

리액션성 대사의 음량이 시청영상 감성평가에 미치는 영향

최형민¹ · 이주환^{2*}¹서울미디어대학원대학교 융합미디어학과 석사과정²서울미디어대학원대학교 융합미디어학과 교수

The Effect of Sound Level of Reactive Dialogue on Emotional Evaluation of Video Content

Hyung-Min Choi¹ · Ju-Hwan Lee^{2*}¹Master's Course, Department of Convergence Media, Seoul Media Institute of Technology, Seoul 07590, Korea²*Professor, Department of Convergence Media, Seoul Media Institute of Technology, Seoul 07590, Korea

[요약]

영상 사운드는 시각과 결합하여 사운드의 각 요소들이 정보와 감정, 그리고 의미를 관객들에게 전달하고 있으며 내러티브 요소의 역할로서 자리 잡았다. 영상 콘텐츠의 사운드 디자인 작업 과정에서 정확한 정보 전달을 위해 일정한 음량을 유지하는 것이 일반적이지만, 다른 사운드 요소인 배경음악의 경우, 격렬한 감정이나 평온한 감정을 표현하기 위해 연주 형태나 음량을 변화시키는 방식으로 작업한다. 영상 콘텐츠 사운드에서도 시청자의 감성적 공감이나 몰입 경험을 제공하기 위해 대사의 음량을 조절할 수 있을 것이다. 따라서 본 연구에서는 내러티브 영상 콘텐츠에서 등장인물의 반응과 호흡을 포함하는 리액션성 대사의 음량이 시청자 감정에 어떠한 영향을 미치는지 평가하였다. 그 결과로 콘텐츠 내러티브와 시청자의 감성적 공감상태 사이의 음향적 특징을 살펴보고, 미디어에서 인간이 감성적으로 몰입할 수 있는 음향이 무엇인지 평가해서 영상 사운드 제작 및 사운드가 필요한 다양한 분야에 적용 가능성을 논의한다.

[Abstract]

The sound in the video is combined with the visual, and each element of the sound delivers information, emotion, and meaning to the audience, and has established itself as a narrative element. In the process of sound design for video content, it is common to maintain a constant volume to deliver accurate information, but in the case of background music, which is another sound element, the performance type or volume is changed to express emotions. In the sound of video content, the volume of dialogue elements can be adjusted to provide an emotional empathy or immersion experience for viewers. Therefore, in this study, we evaluated how the sound level of reactive dialogue including the vocal reaction and breathing of characters in narrative video content affects viewers' emotions. As a result, we investigated the acoustic characteristics between the content narrative and the emotional empathy of viewers, and evaluating what kind of sound can be emotionally immersed in the media, the possibility of application to various fields requiring sound production and sound will be discussed.

색인어 : 사운드 디자인, 내러티브, 감정, 오디오 레벨, 리액션성 대사**Keyword** : Sound design, Narrative, Emotion, Audio level, Reactive dialogue<http://dx.doi.org/10.9728/dcs.2021.22.12.1997>

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 19 November 2021; Revised 07 December 2021

Accepted 07 December 2021

*Corresponding Author, Ju-Hwan Lee

Tel: +82-2-6393-3240

E-mail: jhlee@smit.ac.kr

I. 서론

초기 영상 콘텐츠는 무성영화의 형태로, 촬영 당시의 실제 소리는 관객에게 전달되지 않고, 움직이는 이미지를 시각적으로 전달하는 방식만으로 극장에서 상영되었다. 하지만 대사과 분위기 조성에 대한 필요성이 대두되면서 그에 대한 다양한 기법 연구가 진행되었다. 출연 배우의 발성을 전달하기 위해 영화가 상영될 때 자막을 활용하거나, 극장 스크린 뒤에서 말을 하는 해설자가 존재하였고, 극에 등장하는 효과음을 표현하기 위해 악기를 연주하는 방식 등이 있었다. 하지만 이것은 촬영현장이 아니라 극장 내에서 직접 말하거나 연주하는 형태였기 때문에 영상과 소리가 완벽하게 맞을 수 없다는 단점이 존재하였다. 이후 음향기술의 발달로 무성영화의 시대를 지나 유성영화로 넘어 오면서 화면과 음향이 동기화된 영화를 관람할 수 있게 되었다. 1927년에 개봉한 영화 ‘재즈 싱어(Jazz Singer)’에서’는 화면과 동기화된 배우의 발성을 들을 수 있었고, 이는 관객과 제작자들에게 큰 충격이었다. 이후 스테레오(stereo), 5.1 서라운드 음향(surround sound) 등 다중 채널 음향(multi-channel sound)으로 발전하면서 다양한 사운드 표현으로 관객들에게 영상 콘텐츠에 보다 깊은 몰입감(immersive experience)을 제공할 수 있게 되었다[1]-[3].

그러나 현재까지 수많은 유성영화가 제작되고 있음에도 불구하고 시각적인 화면 요소에 비해 영상 콘텐츠 사운드에 대한 연구는 부족할 뿐만 아니라 오랜 시간동안 큰 관심을 받지 못하였다. 하지만 영상 속 사운드는 시각과 결합하여 사운드의 각 요소들이 정보와 감정, 그리고 의미를 관객들에게 전달하고 있으며 내러티브 요소(narrative element)의 역할로서 자리 잡았다. 특히 대사는 내러티브를 전달하기 가장 쉬운 사운드 요소로 촬영 현장에서 진행되는 인물의 모든 구두적 표현(verbal expression)을 말하며, 인물 간의 대화를 통해 관객들에게 정보를 전달한다. 그리고 대사(dialogue)는 배우의 표정, 행동 등과 결합하여 인물의 감정이나 상태를 표현하기도 한다. 대사의 기능은 내러티브의 기본 텍스트를 명확하게 전달하는데 있다[4].

이러한 배경에서 영상 콘텐츠의 사운드 디자인 작업 공정상 정확한 정보 전달을 위해 사운드 요소들을 일정한 음량(sound level)으로 들리도록 유지시키는 것이 일반적이다. 하지만 다른 사운드 요소인 배경음악(BGM)의 경우, 격렬한 감정을 표현하기 위해서는 강한 연주 형태와 음량을 키우고, 비교적 평온한 감정을 표현하기 위해서는 부드러운 연주 형태와 음량을 낮추는 등의 방식으로 작업하게 된다. 이러한 측면에서 영상 속 사운드에서도 영상 콘텐츠에 대한 시청자의 감성적 공감이나 몰입 경험을 제공하기 위해서 대사의 음량을 적절히 조절할 수도 있을 것이다. 즉 영상 콘텐츠의 연출 의도나 내러티브 측면에서 특정한 감성적 상태를 표현하기 위해서는 상황에 따른 사운드 요소의 음량 변화가 필요하다. 잘 알려진 배경음악과 같은 상대적으로 재생시간이 많이 필요한 요소와 독립적으로 즉각적인 상황 표현과 감성 유발에

사용되는 사운드 요소로서 등장인물의 대사 가운데 비언어적인 음성(nonlinguistic vocal sound) 요소인 호흡이나 음성적 반응으로 나타나는 소리가 있다. 본 연구에서는 이러한 사운드 요소를 등장인물의 반응적 속성이 강한 측면에서 리액션성 대사(reactive dialogue)로 정의하고, 내러티브를 포함하는 영상 콘텐츠에서 등장인물의 반응과 호흡을 포함하는 리액션성 대사의 소리 크기, 즉 음량(loudness)이 시청자 감정에 어떠한 영향을 미치는지 평가하여 작품의 내러티브, 연출의도, 사운드 믹스, 그리고 시청자가 동일한 감성적 공감 상태를 가지는 음향적 특징을 살펴보고자 한다. 또한 영상 콘텐츠의 사운드 제작뿐만 아니라 리액션성 대사가 많이 포함되는 액션 장르의 게임 등 사운드 요소가 필요한 다양한 콘텐츠 분야에서 그 적용 가능성을 논의하고자 한다.

II. 본론

2-1 영상 콘텐츠 사운드의 가치

영상 콘텐츠의 초기 형태인 무성영화가 촬영될 시기에도 등장인물의 목소리는 실재하였고, 에디슨의 축음기와 같은 기술로 소리를 기록하고 재생할 수 있었다. 다만, 당시 기술적 여건상 음향과 영상의 동기화(synchronization)가 어려웠기 때문에 콘텐츠를 상영하는 극장 현장에서는 음향적 요소들을 관객에게 전달할 수 없었다. ‘재즈 싱어’를 통해 유성영화 시대가 열렸으나 영화 제작자들은 상당한 회의감을 갖기도 하였다. 현장에서 연기하는 배우의 목소리를 녹음하기 위해서는 필수적으로 조용한 환경이 필요했기 때문이다. 무성영화가 촬영될 때보다 시간적, 기술적, 장소적인 제약이 따를 수밖에 없었다. 이러한 우려는 영화의 예술성이 이미지와 사운드의 결합으로 인해 제한되고 파괴된다는 우려로 번지게 되기도 하였다. 사운드의 예술성 부재의 원인에 대해 Rick Altman은 다음과 같이 말했다. 첫째는 유성영화 초기 감독들이 가지고 있던 사운드의 예술적 가치에 대한 회의, 둘째는 이미지 위주의 용어 개발로 사운드에 대한 관심 약화, 셋째로는 영향력 있는 비평가들의 이미지 위주의 문제인식이 사운드에 무관심해지거나 소홀해지게 만들었다고 분석하였다[5]. 하지만 유성영화의 등장은 예술적 잠재능력을 제시하였고, 관객들이 갈망하던 시각과 청각의 결합을 넘어 소리를 통한 내러티브의 기호화를 실현 가능하게 하였다.

2-2 영상 콘텐츠의 음성적 표현

‘듣지 못하는 영화’(무성영화)에서 ‘발성 영화’(유성영화)로 넘어오면서, 극 중 배우의 음성적 표현을 관객이 들을 수 있다는 것은 다음과 같은 의미를 지닌다. 먼저 영상 콘텐츠 속 등장인물의 음성적 연기를 통해 내러티브를 보다 생생하게 느낄 수 있었다. 무성영화가 촬영될 때에도 배우들은 유성영

화처럼 카메라 앞에서 직접 대사를 하며 연기하였다. 단지 대사가 녹음되지 않았을 뿐이었다. 따라서 관객들은 극에서 표현되는 분위기를 통해 인물의 목소리를 떠올리거나, 자막을 통해 표현되는 대사 지문을 보며 인물의 목소리를 상상하였다. 그렇기 때문에 유성영화 시대로 접어들면서 배우의 발성을 들을 수 있게 된 것이 관객들이 상상한 것과 충돌이 일어났고[6], 몇몇 사람들이 유성영화에 대한 회의감을 가지게 된 계기 중 하나가 되었다. 하지만 관객들은 이것에 익숙해졌고, 이는 곧 유성영화의 흥행으로 이어졌다. 두 번째로 자막으로 표현되었던 대사가 음성으로 전달되면서 관객들에게 정보 전달을 더 신속하게 하였다. 당시 교육환경과 문맹률 등을 고려했을 때, 상황과 행동 연기를 통해 내용을 짐작할 수 있었지만, 자막을 통해 전달되는 정보를 유추하기 어려웠을 것이다. 그래서 한국이나 일본의 경우, 변사가 상황에 대해 설명을 하며, 등장인물의 목소리를 흉내내어 관객에게 전달하였다. 하지만 기술발전으로 음향과 영상을 동기화할 수 있게 되면서 관객들은 청각을 통해 배우의 목소리를 들을 수 있었다. 시각적 정보처리를 매개하는 빛의 속도는 소리의 속도보다 약 88만 배 이상 빠르지만 감각적 반응 측면에서는 청각이 시각보다 더 빠른 특성이 있다[7]. 따라서 관객들은 배우의 대사를 통해 전달되는 의미를 보다 더 빠르게 이해할 수 있었다. 세 번째로 등장인물이나 제3의 인물이 시각적 이미지로 전달하기 힘든 내용을 내레이션(narration) 또는 보이스 오버(voice-over)를 통해 전달하는 역할을 담당하였다. 누군가의 목소리만 등장하여 과거를 회상하거나 시대적 상황, 인물의 묘사 등을 관객에게 전달하기도 한다.

2-3 음량과 인간 감정의 관계성

리액션성 대사의 음량과 시청자 감정의 관련성을 찾기 위해서는 악기론과 같은 음악적인 접근이 필요하다. 음악에서는 연주의 세기를 통해 곡의 분위기와 감정을 전달하기도 한다. 이를 효과적으로 표현하는 악기로는 피아노가 대표적이다. 피아노가 탄생하기 이전에는 쳄발로와 하프시코드를 많이 사용하였는데, 이 악기들이 소리를 내는 방식은 연주자가 건반을 눌렀을 때, 악기 내부에 있는 장치가 현을 뜯어서 소리를 내는 방식이었다. 따라서 연주자가 건반을 누르는 세기가 어떠한 동일한 소리를 냈다. 하지만 이탈리아 쳄발로 제작자인 바르톨로메오 크리스토포리가 쳄발로의 몸체를 개조하여, ‘피아노 포르테(Piano Forte)’라는 악기를 만들었다[8]. 현을 뜯는 방식이 아닌 해머로 현을 때리는 방식으로 제작하였고, 이는 피아노의 원형이 되었다(그림 1). 피아노는 쳄발로나 하프시코드와 달리 건반을 누르는 세기에 따라 강약을 조절할 수 있었기 때문에 셈여림(dynamics)을 폭넓게 표현할 수 있었고, 소리의 크기를 통한 자연스러운 감정 전달을 할 수 있게 되었다. 셈여림은 음악을 표현하고 해석함에 있어 중요한 역할을 하는데, 음악이 진행함에 있어 인간의 감정적 긴장과 이완을 통해 음악적 표현을 깊게 한다.

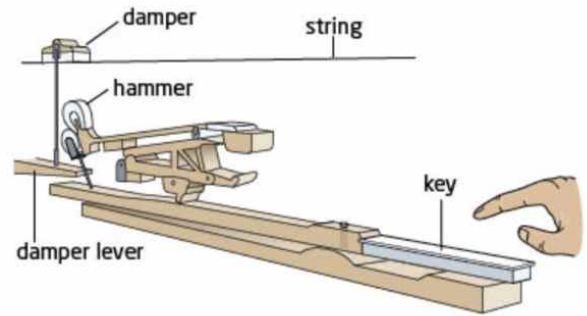


그림 1. 피아노가 소리를 내는 원리와 구조
Fig. 1. The Principle and Structure of Piano Making Sounds

이러한 셈여림의 변화는 음악의 분위기를 결정하는데 큰 영향을 미칠 뿐만 아니라 연주자의 해석에 따라서 음악에 다양한 표정을 담는 중요한 요소가 된다[9]. 셈여림은 기호로써 표현되지만 정확한 수치로 나타나기 보다는 감각적 판단으로 이루어지기 때문이다(표 1). MIDI에서는 벨로시티의 수치로 셈여림을 표현한다. 연주를 하지 않는 0에서부터 가장 여리게 표현하는 1, 가장 세게 연주하는 127까지 셈여림을 표현 가능하다. 셈여림 그리고 벨로시티와 볼륨의 차이점은 연주자의 세기를 반영하는가에 있다. 매우 센 셈여림의 경우 큰 소리를 발생함과 동시에 큰 어택타임(attack time)을 가진다. 반면 볼륨은 소리의 크기만을 표현하며 연주의 세기를 표현하지는 않는다. 즉, 여린 셈여림의 오디오를 센 셈여림의 크기만큼 증폭한다 하더라도 동일한 결과를 얻을 수는 없다. 따라서 볼륨이 크다고 해서 셈여림이 세다고 볼 수 없으나 셈여림은 볼륨과의 관계가 연결되어 있다.

한편, 마준호에 따르면 음악의 소리 크기에 따라 정서 변화에 차이가 있음이 검증되었다[10]. 소리 크기에 따른 정서적 변화는 방송매체에서도 나타났는데, 박덕춘에 따르면 텔레비전으로 드라마를 시청할 때, 볼륨이 큰 음향으로 시청한 시청자가 볼륨이 작은 음향으로 시청한 시청자보다 몰입도가 더 높았다[11]. 하지만 이 결과들은 음원 및 드라마 오디오가 마스터링이 된 결과를 토대로 연구된 것이기 때문에 리액션성 대사의 크기에 따른 시청자 감정의 영향을 분석하기에는 한계가 있다.

표 1. 셈여림 표
Table 1. Dynamics marks

Original	Abbreviation	Meaning
Pianissimo	pp	very softly
Piano	p	softly
Mezzo Piano	mp	moderately softly
Mezzo Forte	mf	moderately loudly
Forte	f	loudly
Fortissimo	ff	very loudly

따라서 본 연구에서는 내러티브가 있는 영상에서 표현되는 음향 요소 중 리액션성 대사의 음량이 시청자에게 미치는 감정적 영향을 탐색하고자 한다.

Ⅲ. 평가실험: 리액션성 대사의 음량수준에 따른 영상 콘텐츠 감성평가

3-1 실험 참가자

본 연구의 평가실험 참가자는 22명의 일반인(남성: 7명, 여성: 15명; 평균 28.3세)으로 영상 콘텐츠 감성평가에 필요한 시각 및 청각 기능이 정상임과 감성평가 대상인 영상콘텐츠에 대한 사전경험이 없음을 확인한 후 참여하였다.

3-2 평가방법 및 절차

본 연구의 평가실험은 영상 콘텐츠 속 리액션성 대사의 음량 수준에 따른 감성평가 차이를 검증하기 위해 독립변인인 리액션성 대사의 음량을 세 가지 수준으로 조건화하여 그에 따른 영상 콘텐츠 감성평가 점수를 통계적으로 비교 분석하고자 하였다.

실험설계에서 독립변인인 리액션성 대사의 음량수준 세 조건은 첫 번째로 영상 콘텐츠에서 리액션성 대사가 제거된 조건(요소없음)으로, 이는 리액션성 대사에 의한 영상 콘텐츠의 감성평가를 이해하는 기준(baseline)이 된다. 두 번째는 영상 콘텐츠의 음향적 요소를 최적화하는 작업인 사운드 디자인 과정이 완료된 음량 수준으로 국내방송 표준이 적용되는 -24 LUFS를 마스터 오디오 레벨로 설정하여 사용하는 조건(기본 음량)이다[12]. 세 번째는 기본음량에서 리액션성 대사의 음량을 증폭하여 영상 콘텐츠의 감성평가에 미치는 영향을 살펴보기 위한 6dB(데시벨) 증가 조건인데, 소리의 특성상 물리적 크기 변화보다는 청각적 감각기준으로 두 배 크기의 소리를 만들어주는 6dB 증가를 통해 충분한 크기 변화를 인식 시키기 위함이다[13].

표 2. 집단별 영상 시청 분류

Table 2. Classification of video viewing by group

Group	Video A			Video B			Video C		
	Non	Basic	+6dB	Non	Basic	+6dB	Non	Basic	+6dB
A	○				○				○
B		○				○	○		
C			○	○				○	

또한 실험설계에서 종속변인에 해당하는 영상 콘텐츠의 감성평가는 실제 평가대상인 드라마에 포함된 등장인물들 사이의 액션씬에서 발생하는 흥분감과 고통감정의 정도를 10점 척도로 평가하였다. 이와 함께 다양한 콘텐츠에 대한 감성평가에서 기본적으로 감성특성을 이해하는데 도움을 주는 Russell의 감정차원모델(circumplex model)과 Thayer의 감정모델(2D emotion model) 감정가-각성수준(valence-arousal) 2차원 감성모델에 근거한 감성평가(각 차원별 5점 척도)를 진행하였다[14,15].

평가실험에서는 감성평가의 대상인 영상 콘텐츠로서 몇 가지 리액션성 대사를 포함하고 있는 국내 드라마 콘텐츠를 선정하였다.



그림 2. 평가실험에 사용된 영상 콘텐츠(드라마 작품)의 스크린샷; SBS의 굿캐스팅(위), 배가본드(가운데), JTBC의 언더커버(아래)

Fig. 2. Screenshots of video content(drama) used in the evaluation; SBS's 'Good Casting'(top), 'Vagabond'(middle), JTBC's 'Undercover'(bottom)

평가실험에서 동일한 참가자에게 3가지 수준의 음량을 평가 받아야하므로 같은 영상 콘텐츠의 반복 시청의 문제가 발생하는데, 이를 해결하기 위해 표 2와 같이 서로 다르지만 유사한 리액션성 대사가 포함된 3개 드라마를 이용하여 참가자별로 조건과 영상의 짝(pair)을 달리하는 방식, 즉 라틴스퀘어(Latin Square) 설계 방식으로 진행하였다. 선정된 영상 콘텐츠(드라마)는 세 가지 영상 모두 액션씬(action scene)이 포함된 장면으로, 다른 장면에 비해 빠른 움직임이나 강한 힘을 사용하는 부분이 많아서 배우의 비언어적 대사인 리액션성 대사가 비교적 잘 표현되었다. 선정된 구체적인 작품 정보는 SBS의 ‘굿캐스팅’과 ‘배가본드’, 그리고 JTBC의 ‘언더커버’ 등이다(그림 2 참조). TV 드라마는 극의 내러티브를 진행하기 좋은 매체이며 비교적 쉽게 접할 수 있을 뿐만 아니라, 인간의 감정적 반응을 자연스럽게 이끌어 낼 수 있으며, 유사한 장면이 가지는 시청자 감정의 변화를 파악하기에 적합한 것으로 판단되었다. 다만 평가실험에서는 드라마의 일부만 시청하기 때문에 참가자가 영상 콘텐츠의 맥락(context)을 파악하기 어려울 수 있고, 그 결과로 내러티브 요소가 전달되지 않을 수 있다. 그래서 본 연구에서는 참가자가 영상 콘텐츠를 시청하기 전 내용 맥락을 파악하는데 도움을 주는 줄거리 설명 정보를 제공하였다. 또한 앞서 소개한 바와 같이 본 연구의 목적이 영상 콘텐츠의 여러 가지 음향적 요소 가운데 리액션성 대사의 음량만을 주요 변수로 고려하기 때문에 평가대상인 영상 콘텐츠의 음향적 요소 중 배경음악이나 언어적 대사 등은 제외하였다.

평가실험 참가자의 영상 콘텐츠 시청환경은 LED TV와 기준음량 79dB로 설정하며, 시청환경의 음량측정은 1kHz 및 C-Curve를 적용하였다. 이는 시각적인 측면에서는 일반적인 환경이지만, 청각적인 측면에서는 사운드 디자이너가 음향 제작을 할 때 크기와 유사한 환경에서 참가자에게 들려주기 위함이다. 참가자의 실험 참여과정은 다음의 순서와 같았다.

- ① 실험에 대한 설명과 실험참험자의 참가 동의 확인
- ② 영상 시청 전 참가자에게 A영상에 대한 줄거리 설명
- ③ 영상 시청 및 A영상에 대한 감성평가 진행(집단마다 다른 조건이 적용된 영상 시청)
- ④ 영상 시청 전 참가자에게 B영상에 대한 줄거리 설명
- ⑤ 영상 시청 및 B영상에 대한 감성평가 진행(집단마다 다른 조건이 적용된 영상 시청)
- ⑥ 영상 시청 전 참가자에게 C영상에 대한 줄거리 설명
- ⑦ 영상 시청 및 C영상에 대한 감성평가 진행(집단마다 다른 조건이 적용된 영상 시청)
- ⑧ 평가실험 종료 및 참가자 퇴실 안내

IV. 실험결과 및 분석

4-1 영상 콘텐츠의 내러티브 감성에 대한 공감 결과 분석

영상 콘텐츠의 내러티브가 포함하는 감성에 대한 공감(흥분감정, 고통감정) 평가결과 분석절차는 리액션성 대사의 음량조건(요소없음, 기본음량, 6dB증가)을 세 가지 수준으로 하는 반복측정 변량분석(ANOVA)을 통해 진행되었다.

그림 3과 4의 결과 그래프에서 볼 수 있듯이, 참가자들은 전반적으로 리액션성 대사가 없을 때보다는 기본음량 이상으로 있을 때 영상 콘텐츠의 내러티브 감성인 흥분감정(excited feeling)을 더 크게 평가하였다($F(2,42)=4.385, p=.019$).

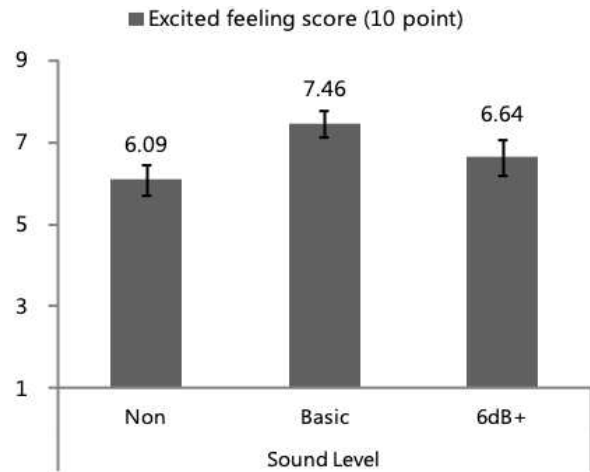


그림 3. 리액션성 대사의 음량조건(요소없음, 기본음량, 6dB증가)에 따른 흥분감정 평가점수 분석결과
 Fig. 3. The statistically analyzed results of the excited feeling score according to the sound level conditions(Non, Basic, 6dB+) of reactive dialogue

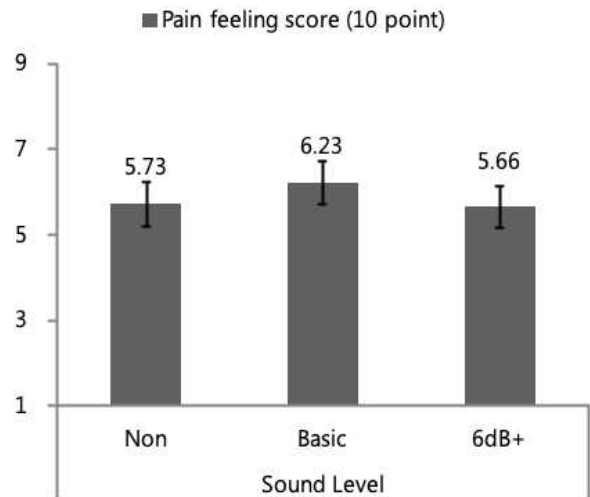


그림 4. 리액션성 대사의 음량조건(요소없음, 기본음량, 6dB증가)에 따른 고통감정 평가점수 분석결과
 Fig. 4. The statistically analyzed results of the pain feeling score according to the sound level conditions(Non, Basic, 6dB+) of reactive dialogue

이러한 차이는 요소없음($M=6.09$)과 기본음량($M=7.46$) 사이에서 특히 통계적 차이가 나타났다($t=2.942, p=.016$). 그러나 고통감정(pain feeling)에 대한 평가에서는 리액션성 대사의 음량조건 세 가지 수준에 따른 차이가 통계적으로 유의미하지는 않았다($F(2,42)=0.490, p>.05$). 이와 같은 결과는 리액션성 대사의 음량조건이 영상 콘텐츠의 내러티브 감성 측면에서 등장인물의 고통에 대한 공감 보다는 내러티브의 분위기적 요소로서 흥분적 상황에 대한 공감이 커지는데 역할을 하는 것으로 이해될 수 있다.

4-2 기본 감성차원에 따른 감정가/각성수준에 대한 결과 분석

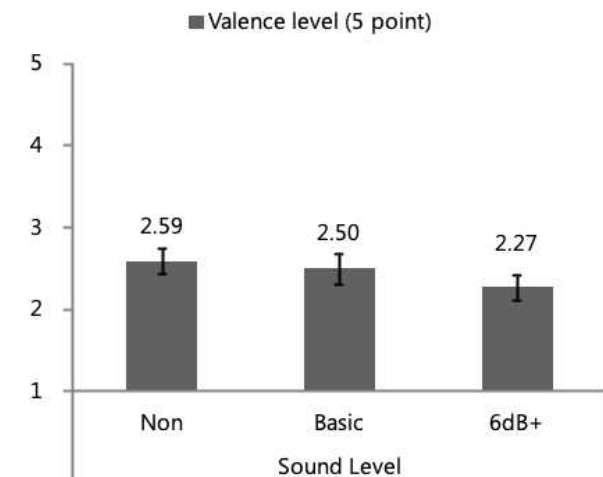


그림 5. 리액션성 대사의 음량조건(요소없음, 기본음량, 6dB증가)에 따른 감정가(유쾌/불쾌) 평가수준 분석결과
 Fig. 5. The statistically analyzed results of the valence level according to the sound level conditions(Non, Basic, 6dB+) of reactive dialogue

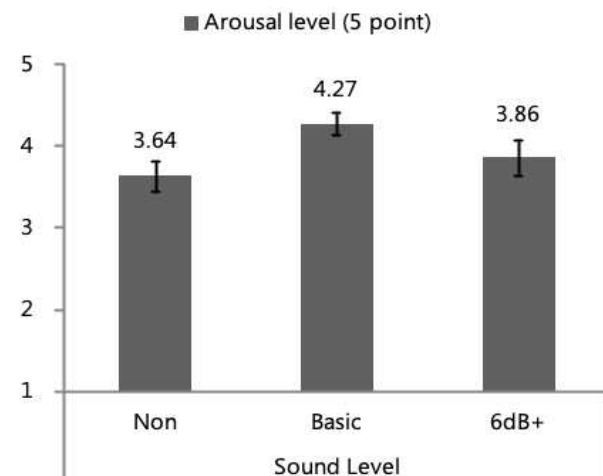


그림 6. 리액션성 대사의 음량조건(요소없음, 기본음량, 6dB증가)에 따른 각성 평가수준 분석결과
 Fig. 6. The statistically analyzed results of the arousal level according to the sound level conditions(Non, Basic, 6dB+) of reactive dialogue

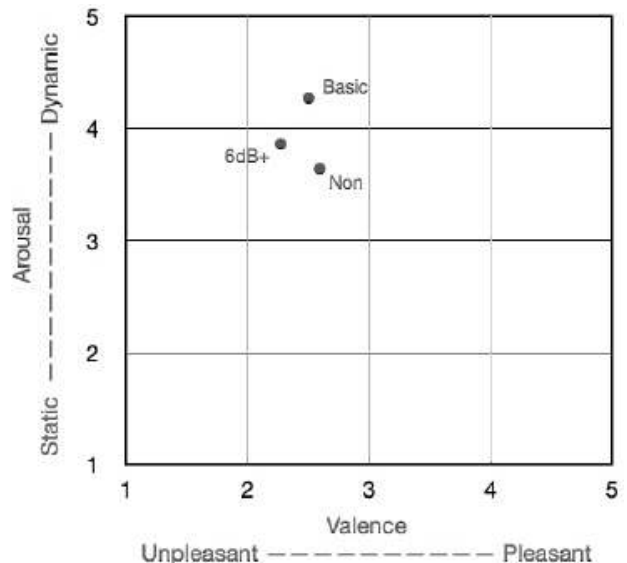


그림 7. 기본적인 감성모델의 2차원(감정가-각성수준) 위 리액션성 대사의 음량조건(요소없음, 기본음량, 6dB증가)의 공간 포지셔닝
 Fig. 7. The positioning of the sound level conditions(Non, Basic, 6dB+) of reactive dialogue on basic 2D emotion model

평가대상인 영상 콘텐츠에 대한 전반적인 감성특성을 평가하기 위한 감정가-각성수준(valence-arousal) 2차원 모델 감성평가 결과분석 또한 리액션성 대사의 음량조건(요소없음, 기본음량, 6dB증가)을 세 가지 수준으로 하는 반복측정 변량 분석(ANOVA)을 통해 진행되었다.

먼저 그림 5의 결과 그래프를 살펴보면, 감성차원의 감정가(유쾌/불쾌)에 대한 평가결과에서는 리액션성 대사의 음량조건 세 가지 수준에 따른 통계적 차이가 나타나지 않았다 ($F(2,42)=1.675, p>.05$).

즉 영상 콘텐츠에 대한 시청자의 감정가(유/불쾌)는 리액션성 대사의 음량조건에 영향을 받지 않은 것으로 보인다. 반면 그림 6의 결과 그래프에서 확인할 수 있는 바와 같이, 감성차원의 각성수준(정적/동적)에 대한 평가결과에서는 리액션성 대사의 음량조건 세 가지 수준에 따른 통계적 차이가 유의미하게 나타났다($F(2,42)=4.783, p=.013$). 즉 참가자들이 리액션성 대사의 음량에 영향을 받아 영상 콘텐츠에 대한 각성수준을 상대적으로 높게 평가한 것이다. 이러한 차이는 요소없음($M=3.64$)과 기본음량($M=4.27$) 사이에서 특히 통계적 차이가 나타났다($t=3.052, p=.012$). 이러한 결과는 앞선 영상 콘텐츠의 내러티브 감성에 대한 공감 측면에서 나타난 흥분감정의 공감효과와 일맥상통하는 것으로 이해될 수 있다. 즉 영상 콘텐츠 속 등장인물들의 상황적 분위기에 공감하는 흥분감정의 고조가 영상 콘텐츠에 대한 전반적인 감성평가로서의 각성수준과 일치되어 리액션성 대사의 음량조건에 따라 달리 영향 받았다는 것이다.

그림 7은 기본적인 2차원 감성모델에서 리액션성 대사의 음량조건 세 수준이 어떤 공간적 위치를 갖는지 확인하고 비교할 수 있는 포지셔닝(positioning)의 결과이다. 즉 감정가(valence)와 각성수준(arousal)의 두 가지 차원을 동시에 고려하여 리액션성 대사의 음량조건이 갖는 효과 차이를 시각적으로 확인할 수 있는 것이다. 이를 통해 앞서 논의한 바와 같이, 리액션성 대사의 음량이 갖는 효과는 가로축인 감정가(valence)에서 보다는 세로축인 각성수준(arousal)에서 확인되었다.

V. 결 론

본 연구에서는 영상 콘텐츠의 음향적 요소 가운데 등장인물의 비언어적 대사로써 리액션성 대사가 영상 콘텐츠의 내러티브 감성에 대한 공감과 전반적인 감성평가에서 어떤 영향 관계에 있는지를 탐색하기 위해 서로 다른 세 가지 수준의 음량조건에서 시청자로서의 참가자가 느끼는 감성을 평가하였다. 평가실험의 전반적인 결과는 리액션성 대사의 음량조건 세 수준에 따라 영상 콘텐츠의 흥분적 분위기에 대한 공감과 각성수준의 감성차원에서 증가하는 영향 관계가 나타났다. 반면, 등장인물들의 고통감정이나 감정가의 감성차원에서는 그 영향이 크게 나타나지 않았는데, 이러한 결과는 리액션성 대사의 본질적인 특징이 감정의 다양성 보다는 강약이나 깊이 차원과 연결된 것으로 추측해볼 수 있다. 유/불쾌를 포함한 감정의 다양성 측면에서는 등장인물의 비언어적 리액션성 대사보다는 전통적인 음향요소로 사용되는 배경음악이나 구체적인 정보를 담은 언어적 대사 등이 그 역할을 담당하는 것으로 이해될 수 있을 것이다.

제한적이지만 본 연구의 결과를 통해 이해할 수 있는 것은 일반적으로 언어적 대사로 등장인물의 감정을 시청자에게 전달하지만, 리액션성 대사의 음량으로도 시청자의 공감을 이끄는 감정 전달이 어느 정도 가능하다는 것을 확인하였다. 따라서 작품의 장르에 따라 사운드 디자인 제작 과정에서 어떠한 방식으로 감정을 전달할지에 대해 연출자와 음향감독이 선택할 수 있는 가능성이 확인된 것이다. 특히 평가대상 영상 콘텐츠와 같은 액션 장르에서 시청자의 감정변화나 공감을 위해 리액션성 대사의 음량을 적절하게 활용할 수 있을 것이다.

그리고 사운드 포스트 프로덕션의 비중이 소폭 증가할 수 있다. 프로덕션 단계에서 연출자의 창작의도에 의해 주로 이미지 위주로 촬영이 진행된다. 대표적으로 강한 액션이 포함된 장면을 촬영할 때를 예시로 들면, 빠른 움직임을 담아내기 위해 시각적인 비중을 높이고 청각적인 요소를 적게 생각할 수 있다. 이 과정에서 등장인물의 리액션성 대사가 움직이는 소리, 스텝들의 소리 등에 의해 제한될 수 있으며, 등장인물의 연기가 좋아도 포스트 프로덕션 단계에서 비중 있게 살려낼 수 없는 경우가 있다. 하지만 연출자 또는 음향감독이 내러티브 진행과 감정전달을 위해 리액션성 대사에 대한 ADR(automatic dialogue replacement), 즉 후시녹음을 제

안하여 작품의 질을 높일 수 있다. 다만, 본 연구의 평가실험에서는 액션 장면과 강한 리액션 위주로 다루었기 때문에 다른 장르와 비교적 작은 음량의 리액션성 대사에서 어떻게 적용할 수 있는지는 추가적으로 검토해야 할 것이다.

드라마와 같은 영상 콘텐츠뿐만 아니라 게임 콘텐츠 분야에서도 리액션성 대사는 적극 활용될 수 있다. 특히 FPS, RPG 장르의 게임에서 캐릭터 설정과 성우들의 캐릭터 연구 등을 기반으로 대사를 녹음하여 적용하고 있다. 따라서 녹음된 대사를 최종적으로 게임에 접목할 경우 밸런스가 무너지지 않는 선에서 리액션성 대사의 레벨을 높여 플레이어의 심리상태를 변화시킬 수도 있을 것이다. 또한 TV 드라마 음량의 경우 디지털 텔레비전 방송프로그램 음량 등에 관한 기준에 의해 제한되고 있다. 이러한 측면에서 리액션성 대사의 음량을 무조건적으로 높이기보다는 해당 법령에 맞게 적용하며 사운드 믹스 방법도 그에 맞게 적용할 필요가 있다.

끝으로 본 연구에서 포함한 평가실험은 평가대상인 영상 콘텐츠로서 다양한 장르를 다루지 못한 한계를 지닌다. 장르가 지니는 특성으로 인해 포함된 리액션성 대사의 종류가 다를 수 있기 때문에 멜로, 코미, 스릴러 장르 등에서 어떤 영향 관계가 나타나는지를 비교 분석할 필요가 있을 것이다. 또한 COVID-19 확산이라는 외부적 요인으로 사회적 거리두기 단계가 상향됨에 따라 오프라인 현장에서 진행되는 평가실험 참가자 모집 및 평가진행 등에 어려움이 있었다. 추후 후속 연구가 진행된다면 연구내용과 범위, 그리고 진행과정 등 여러 요인들을 적절하게 고려한 연구진행이 필요할 것이다.

감사의 글

본 연구는 2021년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2020S1A5A2A03045921).

참고문헌

- [1] D. Shin and F. Biocca, "Exploring immersive experience in journalism," *New Media & Society*, Vol. 20, No. 8, pp. 2800-2823, August 2018.
<https://doi.org/10.1177/1461444817733133>
- [2] D. Shin, "How does immersion work in augmented reality games? A user-centric view of immersion and engagement," *Information, Communication and Society*, Vol. 22, No. 9, pp. 1212-1229, July 2019.
<https://doi.org/10.1080/1369118X.2017.1411519>
- [3] D. Shin, "How do users experience the interaction with an immersive screen?" *Computers in Human Behavior*, Vol. 98, pp. 302-310, March 2019.
<https://doi.org/10.1016/j.chb.2018.11.010>

- [4] D. H. Lee, "Narrative functions of sound design in films," *The Journal of Korea Contents Association*, Vol. 13, No. 12, pp. 626637, December 2013.
<https://doi.org/10.5392/JKCA.2013.13.12.626>
- [5] R. Altman, *Sound Theory and Sound Practice*, Routledge, 1992.
- [6] M. Chion, *The Voice in Cinema*, Columbia University Press, 1999.
- [7] S. Horowitz, *The Universal Sense: How Hearing Shapes the Mind*, Bloomsbury USA, 2013.
- [8] D. S. Kim, and G. W. Park, *Instrumentation*, Sekwang Music Publisher, pp. 125-130, 1988.
- [9] C. K. Kim, S. Y. Lee, Y. J. Lee, J. J. Jeong, and W. Y. Choi, Dynamics in music therapy, in *Counseling Dictionary*, Hakjisa, 2016.
- [10] J. H. Ma, A study of relations between music and trance-based on loudness and emotional change, Master's thesis, Sangmyung University, Seoul, 2017.
- [11] D. C. Park, "Effect of watching environment of TV on audience's engagement," *Journal of Communication Design*, Vol. 50, pp.126-132, January 2015.
- [12] International Telecommunication Union, Recommendation ITU-R BS.1770: Algorithms to measure audio programme loudness and true-peak audio level, 2015.
- [13] S. H. Kang, *Sound Production Technology*, Sound Media, 2013.
- [14] J. A. Russell, "A circumplex model of affect," *Journal of Personality and Social Psychology*, Vol. 39, No. 6, pp. 1161-1178, December 1980.
<https://doi.org/10.1037/h0077714>
- [15] R. E. Thayer, *The Biopsychology of Mood and Arousal*, Oxford University Press: New York, NY, USA, 1990.



최형민(Hyung-Min Choi)

2018년 : 한국영상대학교 (음향제작학 학사)
2021년 : 서울미디어대학원대학교 (융합미디어전공 석사과정)

2019년~현 재: (주)모비사운드 Dialogue Editor, ADR Recordist
2018년~현 재: 서울미디어대학원대학교 융합미디어학과 석사과정
※관심분야: 대사 연출(Dialogue Direction), 사운드 편집(Sound Edit), 몰입형 사운드(Immersive Sound) 등



이주환(Ju-Hwan Lee)

2003년 : 연세대학교 대학원 (인지공학석사)
2007년 : 연세대학교 대학원 (인지공학박사-HCI)

2000년~2007년: 연세대학교 인지과학연구소 연구원/전문연구원
2007년~2009년: 영국 옥스퍼드대학교 Crossmodal Research Lab 박사후연구원
2009년~2010년: 성균관대학교 인터랙션사이언스학과 연구교수
2010년~현 재: 서울미디어대학원대학교(SMIT) 뉴미디어학부 융합미디어전공 부교수
※관심분야: 다중감각 사용자 인터페이스(Multisensory User Interfaces), 인간-컴퓨터 상호작용(HCI), 가상 & 증강현실 인터랙션(VR & AR Interaction) 등