PERCEPTUAL DIMENSIONS OF LAUGHTER AND THEIR ACOUSTIC CORRELATES

SHIRO KORI

Osaka University of Foreign Studies
Minoo, Osaka, JAPAN 567

ABSTRACT

This study described the acoustic correlates of two perceptual factors that were found to determine the recognition of laughter. 16 tokens of laughter and a non-laughter control token simulated by a Japanese male performer were presented to 10 Japanese who rated the appropriateness of the tokens to each of 12 labels of laughter.
A factor analysis of the appropriateness scores yielded two factors that have been labeled pleasant-unpleasant and superior-inferior, respectively. Correlation analysis of the factor scores and the acoustic data showed that pleasantness vs unpleasantness has significant correlation with the long vs short duration of the strong expiratory noise that may occur at the beginning of laughter, and with the large vs small rate of overall diminishment of the vowel amplitude. Superiority vs inferiority was highly correlated with the long vs short interval between vowels, the high vs low F0 max. or mean value, and the small vs large rate of overall vowel amplitude diminishment.

INTRODUCTION

People laugh for various reasons. Funny or 'incongruous' situations have been considered to be the most characteristic stimulator of laughter. From a communicative point of view, however, laughter is most frequently a signal of the well-being and friendliness of the person who is laughing. It may also used either as a sign of superiority, a device to communicate contempt, or as a submissive vocal gesture.
In most cases the situational cues help us to interpret why a man is laughing. But we are also capable of more or less inferring the reason for the laughter by only listening to the sound. However, very little is known about the acoustic correlates of laughter.
In this paper I will first check the transmissibility of laughter content through the auditory channel, then describe the acoustic correlates of the factors which account for the results of an auditory recognition test on several tokens of laughter.

MATERIAL AND METHOD OF THE RECOGNITION TEST

A thirty year old Japanese male simulated a set of laughters in an anechoic chamber imagining various kinds of laughter in his mind. 16 tokens which seemed to cover a wide range of laughter types were selected from the recordings. These tokens were used in an auditory recognition test along with a non-laughter [hahahahaha]utterance as a control stimulus.
The auditory recognition test was designed to obtain a perceptual characterization of the tokens and to examine the extent to which the performer's intent in laughter transmits through the auditory channel. 17 cassette tapes, each containing more than 50 repetitions of one of the 17 tokens, were prepared and presented to 10 Japanese subjects. The subjects were 20-40 years and 4 of them were male. They were instructed to listen to each of the cassette tapes through headphones as many times as they needed, and to judge, in the first place, whether the token was laughter or not. When the subjects judged the token to be laughter, they were then asked to describe freely the type of laughter and judge whether it was spontaneous laughter or forced laughter. After the judgment of spontaneousness, the subjects were invited to rate the degree of appropriateness of the sound stimulus to each of the 12 Japanese labels of laughter on a 3-point scale. A typical question given in the questionnaire was as follows: "Do you think that this is happy laughter? Please check one. / No. Somewhat. Very much." When the subjects did not judge the token as laughter, no more judgments or ratings were made on that token and the next token was presented to them. The order of the presentation of tokens was random and it was different from one subject to another.
The 12 laughter lablels used in the test were tentative ones. They were chosen by the present author from a list of some forty idiomatic or nearly idiomatic descriptions of laughter collected mainly in a questionnaire of several university students. An attempt was made to include labels which could cover as wide a range of laughter as possible. Most of the labels used here consisted of an adjective and a noun warai (laughter), or of a compound noun whose last element was warai.

They can be translated as follows: /1/ happy laughter (ab. happy), /2/ laughter for funny situation (funny), /3/ mocking laughter (mocking), /4/ ingratiating laughter or friendly laughter (ingratiating), /5/ triumphant laughter (triumphant), /6/ boisterous laughter or heroic laughter (boisterous), /7/ bawdy laughter (bawdy), /8/ laughter to cover one's awkwardness or shy laughter (awkward-covering), /9/ self-deprecating laughter (self-deprecating), /10/ cold-hearted laughter (cold-hearted), /11/ embarrassed laughter or uncomfortable laughter (embarrassed), and /12/ defiant laughter or challenging laughter (defiant). Since many of the labels did not seem to have equivalent expressions in everyday English, the translations given here are not necessarily 100% accurate. In the following discussion, however, I will use these English approximation for convenience's sake.

RESULTS AND DISCUSSION

Appropriateness scores of the tokens for each of the 12 labels were calculated in the following manner. Responses of maximum and medium favor were given 1 and 0.5 points, respectively. Negative responses were given 0 points. Sums of these scores across the subjects were then calculated and divided by the total number of the subjects and transformed into percentages. Tab.1 shows the obtained scores together with the spontaneousness scores computed in a similar way. The performer's intent is indicated by underlining. Note that a token which was judged to be laughter by X% of the subjects could at most obtain the score 'X' for any of the 12 labels, even though all these subjects judged it with maximum favor, because the subjects who did not judge it to be laughter did not make any ratings about the content of the laughter in that token. The results of the free description and the spontaneousness scores will not be discussed in this paper.

Transmissibility of laughter content through the auditory channel

12 tokens out of 17 obtained favorable judgment as laughter (≥80%) and 2 tokens registered very unfavorable scores (≤20%). One of the two tokens which were not heard as laughter was the non-laughter control stimulus (No.7 in Tab.1), and another one was intended and judged to be laughter by the

Table 1

Recognition scores of laughter
Percent appropriateness of the labels to the tokens

label / token	L	S	1	2	3	4	5	6	7	8	9	10	11	12
(1)	60	50	35	35	20	25	10	5	20	10	5	10	10	20
(2)	100	70	30	45	35	20	5	0	40	30	20	20	25	20
(3)	50	20	10	15	30	5	40	30	0	0	5	20	5	45
(4)	100	40	25	30	30	50	10	10	15	15	5	25	10	20
(5)	20	0	0	0	10	15	0	0	5	5	5	10	5	0
(6)	100	100	75	60	20	5	10	20	5	10	10	15	15	15
(7)	10	0	0	0	0	10	0	0	0	0	0	10	0	0
(8)	100	30	20	25	75	10	25	5	15	0	15	60	0	55
(9)	80	20	0	5	70	10	10	0	5	15	30	65	25	35
(10)	90	70	35	45	50	15	35	10	10	15	20	55	10	50
(11)	80	30	15	35	50	0	50	25	20	5	15	35	5	50
(12)	80	30	10	5	45	35	20	5	35	25	40	40	45	25
(13)	90	50	25	40	55	5	20	10	10	5	25	20	5	30
(14)	80	20	15	20	20	40	10	0	5	25	20	25	35	5
(15)	100	50	40	50	35	35	5	10	5	20	15	25	25	20
(16)	100	100	80	85	15	10	20	40	15	20	0	5	10	25
(17)	50	10	15	10	15	0	10	5	0	0	0	30	5	20

L : whether or not the token expressed laughter
S : spontaneousness

performer. These two tokens has two unique acoustic characteristics: the respiratory noise between vowel-like component was all voiced, and the range of vowel amplitude change was very small.
Of the 12 tokens which were heard as laughter, 3 showed a strong agreement between the performer's intent and the judged content by the subjects (≥75%, No.6,8,16). They were two tokens of happy and one token of mocking laughter. The happy tokens were also judged to be funny. The mocking token was also heard as cold-hearted. There was one token (No.9) in which the judgment of the subjects (mocking) opposed to the performer's intent and perception (funny).
These results show that it is possible, though not always, to identify the content of laughter through the auditory channel at least for happy and mocking laughter. However, the results suggest at the same time that there are individual differences in the way of laughing that might prevent the hearer from detecting its real intent: funny laughter for one person might be mocking one for others. In other words, while there are individually independent expressions with respect to different types of laughter, there is also room for the individuals to express their feeling and attitude in their own way.
Other types of laughter failed to be recognized correctly. Some of them were given completely different interpretations by the listeners (e.g. No. 10, funny or happy and cold-hearted or mocking). This may either be explained by the bad quality of performance, or this may suggest that laughter itself does not necessarily have the function of communicating such intents. This may highlight the importance of contextual cues. Moreover, the intent of laughter may also be hidden intentionally. For example, the intention of the person who laughs to please someone on purpose is presumably to express apparent happiness or friendliness regardless of his actual disposition. In such a case it would be only the artificiality of the laughter or the situational cues that could signal the real intent. Anyway, we can conclude that the content of laughter is not always unambiguously encoded in its vocal output. Facial expression and body movement may provide cues to detect the laughter content, although laughter is most characteristically manifested in the voice. At this point, however, we cannot make any further assumptions about the relative importance of different channels in decoding laughter content.

Perceptual dimensions of laughter

In judging the content of laughter, such pairs of labels as happy and funny, mocking and cold-hearted and triumphant and defiant were used similarly. This suggests that the two laughter-types of each pair are realized similarly. Evidently, the two labels of each pair share a common psychological meaning or 'factor'. Most likely, the judgment of the laughter content of the sound tokens was made according to such underlying psychological factors. In order to extract such factors that determined the listeners' strategy in judging the tokens in terms of appropriateness to the laughter labels,

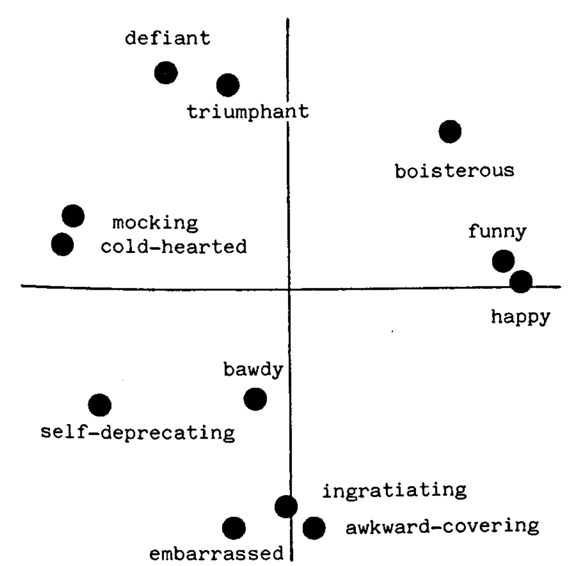
the appropriateness scores were analyzed in a factor analysis.
Of the 17 tokens used in the recognition test, two tokens, including the control token, were excluded from the data set for the factor analysis, because they obtained very unfavorable judgment as laughter (≤20%).
Principal factor analysis using squared multiple correlations for the prior communality estimates yielded two major factors. These two factors together accounted for 70% of the total variance. They were then rotated using the varimax procedure. Fig. 1 presents the rotated factor loading pattern. Both factors turned out to be bipolar. The first factor has been labeled pleasant-unpleasant, because it loaded happy, funny, and boisterous on one side and cold-hearted, mocking and self-deprecating on the other side. The second factor has been named superior-inferior, because triumphant and defiant were opposed to embarrassed, awkward-covering and ingratiating. These two factors correspond to the two fundamental psychological dimensions of interpersonal behavior (love-hostility and dominance-submission). Pleasant-unpleasant is also the most important dimension in the recognition of emotionss. It is labeled also as evaluative or friendly-hostile.
As a final step of factor analysis, the factor scores of the tokens were calculated for each of the two factors.

Correlation analysis of the perceptual factors and the acoustic variables

15 acoustic measures were defined and calculated for each of the tokens and then correlated with the factor scores. These variables consisted of 5 durational, 4 FO and 3 amplitude characteristics

Figure 1

Rotated factor pattern



as well as the first three formant frequencies.
Durational variables are: /1/ duration of the initial strong expiratory noise, /2/ duration of the initial high FO (500-1000Hz), /3/ number of alternations of respiratory noise ([h]) and vowel, /4/ mean duration of vowels and /5/ mean interval from the end of a vowel to the beginning of the next one (ab. mean vowel interval). See spectrograms for the variables /1/, /2/ and /3/.
FO variables are /6/ the highest FO value of the vowels in the token (ab. FO max.), /7/ the mean of the vowel maximum FO values on a logarithmic scale (ab. FO mean), /8/ the range of FO movement on a logarithmic scale, defined as the difference between the highest and the lowest vowel maximum values, /9/ the normalized FO range defined as the measure /8/ divided by the total duration of the noise-vowel alternations in the token. Since the vowel maximum FO values monotonously declined from the beginning to the end of the noise-vowel cycles, this measure roughly correspond to the rate of overall FO declination. The pattern of FO movement during the noise-vowel alternations was not included in the variables to be correlated with the factor scores, because in these tokens it declined almost monotonously from the beginning to the end.
Amplitude variables are: /10/ the range of amplitude change in the token, defined as the difference between the highest and the lowest vowel maximum velue, /11/ the normalized amplitude range of vowel defined as the measure /10/ divided by the total duration of the noise-vowel alternations, and /12/ the mean amplitude difference between the maximum value of the fricative noise and that of the following vowel (ab. noise-vowel amp. difference).
Since in most cases the vowel maximum amplitude diminished monotonously from the beginning to the end of laughter, the measure /11/ roughly correspond to the rate of overall vowel amplitude diminishment.
Formant frequencies are the mean values of steady state portions (/13/ F1, /14/ F2, /15/ F3).

Table 2

Correlation coefficients between 15 acoustic variables and factor scores

accustic variable	Factor 1		Factor 2	
/ 1/	.646	**	-.166	
/ 2/	.148		.270	
/ 3/	.185		.419	
/ 4/	-.351		.469	
/ 5/	-.157		.761	**
/ 6/	-.091		.570	*
/ 7/	-.249		.558	*
/ 8/	.267		.081	
/ 9/	.425		-.290	
/10/	.322		.090	
/11/	.570	*	-.555	*
/12/	-.084		.511	
/13/	-.102		.445	
/14/	-.348		.210	
/15/	-.234		.507	

* $p < .05$ ** $p < .01$ HO : r=0

In tab.2 the correlations between the two sets of factor scores and the 15 acoustic variables is presented together with the statistical significance (H0: r=0). Since one of the tokens consisted of only one vowel-noise cycle (No.9), it was excluded from the calculation of correlation coefficients for the variables /5/, /8/, /9/, /10/ and /11/. Acoustic variables which showed significant correlation with the first perceptual factor (pleasant-unpleasant) were /1/ initial expiratory noise duration (r=.646) and /11/ normalized vowel amplitude range (r=.570). These two variables were significantly correlated with each other (r=.750, p<.01). The next highest correlation was found in /8/ F0 range (r=.425), but it did not result statistically significant.

The second perceptual factor (superior-inferior) was significantly correlated with /5/ mean vowel interval (r=.761), /6/ F0 max. (r=.570), /7/ F0 mean (r=.558) and /11/ normalized vowel amplitude range (r=-.555). /12/ noise-vowel amplitude difference (r=.511) and /15/ F3 (r=.507) had relatively high correlation though they were slightly lower than the significance level. Of these variables, /6/ F0max and /7/ F0 mean (r=.951, p<.001), /5/ mean vowel interval and /11/ normalized vowel amplitude range (r=-.671, p<.01), /5/ mean vowel interval and /15/ F3 (r=.554, p<.05), /7/ F0 mean and /15/ F3 (r=.560, p 05), /11/ normalized vowel amplitude range and /15/ F3 (r=-.659, p<.05) showed significant inter-correlation.

These results suggest that the pleasantness vs. unpleasantness of laughter was acoustically characterized in part by the long vs short duration of the initial expiratory noise (42%) and by the large vs small normalized amplitude range, i.e. the rate of amplitude diminishment from the beginning to the end of the noise-vowel cycles (32%). Since these two variables were highly correlated with each other, they accounted only for 44% of the total variance. Even adding the normalized F0 range (/9/) which showed the next highest correlation, did not improve the R^2. On the other hand, the superiority vs inferiority of laughter was determined well by the long vs short interval from vowel to vowel (58%), the high vs low maximum or mean F0 value of the noise-vowel reiteration (32, 31% respectively) and by the small vs large rate of amplitude diminishment (31%). These three variables accounted for 85% of the total variance.

The present results, however, do not ensure that those acoustic variables found to be correlated with the hypothetical perceptual factors are the 'real' perceptual correlates. A research using synthesized stimuli will be necessary in order to evaluate the perceptual effect of those acoustic variables.

Figure 2

Wide-band spectrograms of a typical happy or funny laughter (No.16) and a mocking laughter (No.8)