 경우 그 좌표와 방향을 함께 전송했다. 노출 시간을 측정하는 방법은 비디오 실험에서 관찰한 결과를 토대로 한 가지 방법을 더 추가하였다. 영상 촬영 실험 결과, 엄지 손가락이 화면의 하단을 가린다는 점에 주목하여 엄지손가락이 닿는 부분의 상단에 위치한 문장들의 노출 시가만 계산하는 방법을 추가하였다(그림 6, above thumb).

결과에 대한 간단한 평가도 함께 이루어졌다. 어플리케이션을 통해 수집한 데이터가 생성한 하이라이트가 글의 실제 중요도와 어떤 관계를 가지는지 확인하기 위해 글의 중요도 평가와 비교해보았다. 중요도 평가는 26명을 따로 평가자를 모집하여 진행하였다. 각각의 글에 대해 가장 중요한 문장을 5개 내외로 표시하라는 지시를 주었다. 한 번 표시가 된 문장에 1점씩 부여하고 각 문장에 대한 누적 점수를 그 문장에 대한 중요도 점수로 사용하였다.

참여자들은 스마트폰에 어플리케이션을 다운로드 받아서 실험을 진행하였다. 실험이 시작되기 전해 전화 통화를 통해 상세한 지시사항을 일러주었으며, 실험을 마친 후에도 전화를 걸어 실험이 정확히 끝났음을 확인하였다. 실험 어플리케이션은 간단한 인구통계 정보를 묻는 것으로 시작하여 모바일 화면에서 얼마나 자주 글을 읽는지, 얼마나 길게 읽는지를 조사하는 설문이 포함되었다. 설문을 마치고 참여자들은 3개의 글을 읽게 되며, 하나의 글을 다 읽을 때마다 질문지를 삽입하여 응답하도록 하였다.

### 5.3.2. 결과

실험에서 51명 중 읽기 과정을 성실하게 수행하지 못한 참여자 1명과 기술적인 문제로 데이터가 전송되지 않은 참여자 1명을 제외한 49명을 대상으로 결과를 분석하였다. 77.6%의 참여자는 하루에 5회 이상모바일 기기를 통해 글을 읽으며, 75.5%는 하루에 30분 이상을 모바일기기의 화면을 통해 글을 읽는 다고 응답하였다. 평균적으로 참여자들은하루에 17.1회, 63.8분을 모바일 기기에서 글을 읽는 데 보낸다고응답하였다.

독자들이 화면을 스크롤 할 때, 손가락이 닿는 부분은 평균적으로 화면의 상단으로부터 80.8%되는 지점이었다(중간값=82.8%, 최소값=28.3%, 최대값=98.7%). 이러한 결과는 비디오 실험에서 관찰했던 것과 일치함을 확인하였다. 참여자들은 대부분 화면의 하단부를 스크롤 할 때 사용하며, 이 부분은 보통 손가락에 의해가려지게 된다.

그림 8에는 선 그래프가 4가지 서로 다른 방법으로 측정한 노출 시간을 그리고 있다. 세 개의 그래프는 각각 실험에 사용한 글 3개를 나타낸다. 평가자를 통해 계산한 각 문장의 중요도를 막대 그래프로 함께 나타내었다. 중요하게 평가된 문장과 누적된 노출 시간을 비교해 볼 수 있다. 그러나 중요도 수치와 어플리케이션으로 수집한 노출 시간 사이에서 유의미한 상관관계를 발견하지는 못하였다.

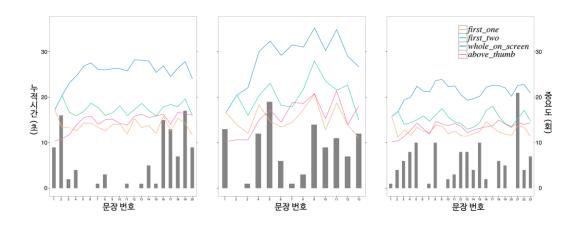


그림 8. 프로토타입으로 수집한 각 문장의 평균 노출시간(선 그래프)과 문장 중요도(막대 그래프)

## 6. 논의

세 차례의 실험을 통해 온라인 문서에서 하이라이트의 효과 및 생성 방법에 관한 연구문제의 답을 찾고 가설을 확인하는 과정을 거쳤다. 실험의 과정은 세 개의 연구문제에 대한 답을 찾고 가설을 검증하는 과정이었다. 먼저 스킴 리딩 상황에서 하이라이트의 효과를 검증하고, 하이라이트의 생성 방법에 따른 독자의 반응의 차이를 알아보았다. 마지막으로 하이라이트를 생성하는 어플리케이션을 제작해 봄으로써 연구 결과를 활용하기 위한 준비 단계를 수행했다.

첫 번째 연구문제 "스킴 리딩 상황에서 글의 중요 부분을 하이라이트를 통해 강조하여 보여 주는 것이 효과적인가?"에 대해 스킴 리딩과 정독 이라는 두 가지 읽기 방법에서 각각 중요한 부분에 잘 표시된 하이라이트(GH조건), 중요하지 않은 부분에 표시된 하이라이트(BH조건), 하이라이트가 표시되지 않은 조건(NH조건) 간에 어떤 차이가 있는지 확인하였다. 특히 단어를 기억하는 데 있어서 효과가 잘 나타났는데, 스킴 리딩 조건에서 하이라이트의 효과가 더 많이 관찰되는 것을 확인하였다. 그리고 하이라이트의 영향을 받아서 중요하지 않은 단어를 중요한 단어로 받아들이는 경우를 관찰하였다. 이를 통해 하이라이트가 스킴 리딩 조건에서 내용을 파악하는 영향을 준다는 가설을 증명할 수 있었다.

글과 상관없는 EX단어를 고르는 비율을 분석한 결과를 보면하이라이트가 존재하는 조건에서는 글에 없는 단어를 중요하다고생각하지는 않는다. 그 하이라이트가 중요하게 표시되었건 중요하지않은 부분에 표시되었건, 문서 자체에 시선을 끄는 부분이 존재하기때문에 글에 더 집중하게 된다. 그 결과, 스킴 리딩을 할 때에하이라이트가 있으면 시선을 끌게되어 읽기 활동에 도움이 된다고 볼수있을 것이다. 이 결과는 실험 설계 시 예상하지 못한 내용이어서흥미롭게생각된다.

하이라이트의 효과에 따른 이해도의 측정은 단어와 문장, 그리고 고난도의 문제로 이루어진 질문으로 진행되었다. 하이라이트의 효과가 단어보다 문장에 덜 나타난 이유에 대해 살펴볼 필요가 있다. 단어와 문장의 차이는 기억하는 단위의 크기에서 비롯된다. 문장의 경우선택지에 포함된 내용의 단위가 단어에 비해 크기 때문에 문장의 일부만 기억하더라도 답안을 고를 수 있을 것으로 생각된다. 우리가 실제로 글을 읽을 때 기억해야하는 내용이 어떠한 단위로 이루어지는가를 탐구한다면, 이러한 평가 기준을 더욱 정교화할 수 있을 것이다. 이러한 맥락에서 하이라이트를 표시하는 단위 크기도 어떻게 정해야 할 지결정하는 연구도 필요할 것이다. 글자, 단어, 구(句), 문장, 문단 등으로 달리하여 그 효과를 측정하는 실험해 보는 것도 필요할 것이다.

단순히 중요한 단어나 문장을 고르는 문제에 비해 고난도의 문제에서는 하이라이트 조건에 따른 효과가 거의 나타나지 않았다. 그러나 읽기 방법에 따라 스킴 리딩 그룹과 정독 그룹에서 확실한 점수의 차이를 보였다. 이 결과는 "하이라이트를 표시하는 대상이 어떤 종류의 글이어야 하는가"에 대한 물음에 답하는 데 도움을 줄 수 있을 것이다. 전문적인 지식을 얻거나 깊은 이해를 요구하는 글은 모든 부분을 꼼꼼하게 정독해야만 이해할 수 있기 때문에, 이런 상황에서는 하이라이트를 표시하는 것이 크게 도움이 되지 않을 것이다. 반면에 앞서 주요 단어를 기억하는 등의 단순한 정보전달이 목적인 글에서 더욱 효과를 나타낼 것임을 시사한다. 하이라이트의 활용도를 높이기 위해서는 모든 종류의 글에 같은 방법으로 하이라이트를 적용할 것이 아니라 글의 종류나 글이 목표로 하는 바를 고려해야 할 것이다.

이와 더불어 전체 글에서 하이라이트가 표시된 문장이 어느 정도의 비율일 때 가장 효과적인지 살펴보는 것도 필요할 것이다. 하이라이트가 너무 적다면, 지나치게 제한된 분량의 내용만을 강조할 수 있기 때문에 큰 도움이 되지 않을 수도 있고, 반면에 너무 많은 양의 하이라이트가 표시되어 있다면, 하이라이트 본연의 효과가 덜어질 수 있을 것이다.

하이라이트가 표시되는 '중요한 부분'이 과연 어디인지에 대한 의문도 실험을 통해 꾸준히 제기되었다. 글을 읽는 사람마다 같은 글을 읽고서도 다른 부분이 중요하다고 생각할 수도 있다. 개인의 성향 차이가 실험 결과에 영향을 주는 것을 방지하기 위해 실험에서는 가급적 다양한 전공을 가진 참여자를 모집하였으나, 애초에 글을 읽는 목적이 다르다면 하이라이트도 달라질 가능성이 있다. 같은 뉴스 기사를 읽더라도 누군가는 어려운 용어를 설명해 주는 부분이 중요하다고 생각할 수도 있고, 누군가는 기존에 알려진 사실이 아닌 새로운 부분을 찾는 것이 더중요하다고 생각할 수도 있을 것이다. 하이라이트의 목적은 글의 내용을 파악하는 데 도움을 주는 것이므로 목적에 따라 다른 하이라이트를 원할수도 있다는 관점을 앞으로 반영하는 것도 필요할 것으로 보인다.

연구가설 1-3을 검증하기 위해 아이트래커를 가지고 시선 추적실험을 진행했다. 하이라이트에 영향을 받아 중요하지 않은 부분에하이라이트가 표시되어 있더라도 스킴 리딩 상황에서는 이를 오래바라본다는 것을 확인하였다. 그러나 기술적인 문제로 실험 과정에서내용의 이해도와 시선의 머문 시간의 관계를 분석하지는 못하였다. 시선추적 실험의 경우 데이터가 온전히 모이지 못하는 경우가 많이 있다. 본실험에서도 시선 추적 데이터를 가져오는데 실패하여 모든 사용자의데이터를 활용하지는 못하였다. 글을 읽다가 자세를 바꾸거나 눈을가늘게 뜨는 경우 추적에 실패하게 된다. 특히 정독 그룹의 경우 이러한경우가 두드러졌는데, 글을 읽는데 충분한 시간이 주어지기 때문에 몸의움직임 등이 잦아 추적에 실패하는 경우가 많이 있었다. 더 정교한기기를 사용하거나 자세를 제어할 수 있는 방법들을 활용한다면, 시선이

많이 머문 단어를 이해도 테스트에서 선택하는지 확인하는 방법을 통해 더 완성도 있는 실험 결과를 낼 수 있을 것으로 생각된다.

두 번째 연구문제 "하이라이트를 추출하는 두 가지 소스(source)에 따른 심리적 차이가 하이라이트에 대한 평가에 영향을 미치겠는가?"에 대한 답을 찾기 위해 실험을 통해 컴퓨터 프로그램, 프로그램의 사용자까지도 커뮤니케이션의 요소로 간주하고 커뮤니케이션의 모델에 따라, 모든 요소를 소스로 해석하였다. 소스의 측면에서 보았을 때컴퓨터 알고리즘은 기술적인 소스이며, 크라우드소싱은 수용자와 같은 상황에 놓인 다른 수용자들을 소스로 간주한 것이다. 여기서 크라우드소싱을 통한 하이라이트에 대한 신뢰도가 높게 나온 것은, 하이라이트 생성과 같은 인간의 인지과정이 중요한 독자들이 더 높은 신뢰를 보임을 의미한다. 같은 하이라이트를 두고도 더 신뢰가 가는 결과물로 인지한 것으로 보아 어떤 이러한 어플리케이션의 결과물을 평가하고 인지함에 있어서도 심리적인 요인이 중요하게 작용함을 알 수 있었다.

신뢰도에 비해 대표성과 효용성에서는 두드러지는 효과가 관찰되지 않은 이유에 대해서도 탐구해볼 필요가 있다. 유의미한 차이가 있는 것으로 드러난 신뢰도를 다른 인지의 요소로 측정한 대표성과 효용성에 비교해보면, 대표성의 경우에는 문항의 항목은 중요한 부분인지, 유용한 정보가 포함되었는지, 적절한지를 묻는 항목으로 대체로 하이라이트

자체를 객관적으로 보는 요소가 담겨있다. 반면, 신뢰도 항목은 하이라이트가 믿을 만한지, 확신이 가는지, 정확성이 있는지를 묻는 항목으로 이루어져 비교적 주관성을 갖는 문항으로 이루어져 있기 때문에 더 유의미한 차이를 보여주었을 것으로 보인다. 또 다른 인지요소인 효용성의 경우에도 하이라이트 생성을 실제로 사용했을 때도움이 되는 지의 측면을 내포하고 있기 때문에 중요한 요소라고 볼 수 있다. 효용성은 실험 표본의 숫자가 그렇게 크지 않아 통계적 유의성을 관찰하지는 못했지만, 추후 연구에서 효용성을 인지하는데 영향을 주는요소를 발견한다면 사람들이 크라우드소싱을 통한 방법으로 생성된하이라이트가 더 가치 있다고 생각한다는 근거로 가치 있게 사용될 수 있을 것이다.

여기서 신뢰도, 대표성, 효용성은 어떤 하이라이트가 실질적으로 가지고 있는 특성이 아니라 수용자가 심리적으로 인식하는 인지요소로서, 인지된 신뢰도(perceived credibility), 인지된 대표성(perceived representativeness), 인지된 효용성(perceived usefulness)을 의미한다. 이 실험에서는 이렇게 3가지 인지 요소만을 측정하였는데 어떤 인지 요소가 실제로 어플리케이션을 사용할 때사용자에게 영향을 주는지 면밀하게 측정한다면, 컴퓨터 프로그램에 대해 사용자가 느끼는 심리적 요인을 깊이 탐구할 수 있어 하이라이트

생성 연구뿐만 아니라 컴퓨터 공학 및 HCI(Human-Computer Interaction) 분야의 연구에도 도움을 줄 수 있을 것이다.

앞선 실행을 통해 얻은 하이라이트의 효과와 하이라이트 생성 방법에 대한 결과를 바탕으로 어플리케이션을 실제로 제작해 보았다. 자연스러운 읽기 행동의 데이터를 수집한 것만을 가지고도 패턴을 추적하여 사용자가 의식하지 않고도 자동 하이라이트를 생성하는 실험을 수행하였다. 이 데이터를 활용하여 실제로 하이라이트를 제공하기 위해서는 몇 가지 과제가 더 남는다. 앞서 논의한 바와 같이 전체의 글 중 얼마의 비율로 하이라이트를 선택하느냐가 중요한 문제가될 것이다. 이러한 비율이 결정이 되면, 누적 시간 기준으로 상위 몇 퍼센트의 문장에 하이라이트를 제공하는 방법, 누적시간의 변화율을 고려하여 주변의 다른 문장에 비해 유독 높은 문장을 하이라이트로 표시하는 방법 등을 시도해 볼 수 있다.

사용자의 읽기 패턴를 가지고 어플리케이션을 실제로 제작하고 이를 사용자에게 직접 테스트 해본 것은 실용화를 전제로 한 것으로 의미있는 것이다. 그러나 어플리케이션의 프로토타이핑 과정이 정교하게 완결된 실험으로 정리될 수 없었던 것은 하이라이트 결과물의 평가가 면밀하게 이루어지지 못한 점을 지적할 수 있다. 평가자를 별도로 모집하여 평가한 문장의 중요도와 어플리케이션으로 수집한 노출 시간 사이에는 유의미한 상관관계를 발견하지 못하였다. 사람들이 더 많이 읽는 부분이 글에서 중요하다고 생각되는 부분과 직접적인 관련성이 없을 수도 있음을 시사한다. 다른 원인들을 생각해볼 수 있는데, 글에서 이해하기 어려운 부분, 글의 구조상 눈에 띄는 부분 등 다른 요소들과 읽기 패턴 과의 관계를 분석해 보는 것이 필요하다. 그리고 하이라이트가 어디에 표시되는 것이 좋은 방법인지 더 탐구해 볼 필요가 있을 것이다. 앞서 지적한대로 읽는 사람마다 혹은 읽는 목적에 따라 중요하다고 생각하는 부분이 다르다는 점을 생각해볼 수 있다. 이를 보완하여 정교한 평가 방법을 찾아야 하는 과제가 남아 있다.

# 7. 결론

온라인에서 글을 읽는 많은 사람들은 어떤 정보가 필요한 정보인지 판단하고 받아들이는 과정에서 정보의 과잉으로 인한 어려움를 겪고 있다. 스킴 리딩은 이미 많은 온라인 독자들이 자연스럽게 행하고 있는 읽기 습관이고, 이 과정이 효율적으로 진행되도록 돕는 도구가 있다면 시간 활용이나 정보 습득에서 유리한 방법을 제공할 것이다.

본 연구에서는 하이라이트의 효과를 실험과 생성 방법에 따른 심리적 차이를 알아보는 실험을 통해 하이라이트가 스킴 리딩 상황에서 내용을 효율적으로 파악하는데 도움을 줄 수 있음을 밝혔으며, 크라우드소싱 방법이 수용자에게 더 높은 신뢰를 주고 있음을 증명하였다. 이 결과를 응용하여 실제로 하이라이트를 생성하는 실용적 어플리케이션의 프로토타입도 작성해 보았다. 첫 번째 실험에서 하이라이트의 영향이 스킴 리딩 상황에서 효과를 발휘하는 것을 확인함과 동시에 아이트래커를 통해 측정한 시선의 움직임도 하이라이트의 영향을 받음을 알 수 있었다.

본 연구는 무엇보다 컴퓨터 활용에 인간적 요소(human factor)를 중점적으로 고려한 결과라는 데 의의가 있다. 기존에 하이라이트에 대한 연구는 주로 공학의 분야에서 다루어 왔고. 알고리즘의 성능 분석에 치중한 연구가 대부분이었다. 이 논문에서는 사회 과학 분야에 속하는 커뮤니케이션학에서 다루는 소스의 개념을 하이라이트 생성에 적용하여, 다른 수용자가 소스로 작용하였을 때 신뢰감을 준다는 것을 실험적으로 밝혔다. 또한 크라우드소싱이라는 비교적 새로운 방법을 응용하여 실용화 할 수 있는 가능성을 보여준다는 목적에도 부합하는 연구라고 할 수 있다. 본 연구의 목표는 궁극적으로는 크라우드소싱을 활용하여 문장의 중요한 부분을 추출해 주는 어플리케이션 제작에 도움이 될 기초 연구를 수행하는 것이다. 이를 위해 기본적으로 사용자의 입장이 고려한 기초 자료를 만들고자 하였으며, 하이라이트 효과와 추출 방법에 대해 소스에 의한 심리적 영향을 연구문제로 제시하는 등 공학과 사회과학적 요소를 함께 사용한 것에도 의의가 있다.

### 8. 참고 문헌

- 1. Ahmed, F., Borodin, Y., Puzis, Y., & Ramakrishnan, I. V. (2012, April). Why read if you can skim: towards enabling faster screen reading. In *Proceedings of the International Cross–Disciplinary Conference on Web Accessibility* (p. 39). ACM.
- 2. Alam, H., Hartono, R., Kumar, A., Rahman, A. F. R., Tarnikova, Y., & Wilcox, C. (2003, August). Web Page Summarization for Handheld Devices: A Natural Language Approach. In *ICDAR* (Vol. 3, p. 1153).
- 3. Alter, Alexandra. (2012). "Your E-Book Is Reading You." The Wall Street Journal 19.7.
- 4. Blignaut, P., & Beelders, T. (2012, March). TrackStick: a data quality measuring tool for Tobii eye trackers. In Proceedings of the Symposium on Eye Tracking Research and Applications (pp. 293–296). ACM.
- 5. Biedert, R., Dengel, A., Buscher, G., & Vartan, A. (2012, March). Reading and estimating gaze on smart phones. In Proceedings of the Symposium on Eye Tracking Research and Applications (pp. 385–388). ACM.
- 6. Bower, G. H. (1978). Experiments on story comprehension and recall\*. Discourse Processes, 1(3), 211–231.
- 7. Buyukkokten, O., Kaljuvee, O., Garcia-Molina, H., Paepcke, A., & Winograd, T. (2002). Efficient web browsing on handheld devices using page and form summarization. *ACM Transactions on Information Systems*, 20(1), 82–115.
- 8. Brown, B., McGregor, M., & Laurier, E. (2013, April). iPhone in vivo: video analysis of mobile device use. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1031–1040). ACM.

- 9. Chaffee, S. H. (1982). Mass media and interpersonal channels: Competitive, convergent, or complementary. Inter/media: Interpersonal communication in a media world, 57–77.
- 10. Chi, E. H., Gumbrecht, M., & Hong, L. (2007). Visual foraging of highlighted text: An eye-tracking study. In Human-Computer Interaction. HCI Intelligent Multimodal Interaction Environments (pp. 589–598). Springer Berlin Heidelberg.
- 11. Chi, E. H., Hong, L., Gumbrecht, M., & Card, S. K. (2005, January). ScentHighlights: highlighting conceptually-related sentences during reading. In Proceedings of the 10th international conference on Intelligent user interfaces (pp. 272–274). ACM.
- 12. Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS quarterly, 319–340.
- 13. Davis, J. R., & Huttenlocher, D. P. (1995). Shared annotation for cooperative learning. In *The first international conference on Computer support for collaborative learning* (pp. 84–88). L. Erlbaum Associates Inc..
- 14. Donley, K. K., & Reppen, R. (2001). Using corpus tools to highlight academic vocabulary in SCLT. *TESOL journal*, 10(2, 3), 7–12.
- 15. Duggan, G. B., and Payne, S. J. (2008). Knowledge in the head and on the web: using topic expertise to aid search. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '08), 39–48.
- 16. Duggan, G. B., and Payne, S. J. (2009). Text skimming: The process and effectiveness of foraging through text under time pressure, *Journal of Experimental Psychology*: Applied, Vol. 15(3), 228–242.
- 17. Dyson, M. C. and Haselgrove, M. (2001). The influence of reading speed and line length on the effectiveness of reading from

- screen, *International Journal of Human-Computer Studies*, Vol. 54(4), 585–612.
- 18. Fitchett, S., & Cockburn, A. (2009, November). Evaluating reading and analysis tasks on mobile devices: a case study of tilt and flick scrolling. In *Proceedings of the 21st Annual Conference of the Australian Computer–Human Interaction Special Interest Group: Design: Open 24/7* (pp. 225–232). ACM.
- 19. Flanagin, A. J., & Metzger, M. J. (2007). The role of site features, user attributes, and information verification behaviors on the perceived credibility of web-based information. New Media & Society, 9(2), 319–342.
- 20. Gillick, D., & Liu, Y. (2010, June). Non-expert evaluation of summarization systems is risky. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk* (pp. 148–151). Association for Computational Linguistics.
- 21. Hahn, U., & Mani, I. (2000). The challenges of automatic summarization. Computer, 33(11), 29–36.
- 22. Hovy, E., Marcus, M., Palmer, M., Ramshaw, L., & Weischedel, R. (2006, June). OntoNotes: the 90% solution. In *Proceedings of the human language technology conference of the NAACL, Companion Volume: Short Papers* (pp. 57–60). Association for Computational Linguistics.
- 23. Hu, Y., & Sundar, S. S. (2010). Effects of online health sources on credibility and behavioral intentions. Communication Research, 37(1), 105–132.
- 24. Hunt, R. R. (1995). The subtlety of distinctiveness: What von Restorff really did. *Psychonomic Bulletin & Review*, 2(1), 105–112.
- 25. Jacoby, J., Jaccard, J. J., Currim, I., Kuss, A., Ansari, A., & Troutman, T. (1994). Tracing the impact of item-by-item

- information accessing on uncertainty reduction. *Journal of Consumer Research*, 291–303.
- 26. Jing, H. (2000, April). Sentence reduction for automatic text summarization. In *Proceedings of the sixth conference on Applied natural language processing* (pp. 310–315). Association for Computational Linguistics.
- 27. Kaptelinin, V., Mäntylä, T., & Åström, J. (2002, April). Transient visual cues for scrolling: an empirical study. In CHI'02 extended abstracts on Human factors in computing systems (pp. 620–621). ACM.
- 28. Kol, S., & Schcolnik, M. (2000). Enhancing screen reading strategies. Calico journal, 18(1), 67–80.
- 29. Liu, Z. (2005). Reading behavior in the digital environment: Changes in reading behavior over the past ten years, *Journal of Documentation*, Vol. 61(6), 700–712.
- 30. Lloret, E., Plaza, L., & Aker, A. (2013). Analyzing the capabilities of crowdsourcing services for text summarization. *Language Resources and Evaluation*, 1–33.
- 31. Lombard, M., & Dittion, T. At the heart of it all: The concept of presence. *Journal of Computer Mediated Communication*, 3, 2 (1997).
- 32. Mani, I. (2001). Automatic summarization (Vol. 3). John Benjamins Publishing.
- 33. Moens, M. F., & Dumortier, J. (2000). Use of a text grammar for generating highlight abstracts of magazine articles. *Journal of documentation*, 56(5), 520–539.
- 34. Morris, A. H., Kasper, G. M., & Adams, D. A. (1992). The effects and limitations of automated text condensing on reading comprehension performance. Information Systems Research, 3(1), 17–35.

- 35. Morris, M., & Ogan, C. (1996). The Internet as Mass Medium. Journal of Communication, 46(1), 39–50.
- 36. Muter, P. and Maurutto, P. (1991). Reading and Skimming from Computer Screens and Books: The Paperless Office Revisited?, *Behaviour and Information technology* 10, 257–266.
- 37. Nagamatsu, T., Yamamoto, M., & Sato, H. (2010, April). MobiGaze: Development of a gaze interface for handheld mobile devices. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems* (pp. 3349–3354). ACM.
- 38. Öquist, G., & Lundin, K. (2007, December). Eye movement study of reading text on a mobile phone using paging, scrolling, leading, and RSVP. In *Proceedings of the 6th international conference on Mobile and ubiquitous multimedia* (pp. 176–183). ACM.
- 39. Poetz, M. K., & Schreier, M. (2012). The value of crowdsourcing: can users really compete with professionals in generating new product ideas?. *Journal of Product Innovation Management*, 29(2), 245–256.
- 40. Reader, W. R., & Payne, S. J. (2007). Allocating time across multiple texts: Sampling and satisficing. Human Computer Interaction, 22(3), 263–298.
- 41. Sabou, M., Bontcheva, K., & Scharl, A. (2012, September). Crowdsourcing research opportunities: lessons from natural language processing. In *Proceedings of the 12th International Conference on Knowledge Management and Knowledge Technologies* (p. 17). ACM.
- 42. Sarker, S., & Wells, J. D. (2003). Understanding mobile handheld device use and adoption. *Communications of the ACM*, 46(12), 35–40.
- 43. Shen, H., & Sun, C. (2004). Improving real-time collaboration with highlighting. Future Generation Computer Systems, 20(4), 605–625.

- 44. Sundar, S. S. and Nass, C. (2001), Conceptualizing sources in online news. Journal of Communication, 51, 52 72.
- 45. Tobii Technology. (2011). Accuracy and Precision Test Method for Remote Eyetrackers. Tobii Technology AB. Retrieved from http://www.tobii.com.
- 46. Walter, T. P., & Back, A. (2013, January). A Text Mining Approach to Evaluate Submissions to Crowdsourcing Contests. In *System Sciences (HICSS)*, 2013 46th Hawaii International Conference on (pp. 3109–3118). IEEE.

### **ABSTRACT**

# The Effectiveness of Highlight on Online Texts and its Generation Methods

Oh, JongHwan Department of Communication The Graduate School Seoul National University

The amount of accessible information online is rapidly expanding. As a result, online reading often requires skim reading rather than full—text reading. This increases the demand for supporting tools of improving effectiveness of reading as a strategy for understanding text in a relatively short time. This research hypothesizes highlight could be of help to understand better and to allocate attention selectively on important phrases. In the first experiment, participants were divided into two groups; skim readers and full—text readers. They were asked to read text under three different highlight conditions. At the same time, eye—tracking technology was used to record where they are looking at in the given text, and they also took a set of comprehension tests. Results indicate that highlight plays more

important role under skim reading than full-text reading condition.

Highlight captures longer fixation duration during skim reading. It

also leads readers to understand the text better. In the second

experiment, this research attempts to investigate psychological factors

that affect the perception of highlight by proposing a concept of

communication sources that would apply not only to media studies

but also to generation method of highlight. As a result of the

experiment, it is investigated that the credibility of the crowdsourcing

method soliciting contribution from a large group of people is higher

than that of computer algorithm method. Furthermore, as a

prototype, the present study suggests a novel method of collecting

reading patterns from readers through the use of mobile devices so

that this data can be utilized to automatically generate highlighted

text. Through preliminary experimentation, data can be collected

while participants read on their mobile devices. The data includes the

location of sentences on the screen, the position of scrolling fingers,

and the exposure time of each sentence on the screen. Explanations

for the results and implications for future research are discussed.

Keywords: highlight, reading behavior, source, skim reading, online

text

Student Number: 2012-22955

83