

인공지능 속성에 대한 고객 태도 변화: AI 스피커 고객 리뷰 분석을 통한 탐색적 연구¹

Customer Attitude to Artificial Intelligence Features: Exploratory Study on Customer Reviews of AI Speakers

이 홍 주 (Hong Joo Lee) 가톨릭대학교 경영학부²

ABSTRACT

AI speakers which are wireless speakers with smart features have released from many manufacturers and adopted by many customers. Though smart features including voice recognition, controlling connected devices and providing information are embedded in many mobile phones, AI speakers are sitting in home and has a role of the central entertainment and information provider.

Many surveys have investigated the important factors to adopt AI speakers and influencing factors on satisfaction. Though most surveys on AI speakers are cross sectional, we can track customer attitude toward AI speakers longitudinally by analyzing customer reviews on AI speakers. However, there is not much research on the change of customer attitude toward AI speaker. Therefore, in this study, we try to grasp how the attitude of AI speaker changes with time by applying text mining-based analysis.

We collected the customer reviews on Amazon Echo which has the highest share of AI speakers in the global market from Amazon.com. Since Amazon Echo already have two generations, we can analyze the characteristics of reviews and compare the attitude according to the adoption time. We identified all sub topics of customer reviews and specified the topics for smart features. And we analyzed how the share of topics varied with time and analyzed diverse meta data for comparisons.

The proportions of the topics for general satisfaction and satisfaction on music were increasing while the proportions of the topics for music quality, speakers and wireless speakers were decreasing over time. Though the proportions of topics for smart features were similar according to time, the share of the topics in positive reviews and importance metrics were reduced in the 2nd generation of Amazon Echo. Even though smart features were mentioned similarly in the reviews, the influential effect on satisfaction were reduced over time and especially in the 2nd generation of Amazon Echo.

Keywords: Customer review, Topic modeling, AI speaker, Artificial intelligence, Topic distribution

1) 본 연구는 2017년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구이며(NRF-2017S1A3A2066740), 2019년도 가톨릭대학교 교비연구비의 지원으로 이루어졌음.

논문접수일: 2019년 1월 22일; 1차 수정: 2019년 2월 28일; 게재확정일: 2019년 5월 15일

2) 제 1저자(hongjoo@catolic.ac.kr)

1. 서론

스마트 스피커는 스마트폰에 탑재되는 인공지능 비서와 같은 기능이 포함되어 사람이 목소리로 다양한 명령을 내릴 수 있으며, 연결된 기기의 제어와 연계된 다양한 오프라인 서비스를 이용할 수 있다. 인공지능 기술 기반의 서비스와 연계된 제품 중에서 실제로 많이 보급되고 있는 것이 AI스피커이다(Kabir 2018). 기존의 Wi-Fi나 Bluetooth 기반의 인터넷 연결성을 가진 포터블 스피커에 인공지능 속성이 포함되어 있다. 대표적인 인공지능 속성은 음성 인식, 정보제공, 연결된 앱의 조절 및 제어, 연결된 기기의 조절 및 제어 등이다. 초기에는 음성 인식을 통한 정보제공에 초점이 맞추어져 있었으나, 제품이 성숙하면서 다른 앱 및 기기와의 연동성이 증대하고 있다.

국내에는 2018년 연말까지 대략 300만대가 보급될 것으로 예측되며, 이는 전세계에 보급된 AI 스피커 중 3% 정도에 해당한다(연합뉴스 2018).

국내 보급 속도는 빨라지고 있지만 AI 스피커를 직접 구매하는 비율보다는 다른 서비스의 패키지 상품이나 판촉 캠페인으로 인해 구매하는 경우가 많으며 이용만족률은 49%에 머무르고 있다(컨슈머인사이트 2018). 설문조사 패널을 대상으로 한 이메일 설문을 통해 밝혀진 주요 활용 기능은 음악 재생, 날씨 정보 확인, 블루투스 스피커 기능이며, 불만족 요인은 낮은 음성인식률과 실망스러운 대화수준, 부정확한 정보제공, 비용, 연결 기능 제어 부족 순이었다(컨슈머인사이트 2018).

텍스트 마이닝 기반의 고객 리뷰 분석을 통해 제품의 기능이나 속성에 대한 고객 태도, 감성 분석 및 제품 평가에 미치는 영향을 파악하는 방안이 많이 제시되어 왔다(박명석 등 2018; Cao et al. 2011; Humphreys and Wang 2017). 제품 리뷰에서 속성이나 기능에 해당하는 단어를 추출하여 감성분석을 수행하거나 평가에 미치는 영향을 분석하여 왔다(이홍주 2018b; Cao

et al. 2011; Ghose et al. 2012). 리뷰에서 토픽을 추출한 후 평가에 미치는 영향을 분석하거나(Büschken and Allenby 2016), 유사한 제품들의 시장 경쟁 상황을 가시화하는 방안도 제시되었다(Chen et al. 2015; Netzer et al. 2012). AI 스피커 사용자들의 리뷰를 텍스트 마이닝을 통해 분석하는 연구도 수행되었으며, 주요하게 언급되는 속성 파악, 속성에 대한 감성분석 및 인공지능 속성이 제품 만족에 미치는 영향을 분석하였다(이홍주 2018a).

텍스트 마이닝을 통해 고객의 리뷰를 분석하는 경우에는 리뷰 작성 시점에 대한 메타 데이터가 존재하므로 통시적인 분석이 가능하다. 하지만 AI 스피커의 수용에 대해서 시기의 변화에 따른 고객 태도 변화에 대한 연구는 많지 않다. AI 스피커는 기존에 존재하던 인터넷 연결이 가능한 포터블 스피커에 인공지능 속성이 추가된 형태로 볼 수 있다. 인공지능 기능을 갖춘 다양한 제품들이 출시되고 있지만, 이러한 속성들이 고객들에게 어떻게 인식되고 있는 지를 시간에 따라 분석한 연구는 부족하다. 따라서, 본 연구에서는 AI 스피커의 인공지능 속성에 대한 태도가 시간이 흐름에 따라 어떻게 변화하는 지를 텍스트 마이닝 기반 분석을 통해 탐색적으로 파악하고자 한다.

글로벌 시장에서 AI 스피커에서 가장 높은 점유율을 가지고 있는 아마존 에코(Amazon Echo)에 대한 리뷰를 수집하였다. 시기에 따른 리뷰의 특성을 분석하였으며, 인공지능 속성에 대한 토픽을 파악하여 토픽의 점유율이 시기에 따라 어떤 변화를 보였는 지를 분석하였다.

2장에 관련연구를 정리하였으며, 3장에서 본 논문이 활용한 자료에 대해 기술하였다. 4장에서 시기에 따른 고객 태도의 변화를 탐색하였으며, 5장에서 본 연구의 결론을 제시하였다.

2. 관련연구

사용자 생성 콘텐츠(UGC)와 고객 리뷰의 텍스트를 분석하는 연구들이 많이 수행되어 왔으며(김정훈 등 2015), 마케팅 연구에서도 다양한 목적으로 활용되었다. Büschken and Allenby(2016)은 텍스트 기반 연구가 활용된 영역을 1) 감성이 판매에 미치는 영향, 2) 고객 리뷰의 상대적 중요도, 3) 고객 리뷰의 감성적인 내용의 변화에 따른 구매전환율의 변화, 4) 판매 예측(Ghose et al. 2012), 5) 속성에 대한 선호도 파악(Archak et al. 2011; Lee and Bradlow 2011), 6) 시장 구조 파악(Lee and Bradlow 2011; Netzer et al. 2012)으로 구분하였다. 위의 연구들은 주로 텍스트 데이터에서 정보를 추출하고, 추출된 정보들을 텍스트 데이터의 다른 속성과 연관관계를 측정하여 수행되었다(Büschken and Allenby 2016). 또한, 텍스트 마이닝을 통한 고객 연구는 접근방법에 따라 사전 기반, 분류 기반, 토픽 파악으로 분류할 수 있다(Humphreys and Wang 2017). 사전 기반 접근 방안은 연구자가 분석에 사용하고자 하는 단어들의 사전을 미리 정의하고 이에 기반하거나 규칙을 미리 정의하여 문서들간의 비교나 연관분석, 예측을 수행한다. 텍스트의 감성을 파악하는 많은 연구들이 이러한 접근 방안을 사용하여 왔으며(Mudambi and Schuff 2010), LIWC, AFINN 등의 사전을 활용하여 감성을 측정하여 왔다. LIWC와 AFINN 등의 사전이 일반적인 단어의 감성을 표현하고 있기에 도메인에 특화된 사전을 구축하여 감성을 측정하는 방안들도 활용되어 왔다(Lee et al. 2017).

분류기반과 토픽 파악은 사전 기반 접근 방안과 다르게 정의된 규칙이나 사전없이 데이터로부터 학습하여 분류규칙을 생성하거나 토픽을 파악하는 접근방안이다. 분류기반 방안은 이미 분류가 된 학습 데이터를 통해 적합한 분류자(classifier)를 형성하고 이를 통해 학습에 활용되지 않은 데이터를 분류한다. 리뷰의 문장이

긍정인지 부정인지를 이와 같이 학습한 모형에 기반하여 분류할 수 있다. 기계 학습에 기반한 분류자를 활용하는 경우는 아니지만 문서에 빈번하게 함께 등장하는 단어들을 네트워크로 표현하여 다차원 척도로 시각화하는 방안도 이러한 접근방안에 해당된다(Netzer et al. 2012). Lee and Bradlow(2011)은 고객 리뷰로부터 제품 속성을 추출하여 대응분석을 통해 디지털 카메라 브랜드들의 시장 구조를 시각화하였다.

토픽 파악은 전체 텍스트에 등장하는 여러개의 하위 토픽을 추출하는 방안이며, Latent Semantic Analysis(LSA)나 Latent Dirichlet Allocation(LDA)가 사용되어 왔다(Blei 2012). Palese and Usai(2018)은 weakly supervised LDA 방안을 활용하여 고객 리뷰에서 SERVQUAL 차원의 토픽을 추출하였으며, 각 토픽이 평점에 미치는 영향을 회귀분석을 통해 파악하였다. 트위터 데이터를 활용하여 의료 서비스에 대한 생각을 토픽 모델링을 활용하여 SERVQUAL 차원으로 분류하고 각 차원에 대한 감성을 측정하는 방안이 제시되었다(Lee et al. 2017).

3. 자료

아마존 에코는 인공지능 서비스가 포함된 스마트 스피커로 최초로 출시된 제품이며, 2015년 1세대 제품이 출시되었고 2017년에 2세대 제품이 출시되었다. 기본적으로 블루투스를 통해 인터넷에 연결된 스피커이며, 제공되는 추가 기능은 음성 인식, 대화, 음성을 통한 제어(연결된 기기, 전화, 음악 스트리밍, 라디오, 스피커, 알람, 타이머, 리스트 관리 등), 정보제공(날씨, 시간, 일정 등) 이다. 운영 소프트웨어는 지속적으로 갱신되고 있으며, 소프트웨어 업그레이드를 통해 다양한 기능을 추가하고 있다. 아마존 에코의 성공으로 에코 닷, 에코 룩, 에코 쇼, 에코 스팟 등의 다양한 연관기기들이 출시되

<Table 1> Customer Review Data

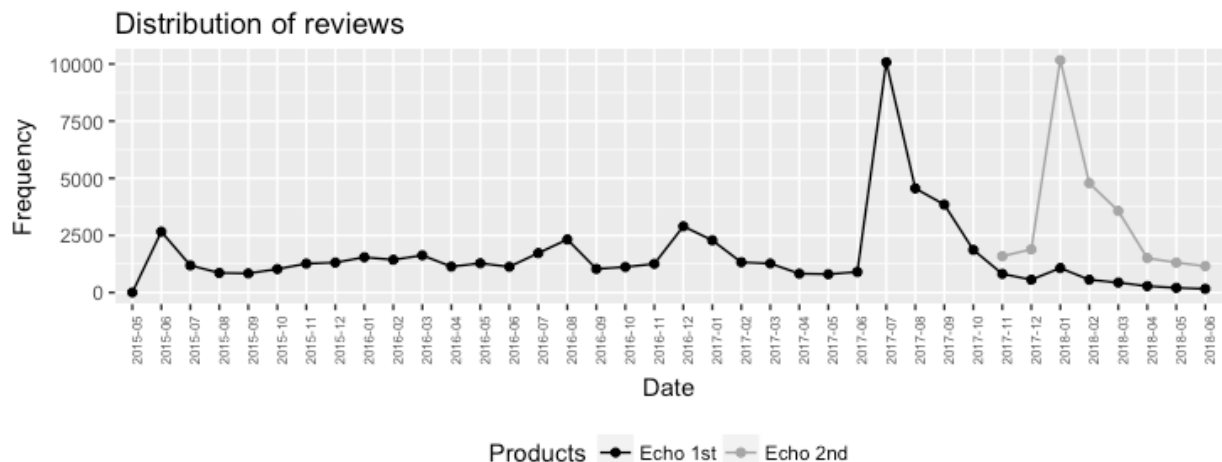
	Echo 1st Generation	Echo 2nd Generation
Number of reviews	59,490	25,970
Date of reviews	May 18, 2015 ~ June 27, 2018	Nov. 2, 2017 ~ June 26, 2018
Average number of stars	4.340	4.439
Standard deviation of stars	1.179	1.074
Average number of characters of reviews	225.283	147.636

었다. 제품 수용 시기에 따른 인공지능 속성에 대한 태도를 탐색하기 위해서 1세대와 2세대 에코에 대한 고객 리뷰를 Amazon.com에서 수집하였으며, 수집된 리뷰에 대한 기술적인 설명은 <Table 1> 과 같다. 2세대 에코가 출시된 지 8개월 정도 지난 시점에 리뷰가 수집되었기에 1세대 에코에 비해 고객 리뷰 수가 적지만, 2만 개 이상의 리뷰를 수집하였기에 분석에 활용하였다.

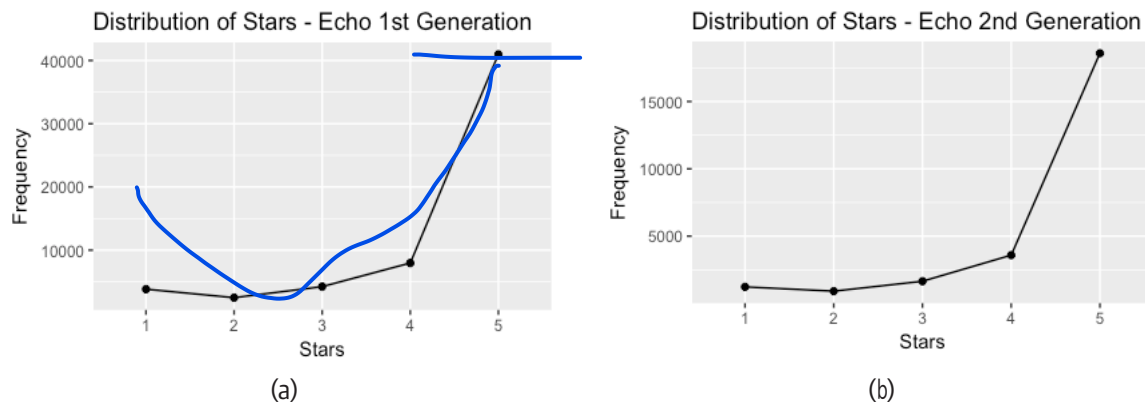
<Figure 1>은 고객 리뷰의 월별 분포를 보여준다. 1세대 에코 출시 이후에 꾸준히 리뷰가 작성되고 있으며, 1세대 에코 제품에 대한 리뷰가 가장 많이 작성된 달은 2017년 7월이었다. 2017년 7월부터 아마존 프라임 회원을 대상으로 에코의 가격을 \$179에서 \$89.99로 인하하였고, 2017년 8월부터는 모든 고객을 대상으로 \$99로 가격을 인하하였다. 이러한 가격 인하가 판매에

긍정적인 영향을 미친 것으로 볼 수 있다. 대략 월별로 1천에서 2천개 정도의 리뷰가 작성되어왔지만 2017년 7월에는 1만개 정도의 리뷰가 작성되었다. 1세대 에코의 출시 이후 리뷰의 작성 추세는 폭발적이기 보다는 꾸준한 편이지만, 1세대와 달리 2세대 에코에 대한 리뷰는 출시 이후에 폭발적으로 작성되었으며 2018년 1월에 가장 많이 작성되었다. 2017년 7월에 1세대 에코에 대한 폭발적인 리뷰 작성 패턴과 매우 유사한 모습을 보여주고 있다.

<Figure 2>는 1세대와 2세대 에코에 대한 고객 리뷰 평점의 분포이다. 고객 리뷰의 평점 분포는 일반적으로 5점 만점의 경우 5점을 준 리뷰의 빈도가 가장 많고 4, 3, 2점으로 갈 수록 빈도가 줄다가 최저점인 1점을 준 리뷰가 2점이나 3점을 준 리뷰보다 많은 J자 형태를 띄



<Figure 1> Distribution of reviews



<Figure 2> Distribution of stars

는 것으로 알려져 있으며(Hu et al. 2009), 수집한 에코에 대한 리뷰도 1점의 빈도가 2, 3점에 비해 매우 높지는 않지만 J자 형태를 띄었다.

1세대 에코와 2세대 에코에 대한 고객 리뷰의 평균 평점을 t 테스트를 통해 비교하였더니, t 통계량이 -12.019이고 p 값이 0.05 유의 수준에 비해 매우 작았다. 또한 95% 신뢰구간이 -0.115에서 -0.082이기에 두 평균 값이 같다고 볼 수 없다. 2세대 에코에 대한 평점 평균이 더 높다고 할 수 있다. 소프트웨어는 동일하지만 하드웨어가 개선된 제품이기 때문으로 판단된다.

1세대와 2세대 에코에 대한 리뷰에 빈번하게 등장하는 30개 단어는 매우 유사하며 빈도수에 따른 순위가 상이할 뿐이다. 1세대 에코 리뷰에는 빈번하게 등장하지만 2세대 에코 리뷰에는 많이 등장하지 않는 단어는 question, home, now, say이며, 2세대 리뷰에 많이 등장하지만 1세대 리뷰에는 빈번히 등장하지 않는 단어는 qualiti, better, dot, easi 이다. dot은 Amazon에서 판매하는 AI 스피커의 다른 종류인 Echo Dot을 지칭하는데 사용되는 단어로 보인다. 2세대 제품을 1세대 제품이나 다른 AI 스피커와 비교하기 위한 단어들이 많이 사용되었다. 제품 리뷰에 등장하는 단어들의 빈도수 순위가 2세대 제품에서 5단계 이상 상승한 단어는 sound, good, speaker이며, 5단계 이상 하락한 단어는 amazon, thing, product, realli이다. 2세대 제품이 1세

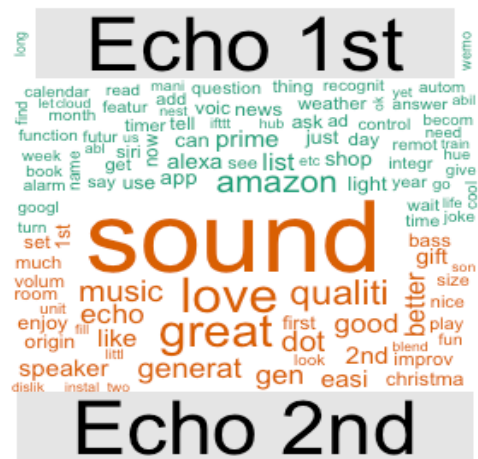
<Table 2> Top 30 Frequent Words

Top 30 Frequent Words	
Echo 1st	Echo 2nd
echo	love
love	echo
alexa	sound
use	music
music	alexa
can	great
great	use
amazon	like
get	can
play	play
like	work
work	get
just	one
one	good
ask	just
thing	amazon
time	qualiti
devic	devic
sound	speaker
product	ask
question	time
good	thing
realli	better
much	much
home	set
set	product
now	dot
say	easi
speaker	fun
fun	realli

대 제품에 비해서 음질(sound)이 개선이 되었다는 내용이 리뷰에 많이 등장하였다.

<Figure 3>은 1세대와 2세대 에코 리뷰 각각에 더 많이 등장하는 단어 모음을 보여준다. 1세대에 비해 2세대 에코에 대한 리뷰에 음악(music), 사운드(sound), 품질(qualiti) 등의 음질관련 단어가 더 많이 등장하고 1세대와 비교하기 위한 세대(gen, generat, 1st, 2nd) 등의 단어도 더 많이 등장하였다. 2세대가 더 좋은 면을 언급하는 Better, good, improv 등의 단어도 더 많이 등장하였다. 1세대 에코의 리뷰에는 처음으로 소개되는 제품이기 때문에 자체의 속성과 기능에 대한 언급이 더 많이 보인다. 음성인식과 관련한 recognit, ask, tell, voic, question, read 등의 단어가 보이고, 기능에 해당하는 light, control, timer, list, news, weather, alarm 등의 단어가 더 많이 등장하였다.

1세대와 2세대 리뷰에 등장하는 단어 빈도의 통계



<Figure 3> Word cloud of the two products

적인 차이를 파악하기 위하여 Keyness comparison을 수행하였다(Scott and Bondi 2010). Keyness comparison은 단어 빈도간의 Chi-square 값을 계산하여 통계치가 양의 방향과 음의 방향으로 큰 단어들

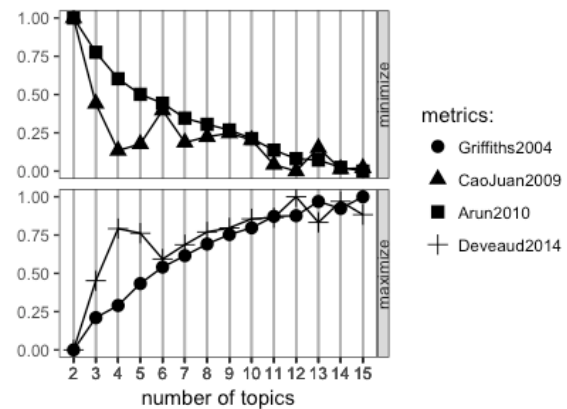
<Table 3> Top 20 Keyness Words

Top 20 Keyness Words			
Words appeared more in 2nd Generation	Chi-Square Statistics	Words appeared more in 1st Generation	Chi-Square Statistics
sound	8006.0637	prime	-454.5798
gen	3008.1544	list	-382.2344
generat	2854.5174	amazon	-363.5185
2nd	1933.8993	shop	-291.7581
qualiti	1509.0815	siri	-232.2698
1st	1201.9707	remot	-167.2874
dot	974.0094	app	-164.4938
great	934.1278	light	-157.9249
origin	812.1727	news	-155.5329
bass	698.2404	integr	-155.5329
love	687.4695	ad	-145.7805
better	492.6431	futer	-129.5339
fabric	478.5114	timer	-122.8234
good	381.9684	see	-122.1274
easi	373.0623	ifttt	-116.0247
size	372.1522	calendar	-114.5403
gift	364.4212	add	-113.8547
smaller	356.5017	wait	-108.6892
christma	327.7467	recognit	-107.0327
improv	249.8625	wemo	-104.7588

을 선별하여 보여준다. <Table 3>은 각 제품의 리뷰에 상대적으로 더 많이 등장하는 20개의 단어와 Chi-square 통계치를 보여준다. 왼쪽 컬럼이 2세대 에코에 더 빈번히 등장하는 단어이며, 오른쪽 컬럼이 1세대 에코에 더 빈번히 등장하는 단어들이다. 2세대 에코의 리뷰에 상대적으로 더 많이 등장하는 단어는 음질(sound, quality, bass), 제품비교(gen, generat, 2nd, 1st, dot, improv), 상대비교(great, love, better, good, easi)라고 볼 수 있다. 반대로 1세대 에코의 리뷰에 상대적으로 더 많이 등장하는 단어는 아마존(amazon, prime), 제어(app, light, timer), 정보(news, calendar), 음성인식(recognit) 등의 단어로 볼 수 있다. 위에서 언급하였던 것처럼 첫 제품이기 때문에 가지고 있는 속성과 기능에 대한 언급이 많았다고 볼 수 있다.

4. 분석 및 결과

인공지능 속성에 대한 고객 태도 변화를 분석하기 위해서는 리뷰에 포함된 다양한 하위 토픽 중에서 인공지능 속성에 대해 주로 언급하고 있는 토픽을 파악하여야 한다. 이를 위해 LDA 기법을 활용하여 리뷰에 존재하는 하위 토픽을 파악하였다. LDA 기법은 분석자가 정해진 수 만큼의 하위 토픽을 파악하여주고, 분석자가 추출된 토픽에 대해 이해하고 분석대상에 가장 적합한 토픽의 수를 찾는 방안이 일반적이다(Ramage et al. 2009). 하지만 **토픽의 수 선정**에 있어서 자의적인 면이 많기에 적합한 토픽의 수를 찾기 위한 다양한 지표들이 제시되어 왔다(Arun et al. 2010; Juan et al. 2009; Deveaud et al. 2014; Griffiths and Steyvers 2004). **R의 ldatuning 패키지**(Murzintcev 2016)를 활용하여 토픽의 수를 증가시켜가며, 지표 값들이 변화하는 추세를 파악하였다. <Figure 4>가 토픽의 수에 따른 각 지표의 값을 보여준다.

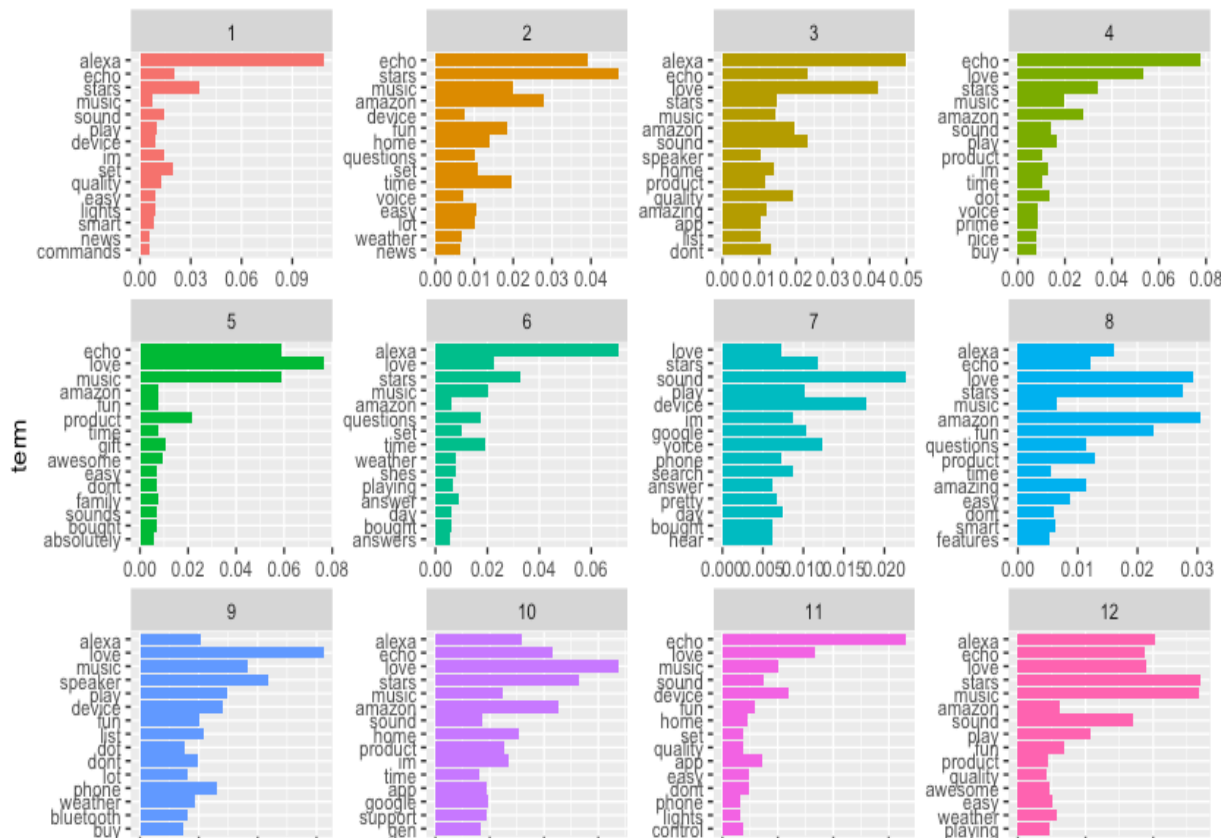


<Figure 4> Number of topics and metrics

기본적으로 지표가 토픽 간의 차이가 명확한 지와 많은 부분을 포함하고 있는 지를 측정하기에 토픽의 수가 많아지면 최소화 하는 지표와 최대화 하는 지표 모두 좋은 값을 띄게 되다가 토픽의 수가 너무 많아지면 지표 값이 다시 나빠지는 패턴을 띈다. 본 연구에서는 최소화 지표 값과 최대화 지표 값이 안정적인 패턴을 보이기 시작하는 토픽의 수인 12개를 선택하여 토픽 분석을 수행하였다.

토픽의 수 12개로 LDA를 시행하여 얻은 토픽들을 <Figure 5>에 표시하였으며, 토픽별로 토픽을 대표하는 단어들과 이들의 베타 값을 가시화 하였다. 베타 값은 해당 단어가 특정 토픽에 속하는 확률이다(Blei 2012).

토픽의 수가 12개로 많은 편이기에 토픽간의 차이가 명확하지 않은 것이 존재하며, 어느 정도는 겹치는 부분이 있는 것으로 보이지만 <Table 4>처럼 토픽에 명칭을 부여하였다. 다른 토픽에도 alexa라는 단어가 많이 포함되지만 이는 echo, music, amazon 처럼 리뷰에 빈번히 등장하는 단어이기 때문이다. 음성을 통한 제어와 대화, 정보 획득에 대한 단어를 많이 포함하고 있는 것은 2(음성명령+정보), 6(음성명령+제어), 8번(스마트 속성) 토픽이다.



<Figure 5> Terms and beta values of topics

<Table 4> Topic Labels

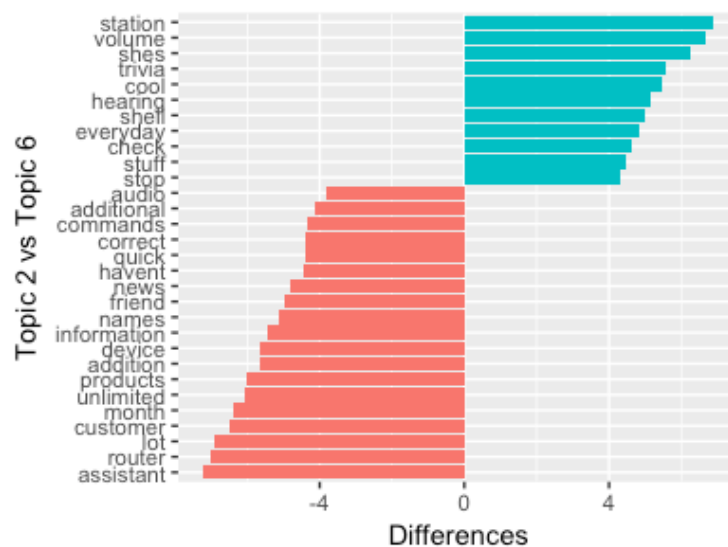
Topic Labels	
Topic Number	Labels
1	음악 재생 + 음질
2	음성 명령 + 정보
3	음악 + 스피커
4	아마존 프라임 + 음악 재생
5	만족
6	음성 명령 + 제어
7	구매경험 + 음성대화
8	스마트 속성
9	무선 스피커
10	구글 홈과의 비교
11	설치
12	음질 만족

2, 6, 8번 토픽을 서로 비교하여 보았다. <Figure 6> 과 <Figure 7>이 2번 토픽과 6번, 8번 토픽을 각각 비교한 그림이다. 그림의 오른쪽 상단에 베타 값이 표시된 단어들이 토픽 2번에서 베타 값이 더 큰 단어들이며, 왼쪽 하단에 나타난 단어들이 각각 토픽 6번과 8번에서 베타 값이 더 큰 단어들이다.

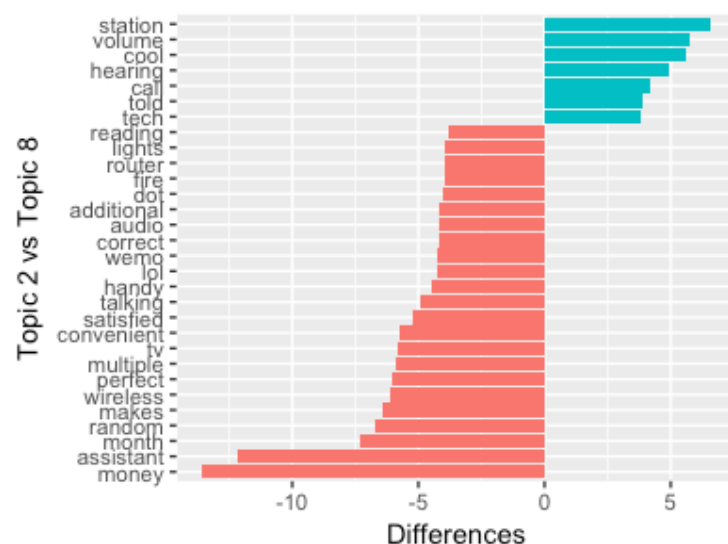
토픽2번에 비해 토픽 6번에 더 많이 연관된 단어들 중에 commands, news, information, correct, quick 등

의 단어가 존재한다. 반면 토픽 2번에는 hearing 이나 volume 등의 스피커나 음성 관련 단어들이 존재한다.

토픽 8번과 토픽 2번의 비교에서도 마찬가지로 토픽 2번에는 volume이나 hearing 등의 스피커나 음성 관련 단어들이 많고, 토픽 8번에는 스마트 속성인 reading, talking 단어와 satisfied, convenient, perfect등 만족에 대한 단어들이 존재한다. 토픽 2, 6, 8번이 서로 상이한 부분이 존재하지만 음성명령, 정보



<Figure 6> Topic comparison between Topic 2 and Topic 6



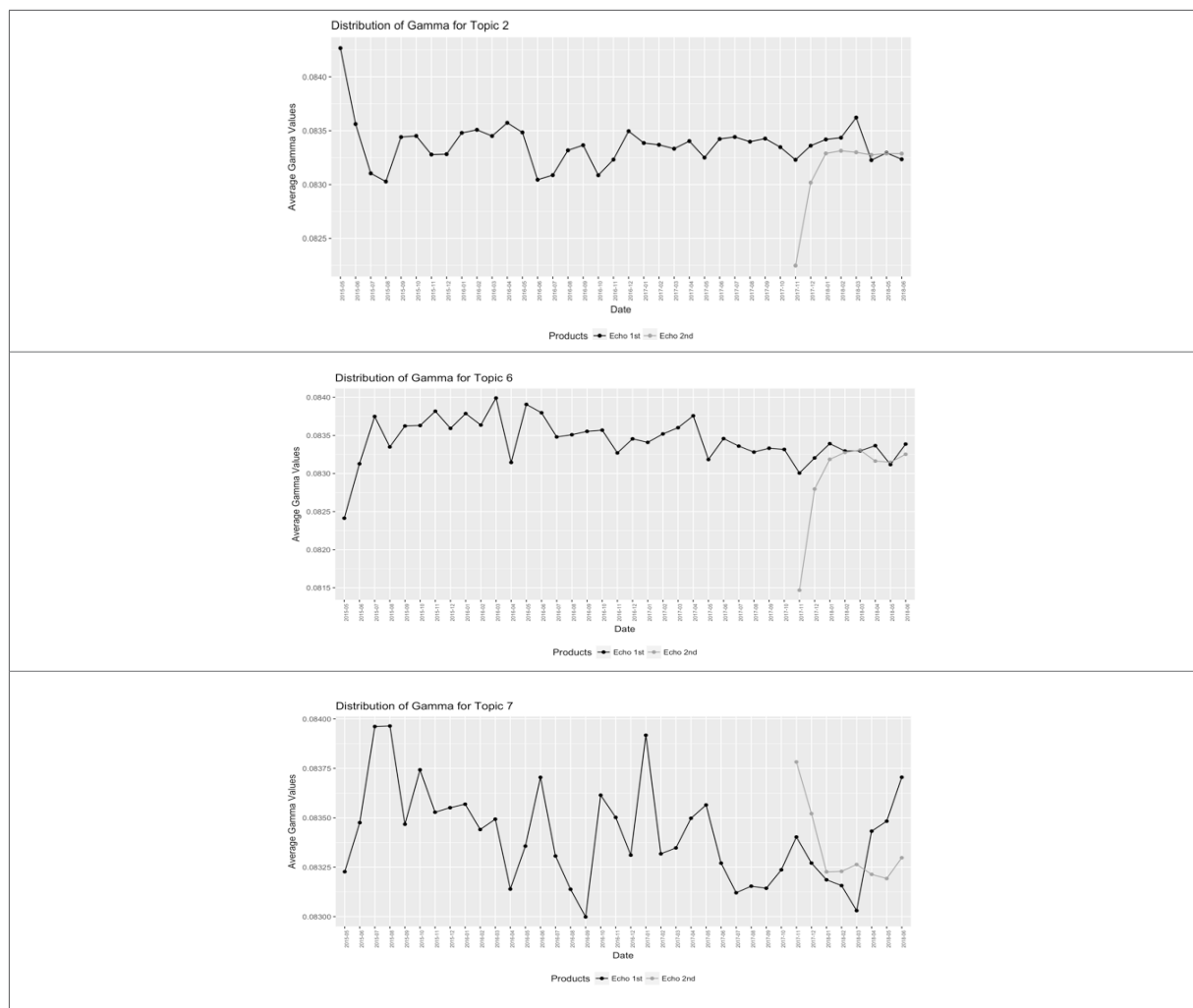
<Figure 7> Topic comparison between Topic 2 and Topic 8

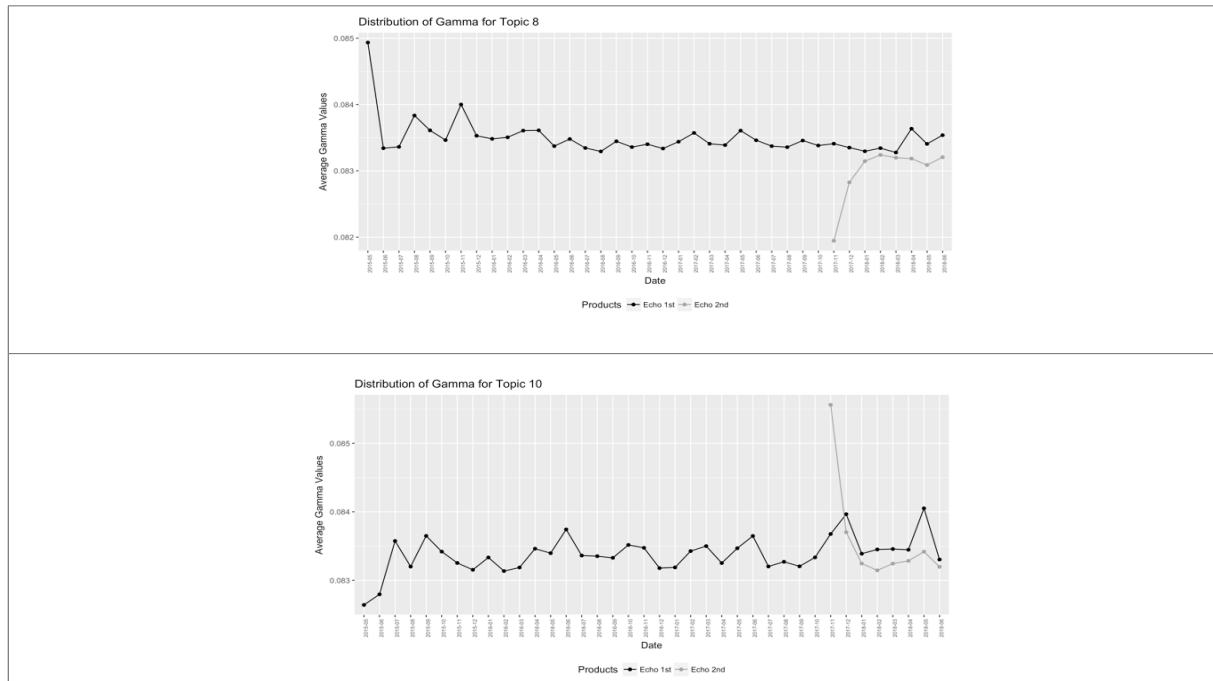
제공, 제어와 같은 스마트 속성에 대한 내용을 많이 포함하고 있는 것으로 보여 이 세개의 토픽의 분포가 시간의 흐름에 따라 어떻게 변화하는 지를 중점적으로 분석해 보고자 한다.

고객 리뷰에서 시간의 흐름에 따른 토픽 분포를 파악하기 위하여 월별로 리뷰에 포함된 토픽의 감마(Gamma) 값의 평균을 구하였다. 토픽의 감마 값은 문서에 포함된 단어들이 해당 토픽에 속하는 확률이다 (Blei 2012).

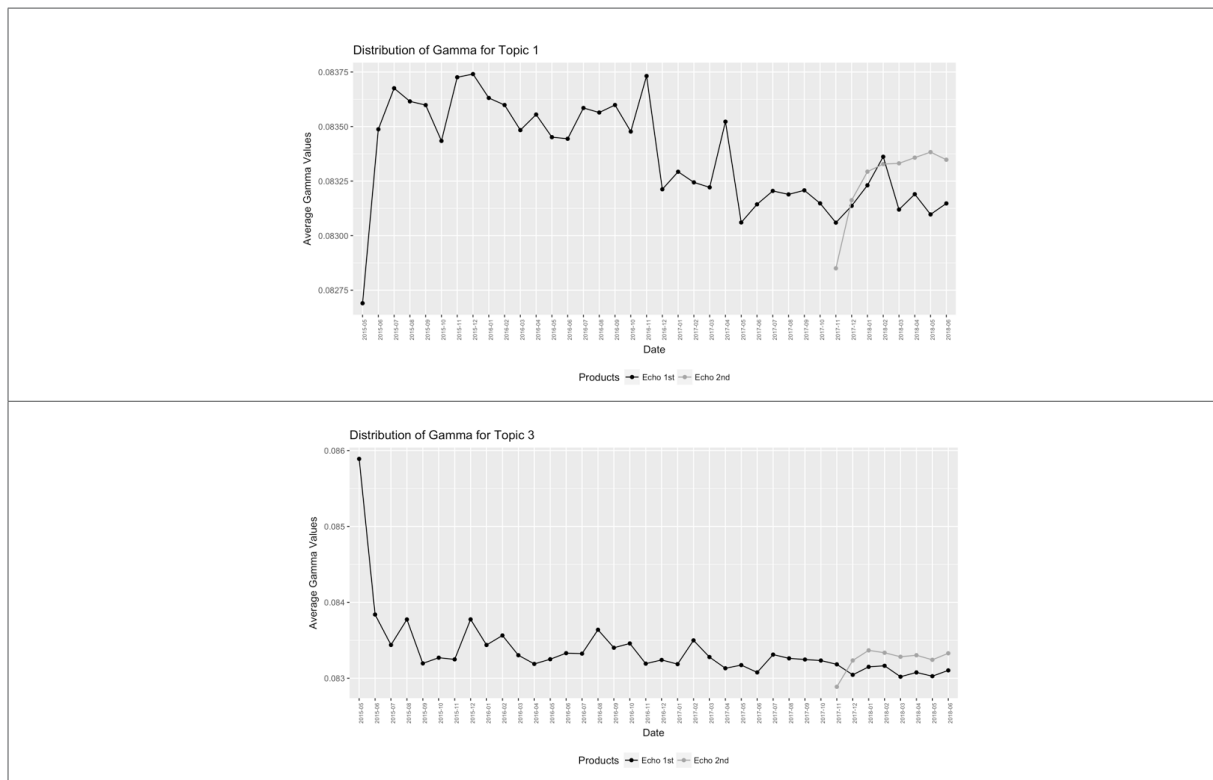
<Figure 8>, <Figure 9>, <Figure 10>은 월별 리뷰에 포함된 토픽별 감마 값의 평균이다. <Figure 8>은 시간이 지나도 월별 감마 값의 평균이 큰 변화가 없는

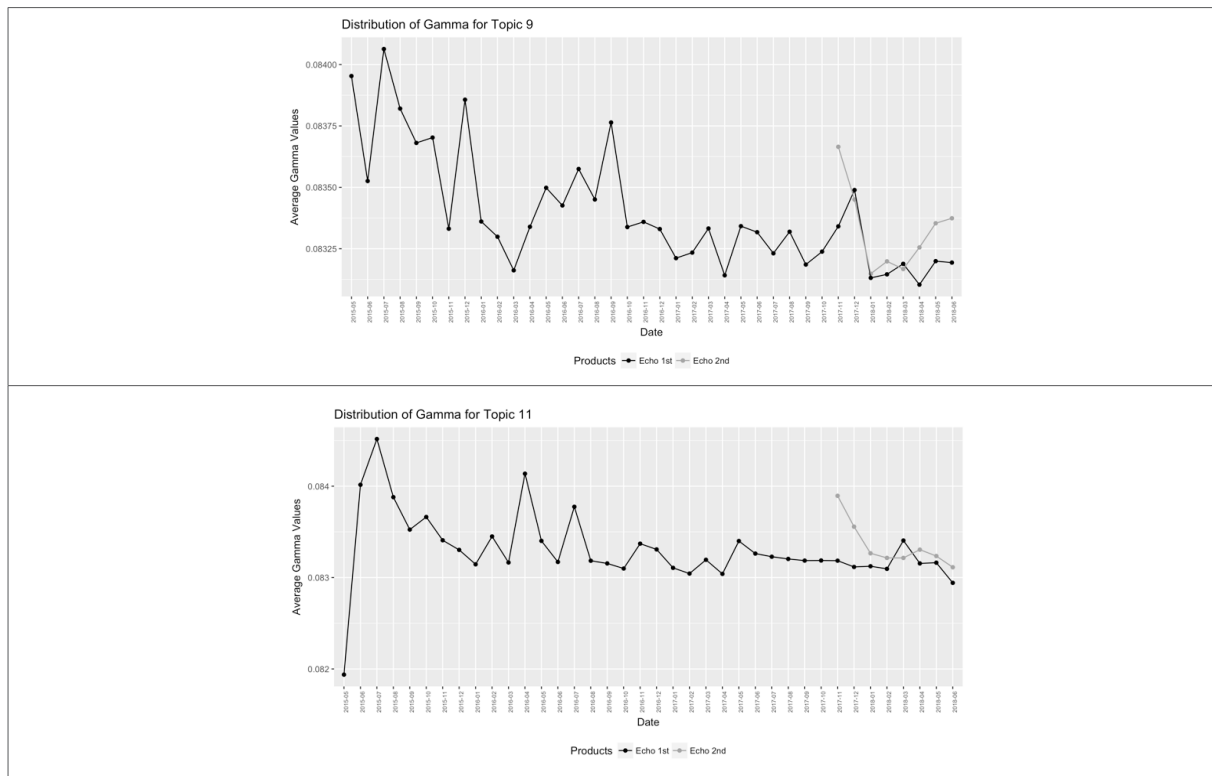
토픽 집합이다. 토픽 7 (구매경험 + 음성대화)의 그래프는 진동이 있어 보이지만 y축 값의 범위가 넓지 않다. 토픽 10 (구글 홈과의 비교)도 일정한 분포를 보이고 있다. 위에서 스마트 속성에 대한 토픽으로 분류한 토픽 2, 6, 8은 모두 큰 변화를 보이지 않았다. 시간이 흘러도 스마트 속성에 대한 언급 비율이 일정하게 유지되는 것으로 볼 수 있다. 다만 2세대 에코에 대한 리뷰에서는 1세대 에코에 대한 리뷰에 포함된 토픽 비율보다 낮은 비율을 보였다. 제품의 하드웨어적인 진화가 이루어진 새 제품에서는 스마트 속성에 대한 언급보다는 개선된 부분에 대한 언급이 늘어난 것으로 볼 수 있다.





<Figure 8> Gamma value distribution over time (Topic 2, 6, 7, 8 and 10)



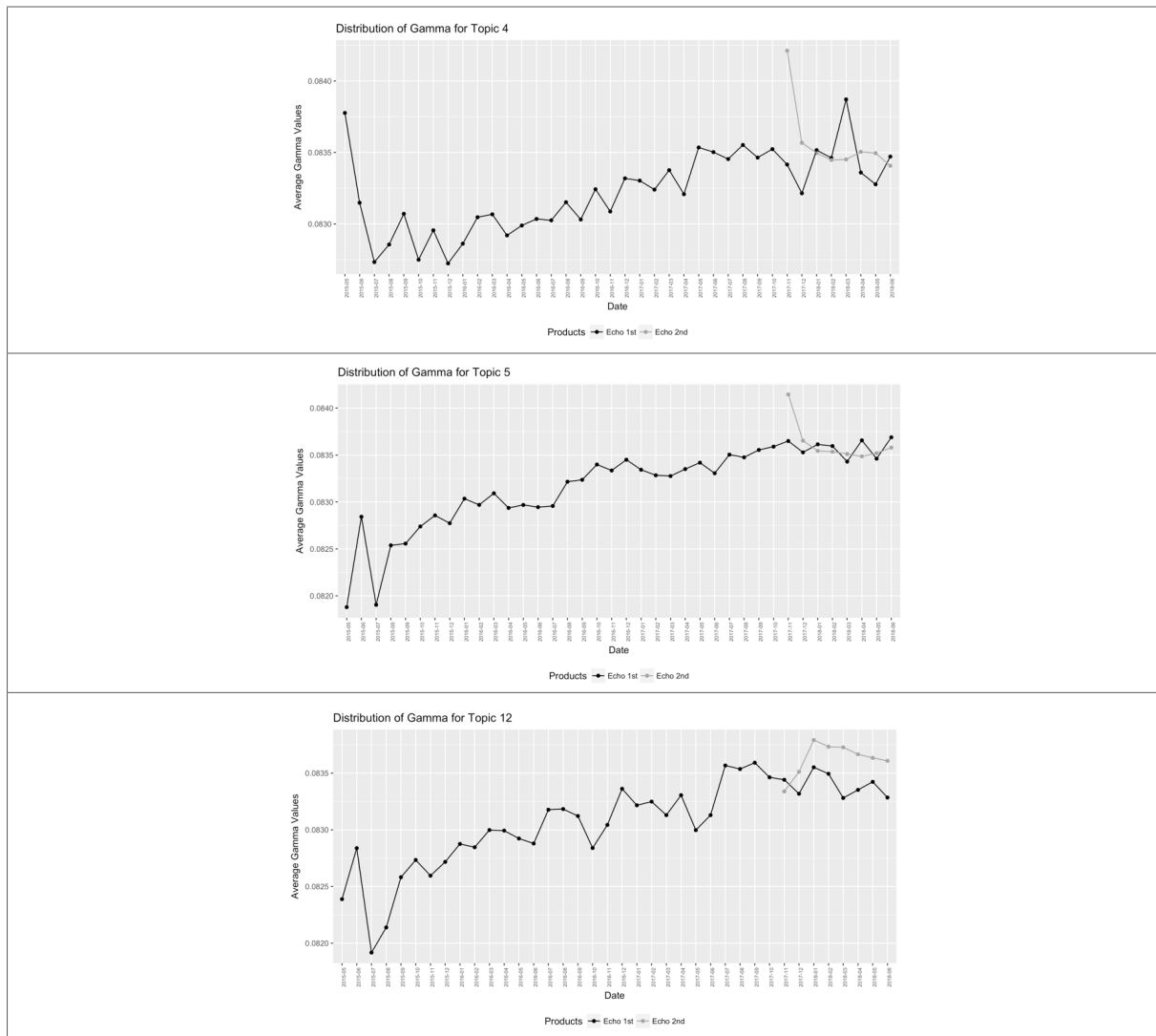


<Figure 9> Gamma value distribution over time (Topic 1, 3, 9 and 11)

<Figure 9>는 월별 토픽의 감마 값 평균이 작지만 감소하는 추세를 보이는 토픽들이다. 리뷰에 포함된 토픽이 균등하게 분포되어 있어 토픽의 평균 감마 값에 큰 변화는 없지만 미세하게 감소하는 추세를 보이는 토픽들이다. 토픽 1(음악재생 + 품질), 토픽 3(음악 + 스피커), 토픽 9(무선스피커)은 1세대 에코의 리뷰에서는 지속적으로 평균 감마 값이 낮아지는 추세였으나, 2세대 에코의 출시이후에는 해당 토픽의 월별 감마 값 평균이 1세대에 비해서 소폭 상승한 것으로 볼 수 있다. 2세대 에코가 하드웨어적인 측면에서 음질과 음악재생에서 개선된 것으로 리뷰에서 평가받고 있기에 이러한 내용이 토픽의 분포에도 영향을 미친 것으로 볼 수 있다. 토픽 11(설치)는 1세대 에코와 2세대 에코에서 모두 시간이 흐를수록 감마 값 평균이 낮아지는 것을 볼 수 있다. 설치관련해서는 언급되는 비율이 지속적으로 감소한다고 볼 수 있으며, 제품의 개선으로 인한 감소도 있지만

사용자들의 학습효과와 고객 리뷰 및 다양한 원천으로부터 획득한 정보가 설치에 대한 언급을 줄이게 된 것으로 보인다.

<Figure 10>에 보여지는 토픽 4(아마존 프라임 + 음악 재생), 5(만족), 12(음질만족)는 시간이 흐를수록 언급되는 비율이 높아지는 토픽들이다. 특히 토픽 5와 12는 상대적으로 감마 값 평균이 많이 높아지고 있으며, 특히 토픽 12는 2세대 에코 출시 이후에 리뷰에 포함된 비율이 1세대 보다 더 높아졌다. 일반적인 만족도에 해당하는 토픽 5의 비율이 지속적으로 커지는 것은 제품에 대한 만족도가 높아지는 것으로 볼 수 있으며, 특히 음질에 대한 만족(토픽 12)이 2세대 에코가 출시된 이후에 더 많이 언급되었다. AI 스피커의 중요한 기능인 음악재생과 음질에 대한 만족이 더 많이 언급되고, 고객 리뷰의 평균 평점이 4.34에서 4.43으로 증가하였다. 이는 일반적인 만족도와 음질 및 음악 재생에 대한 만



<Figure 10> Gamma value distribution over time (Topic 4, 5 and 12)

족도가 증가하였다고 볼 수 있다.

<Figure 8>, <Figure 9>, <Figure 10>은 토픽별 리뷰에 언급되는 분포를 시간의 흐름에 따라 가시화한 것이다. 좀 더 심층적인 분석을 위해 모든 리뷰에 언급되는 분포가 아닌, 리뷰에 따라 언급되는 분포가 차이가 있는 지를 파악하고자 한다. 고객 리뷰는 기본적으로 제품에 대한 만족을 표시하는 별표를 1개부터 5개까지 부여하게 되어 있다. 일반적으로 별 1, 2개를 받은 리뷰는 부정적인 것으로, 별 4, 5개를 받은 리뷰는 긍정적

인 것으로 분류한다. <Table 5>와 <Table 6>은 토픽의 분포와 리뷰에 부여된 별의 개수를 고려하여 만족도에 따른 토픽의 평균 분포를 계산하였다.

<Table 5>는 1세대와 2세대 에코에 대한 리뷰에 포함된 토픽 별 평균 감마 값(Average Gamma Value)과 리뷰에 부여된 만족도에 토픽의 분포 비율을 곱하여 계산한 Information Score의 평균과 표준편차를 보여주고 있다. Information score는 토픽이 똑같은 비율로 리뷰에 존재하더라도 더 긍정적인 리뷰에 빈번히 등장

<Table 5> Information Scores of Topics

Topic	Average Gamma Value of Echo 1st	Information Score of Echo 1 st (Standard deviation)	Ranking	Average Gamma Value of Echo 2 nd	Information Score of Echo 2 nd (Standard deviation)	Ranking
1	0.0833	0.3620 (0.1015)	4	0.0832	0.3698 (0.0902)	5
2	0.0833	0.3619 (0.0997)	5	0.0831	0.3695 (0.0902)	7
3	0.0833	0.3620 (0.1007)	3	0.0832	0.3700 (0.0905)	4
4	0.0832	0.3615 (0.1004)	7	0.0835	0.3708 (0.0907)	3
5	0.0831	0.3615 (0.1000)	8	0.0836	0.3710 (0.0904)	2
6	0.0834	0.3622 (0.1006)	1	0.0829	0.3690 (0.0901)	11
7	0.0834	0.3608 (0.0979)	11	0.0833	0.3691 (0.0887)	10
8	0.0835	0.3621 (0.1011)	2	0.0829	0.3689 (0.0901)	12
9	0.0833	0.3613 (0.0989)	10	0.0833	0.3692 (0.0895)	9
10	0.08337	0.3608 (0.0985)	12	0.0836	0.3695 (0.0898)	8
11	0.0833	0.3614 (0.0991)	9	0.0833	0.3696 (0.0902)	6
12	0.0830	0.3615 (0.1002)	6	0.0836	0.3718 (0.0910)	1

한다면 값이 커지고, 반대로 부정적인 리뷰에 더 빈번히 등장한다면 값이 작아진다. 평균 감마 값이 토픽 별로 대동소이하기 때문에 Information score를 통해 만족에 미치는 영향을 파악할 수 있다. 1세대 에코에서는 스마트 속성에 대한 토픽인 토픽 2, 6, 8의 영향력이 각

각 5, 1, 2 등으로 높은 영향을 미쳤다고 볼 수 있다. 반면 2세대 에코에서는 토픽 2, 6, 8의 영향력이 각각 7, 11, 12 등으로 낮아졌다.

스마트 속성에 대한 토픽의 영향력이 낮아진 이유를 분석하기 위해 <Table 6>과 같이 아마존 에코 세대별

<Table 6> Average Gamma Values for Positive and Negative Reviews

Topic	Echo 1st				Echo 2nd			
	Average Gamma Values for 5 Stars	Ranking	Average Gamma Values for 1 Star	Ranking	Average Gamma Values for 5 Stars	Ranking	Average Gamma Values for 1 Star	Ranking
1	0.08346	5	0.08237	10	0.08331	6	0.08249	10
2	0.08344	7	0.08323	7	0.08333	5	0.08280	8
3	0.08350	2	0.08215	11	0.08345	4	0.08233	12
4	0.08341	8	0.08287	8	0.08359	3	0.08331	6
5	0.08346	6	0.08327	6	0.08362	2	0.08392	3
6	0.08349	3	0.08280	9	0.08316	9	0.08236	11
7	0.08284	12	0.08462	2	0.08299	12	0.08464	2
8	0.08352	1	0.08350	5	0.08318	8	0.08287	7
9	0.08313	10	0.08418	3	0.08313	10	0.08384	4
10	0.08300	11	0.08508	1	0.08308	11	0.08543	1
11	0.08321	9	0.08380	4	0.08322	7	0.08340	5
12	0.08348	4	0.08208	12	0.08387	1	0.08256	9

로 5점을 준 리뷰와 1점을 준 리뷰의 토픽 별 분포를 살펴 보았다. 1세대 에코에서는 5점을 준 리뷰에 스마트 속성에 대한 토픽의 분포 비율이 각각 7, 3, 1등이었으나 2세대 에코에서는 각각 5, 9, 8 등으로 낮아졌으며, 1점을 준 리뷰에서도 1세대에서는 각각 7, 9, 5등이었으나 2세대에서는 8, 11, 7 등으로 낮아졌다. 긍정적인 리뷰에 포함된 비율이 부정적인 리뷰에 포함된 비율에 비해 상대적으로 높기는 하지만 제품의 세대가 변하면서 스마트 속성에 대한 토픽이 언급되는 비율이 긍정과 부정 리뷰에서 모두 줄었다. 이로 인해 <Table 5>와 같이 스마트 속성에 대한 토픽이 만족도에 미치는 영향이 감소한 것으로 볼 수 있다.

5. 토의 및 결론

AI 스피커의 스마트 속성에 대한 고객의 태도를 통시적으로 분석하기 위해 가장 많이 보급된 AI 스피커인 아마존 에코의 고객 리뷰를 수집하여 분석하였다. 아마존 에코는 2세대 제품까지 출시된 상황이며 1세대 제품이 출시된 지 3년이상이 경과되었다. 1, 2 세대 에코에 대한 고객 리뷰를 아마존으로부터 수집하여 기술적인 분석을 통해 가장 빈번히 등장하는 단어들의 변화를 살펴보았다. 1, 2세대 에코에 대한 리뷰에 빈번하게 등장하는 30개 단어는 매우 유사하였으며, 빈도수에 따른 순위가 서로 상이할 뿐이었다. 2세대에 대한 리뷰에는 2세대 제품을 1세대 제품이나 다른 AI 스피커와 비교하기 위한 단어들이 많이 사용되었으며, 음질이 개선과 관련된 단어들이 리뷰에 많이 등장하였다.

LDA 기법을 활용하여 고객 리뷰에 존재하는 하위 토픽들을 추출하였으며, 하위 토픽의 숫자는 최적의 토픽 수를 결정하는 지표를 활용하여 정하였다. 추출된 토픽들에 포함된 단어와 베타 값을 고려하여 각 토픽의 의미 부여하였으며, 이를 통해 AI 스피커가 가지고

있는 스마트 속성에 대한 토픽을 파악하였다. 토픽들의 분포 비율을 월별로 시각화 하였으며, 스마트 속성에 대한 토픽들의 분포 비율은 시간이 흐름에도 안정적으로 유지되는 것으로 보였다. 단 2세대 에코에 대한 리뷰에서는 스마트 속성에 대한 토픽들이 차지하는 비율이 1세대 에코에 비해 감소하였다. 5점 리뷰에 분포하는 비율도 2세대에서 감소하였으며, 마찬가지로 1점 리뷰에 분포하는 비율도 감소하였다. 스마트 속성에 대한 토픽의 분포 비율 감소로 인해 리뷰의 평점과 토픽의 값의 곱인 Information Score의 평균이 1세대보다 2세대에서 감소하였다. 긍정이나 부정 리뷰에서 언급되는 정도가 모두 감소하였기에 스마트 속성이 제품의 만족에 대한 기여도가 감소한 것으로 볼 수 있다.

2세대 에코가 출시되면서 개선된 하드웨어와 다른 제품과의 비교에 대한 내용이 리뷰에서 더 많이 차지하게 된 것으로 볼 수 있으며, 스마트 속성이 지속적으로 추가되고 있지만 초기에 비해서 리뷰에 덜 포함되고 있다.

본 연구는 다음과 같은 한계가 있다. 첫째, 다양한 AI 스피커 중에서 하나의 제품인 아마존 에코에 대해서만 고객 리뷰를 수집하여 분석하였기에 결과를 일반화하기에는 주의가 따른다. 하지만 아마존 에코가 가장 많이 판매된 AI 스피커이며, 2세대까지 출시된 제품이기 때문에 통시적으로 고객의 태도를 분석하는 데 적합하다. 둘째, 본 연구에서 분석된 제품의 출시 이후에 다른 제품들이 출시되면서 다양한 기능들이 부가되었고, 아마존 에코와의 비교가 이루어졌다. 이를 통해 스마트 스피커 속성의 중요도에 대한 인식이 변화했을 수 있다. 따라서, 다른 제품들의 출시를 고려하여 이러한 부분에 대한 세심한 연구가 필요하다. 셋째, 토픽의 분포 비율과 리뷰의 평점만을 활용하여 지표를 구성하였기에 텍스트 분석에서 많이 활용되는 텍스트의 감성을 고려하지 못하였다. 텍스트의 감성을 고려한 지표의 추가를 통해 스마트 속성에 대한 고객 태도를 파악하는 연구

가 필요하다. 기본적으로 리뷰 평점에 전반적인 제품에 대한 감성이 포함되어 있기에, 현재 지표에서도 간접적인 제품에 대한 고객의 감성이 반영되었다고 볼 수 있다.

참 고 문 헌

[국내 문헌]

1. 김정훈, 송영은, 진윤선, 권오병 2015. “텍스트마이닝을 통한 댓글의 공감도 및 비공감도에 영향을 미치는 댓글의 특성 연구,” *한국IT서비스학회지* (14:2), pp. 159-176.
2. 박명석, 권영진, 이상용 2018. “댓글이 음원 판매량에 미치는 차별적 영향에 관한 텍스트마이닝 분석,” *지식경영연구* (19:2), pp. 91-108.
3. 연합뉴스 2018. “한국 AI스피커 시장 급성장...올해 세계 톱 5 전망”, 2018/07/12, <http://www.yonhapnews.co.kr/bulletin/2018/07/11/0200000000AKR20180711151200017.HTML?sns=copy>
4. 이홍주 2018a. “A Ghost in the Shell? 고객 리뷰를 통한 스마트 스피커의 인공지능 속성이 평가에 미치는 영향 연구,” *한국IT서비스학회지* (17:2), pp. 191-205.
5. 이홍주 2018b. “헬스케어 서비스 리뷰를 활용한 서비스 품질 차원 별 중요 단어 파악 방안,” *지식경영연구* (19:4), pp. 171-185.
6. 컨슈머인사이트 2018. “뜨거운 AI스피커 시장, 차가운 소비자 평가”, 이동통신리포트, 2018/07/09, https://www.consumerinsight.co.kr/voc_view.aspx?no=2924&id=ins02_list&PageNo=1&schFlag=0

[국외 문헌]

1. Archak, N., Ghose, A., and Ipeirotis, P. G. 2011. “Deriving the Pricing Power of Product Features by Mining Consumer Reviews,” *Management Science* (57:8), pp. 1485-1509.
2. Arun, R., Suresh, V., Veni Madhavan, C.

- E., and Narasimha Murthy, M. N. 2010. "On finding the natural number of topics with latent dirichlet allocation: Some observations," *In Advances in knowledge discovery and data mining*, Mohammed J. Zaki, Jeffrey Xu Yu, Balaraman Ravindran and Vikram Pudi (eds.). Springer Berlin Heidelberg, 391-402. http://doi.org/10.1007/978-3-642-13657-3_43
3. Blei, D. M. 2012. "Probabilistic Topic Models," *Communications of the ACM* (55:4), pp. 77-84.
 4. Büschken, J., and Allenby, G. M. 2016. "Sentence-Based Text Analysis for Customer Reviews," *Marketing Science* (35:6), pp. 953-975.
 5. Cao, Q., Duan, W., and Gan, Q. 2011. "Exploring determinants of voting for the 'helpfulness' of online user reviews: A text mining approach," *Decision Support Systems* (50:2), pp. 511-521.
 6. Chen, K., Kou, G., Shang, J., and Chen, Y. 2015. "Visualizing market structure through online product reviews: Integrate topic modeling, TOPSIS, and multi-dimensional scaling approaches," *Electronic Commerce Research and Applications* (14:1), pp. 58-74.
 7. Deveaud, R., Juan, É. S., and Bellot, P. 2014. "Accurate and effective latent concept modeling for ad hoc information retrieval," *Document numérique* (17:1), pp. 61-84. <http://doi.org/10.3166/dn.17.1.61-84>
 8. Ghose, A., Ipeirotis, P. G., and Li, B. 2012. "Design Ranking Systems for Hotels on Travel Search Engines by Mining User-Generated and Crowdsourced Content," *Marketing Science* (31:3), pp. 493-520.
 9. Griffiths, T. L., and Steyvers, M. 2004. Finding scientific topics. *Proceedings of the National Academy of Sciences* (101: suppl 1), pp. 5228-5235. <http://doi.org/10.1073/pnas.0307752101>
 10. Hu, N., Pavlou, P. A., and Zhang, J. 2009. "Overcoming the J-shaped Distribution of Product Reviews," *Communications of the ACM* (52:10), pp. 144-147.
 11. Humphreys, A., and Wang, R. J. 2018. "Automated Text Analysis for Consumer Research," *Journal of Consumer Research* (44: 6), pp. 1274-1306.
 12. Juan, C., Tian, X., Jintao, L., Yongdong, Z., and Sheng, T. 2008. "A density-based method for adaptive LDA model selection." *Neurocomputing — 16th European Symposium on Artificial Neural Networks* (72:7-9), pp. 1775-1781. <http://doi.org/10.1016/j.neucom.2008.06.011>
 13. Kabir, K. 2018. "Smart speakers – everything you need to know," What Hi-Fi?, Available at <https://www.whathifi.com/advice/smart-speakers-everything-you-need-to-know> (Accessed February 14, 2018)
 14. Lee, H. J., Lee, M., and Lee, H. 2017. "Compilation of Tweets Sentiment into SERVQUAL for Tracking Social Perception on Public Service," The 19th International Conference on Electronic Commerce, 2017, Pangyo, Korea.

15. Lee, T. Y., and Bradlow, E. T., 2011. "Automated marketing research using online customer reviews," *Journal of Marketing Research* (48:5), pp. 881-894.
16. Mudambi, S. M., and Schuff, D. 2010. "What makes a helpful online review? A Study of Customer Reviews on Amazon.com," *MIS Quarterly* (34:1), pp. 185-200.
17. Murzintcev, N. 2016. "ldatuning," R package. <https://cran.r-project.org/web/packages/ldatuning/ldatuning.pdf>
18. Netzer, O., Feldman, R., Goldenberg, J., and Fresko, M. 2012. "Mine Your Own Business: Market-Structure Surveillance Through Text Mining," *Marketing Science* (31:3), pp. 521-543.
19. Palese, B., and Usai, A. 2018. "The relative importance of service quality dimensions of E-commerce experiences," *International Journal of Information Management* (40), pp. 132-140.
20. Ramage, D., Rosen, E., Chuang, J., Manning, C. D., and McFarland, D. A. 2009. Topic Modeling for the Social Sciences. Presented at the Neural Information Processing Systems (NIPS) Workshop on Applications for Topic Models: Text and Beyond, Whistler, Canada.
21. Scott, M., and Bondi, M. 2010. *Keyness in Texts*. Amsterdam, Philadelphia: John Benjamins, pp. 21-42.

● 저 자 소 개 ●



이 홍 주 (Hong Joo Lee)

현재 가톨릭대학교 경영학부 교수로 재직 중이다. KAIST 테크노경영대학원에서 경영학 박사 학위를 취득하였다. 주요 관심분야는 데이터 분석, 지능형 정보시스템, 온라인 사용자들의 상호작용 등이다. 지금까지 *Journal of Electronic Commerce*, *Cyberpsychology, Behavior, and Social Network*, *Technological Forecasting & Social Change*, *International Journal of Electronic Commerce* 등 주요 학술지에 논문을 발표하였다.