

컴퓨터 합성음성 경보의 주관적 위급도 정량화

박 경수¹ · 장 필식² · 이 경태³

¹한국과학기술원 산업공학과, ²대불대학교 컴퓨터학부, ³대불대학교 기계산업·시스템공학부

ABSTRACT

This paper presents an experimental study of the relationship between sound parameters of synthesized voice warning and perceived(psychoacoustic) urgency. Twenty four subjects participated in two experimental sessions to evaluate and quantify the effects of the voice parameters. Experiments showed that speech rate, fundamental frequency, fundamental frequency contour types and voice types have clear and consistent effect on perceived urgency. The results of these experiments can be applied to the improvement of existing auditory warning systems and the design of new systems.

1. 서론

청각적 표시 장치는 청각 계통의 고유한 특성 때문에 다른 感官(sensory modality)을 이용하는 표시 장치들과 비교해서, 특정한 상황에 있어 여러 가지 장점을 가진다. 이 때문에 청각적 표시 장치는 위급한 상황에 대한 정보를 전달하는 기능을 가지는 경계 및 경보 신호를 위한 용도로 많이 사용된다. 하지만 기존의 경계 및 경보 기구들, 즉 부저(buzzer), 경적(horn), 호각(whistle), 사이렌(siren), 전종(bell) 등에 의한 경보에 대해서 최근, 여러 가지 문제점들이 지적되고 있다. 이러한 단점들에 대처하고, 기존의 청각적 경보를 개선시키거나 새로운 청각적 경계, 경보를 설계하기 위한 접근 방법 중의 하나가 합성 음성(synthesized speech)을 경보 체계에 이용하는 방법이다. 합성 음성은 합성 음성 자체로써 경보의 기능을 가질 수 있으며, 기존의 청각적 경보 체계와는 달리 학습 없이 인간에게 친숙한 형태로 정보를 전달할 수 있다는 장점 때문에 최근, 새로운 청각적 경보의 형태로써 많은 연구들을 통해 활용 범위가 넓어지고 있다.

현재 대부분의 상업용 음성 합성 S/W 들은 기초진동수와 같은 여러 가지 합성음성 파라미터들을 사용자가 조정하거나 선택할 수 있도록 하고 있다. 합성음성경보의 설계에 있어 이러한 파라미터들의 적정 수준은 여러 가지 인간공학적 검토를 통해 이루어 져야 한다. 본 연구에서는 합성음성 경보와 인간이 경보를 듣고 주관적으로 느끼는 위급한 정도, 즉, 주관적 위급도(perceived urgency)간의 관계를 고찰하기 위하여, 주관적 위급도에 영향을 미치는 합성음성 요소들을 실험을 통해 찾아내었으며, 각 요소들의 변화에 따른 주관적 위급도를 측정하였다. 또한, 이러한 요소들 중 정량적 요소들에 대해서는 Stevens의 power law를 이용하여 주관적 위급도와와의 관계를 정량화 하였다.

2. 합성 음성 경보에의 주관적 위급도 측정, 정량화

본 연구에서는 합성 음성 경보 체계의 성능을 향상시킬 수 있는 지침을 마련을 위한 기초연구로써 인간의 주관적 위급도에 영향을 미치는 합성 음성의 요소들을 실험을 통해 찾아내고, 이러한 요소들이 어느 정도 주관적 위급도에 영향을 미치는 지를 평가하였으며, 경보가 발생하는 상황의 위급한 정도와 주관적으로 느끼는 위급도와 의 관계를 정량화 하였다. 본 연구에서 인간의 주관적 위급도에 영향을 미치는 합성 음성 요소들을 찾아내기 위해 합성 음성 경보에 사용되는 기초 진동수(Fundamental Frequency, Pitch), 음성 경보의 빠르기(speech rate), 합성 음성 경보의 음색 등을 고려하였다.

이러한 요소들 중 어느 요인이 인간의 주관적 위급도에 영향을 미치며, 어느 정도 영향을 미치는 지를 평가하기 위해서 Magnitude Estimation 기법을 사용하여 측정하였으며, 외부 자극인 상황의 위급도와 인간이 주관적으로 느끼는 주관적 위급도의 관계는 Stevens의 Power Law를 사용하여 표현하였다. 또한 실험후 측정된 자료들에 대한 타당성을 검증하기 위하여 Cross-Modality Matching Paradigm을 이용하였다.

3. 실험

실험은 예비실험과 2개의 본실험으로 이루어졌다. 예비실험은 magnitude estimation에 필수적인 피실험자의 비율판단 능력을 검증하고 피실험자를 선별하며, 본실험에서의 측정치를 보정하기 위하여 실시되었다. 예비 실험에 합격한 피실험자들을 대상으로 하여 실험 1이 주관적 위급도에 유의한 영향을 미치는 합성음성 요소를 찾아내기 위해 실시되었다. 두 번째 실험에서는 주관적 위급도에 영향을 미치는 합성음성 요소들의 수준(level)을 세분화하여 실험을 통해 정량적 요소들에 대해서는 주관적 위급도와의 관계를 정량화 하였다.

실험에는 19세에서 31세까지의 대학원생과 대학생 24명이 참여하였는데, 이들 중 11명은 영어를 모국어로 사용하는 외국인이며, 13명은 한국인이었다. 모든 피실험자가 개인용 컴퓨터를 다루어 본 경험이 있었고, 실험전 청각의 이상 유무를 확인하였다. 이들 중 20명이 예비실험에 합격하여 본 실험에 참여하였다.

3.1 실험 장치

실험에 사용되는 합성 음성 경보는 개인용 컴퓨터에서 사용 가능한 sound board(Sound Blaster AWE32)와 이 board를 사용하는 음성합성 소프트웨어인 Text Assist(Creative Labs)를 사용해 제작하였다. 합성 음성은 같은 사양의 sound board를 장착한 두 대의 PC를 연결한 뒤, sampling rate 20kHz, 16bit WAV file format 형태로 저장하여 실험에 사용하였다. sound board를 제어하고, 실험을 진행시키는 프로그램은 Visual Basic 으로 구현하였으며, 주위의 잡음을 차단하기 위하여 sound board에 헤드폰(Philips SBC 3375)을 연결하여 사용하였다. 기초 진동수는 16비트 20kHz sampling rate 형태로 저장된 합성 음성을 8비트 10 kHz sampling rate 형태로 변환하여 SIFT(Simplified Inverse Filter

Tracking) algorithm을 사용하여 측정하였다. 실험 1과 2에서 사용된 warning message 는 피 실험자가 자동차를 운전하고 있을 때 자동차 상태에 따라 message 가 들려진다고 가정하고 “Brake Oil Leakage” 라는 warning message를 들려주었다. 상대적인 주관적 위급도를 선의 길이와 수치로 표시하도록 하였다. 자극으로 주어지는 합성음성 경보에 대한 피실험자의 주관적 위급도는 한 자극에 대하여 선의 길이 (Line Production: LP)와 수치(Numerical Estimation: NE) 두 가지로 평가되는데, 이 두 가지 평가치를 통합하여 실험의 종속변수로 사용하게 된다. 각 실험 조건에서 얻게 되는 이 두개의 평가치에 대해서는 Regression Bias를 교정하기 위하여 Lodge가 제안한 다음과 같은 식을 이용하여 통합하였다.

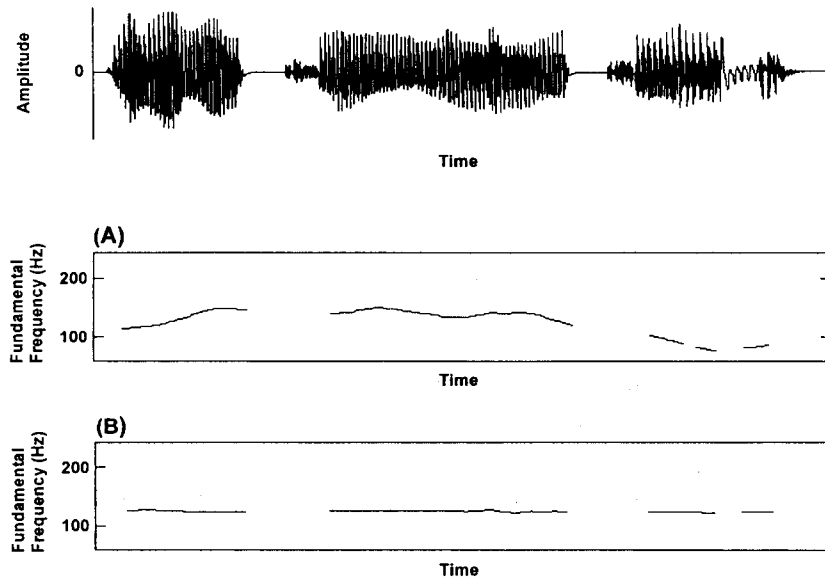
$$IM = (L^{1/n1} \cdot N^{1/n2})^{0.5} \left\{ \begin{array}{l} IM : \text{통합된 측정치(Integrated Measure)} \\ L : \text{선의 길이(Line Production)} \\ N : \text{수치(Numerical Estimation)} \\ n1 : \text{각 피실험자의 선을 긋는 경향에 관한 회귀식 계수} \\ n2 : \text{각 피실험자의 수치 평가 경향에 관한 회귀식 계수} \end{array} \right.$$

3.2 실험 1

총 24명의 피실험자 중 예비 실험에 합격된 피실험자 20명을 대상으로 하여 실험과 결과 분석을 행하였다. 실험 1에서 실험 변수로 사용된 합성 음성 요소들과 각 변수의 수준은 다음 표와 같으며, 각 실험 조건을 2회 반복하여 1인당 총 36회의 실험을 행하였다.

[표 1] 실험 변수

실험 변수(factor)	변수의 수준(level)
평균 기초 진동수(Average F ₀)	125 Hz
	255 Hz
합성 음성 경보의 빠르기(speech rate)	140 words/min
	200 words/min
합성 음성의 음색	Male-sounding
	Female-sounding
F ₀ Contour Type	A
	B



[그림 1] F₀ Contour Type

3.3 실험 1 결과

[표 1]에 보인 합성음성 파라미터들에 따라 피실험자들의 주관적 위급도가 영향을 받는지를 파악하기 위하여 분산 분석을 실시하였다. 분석 결과는 average F₀ ($F = 166.15, p < 0.0001$), speech rate ($F = 683.34, p < 0.0001$), voice type ($F = 93.41, p < 0.0001$), F₀ contour type ($F = 22.76, p < 0.0001$) 등의 요인들이 모두 주관적 위급도에 유의한 영향을 미치고 있음을 보여주었다. 이들 인자들간에 2인자 이상 교호작용은 통계적으로 유의하지 않는 것으로 판명되었다.

3.4 실험 2

두 번째 실험에서는 실험 1에서 인간의 주관적 위급도에 유의한 영향을 미치는 것으로 판명된 요인들 중 정량적 요인인 평균 기초진동수(average F₀)와 speech rate 의 실험 수준을 각각 6 level로 세분화 한 뒤 피실험자 20명의 이에 따른 주관적 위급도를 다음과 같은 Stevens의 Power Law를 이용해 정량화 하였다. 실험에 사용된 warning message와 실험 절차는 실험 1과 동일하며, F₀ contour type 은 [그림 1]에 예시된 유형중 Type (A)를 자극으로 이용하였다. 세분화된 자극들의 수준은 다음 [표 2]와 같다.

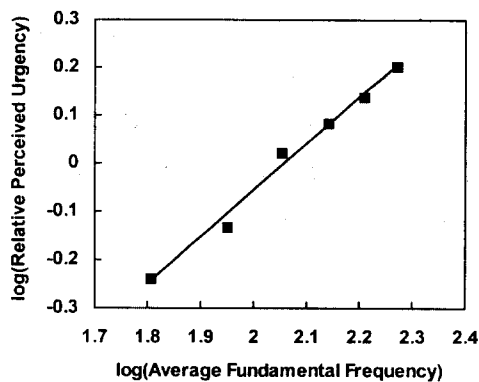
$$S = kO^m \quad \left\{ \begin{array}{l} S : \text{주관적인 지각 강도(Subjective Magnitude)} \\ O : \text{외부 자극 강도(Stimulus Magnitude)} \\ k : \text{비례상수} \\ m : \text{지수(Exponent)} \end{array} \right.$$

[표 2] 실험 변수

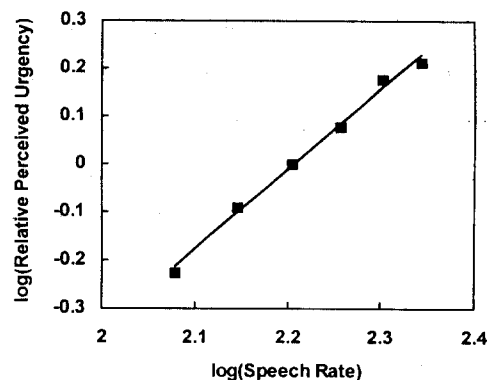
Level	Average F0 (Hz)	Speech Rate (words/min)
1	64	120
2	89	140
3	113	160
4	138	180
5	162	200
6	186	220

3.5 실험 2 결과

실험은 각 피실험자 마다 반복 2회씩 총 24회씩 실시되었다. 실험 1에서 사용된 Lodge의 correction method를 이용하여 실험치를 통합하였으며, 피실험자의 실험치를 기하평균하여 회귀분석을 실시하였다. 이 결과를 log 변환하여 주관적위급도와 각 요인과의 관계를 나타내면 [그림 2]와 같다.



[그림 2] Average F0 & Perceived Urgency



[그림 3] Speech Rate & Perceived Urgency

average F0의 경우, 회귀식 기울기 0.968으로 주관적위급도와 선형적으로 비례하는 것으로 나타났으며, speech rate의 경우는 1.687로 나타났다. 회귀식의 결정계수는 0.989(average F0)와 0.998(speech rate)로써 각 요인과 주관적위급도의 관계를 Stevens의 Power Law가 잘 설명해 주는 것으로 보인다. Stevens의 Power Law에서 지수값 m 은 주어진 자극에 대한 민감도를 나타내는데, 주관적 위급도를 절반으로, 또 두배, 세배로 만들어주기 위하여 average F0와 speech rate를 어느 정도나 증가 또는 감소시켜야 할지를 실험 2를 통해 구해진 지수값을 이용하여 계산하면 [표 3]과 같다.

[표 3] 실험 변수들과 주관적 위급도와와의 관계

Parameter	Exponent	Increment to Increase Perceived Urgency		
		50%	200%	300%
Average F ₀	0.968	48.9%	204.6%	311.1%
Speech rate	1.687	66.3%	150.8%	191.8%

4. 결론 및 연구 방향

본 연구에서는 인간의 주관적 위급도에 영향을 미치는 합성음성의 요소들을 실험을 통해 찾아내고, 이러한 요소들의 영향을 평가하였다. 실험 결과 합성음성 경보의 여러 가지 가변적인 요소 중 평균기초 진동수와 speech rate, 음성의 종류(음색), fundamental frequency contour 등이 인간의 주관적 위급도(perceived urgency)에 유의한 영향을 미치는 것으로 평가되었으며, 이러한 결과를 바탕으로 하여, 정량적 요인들에 대해 Stevens의 Power Law를 이용, 외부 자극인 상황의 위급도와 인간이 주관적으로 느끼는 주관적 위급도의 관계를 정량화하였다. 이를 통해 추후, 음성의 이해도(intelligibility)를 저하시키지 않는 범위 내에서 요인들을 변화시켜, 상황의 위급도와 이에 상응하는 인간의 주관적 위급도를 합성음성 경보음에 암호화하고자 한다. 이렇게 위급도를 암호화한 합성음성 경보는 '정보 전달'과 '경보'의 기능을 가지는 합성음성 자체의 장점을 가지게 될 것이며, 인간의 음성과 구별되는 경보음으로써의 장점을 가지게 되므로, 합성음성 경보 체계의 성능을 향상시킬 수 있으리라 생각된다.

참고 문헌

- [1] Carl E. Williams., Stevens. 1972 'Emotion and Speech: Some Acoustical Correlates 'J.A. Acoust. Soc. Am Vol 52 Number 4 part 2, 1238-1250.
- [2] Edworthy, J., Loxlry, S. and Dennis, I. 1991 'Improving auditory warning design: relationship between warning sound parameters and perceived urgency' Human Factors 33(2), 205-231.
- [3] Hellier, E , Edworthy, J. and Deniss, I. 1993 'Improving auditory warning design : quantifying and predicting the effects of different warning parameters on perceived urgency' Hum Factors. 35(4), 693-706.
- [4] Markel, J.D. et al & Gray. 1976 'Linear Prediction of Speech' Springer-Verlag.

- [5] Kerr, J.H. 1985 'Auditory warnings in intensive care units and operating theatres' *Ergonomics International*, 85, 172-174.
- [6] Lodge, M. 1981 'Magnitude Scaling: Quantitative measurement of opinions' Beverly Hills, CA: Sage.
- [7] Momtahan, K.L., Hetu, R. and Tansley, B.W. 1993 'Audibility and identification of auditory alarms in operating rooms and an intensive care unit' *Ergonomics* 36(10), 1159-1176.
- [8] Rabiner, L. R., M. J. Cheng, A. E. Rosenberg, and C. A. McGonegal. 1976 'A Comparative Performance Study of Several Pitch Detection Algorithms' *IEEE Trans. Acoust., Speech, and Signal Process.*, vol ASSP-24, 399-417.
- [9] Sorokin, R., Kantowitz, B., and Kantowitz, S. 1988 'Likelihood alarm displays' *Human Factors*, 30, 455-459.
- [10] Stevens S. 1957 'On the psychophysical law' *Psychological Review*, 64, 153-181.
- [11] 박경수. 1980 '인간공학', 영지문화사.