

# 온라인 뉴스 기사 헤드라인의 논조에 따른 댓글 양상

서기열, 권가진  
서울대학교, 융합과학부  
skyseo@snu.ac.kr, ggweon@snu.ac.kr

## Comments Complexion by Argument's Tone of Online News Headline

Ki-Yeal Seo, Gahgene Gweon  
Dept. of Transdisciplinary Studies, Seoul National University

### 요 약

온라인 뉴스 소비의 확산과 함께 댓글은 여론 형성에 큰 역할을 담당한다. 그러나 아직 댓글에 영향을 미치는 형식 요소에 대한 실증 데이터 기반의 연구는 미흡하다. 본 연구는 이의 시작으로 온라인 뉴스 기사 소비의 두 가지 중요 요소 즉, 헤드라인과 댓글의 관계에 대해 다루고자 한다. 이를 위해, 헤드라인의 논조 유무에 따른 댓글의 논쟁 활성화 정도 차를 확인하고자 댓글의 수와 길이를 분석하였다. ‘이세돌, 알파고 바둑대결’, ‘최저임금’, ‘북미회담’ 기사로 총 537 건의 헤드라인과 약 85 만개의 댓글을 수집하였다. 그 결과 논쟁 활성화 측면에서 논조가 있는 헤드라인일 때 댓글의 수가 많고 길이가 길어 논쟁이 더 활발한 것을 알 수 있었다. 또, 댓글의 논쟁 주제도 차이가 있어 헤드라인의 논조가 있는 경우에 의견이나 감정을 표출하는 토픽이 더 많았다. 본 연구는 실증 데이터를 통해, 헤드라인의 논조 유무가 댓글의 논쟁의 활성화 정도와 주제에 영향을 주는 요소임을 밝힘으로써 댓글 소비에 대한 새로운 관점을 제시하고, 헤드라인의 형식 요소의 연구의 중요성을 확인한 데 그 의의가 있다.

### 1. 서론

뉴스의 소비 행태가 모바일을 통한 온라인으로 변화하고, 기사의 선택과 의견 생성 과정에서 헤드라인과 댓글이 중요하게 작용한다. 본 연구는 온라인 뉴스 소비의 중요 요소인 헤드라인과 댓글의 관계에 대해 밝히고자 한다. 이를 위해 헤드라인의 논조 유무에 따른 댓글의 논쟁 활성화 정도와 논쟁 주제를 분석하였다. 그 결과, 헤드라인의 논조가 있는 경우 댓글의 논쟁이 더 활발하였으며, 주제 또한 감정적이며 의견 노출이 많다는 점을 확인하였다.

뉴스 소비가 온라인으로, 이 중에서도 모바일 채널의 비중이 높으며[1], ‘훑어 보기’, ‘제목만 읽기’ 경향이 나타나 헤드라인의 중요성을 높아진다[2,3]. 온라인 뉴스 이용 시, 독자는 댓글을 통해 적극적으로 의견을 표출하며, 노출된 댓글에 따라 여론이 달라진다는 연구도 있다[4]. 댓글을 악용한 범죄도 발생하는데, 이는 그만큼 댓글의 중요성을 보여주는 사례이다.

본 연구는 온라인 뉴스 소비에서 중요한 헤드라인의 표현 방식에 따른 댓글의 논쟁 활성화 정도를 확인하고자 한다. 헤드라인의 표현 방식은 논조 유무로 하였으며, 논쟁 활성화 정도는 헤드라인 별 댓글 수와 길이를 통해 판단하였다. 또, 토픽 모델링을 통해 작성된 댓글의 주제 파악하여 헤드라인 논조 유무에 따른 차이를 확인하였다. 데이터는 여론이 편중되지 않아 논쟁이 대등하게 발생되며, 각기 다른 영역에 속하고, 다른 기간에 작성된 뉴스들을 수집하였다.

연구 결과 헤드라인의 논조 유무에 따른 논쟁 활성화 정도와 논쟁 주제는 차이가 있었다. 특히, 논조가 있는 헤드라인에서 많고 긴 댓글이 작성되었다. 이는 의견이 드러나는 헤드라인이 독자도 하여금 더 활발하게 의견을 표출하게 한다는 것을 뜻한다. 토픽 모델링을 통한 논쟁의 주제도 논조 여부에 따라 다르게 나타났다. 헤드라인의 논조가 있을 때, 의견 표출이 더 많고 관점이 다양했으며, 감정 표현이 많았다.

### 2. 관련 연구

온라인 뉴스 기사의 헤드라인과 댓글과의 관계를 확인하기 위해 본 연구에서 다루고 있는 두 가지 주제 즉, 헤드라인과 댓글에 대한 연구들을 살펴보았다. 첫째는, 온라인 뉴스 기사의 소비 환경에서 헤드라인의 영향이고, 둘째는 댓글의 여론 형성에서의 역할이다. 이 두 과정을 통해 헤드라인은 댓글에, 댓글은 여론 형성에 영향을 미치고 있음을 알고, 올바른 여론 해석을 위해 둘의 영향 여부를 밝히는 본 연구의 중요성을 확인하려 한다. 마지막으로 댓글을 통한 여론 형성 과정에 영향을 미치는 요소를 조사하여, 이후 분석 과정의 기준으로 삼고자 한다.

#### 2.1. 온라인 기사의 헤드라인의 영향력

뉴스 소비에서 가장 먼저 접하는 헤드라인은 기사를 선택 시의 중요 판단 기준이다[3]. 독자들은 중립적 헤드라인이 본문의 전체적 방향을 가장 잘 요약 및 함축하고 있다고 판단하며, 긍정이나 부정적인 제목으로 보도되는 이슈는 더 중요하게 지각하는 경향

이 있다[5]. 헤드라인의 감성이 편향될수록 댓글의 수가 많은 것이 발견되기도 하였으며[5], 궁금증을 유발하는 헤드라인의 조회 수가 더 증가하기도 한다[6].

그러나 이들은 실제 데이터를 대상으로 하지 않거나[3], 실제 데이터를 기반한 경우에는 전체 뉴스를 대상으로 한다[5,6]. 뉴스는 주로 사건이나 사고의 부정적 내용이 많아 전체 뉴스를 대상으로 할 때, 이 부정적 영향을 배제할 수 없다. 따라서 본 연구에서는 기사가 다루는 내용이 아닌 표현 방식에 대한 영향을 파악하기 위해 찬, 반의 논란이 대등하게 발생할 수 있는 주제를 대상으로 삼고자 한다.

### 2.2. 여론 형성 과정에서 댓글의 역할

여론 형성에 영향을 미치는 댓글의 역할이 중요해짐에 따라 댓글로 인한 독자의 의견 변화에 대한 연구들이 진행되었다. 독자는 댓글에 영향을 받아 온라인 뿐 아니라 사회 일반적 여론도 추론하는 경향을 보이며, 기사를 보다 부정적으로 평가한다. 또, 댓글은 기사 자체의 영향력을 감소시키는 조절변인으로 작용하기도 한다[7]. 댓글을 통해 여론이 형성될 수 있으며, 내용의 관여 정도나 사안에 대한 사전 입장(찬성 또는 반대)에 따라 그 변화의 폭이 달라짐을 실험을 통해 확인한 연구도 있다[7].

그러나 이 연구들도 한정된 주제와 소수 자료의 실험 기반으로, 다양한 정보를 통한 실시간 논쟁이 지속되는 온라인의 특성을 배제한 한계점이 있다. 따라서 본 연구는 실제 데이터를 분석하고자 한다.

### 2.3. 독자 행동의 영향 요소

감성 편향의 뉴스에 댓글 수가 많다는 Julio Reis et al.(2015)의 실험 결과[8] 외에도, 댓글 작성 경험자 5명의 사전 인터뷰에서 독자들은 중요 사안일수록 지지 의사를 표현하고자 하였다. 기사의 헤드라인이 사안의 중요도나 시급성을 노출한 경우, 독자의 댓글 참여를 촉진시킬 수 있는 것이다.

가설 1. 헤드라인의 논조가 있는 경우 없는 경우에 비해, 댓글의 논쟁은 더욱 활성화 될 것이다.

William J. McGuire[9]에 의하면 감성적 메시지는 읽은 이의 감정에 호소하므로 인식이나 행동 변화를 이끌어 내는 데 효과적이다. 이를 온라인 뉴스 소비 상황에 대입하면, 감성적 표현의 헤드라인의 경우에 작성되는 댓글의 변화를 일으킬 있다. 또, 어떤 방식으로든 자극을 받은 독자가 작성한 댓글은 그 다루는 내용에 있어서도 차이가 있을 것이라는 가설을 수립할 수 있다.

가설 2. 헤드라인의 논조가 있는 경우 없는 경우에 비해, 감성적인 논쟁 주제가 더 많을 것이다.

## 3. 연구방법

본 연구는 온라인 뉴스 소비와 여론 형성 과정의 실제 데이터를 분석 대상으로 하였다. 우선, 국내에서 가장 많이 이용되는 포털 사이트를 통해 뉴스 기사를 수집하였다. 수집된 기사의 헤드라인에 대한 논

조 유무를 파악하고 각각 댓글의 수와 길이, 논쟁 주제를 비교하였다. 이를 통해 뉴스 기사의 헤드라인 논조가 댓글에 미치는 영향을 확인하고자 한다.

### 3.1. 데이터 수집

데이터는 2016년 3월 1개월 동안의 ‘이세돌, 알파고 바둑 대결’, 2017년 5월부터 1년 동안의 ‘최저임금’, 2018년 4월부터 4개월 동안의 ‘복미회담’ 기사이다. 이들은 뉴스, 미디어 분야 국내 이용 1위 매체인 네이버를 통해 수집하였다[10]. 메인 페이지에 실린 기사를 기준으로 하였는데, 그렇지 않은 경우 노출 기회가 낮아 조회 댓글 수가 적어 의미 있는 분석이 어렵기 때문이다. 연구의 객관화를 위해 기사의 분야, 시기, 이슈 기간이 다른 주제를 선정하였다. 논쟁의 치우침을 방지하고자, 댓글이 찬성이나 반대로 편중되지 않는 총 3개 주제를 확보하였다. 중복 수집된 기사를 제외하고 총 537건의 뉴스 헤드라인과 약 85만개의 댓글을 분석 대상으로 삼았다.

### 3.2. 프로세스

데이터 수집 후, 헤드라인의 논조 유무에 대한 기준을 마련하고, 검증한 후에 코딩 하였다. 논쟁 활성화 정도는 각 헤드라인에 작성된 댓글의 수와 길이로 판단하였으며, 논쟁의 내용으로 댓글의 주제를 파악하였다. 이와 같이 헤드라인 논조 유무에 따른 댓글의 수와 길이, 주제를 분석함으로써 헤드라인의 논조 유무가 댓글에 미치는 영향을 확인하고자 하였다.

#### 3.2.1. 헤드라인의 논조 판단

<표 1> 헤드라인의 논조 유무 구분 기준

논조 없음	개인의 해석, 판단 포함
	감정 상태 표현
	강조, 미사여구 사용
	관용적 표현 사용
논조 있음	수치가 없는 데이터, 예측
	해석, 꾸밈, 과장 표현이 없음
	사건, 현상 기술
	수치를 포함한 데이터

논조에 대한 평가자의 주관적 해석 최소화를 위해 우선 ‘최저임금’ 주제에 대해 기준을 마련하고, 1명의 연구자와 3명의 대학원생이 각 50개의 논조 여부를 판단하였다<표 1>. 다시 연구자 2명이 106건의 헤드라인을 코딩하여 최종 93.8%의 일치와 0.846의 Cohen’s Kappa를 도출하였다. 이후 주 연구자 1명이 나머지 헤드라인의 논조 유무를 코딩하였다.

<표 2> 수집된 헤드라인과 댓글 수

주제	헤드라인 (537건)		댓글 (852,027개)
	논조 없음	논조 있음	
이세돌, 알파고 바둑대결	23건	93건	72,467개
최저임금	74건	183건	590,988개
복미회담	71건	93건	187,572개

헤드라인의 논조 코딩 결과, ‘이세돌, 알파고 바둑대결’, ‘최저임금’, ‘북미회담’의 논조가 없는 헤드라인은 각각 23 건, 74 건, 71 건이며, 논조가 있는 헤드라인은 각각 93 건, 183 건, 93 건으로 세 주제 모두 논조가 있는 헤드라인이 더 많았다<표 2>.

**3.2.2. 논쟁 활성화 및 주제 측정 방법**

논쟁 활성화 정도는 논쟁에 참여하는 댓글의 수와 길이를 통해 판단하였다. 댓글의 수는 논쟁에 참여한 독자들의 규모를 뜻하며, 댓글의 길이는 논쟁의 양을 나타낸다고 볼 수 있기 때문이다.

논쟁의 주제를 파악하기 위해 토픽 모델링을 사용하여 댓글의 내용을 확인하였다. 이는 대량의 텍스트 자료 분석 방법으로, 빈도나 키워드 분석보다 트랜트 분석에 더 효과적이다. 본 연구는 알고리즘이 단순하고 데이터 차원 축소에 용이하며 일관성 있는 토픽 추출이 가능한 LDA(Latent Dirichlet Allocation) 기법을 활용하였다. 한글 처리는 속도의 장점이 있는 트위터 분석기를 사용하고, LDA는 명사 기준으로 하였다.

**4. 연구 결과**

3개 주제의 헤드라인의 논조 유무에 따른 댓글 분석 결과, 논쟁의 활성화와 논쟁 주제 모두 차이가 있는 것을 확인하였다. 특히, 헤드라인의 논조가 있는

경우에 댓글의 수가 많고 길이도 길게 나타났다. 내용 측면에 서도, 모든 주제에서 헤드라인의 논조가 없는 경우에 주제 자체에 대한 토픽이 더 많이 추출되었다. 헤드라인이 논조를 담고 있을 때, 보다 다양한 관점의 토픽이 추출되어 비난, 우려, 기대 등의 의견이나 감정이 더 표출되고 있음을 확인하였다.

**4.1. 헤드라인 논조 유무에 따른 댓글의 논쟁 활성화**

논쟁 활성화 정도 판단을 위해 댓글 수와 길이를 측정하였다. t-test 결과, ‘이세돌, 알파고 바둑대결’, ‘최저임금’, ‘북미회담’ 각각 논조 유무에 따른 댓글 수의 p-value는 ‘0.048, 0.000, 0.007’, 댓글 길이의 p-value는 0.002, 0.011, 0.033으로 0.05 미만으로 분석되어 논조 유무에 따른 논쟁 활성화 정도는 차이가 있음이 밝혀졌다. 특히, 헤드라인의 논조가 있는 경우에 댓글의 수가 많고 길이도 길어 그렇지 않은 헤드라인의에 비해 논쟁이 활발한 것을 확인함으로써 연구 가설 1을 지지하는 것을 알 수 있다<표 3>. 이는 William J. McGuire의 이론[9]과 맥락을 같이하여, 온라인 뉴스 기사에서도 헤드라인에 감정이나 의견이 드러날 경우에 댓글 작성자들도 자기의 의견을 더 구체적이고 자세하게 표현하는 것을 실증 데이터를 통해 증명한 것이다.

<표 3> 주제별 논쟁 활성화 분석 결과

헤드라인 논조 유무	이세돌, 알파고 바둑대결		최저임금		북미회담	
	댓글 수 (평균)	댓글 길이 (평균)	댓글 수 (평균)	댓글 길이 (평균)	댓글 수 (평균)	댓글 길이 (평균)
없음	316 개	51.7 자	1353 개	2607 개	683 개	1496 개
있음	701 개	61.8 자	72.6 자	76.9 자	66.2 자	71.6 자
P-value	0.0481299	0.00200491	0.00021878	0.01053597	0.00670846	0.03310573

**4.2. 헤드라인 논조 유무에 따른 논쟁 주제**

댓글의 논쟁 주제를 파악하기 위해서는 토픽 모델링을 진행하였다. 그 결과, 주제별 헤드라인의 논조 유무에 따라 각 5개 단어로 구성된 5개씩의 토픽이 도출되었다. 구성된 단어들의 해석을 통해 각 토픽을 선정하였으며, 논조 유무에 따른 차이를 분석하였다.

**4.2.1. ‘이세돌, 알파고 바둑대결’ 논쟁 주제**

‘이세돌, 알파고 바둑대결’ 기사의 주제 자체의 토픽이 논조가 없는 헤드라인에 4개, 논조가 있는 헤드라인은 2개의 나타났다<표 4>. 헤드라인의 논조가 있을 때, ‘고가’라는 단어가 포함된 토픽도 있어, 판단이나 해석의 포함되었음을 알 수 있다. 정부 사업에 대한 비난 토픽은 논조 유무 별 각각 1개씩이었다. 이 외, 헤드라인의 논조가 있는 경우에 경기 결과에 대한 예측과 참가자를 응원하는 토픽도 나타났다. 이는 헤드라인의 논조가 있을 때 의견이 드러나는 댓글이 더 많이 작성되었음을 뜻한다.

**4.2.2. ‘최저임금’ 논쟁 주제**

‘최저임금’ 기사의 댓글은 주제 자체에 대한 토픽은 헤드라인의 논조가 없을 때 4개, 논조가 있을

때 2개로 나타났다. 논조가 없는 헤드라인일 때는 정책이나 국가적 차원의 넓은 개념의 단어들이 사용되었다. 반면, 논조가 있는 헤드라인에는 사람이나 급여 등의 내용이 나타나 개인 관점의 논쟁이 일어났음을 알 수 있다. 우려에 대한 토픽이 헤드라인 유무 별 각 1건으로 나타났고, 논조가 있는 헤드라인에 부정 감정의 비난 토픽이 2건 발생했다.

**4.2.3. ‘북미회담’ 논쟁 주제**

‘북미회담’ 기사의 댓글은 주제 자체에 대한 토픽이 헤드라인의 논조가 없을 때 3개, 논조가 있을 때 2개로 논조가 없을 때가 많다. 비난 토픽은 각각 1개, 2개로 나타났는데, 논조가 있을 때 ‘빨갱이, 새끼’ 등의 욕설이 포함되어 부정 감정표현의 정도가 강한 것을 알 수 있다. 논조가 없는 헤드라인에는 한반도, 국민 등의 보다 추상적인 단어를 통해 논쟁이 발생하고 있음을 알 수 있다. 마지막으로 논조가 없는 헤드라인에는 정부나 대통령이 활동에 대한 응원을 통해 적극적으로 구체적인 대상을 지지했다.

<표 4> 논쟁 주제 분석 결과

	토픽	논조가 없는 헤드라인	논조가 있는 헤드라인
		도출 단어	도출 단어
이세돌 알파고 바둑대결	주제 자체	이세돌, 바둑, 사람, 인간, 기사	알파고, 이세돌, 수, 바둑, 고가
		사고, 로봇, 알파고, 인공지능, 때	-
		구글, 바둑, 이세돌, 프로, 알파고	-
		인공지능, 사람, 차원, 수준, 게임	인간, 기계, 사람, 바둑, 컴퓨터
	비난	복지, 놈, 투자, 알파고, 소리	돈, 로봇, 정부, 놈, 한국
	예측	-	인공지능, 세돌이, 인류, 인간, 미래
응원	-	이세돌, 정말, 진짜, 바둑, 화이팅	
최저임금	주제 자체	대통령, 문재인, 정부, 정책, 공약	알바, 자영업자, 돈, 사장, 장사
		한국, 법, 수당, 미국, 일본	만원, 최저임금, 때, 일, 시간
		돈, 사람, 노동자, 일, 비정규직	-
		만원, 시급, 시간, 최저, 알바	-
	우려	최저임금, 물가, 인상, 문제, 임금	물가, 인상, 시급, 최저임금, 지금
	비난	-	문재인, 문재인, 나라, 정신, 정부
-		세금, 돈, 공무원, 월급, 나라	
북미회담	주제 자체	북한, 미국, 철수, 핵, 평화회담	미국, 북한, 핵, 트럼프, 종전
		트럼프, 김정은, 문재인, 회담, 대통령	대통령, 트럼프, 경제, 정부, 국민
		대한민국, 경제, 국민, 국가, 나라	-
	비난	문재양, 김정은, 쇼, 트럼프, 개	문재양, 빨갱이, 개, 트럼프, 새끼
		-	북한, 김정은, 정권, 쇼, 북괴
	기대	평화, 국민, 한반도, 전쟁, 회담	
응원		문재인, 응원, 지지, 정부, 대통령	

5. 토의 및 시사점

본 연구는 온라인 뉴스 기사 헤드라인의 논조에 따라 댓글의 논쟁이 어떻게 달라지는지를 파악하는 데 목적이 있다. 각기 다른 3 개 주제의 헤드라인과 댓글을 분석한 결과, 헤드라인의 논조가 있는 경우에 댓글의 수가 더 많고 길어도 길어 논쟁이 활발한 것을 확인하였다. 토픽 모델링을 통해 파악한 논쟁 주제도 헤드라인의 논조가 있을 때 더 다양한 감정표현과 의견이 나타났다. 이상의 결과로 헤드라인의 논조에 따라 작성된 댓글이 차이가 있음이 검증되었다. 따라서 댓글의 여론 파악에 있어, 헤드라인의 논조에 따른 댓글 논쟁의 정도나 논쟁 내용이 달라질 수 있음을 감안해야 한다는 시사점을 제시한다.

연구에서는 댓글에 영향을 미치는 헤드라인의 표현 형식을 논조의 유무로 삼았으며, 논쟁 활성화 정도는 댓글의 수와 길이를, 논쟁 주제 파악은 토픽 모델링을 사용하였다. 향후에는 헤드라인의 표현 양식의 세분 기준(어법, 표현 방식 등), 게시 시간 등 다양한 요소의 영향도 확인할 수 있을 것이다. 또, 논쟁 활성화 정도를 파악하는데 있어서도 댓글이 수와 길이 외에 찬/반 비율, 감성 분석 등으로 그 분석 범위를 넓힐 수 있다. 댓글은 맞춤법이 틀리거나 줄임말, 비속어 등이 사용되므로 표준어 기준의 분석법이 한계가 있어 이의 처리 방식도 함께 고려되어야 한다.

이처럼 다양한 연구가 진행되면, 기사의 형식과 댓글의

관계 규명이 가능할 것이다. 이를 통해 올바른 온라인 토론 문화가 정착되고, 여론에 대한 일부 요소의 일방적 영향력을 줄일 수 있을 것으로 기대한다.

참고문헌

- [1] 한국언론진흥재단, 언론 수용자 의식조사, 2017
- [2] 조수선, 온라인 신문 기사의 제목과 개요, 2005, 효과 한국언론학보, 49(2), 5-32.
- [3] DMC 미디어, 온라인 뉴스 콘텐츠 소비 행태 조사, 2013
- [4] 문광수 외, 베스트 댓글의 방향성이 일반 댓글의 동조 효과에 미치는 영향, 2013, 한국콘텐츠학회논문지, 13(12), 201-211.
- [5] 유홍식, 기사 제목과 예시가 수용자의 뉴스 가치 평가와 이슈 지각에 미치는 영향, 2009, 한국언론학보, 53(5), 176-198.
- [6] Ori Tenenboim et al., What prompts users to click and comment: A longitudinal study of online news, 2013, Journalism
- [7] 강제원 외, 인터넷 뉴스기사에 달린 댓글의 효과 연구: 이슈 관여도와 의견의 일치성 여부를 고려해서, 2012
- [8] Julio Reis et al., Breaking the News: First Impressions Matter on Online News, 2015, AAAI
- [9] William J. McGuire, "The Nature of Attitudes and Attitude Change" in G. Lindzey and E. Aronson, eds., Handbook of Social Psychology, 2nd ed., Mass.: Addison-Wesley, 1969, 200-203.
- [10] <http://internettrend.co.kr/trendForward.tsp>