Journal of the Ergonomics Society of Korea
Vol. 25, No. 2 pp.63-70, May 2006

연결형 합성음성을 이용한 경보음의 주관적 위급도 정량화

장 필 식[1] · 이 경 태[2]

[1]대불대학교 컴퓨터공학과 / [2]대불대학교 산업정보공학과

Quantifying the Urgency Perception of Voice Alarm Generated by Concatenative Synthesizer

Phil Sik Jang[1], Kyung Tae Lee[2]

[1]Department of Computer Engineering, Daebul University, Youngam, 526-702
[2]Department of Industrial Engineering, Daebul University, Youngam, 526-702

ABSTRACT

This paper presents an experimental study of the factors modulating the urgency perception of voice alarm generated by concatenative synthesizers. Four experiments were conducted using psycho-physical approach in which 105 participants made magnitude estimation for urgency perception of various voice alarm stimuli. Experiment 1 identified 6 acoustic and non-acoustic factors modulating the perceived urgency of synthesized voice alarm. Experiment 2, 3 and 4 quantified the relations between the objective changes in each of the quantifiable parameters and the subjective changes in urgency perception. This research has implications for the design and implementation of synthesized voice alarm systems where urgency mapping is required.

Keyword: Concatenative synthesis, Voice alarm, Urgency perception, Magnitude estimation

1. 서 론

전자공학, 음향공학의 발달과 인간공학적 요구의 증가에 따라 청각적 경보신호(alarm, warning)는 경적(horn), 호각(whistle), 사이렌(siren) 등의 간단한 경보음을 넘어, 좀 더 정교하고 다양한 형태로 발전되고 있다. 특히, Patterson (1982)이 제시한 청각적 경보신호 디자인에 대한 인간공학적 가이드라인 이후로, 청각적 경보신호의 디자인과 이용에 대한 다양하고 광범위한 연구들이 진행되어 오고 있다(Bliss et al., 1995; Burt et al., 1995; Edworthy et al., 1991; Haas and Edworthy, 1996).

이들 연구들은 청각적 경보신호의 평가, 설계에 대한 인지심리적 접근방식을 취하고 있으며, 특히 청각적 경보신호에 대한 주관적 위급도(urgency perception)를 주로 다루고 있다. 즉, 청각적 경보신호의 음향학적, 비음향학적 요소들과 이 경보신호를 청취하였을 때 인간이 주관적으로 인지하는 위급한 정도의 관계를 규명하고, 이를 이용하여 상황에 적합한 정도의 위급함을 전달 할 수 있는 경보신호를 설계하자는 것이다. 이렇게 위급함에 대한 정보를 암호화한 청각적 경보신호는 상황에 맞는 행동을 유도할 수 있으며, 경보에 대한 응답시간을 단축시킬 수 있다(Morris and Montano, 1996).

이러한 청각적 경보음 개선에 대한 시도 이외에, 컴퓨터를

교신저자: 장필식
주 소: 526-702 전남 영암군 삼호읍 72-1, 전화: 061-469-1288, E-mail: phil@mail.daebul.ac.kr

이용한 디지털방식의 합성음성(synthesized voice)을 경보신호로 이용하고자 하는 연구들도 진행되어 오고 있다. 음성을 경보신호에 이용할 경우, '정보전달'이라는 부가적인 기능을 이용할 수 있으며(Barber et al., 1992), 합성음성은 그 자체로서 '경보' 기능을 가지고 있다(Byblow and Corlett, 1989). 따라서 합성음성 경보는 인간에게 친숙한 형태로 경보와 정보의 전달이 한 번에 가능하며, 대부분의 경우, 사전 학습이 불필요 하다. 이러한 장점 때문에 합성음성 경보는 전투기 조종석과 같은 높은 정신부하가 부과되는 상황에 적합하며(Simpson, and Marchionda-Frost, 1984), 조종사의 작업부하를 감소시킬 수 있다는 연구 결과도 있다(Aretz, 1983).

하지만, 이러한 잠재적인 장점에도 불구하고, 실제 응용에 있어 기존의 청각적 경보신호에 비교하여 효용성이 있는지에 대해서는 여러 가지 상반된 연구 결과들이 발표되고 있다. 합성음성 경보 앞에 주의를 환기 시킬 수 있는 경보음(alerting tone)을 발생 시키는 것 보다는 합성음성 자체를 이용하는 것이 더 효과적이며, 피실험자의 응답시간을 단축시킨다는 연구 결과가 있는 반면(Simpson and Williams, 1980; Hakkinen and Williges, 1984; Byblow and Corlett, 1989), 정교한 디자인과 훈련을 이용하면 경보음의 이용이 응답시간과 수행도 측면에서 더 이득이라는 연구도 있다(Wheale, 1983). 합성음성 경보와 경보음(tone warning)의 비교연구에 있어서도, Irving(1981)은 합성음성 경보가 경보음의 경우보다 조종사들의 응답시간을 단축시키며, 정확한 응답을 이끌어 낸다고 주장한 반면, Wheale(1983)은 합성음성 경보가 다른 형태의 청각적 경보신호에 비해 응답속도가 느리며, 따라서 기존의 경보음을 계속 사용하여야 하고 합성음성 경보는 실제 활용을 위한 연구가 더 필요하다고 주장하였다.

합성음성 경보와 주관적 위급도 사이의 관계에 대한 연구는 Park and Jang(1999)에 의해 처음 시도되었는데, 간단한 규칙기반 음성합성기(rule-based synthesizer)를 이용하여 생성한 합성음성 경보의 음향학적 파라미터 중에서 음성경보의 빠르기, 평균기초진동 수 그리고, 억양(intonation)과 같은 음성의 자연스러움 등이 주관적 위급도에 영향을 줄 수 있음을 밝혔다. Park and Jang(1999)의 연구에 이용된 규칙기반 음성합성은 인간의 발성구조를 모델링하고 음소나 음절의 억양과 악센트 등에 여러 가지 규칙을 적용하여 음성을 만들어 내는 방식으로, 음향학적 파라미터들을 쉽게 조정할 수 있으며, 다양한 어휘를 만들어 낼 수 있다는 장점이 있다. 하지만, 최근 대부분의 TTS system들은 연결형 합성(concatenative synthesis)방식을 이용하고 있는데, 그 이유는 미리 녹음, 데이터베이스화된 음절들을 연결하여 합성음성을 만들어냄으로써, 어휘의 수가 규칙기반 음성합성에 비하여 제한되지만 고음질의 더 자연스러운 음성을 생성할 수 있기 때문이다.

Park and Jang(1999)의 연구에서 지적한 바와 같이, 합성음성의 자연스러운 정도는 인간의 주관적 위급도에 영향을 미칠 수 있다. 따라서 사람의 음성에 훨씬 가까우며, 이해도와 명료도가 높은 연결형 합성음성에 대한 주관적 위급도 관계는 새로이 평가되어야 할 것이다. 또한, 주관적 위급도를 합성음성 경보음에 암호화 하기 위해서는 다양한 어휘와 경보메시지 형태 등, 좀 더 다양한 경보음 설계요인들에 대한 평가가 선행되어야 한다. 본 연구에서는 4개의 실험을 통해, 1) 연결형 합성(concatenative synthesis)방식의 합성음을 이용하여 경보음을 제작하기 위해 고려되어야 할 여러 가지 요인들을 파악하고, 2) 경보음을 청취한 후 청취자가 주관적으로 인지하는 위급도 사이의 관계를 규명하고, 이를 정량화하고자 한다.

2. 실험 1

실험 1의 목적은 합성음성의 음향학적 특성과 여러 가지 요인들 중 합성음성 경보자극이 주어졌을 때, 청취자가 주관적으로 인지하는 위급도에 영향을 주는 요인들을 찾아내는 것이다.

2.1 예비실험

예비실험의 목적은 피실험자들의 비율판단 능력을 파악하는 것이다. 본 연구에서 수행된 모든 실험에서, 피실험자가 인지하는 위급도에 대한 주관적 평가는 magnitude estimation 기법을 이용하여 측정하였다. Magnitude estimation 기법은 인간이 외부자극에 대한 인지강도를 비율로 판단할 수 있고, 적절한 형태, 즉 선의 길이나 수치로 표현 가능하다는 가정을 바탕으로 한다. 따라서 본 실험에 앞서 피실험자들을 대상으로 비율판단(ratio judgment)능력에 대한 검증작업을 실시하였다.

실험의 절차는, Lodge(1981)와 Tanenhaus and Murphy (1981) 등이 사용한 실험절차와 지시문을 상황에 맞도록 수정하여 사용하였으며, 선을 긋고, 수치로 평가함으로써 비율판단을 하도록 하였다. 피실험자에게 수치를 제시하고 이에 적합한 길이의 선을 그리도록 함으로써 비율판단을 하는 경우, 이 측정된 선의 길이와 수치자극을 Log 변환하여 선형회귀분석을 실시하면, 회귀식의 기울기는 1에 가까운 값을 가져야 한다. 마찬가지로 선을 제시하고 피실험자로 하여금 수치로 선의 길이를 평가하도록 하는 경우에도 회귀식의

기울기는 1에 가까운 값을 가져야 하며, 1에 가까운 값을 가지지 않으면 피실험자가 비율판단 능력이 떨어지거나, 지시사항을 제대로 이행하지 않는 것으로 판정할 수 있다. 본 실험에서는 회귀식의 기울기가 1 인지를 유의수준 5%에서 검정하였으며, 여기에 합격된 피실험자들의 결과만을 이용하여 실험 결과분석을 행하였다. 또한 합격된 피실험자들의 비율판단 경향을 파악하여 본 실험에서의 평가치들을 통합, 보정하였다. 이 두 가지 회귀식의 계수는 본 실험의 결과분석 시, 그 피실험자의 주관적 위급도 평가치 들을 통합, 보정하는 데에도 이용하였다.

2.2 피실험자

26명의 피실험자가 실험에 참여하였으며, 이들 중 앞에서 언급한 기준을 통과한 23명이 선정되었다. 23명의 피실험자는 19세에서 40세(평균 27.2세, 표준편차 5.13)사이의 남자 13명, 여자 10명으로 이루어졌다.

2.3 실험장비

실험에 이용된 합성음성은 연결형 합성(concatenative synthesis) 방식을 이용하는 Rhetorical rVoice TTS (Text-To-Speech) 시스템과 AT&T Natural Voices를 이용하여 생성하였으며, 16kHz sampling rate, 16bit mono WAV file format으로 저장하였다. 저장된 합성음성은 Praat 4.2.29를 이용하여 여러 가지 실험수준으로 변형되었으며, Pentium 4 PC에 장착된 Philips Seismic Edge PSC705 sound board와 Philips SBC3375 헤드폰을 통하여 피실험자에게 제시되었다.

2.4 실험자극

실험 1에서는 6가지 요인 즉, 합성음성의 음높이를 나타내는 평균 피치(pitch), 합성음성의 빠르기(speech rate: 발성속도), 경보메시지 형태(message format), 메시지의 의미, 합성음성의 성별, 반복되는 경보메시지들 간의 시간 간격 등의 요인들을 평가하였다. 각 요인들의 수준은 표 1과 같이 각각 2개 수준이다.

합성음성은 Rhetorical rVoice TTS 시스템과 AT&T Natural Voices 시스템에서 사용되는 전형적인 남자음성과 여자음성 각 1개씩 총 4개의 화자(speaker)를 선택하여 생성하였다. 합성음의 빠르기(speech rate)는 초당 발음되는 음절(syllable)를 나타내는 것으로, 메시지 음절 수를 발성시간으로 나눈 값이다. 실험수준은 음성합성 시스템에서 생성된 합성음성 빠르기의 85% 및 115% 수준이다. 이 두 가지 수준에 맞추기 위해 Praat 4.2.29의 PSOLA(Pitch Synchronous Overlap And Add) 알고리즘을 이용하여 각 메시지의 발성시간을 조정하였다.

합성음성의 평균 피치는 autocorrelation method(Borsma, 1993)를 이용하여 측정되었으며, 피치 실험수준은 Praat 4.2.29를 이용하여 조정하였다. 조정된 평균 피치 진동 수들의 최대 차이 값은 0.2Hz이다. 80dB의 일정한 크기 순음(pure tone)의 음높이 인지에 있어서, 변화감지역(JND: Just Noticeable Difference) 값은 2Hz 이상(Roederer, 1973)이라고 알려져 있다. 본 연구에 사용된 합성음은 순음이 아닌 복합음이기 때문에 단순비교는 어렵지만, 피치 진동 수 차이 0.2Hz는 변화감지역 미만이라고 가정하였다.

표 1. 실험 1의 실험요인과 수준

실험 요인	수준
합성음 화자(speaker)의 성별	남* 여
평균 피치(pitch)	120Hz* 190Hz
합성음의 빠르기 (speech rate)	4.98 syllables/sec* 3.62 syllables/sec
경보메시지 형태 (message format)	Semantic context* Keyword
경보메시지의 의미	Collision traffic one o'clock* Cabin pressure dropping
반복되는 경보메시지 사이 시간 간격	0.1sec* 0.8sec

* 평가기준 경보음(reference stimuli)

본 실험에서 사용된 경보메시지는 표 2와 같이 가상적인 비행기 조종석 환경을 대상으로, Simpson and Williams (1980)의 실험에 이용되었던 합성음성 경보메시지들 중 두 가지 의미와 메시지 형태(format) 두 가지를 이용하였다. 피실험자들의 실제 생활 환경과는 차이가 있는 비행기 조종석 환경에서의 경보메시지를 대상으로 하므로, 실험 전 피실험자에게 경보가 울리는 위급상황에 대해 충분히 설명하였다.

합성음성 경보음의 크기(amplitude)는 위급도 평가에 영향을 미칠 것으로 예측되지만, 본 연구에서는 제외하였다. 그 이유는 경보음의 크기는 경보음의 이해도와 연관되기 때문에 실험상황이나 실제 응용에 있어 조정할 수 있는 범위가 극히 제한적이기 때문이다. 이해도를 떨어뜨리는 작은 음 크기도 문제이지만, 경보음이 너무 큰 경우, 작업이나 위급한 상황을 처리하는데 방해가 되기 때문에 항공기, 병원 등에서 작업자, 조종사, 운전자가 경보시스템을 꺼버리고 다시 켜지 않는 경향이 있음이 관찰되었다(Rood et al., 1985; Kerr and Hayes, 1983).

표 2. 경보메시지 의미와 형태

	메시지 형태	
	Semantic Context	Keyword
메시지 의미	Collision Traffic One O'clock	Traffic One O'clock
	Cabin pressure Dropping	Pressure Dropping

2.5 실험절차

각 피실험자에는 실험 전, 컴퓨터를 이용한 실험진행과 magnitude estimation을 이용한 평가 방법에 익숙해 질 수 있도록, 충분히 설명하였으며, 여유를 가지고 미리 연습할 수 있도록 하였다. 실험지시문은 Mashour and Hosman (1968)의 지시문을 본 실험에 맞도록 수정하여 사용하였다. 실험은 MS Windows XP 기반의 OS에서 MS Visual Basic 6.0으로 프로그래밍된 소프트웨어를 이용하여 진행되었으며, 각각의 실험자극은 임의의 순서로 Pentium 4 PC에 장착된 Philips Seismic Edge PSC705 sound board와 Philips SBC3375 헤드폰을 통하여 피실험자에게 제시되었다. 피실험자는 제시되는 경보음과, 미리 들려준 평가기준 경보음(reference stimuli)을 비교하여 기준 경보음보다 어느 정도 더 위급하게 느껴지는지를 선을 긋고(line production: LP), 수치 값을 입력함(numerical Estimation: NE)으로써 나타내도록 하였다.

2.6 실험분석 및 결과

각 자극에 대한 피실험자의 평가치는 선의 길이와 수치 값으로 나타나게 되는데, 이 두 개 평가치의 regression bias를 교정하기 위하여 Lodge(1981)가 제안한 다음과 같은 식을 이용하여 통합한 후, 실험의 종속변수로 사용하였다.

$$IM = \sqrt{L^{1/b_L} \cdot N^{1/b_N}} \qquad (1)$$

IM: 통합된 측정치
L: 측정된 선의 길이
N: 수치 평가 값
b_L: 선을 긋는 경향에 관한 회귀식 계수
b_N: 수치 평가 경향에 관한 회귀식 계수

여기에서 b_L과 b_N은 예비실험에서 피실험자를 선별할 때 이용하였던 회귀식의 계수로써, b_L은 각 피실험자가 자극으로 주어진 수치를 선의 길이로 평가했을 때 구해지는 회귀식의 계수이며, b_N은 선의 길이를 수치로 평가했을 때의 회귀식 계수이다.

실험은 전술한 6개 독립변수들에 각각 2개의 수준으로 실시되었으며, 통합된 위급도 평가치를 종속변수로 하여 ANOVA 분석을 실시하였다. 표 3에서 볼 수 있는 것처럼, 6개 요인들 모두 유의한(p<0.0001) 차이가 있는 것으로 나타났으며, 요인들 간의 2인자 이상의 교호작용은 없는 것으로 나타났다. 요인별 각 수준의 주관적 위급도 평가 평균과 표준편차는 표 4와 같다. 위급도 평가 값은 평가기준 경보음에 대한 위급도를 1이라고 하였을 때 이에 대한 상대적인 값이다.

표 3. ANOVA 분석 결과

	Sum of Squares	DF	Mean Square	F	Pr > F
합성음 화자의 성별	96.942	1	96.942	269.94	<0.0001
평균 피치	26.353	1	26.353	73.38	<0.0001
합성음의 빠르기	318.454	1	318.454	886.75	<0.0001
경보메시지 형태	6.961	1	6.961	19.38	<0.0001
경보메시지의 의미	9.651	1	9.651	26.87	<0.0001
경보메시지 사이 시간 간격	84.235	1	84.235	234.56	<0.0001

표 4. 주관적 위급도 평균 및 표준편차

실험요인	수준	평균 (표준편차)
합성음 화자의 성별	남*	1.44(0.83)
	여	1.07(0.57)
평균 피치	120Hz*	1.16(0.65)
	190Hz	1.35(0.80)
경보 합성음의 빠르기	4.98 syllables/sec*	1.58(0.69)
	3.62 syllables/sec	0.93(0.62)
경보메시지 형태	Semantic context*	1.21(0.77)
	Keyword	1.30(0.69)
경보메시지의 의미	Collision traffic one o'clock*	1.31(0.70)
	Cabin pressure dropping	1.20(0.76)
반복되는 경보메시지 사이 시간 간격	0.1sec*	1.42(0.79)
	0.8sec	1.09(0.63)

*평가기준 경보음 (reference stimuli)

표 4에서 볼 수 있는 것과 같이 합성음성 경보음이 빠르게 발성될수록, 평균 피치가 높을수록(음높이가 높을수록) 위급도 평가치는 높아지는 것으로 나타났다. 이 결과는 간단한 규칙기반 음성합성기를 이용하여 생성된 경보음의 위급도 평가실험(Park and Jang, 1999) 결과와 일치한다. 또한 경보메시지가 반복될 때, 메시지 간의 시간 간격이 작은 경

우, 위급도 평가치는 높은 것으로 나타났으며, 경보메시지 형태 중, 단어 수가 적은 keyword 형태가 semantic context 형태보다 높은 평가치를 나타내었다. 그리고 여자음성의 경보음 보다 남자음성 경보음에 대해 높은 위급도 평가 결과가 나타났다. 일반적으로, 남자음성 보다는 여자음성이 높은 주관적 위급도 평가치를 이끌어 낼 것으로 예상하지만(Edworthy et al., 1991), 본 실험에서는 이와는 반대의 실험결과가 도출되었다. 본 실험에서는 남, 여음성의 평균 피치(음높이)를 동일하게 조정하였는데, 음성의 거칠기(roughness), 음색(timbre)과 같은 요인들에 의해 이러한 차이가 발생한 것으로 사료된다. 일부 피실험자들은 높은 음높이를 가지는 남자음성이 상당히 거칠고, 부자연스럽게 들린다고 보고하였다. 물론, 더 많은 음색의 합성음성들을 대상으로 평가되어야 하겠지만, 본 실험 결과는 여성음색을 가지는 합성음성이라고 해서 높은 주관적 위급도 평가 값을 보증하는 것이 아님을 보여준다.

3. 실험 2 ~ 실험 4

실험 1에서는 합성음성 경보의 정성적 요인 3개와 정량적 요인 3개가 인간의 인지 심리적 반응인 주관적 위급도에 영향을 미치는 것으로 파악되었다. 이러한 사실은 이들 요인들을 조정하면, 실제상황의 위급한 정도를 나타낼 수 있는, 즉 인지 심리학적인 적합성을 가지는 경보신호를 생성할 수 있음을 보여준다. 물론 이를 위해서는 유효한 요인들에 대한 상세한 분석이 필요하며, 이를 위하여 실험 2, 3, 4에서는 요인들 중 세부적인 조정이 가능한 요인들에 대하여 정량적인 평가를 수행하였다.

실험 2에서는 정량적 요인인 합성음성의 빠르기(speech rate)에 대하여, 실험 3에서는 평균 피치(pitch)에 대하여, 그리고 실험 4에서는 반복되는 경보메시지 사이 시간에 대하여 각각의 독립적인 실험을 수행하였다.

이들 요인과 주관적 위급도와의 정량적 관계를 파악하기 위하여 Stevens의 멱함수(power law)를 이용하였다. Stevens는 외부자극강도와 주관적인 지각강도 간에 지수적인 관계가 있다고 주장하고, 식2와 같이 정의하였다.

$$\Psi = R = kS^b \qquad (2)$$

Ψ : 주관적인 지각강도
R: 반응강도
S: 외부자극강도
K: 비례상수
b: 지수

이 식에서 b는 주관적인 지각강도와 외부자극강도간의 관계를 나타내는 지수로써, 자극의 종류에 따라 다른 값을 가진다는 것이 여러 연구를 통해 알려져 있다. 예를 들어 인간이 전기적 충격에 대해서는 지수 값이 3.5인데, 이것은 인간이 전기적 충격에 대해서는 매우 민감하다는 것을 나타내며, 소리의 크기에 대해서는 0.67로써 소리의 크기 변화에 별로 민감하지 않다는 것을 보여주고 있다. 실험 2, 3, 4의 목적은 정량적 요인들의 지수 값이 각각 어느 정도인지를 파악하는 것이며, 피실험자들의 반응강도는 magnitude estimation 기법을 사용하여 측정하였다.

3.1 예비실험 및 피실험자

실험 2, 3, 4 전에 각 실험의 예비실험이 이루어졌으며, 실험절차와 피실험자 선별조건은 실험 1의 예비실험과 동일하다. 예비실험을 통과하여 실험에 참여한 피실험자 현황은 표 5와 같다.

표 5. 실험 별 피실험자

실험	실험참여 (명)	예비실험 통과 (명)	남/여 (명)	연령평균, 표준편차
2	31	27	18/9	27.9, 3.9
3	32	28	18/10	28.3, 4.3
4	31	27	18/9	27.9, 3.9

3.2 실험장비 및 실험자극

사용된 실험장비는 실험 1과 동일한 실험장비가 이용되었으며, 실험 1에서 평가되었던 합성음성 경보메시지 중 가장 낮은 위급도 평가를 받은 'Cabin Pressure Dropping' 메시지와 가장 높은 위급도 평가를 받은 'Traffic One O'clock' 메시지를 이용하였다. 이 메시지들은 가장 낮은 위급도 평가를 받은 여성 합성음성(이후 Fe)과 가장 높은 위급도 평가를 받은 남성 합성음성(이후 Ma)을 기준으로 각각 생성(합성)한 후에, 전술한 정량적 요인들을 실험수준에 따라 조정하여 실험자극으로 이용하였다.

실험 2에서는 합성음성 경보음의 빠르기(speech rate)를 세분화 하여 이에 따른 주관적 위급도 평가 변화를 정량화하기 위해, 4개의 합성음성 경보음(메시지 2개 × 음성 2개)의 빠르기(speech rate)를 각각 7개 수준으로 조정하였다. 음성합성기에서 생성된 합성음성의 빠르기를 100%라고 하였을 때, 70~130% 범위에서 10%씩 조정하였으며, 이에 따라 총 28개의 실험자극을 생성하고, 표본화율 16kHz, 16bit mono WAV 파일형식으로 저장하여 실험에 이용하였다.

실험 3에서는 음성합성기에서 생성된 합성음성 평균 피치

의 70~130% 범위에서 10%씩 조정하여 7개 수준, 28개의 실험자극을 생성, 실험에 이용하였다.

실험 4에서는 반복되는 경보메시지 사이 시간 간격을 7개 수준으로 세분화하였는데, 0.25 sec의 7~130% 범위인 0.18~0.32sec에서 10% 간격으로 조정하여 실험자극을 생성하였다.

3.3 실험절차

실험절차는 실험 1과 동일한 절차를 이용하였으며, 각 실험 별로 28개의 자극을 임의의 순서로 들려주고, 각각의 자극이 평가기준(reference) 경보음과 비교하여 어느 정도 위급하게 느껴지는지를 선을 긋고, 수치 값을 입력함으로써 나타내도록 하였다.

3.4 실험분석 및 결과

위급도에 대한 두 가지 평가치(LP, NE)는 Lodge(1981)의 방정식을 이용하여 통합한 후, 기하 평균(geometric mean)을 취하였다. 기하 평균은 magnitude estimation 기법을 이용하여 측정한 자료들의 중심경향을 나타내는 표준적인 측도이다(Lodge, 1981). 기하 평균된 지각강도(주관적 위급도)와 자극강도 값은 각각 log 변환하여, 선형회귀분석을 시행하였다.

표 6은 실험 2에서 합성음성 경보음의 빠르기와, 이에 대한 위급도 지각강도를 각각 log 변환한 후 회귀분석한 결과이다.

표 6. 합성음성 빠르기에 대한 회귀분석

메시지	Traffic One O'clock		Cabin Pressure Dropping	
합성음성	Fe	Ma	Fe	Ma
Exponent	1.099	1.26	1.232	1.162
Pr > |t|	<.0001	0.0002	<.0001	<.0001
Adjusted R^2	0.965	0.939	0.976	0.965

Fe: 여성합성음성, Ma: 남성합성음성

회귀식의 기울기 값은 Stevens의 멱함수에서 지수 값에 해당하는데, Wilks's Lambda를 이용하여 기울기 간에 차이가 있는지를 검증하였다. 네 개의 지수 값들을 두 개씩 쌍으로 6가지 경우에 대하여 검증한 결과, 표 7처럼, 5% 유의수준에서 여성음성(F)으로 합성한 두 개의 메시지 간에 기울기 차이가 있는 것으로 나타났지만, 유의수준 1%에서는 모든 경우에 있어서, 기울기가 다르다고 할 수 없는 것으로 나타났다. 이에 따라 Steven's의 멱함수 지수 값을 통합하면 1.19이며 adjusted R^2는 0.979이다.

표 7. 회귀식 기울기 차이 검정: Wilk's Lamda(Pr>F)

		Traffic One O'clock	Cabin Pressure Dropping	
		Ma	Fe	Ma
Traffic One O'clock	Fe	0.723 (0.225)	0.323 (0.028)*	0.951 (0.634)
	Ma	–	0.988 (0.816)	0.854 (0.399)
Cabin Pressure Dropping	Fe	–	–	0.935 (0.581)

*Significant at 95% confidence interval
Fe: 여성합성음성, Ma: 남성합성음성

표 8은 실험 3의 결과로서, 평균 피치를 변화시킨 자극강도와 기하 평균된 지각강도(주관적 위급도)를 각각 log 변환하여, 선형회귀분석을 시행한 결과이다. 지수 값(기울기) 4개에 대한 기울기 차이 검정 결과는 유의수준 5%의 모든 경우에서 기울기가 다르다고 할 수 없는 것으로 나타났으며, 이에 따라 통합된 Steven's의 멱함수 지수 값은 0.701, adjusted R^2=0.988이다.

표 8. 평균 피치에 대한 회귀분석

메시지	Traffic One O'clock		Cabin Pressure Dropping	
합성음성	Fe	Ma	Fe	Ma
Exponent	0.725	0.749	0.672	0.656
Pr > |t|	0.0002	<.0001	0.0003	<.0001
Adjusted R^2	0.943	0.967	0.927	0.955

Fe: 여성합성음성, Ma: 남성합성음성

표 9는 메시지 간 시간 간격을 변화시킨 자극강도와 기하 평균된 지각강도(주관적 위급도)를 각각 log 변환하여, 선형회귀분석을 시행한 결과이다. 회귀식 기울기의 차이 검정 결과는 유의수준 5% 모든 경우에서 기울기가 다르다고 할 수 없는 것으로 나타났다. 통합된 Steven's의 멱함수 지수 값은 −0.127, adjusted R^2=0.909이다.

실험 결과들을 종합하여 각 요인들과 위급도 지각강도 간의 관계를 나타내면 표 10과 같다. 물론 합성음성의 이해도 등을 고려하여 각 요인들의 조정 가능한 범위에 대한 연구가 더 있어야 하겠지만, 실험된 범위 내에서는 합성음성의 빠르기에 피실험자들의 지각강도가 가장 예민하게 변화되는 것으로 나타났다.

표 10에서 볼 수 있는 것처럼, 각 요인들의 통합된 측정치에 대한 회귀분석의 adjusted R^2 값들은 모두 0.9 이상으

표 9. 메시지 간 시간 간격에 대한 회귀분석

메시지	Traffic One O'clock		Cabin Pressure Dropping	
합성음성	Fe	Ma	Fe	Ma
Exponent	−0.143	−0.135	−0.130	−0.101
Pr > |t|	0.0052	<.0001	0.0008	0.0019
Adjusted R^2	0.780	0.960	0.897	0.965

Fe: 여성합성음성, Ma: 남성합성음성

로써, Stevens의 멱함수를 이용한 모델이 높은 설명력을 가지고 있음을 나타내고 있다. 이러한 관계를 이용하면, 위의 세 가지 정량적 요인들을 변화시킴으로써 합성음성 경보음이 특정수준의 위급한 정도를 나타낼 수 있도록 조정하는 것이 가능할 것이다. 예를 들어 어떤 음성합성 경보음의 위급도 지각강도를 25% 높이고자 한다면, 합성음성의 빠르기를 본래 빠르기의 120.6%($1.25^{(1/1.189)}$)로 조정하거나 평균 피치(음높이)를 원래 피치의 137.5%($1.25^{(1/0.701)}$)가 되도록 높여주면 된다.

표 10. 각 요인들과 위급도 지각강도 간의 관계

파라미터	지수 값	Adjusted R^2	검사 범위	변화 범위
합성음성의 빠르기	1.189	0.979	70~130%	73.2~152.8%
평균 피치	0.701	0.988	70~130%	77.8~120.2%
경보메시지 사이 시간 간격	−0.127	0.909	0.1~2.0sec	134.0~91.6%

4. 결 론

본 연구에서 시행된 일련의 실험들은 합성음성을 이용하는 새로운 형태의 청각적 경보음을 제시하는 데 있어서 상황의 위급한 정도와 같은 심리학적인 적합성(psychological appropriateness)을 부여하는 것이 가능함을 확인시켜 주었다. 연결합성방식의 합성음성을 경보음으로 이용하는 경우, 합성음성의 평균 피치, 경보메시지 형태와 의미, 합성음성의 성별, 음성의 빠르기(발성속도), 반복시의 메시지 사이 시간 간격 등이 경보청취자의 주관적 위급도에 영향을 미칠 수 있으며, 이를 이용하면 상황의 위급한 정도를 경보음에 암호화(coding) 하는 것이 가능하다. 특히, 평균 피치(pitch), 합성음성의 빠르기 그리고 메시지 사이 시간 간격 등의 요소들은 주관적 위급도와 정량적이며, 예측 가능한 관계가 있음이 입증되었다.

본 실험에서는 Simpson and Williams(1980)가 제시한 영어로 된 경보메시지를 그대로 이용하였고, 피실험자들은 한국인들이기 때문에, 경보음에 대한 주관적 위급도 평가가 한국어로 된 경보음에 대한 것과는 차이가 있을 수 있다. 또한, 다양한 위급 상황과, 피실험자들이 피부로 느낄 수 있는 실제 생활 환경에서의 위급도 평가연구가 더 진행되어야 하겠지만, 본 연구는 합성음성을 이용한 새로운 경보음 설계에 효과적으로 활용될 수 있을 것으로 생각된다.

참고 문헌

Aretz, A. J., "A comparison of manual and vocal response medes to the control of aircraft systems", *Proceedings of the 27th Annual Meeting of the Human Factors Society*, (pp. 97-101), Santa Monica. CA, 1983.

Barber, C., Stanton, N. A. and Stockley. A., Can speech be used for alarm displays in 'process control type tasks?, *Behavior & Information Technology*, 11, 216-226, 1992.

Bliss, J. P. and Gilson, R. D., Emergency signal failure: implications and recommendations, *Ergonomics*, 41(1), 57-72, 1998.

Bliss, J. P., Gilson, R. D. and Deaton, J. E., Human probability matching behavior in response to alarms of varying reliability, *Ergonomics*, 38(11), 2300-2313, 1995.

Boersma, Paul., "Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound", *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam*, 17(pp. 97-110), Amsterdam. 1993.

Burt, J. L., Bartolome, D. S., Burdette, D. W. and Comstock, J. R., A psychophysiological evaluation of the perceived urgency of auditory warning signals, *Ergonomics*, 38, 2327-2340, 1995.

Byblow, W. D. and Corlett, J. T., Effects of linguistic redundancy and coded voice warnings on system response time, *Applied Ergonomics*, 20, 105-108, 1989.

Edworthy, J., Loxley, S. and Dennis, I., Improving auditory warning design: Relationship between warning sound parameters and perceived urgency, *Human Factors*, 33, 205-231, 1991.

Haas, E. C. and Edworthy, J., Designing urgency into auditory warnings using pitch, speed, and loudness, *Computing & Control Engineering Journal*, 7(4), 193-198, 1996.

Hakkinen, M. T. and Williges, B. H., Synthesized warning messages: Effects of an alerting cue in single- and multiple-function voice synthesis systems, *Human Factors*, 26, 185-195, 1984.

Irving, A., "An experimental comparison of operator responses to voice and tone system warnings", *Report No. BAE-BT-12051, British Aerospace Dynamic Group*, Bristol, England, 1981.

Kerr, J. H. and Hayes, B., An 'alarm' situation in the intensive care ward, *Intensive Care Medicine*, 9, 103-104, 1983.

Lodge, M., Magnitude Scaling: *Quantitative measurement of opinion*, Sage, Beverly Hills, CA, 1981.

Mashour, M. and Hosman, J., On the new psychophysical law: A validation study, *Perception and Psychophysics*, 3, 367-375, 1968.

Morris, R. W. and Montano, S. R., Response times to visual and auditory alarms during anesthesia, *Anesthesia and Intensive Care*, 24, 682 -684, 1996.

Park, K. S. and Jang, P. S., Effects of Synthesized Voice Warning Parameters on Perceived Urgency, *International Journal of Occupational Safety and Ergonomics*, 5(1), 73-95, 1999.

Roederer, Juan., *Introduction to the Physics and Psychophysics of Music*, The English Universities Press, 1973.

Rood, G. M., Chillery, J. A. and Collister, J. B., Requirements and application of auditory warnings to military helicopters, *Ergonomics International*, 85, 169-172, 1985.

Simpson, C. A. and Marchionda-Frost, K., Synthesized speech rate and pitch effects on intelligibility of warning messages for pilots, *Human Factors*, 26, 509-517, 1984.

Simpson, C. A. and Williams, D. H., Response time effects of altering tone and semantic context for synthesized voice cockpit warnings, *Human Factors*, 22, 319-320, 1980.

Stevens S., On the psychophysical law, *Psychological Review*, 64, 153-181, 1957.

Tanenhaus, J. and W. Murphy., Patterns of public support for the Supreme Court: A panel study, *Journal of Politics*, 43: 324-339, 1981.

Wheale, J. L., Evaluation of an experimental central warning system with a synthesized voice component, *Aviation, Space, and Environmental Medicine*, 54, 517-523, 1983.

◐ 저자 소개 ◑

❖ 장 필 식 ❖ phil@mail.daebul.ac.kr

KAIST 산업공학과 박사
현 재: 대불대학교 컴퓨터공학과 부교수
관심분야: HCI, 감성공학, 음성분석 및 합성

❖ 이 경 태 ❖ ktlee@mail.daebul.ac.kr

KAIST 산업공학과 박사
현 재: 대불대학교 산업정보공학과 부교수
관심분야: HCI, 안전, 작업관리, 작업환경개선

논 문 접 수 일 (Date Received) : 2006년 01월 26일
논 문 수 정 일 (Date Revised) : 2006년 05월 08일
논문게재승인일 (Date Accepted) : 2006년 05월 09일