Extraction of Speech Parameters Relating to the Characteristics of Emotional Expression: Focus on the Attentiveness-related Parameters such as the Nursing Experience

Gao Yahan¹, Yuko Ohno¹, Shima Okada²

1) Department of Mathematical Health Science, Graduate School of Medicine, Osaka University
2) Department of Robotics, Ritsumeikan University

Life experiences and the roles people play in society affect the way people express emotion. Nurses must be skilled communicators because they interact continually with all sorts of people—different personalities, different occupations, varying states of health—so the nature of the job affects the way nurses express emotion, especially on the attentiveness. In this study, the attentiveness toward others (patients, younger siblings, and their own children) was noticed and the relationship between attentiveness-related parameters and speech-related parameters were investigated. For the study, we created a meaningless eight-syllable sample word, then had the participants enunciate the made-up word simulating five emotional states: neutral, joy, sadness, anger, and surprise. We conducted a sound analysis of the different emotions by measuring the fundamental frequency F0 (Hz), the max sound pressure frequency Fmax (Hz), max sound pressure MaxdB (dB), and the average sound pressure AvedB (dB), we also measured the phonation time in seconds. We then statistically analyzed the variation from the neutral state in each emotion. Twenty-two women took part in the study (mean age 29.6 ± 6.4), including 12 women with nursing experience (mean age 32.4 ± 6.0) and ten women without nursing experience (mean age 26.1 ± 5.1). The results revealed significant differences in speech parameters correlated with differences in attentiveness-related parameters—nursing experience, and whether the participant had younger siblings and/or child/children—thus suggesting a relationship between nursing experience and emotional expressiveness.

Key Words
Speech parameter analysis, Emotion, Emotional expression, Occupation characteristic, Nursing, Attentiveness

1. Introduction

Professional actors and actress play various emotions very smart and their recitation affect audience very much. That is why the most studies analyzing the relationship between speech and emotion have used sample utterances spoken by professional actors and actresses for the expressions of emotion.¹⁴ Meanwhile, the characteristics of emotional expression in everyday life vary significantly depending on one's life experiences, role in society, and other factors. This suggests to us that take into account of occupational or societal roles of the participants are important when analyzing the features of speech parameters in emotional expression. Given the nature of their occupation, nurses must interact with all kinds of people from various walks
of life and in different states of health, so they must be skilled communicators. In other words, it is an occupational requirement for nurses to exercise discipline in their speech representation. So far no study has examined the speech representation or features of emotional expression of nurses. Here we noticed on the relationship between emotional expression and speech representation by focusing specifically on the practical work experience of nurses who exhibit attentiveness toward patients. Relating to the attentiveness, there would be some variations such as the attentiveness toward patients, younger siblings, and their own children. In this study, we presented attentiveness-related parameters and investigated the relationship between attentiveness-related parameters and speech-related parameters.

Connections between speech and emotion have been extensively studied in recent years, and previous research has shown that the fundamental frequency F0 is affected by changes in emotion \(^5\), and the fundamental frequency F0 and sound pressure differ for different emotions \(^6\). While many of these studies have compared differences in the fundamental frequency for different emotions in speech, no one has investigated speech parameters of emotions varying from a neutral state. We assume that the change from neutral for each emotion an individual feels is related to the person’s ability to express emotion, and this study will analyze the change in the fundamental frequency from neutral for a range of emotions.

2. Materials and Methods

(1) Subjects
After being briefed on the purpose of the study and understanding that the participants would not be identified or compared, and other experimental conditions, 22 candidates approved of the study's objectives and volunteered to participate. The speech parameters for men and for women are quite different \(^7\), so only women subjects were used in this study. We collected basic data on the participants prior to the experiment including nursing work experience, age, and family structure.

(2) Experimental Conditions
a. Speech Samples
Most past studies have used actual words for speech samples \(^1\)-\(^4\), but the associative implications of real words can affect the emotional content \(^8\), so we sought to eliminate this effect to make sure that we were measuring the true expression of emotion. To eliminate any potential bias from the associative implications of an actual word, we created a nonsensical word to compare parameters among the emotions studied. To create the speech sample, we arranged 50 sounds in random order to create eight syllables' word, then created the meaningless words and selected "ko-ne-yo-chi-yu-e-ho-te" that were evaluated as the easiest to pronounce by two examinees.

Note that most past studies have used actual words for speech samples \(^1\)-\(^4\), but the associative implications of real words can affect the emotional content. If you use an actual word or phrase, it is really impossible to tell if the expressed effect derives from the cadence of word or from the meaning of the word \(^8\), and this concern led us to adopt a meaningless word for the speech sample in this study.

b. Types of Emotion
Based on the previous report \(^9\), we considered four basic emotional states: joy, sadness, anger, and surprise in our study, and added neutral as a control state.
c. Measurement

Recording was done in a quiet room using an ICR-PS1000M digital sound recorder (Sanyo Audio), and the microphone maintained at a distance of 4-6 cm from the subjects' mouths to compare sound pressure. Figure 1 shows an overview of the measurement setup. The recording procedure was as follows. First, we showed the subjects a visual cue card with a facial expression depicting an emotional state and a written character corresponding to the emotional state. Next, we presented a card showing the meaningless speech sample word "ko-ne-yo-chi-ya-e-ho-te," and had the subjects read the word (taking as much time as they wanted) expressing the emotion presented while we recorded the speech. Each emotion was recorded three times, and the emotions were presented in the following order: neutral(3 times), joy(3 times), neutral(1 time), sadness(3 times), neutral(1 time), anger(3 times), neutral(1 time), and surprise(3 times). In order to reset or return to a normal state between emotions, we had the subjects read the made-up speech sample word one time while showing them the neutral emotion cue card. Because it is necessary to acquire the sound data of the neutral, we record this three times in the first of the measurement and used only the third one. This is similar to the other emotions. We concluded that the third recorded reading was a more stable expression of emotion than the previous two, so we discarded the first two samples and used only the third recorded speech samples for the emotion analysis to follow.

d. Sound Analysis

Features of speech emotion are said to derive from prosodic changes, so here we defined four speech feature parameters — speech duration change, frequency change, sound pressure change, and speech variations — and calculate the values of each parameter. For the speech duration change, we measured the total time (TT) needed to speak the eight-syllable sample word for each emotion. In pronouncing the word, a natural break occurs after the first four syllables, so we measured the first half time (FT) to say the first four syllables and the last half time (LT) to speak the last four syllables.

Connections between speech and emotion have been extensively studied in recent years. Previous research has shown that the fundamental frequency F0 is affected by changes in emotion, and the fundamental frequency F0 and sound pressure differ for different emotions. In this study, the fundamental frequency Hz (F0) is thought to accurately reflect the distinctive features of an individual's speech, so we included F0 as one of the feature parameters in the study. We investigated the max sound pressure frequency in Hz (Fmax) to determine how the frequency changes when the voice was raised to its highest level in a state of emotion. For the sound pressure parameter, we measured the average...
sound pressure in dB (AvedB) for each emotion, and the max sound pressure in dB (MaxdB) when the voice was raised to its highest level in a state of emotion.

Although many of previous studies have compared differences in the fundamental frequency for different emotions in speech, no one has investigated speech parameters of emotions while varying from a neutral state. It is supposed that the change from neutral for each emotion of an individual feels is related to the person's ability to express that emotion. So this study analyzed the change in the fundamental frequency from neutral for a range of emotions. In order to measure the speech variations parameter for each emotion, we took the value of speech parameter for the neutral phonation as standard, calculated the ratio of each emotion to the standard for the frequency and phonation time, and then calculated the differences of each emotion from the standard for the sound pressure. The analysis of the speech samples was done using Audacity 1.3.5.

e. Subject Characteristics

The characteristics of the participants were defined first in terms of number of years nursing experience, and then based on an attentiveness score that factored in three kinds of experience. These experiences are indicative of attentiveness toward others: whether they had children, whether they had younger siblings, and whether they had experience as a nurse. We further assume that years of attentiveness toward patients are related to the number of years of the participants had worked as nurses. Finally, assume that the richer life experience of the older participants would make them better at expressing emotion, so we split the data into two group based on the medium age of the subjects, the younger group ranged in age from 22 to 27, and the older group ranged in age from 28 to 40.

We then divided the participants into two groups based on the attentiveness-related parameters — whether they had children, whether they had younger siblings, whether they had nursing experience, whether they exhibited or did not exhibit attentiveness, and whether they were in the younger or older age group. We performed t-tests on each speech parameter across all groups. We also applied Spearman's rank correlation coefficient to analyze the relationship between the attentiveness parameters (years of nursing experience and attentiveness score) and the speech parameters.

3. Results

We analyzed the relations between attentiveness parameters and speech parameters, and the statistically significant items are presented in Tables 1 and 2.

First regarding the emotion joy, in the first line of the speech parameter in Table 1, the result shows that the participants who exhibited attentiveness (0.27) had significantly higher scores than the participants who did not exhibit attentiveness (-2.77) in the speech parameter of MaxdB(joy)-MaxdB(neutral) (show it in Figure 2). As show in Table1, the participants who were older, had children had significantly higher scores than the participants who were younger, without children in the speech parameter of MaxdB(joy)-MaxdB(neutral). The older participants scored higher than the younger participants in the speech parameter of MaxdB(joy)-AvedB(joy). From Table 2, it was found that there were a positive coefficient of correlation of 0.44 between MaxdB(joy)-MaxdB(neutral) and the attentiveness score (show it in Figure 3). These finding suggest that when expressing the emotion joy, the older participants with children and greater attentiveness express themselves...
with louder voices. In terms of phonation time as show in Table 1, the FT(joy)/FT(neutral), LT(joy)/LT(neutral), and TT(joy)/TT(neutral) values were significantly shorter for the participants with children than for those without children, suggesting a tendency toward shorter phonation time when expressing these emotions (joy and neutral).

For correlation of phonation time and attentiveness score, Negative significant correlations were observed of -0.46 between the attentiveness score and LT(joy)/LT(neutral) and -0.43 between the attentiveness score and TT(joy)/TT(neutral) in Table2, revealing that the participants with high attentiveness scores shortened the time to speak the
last 4 syllables of the sample word, so the overall length of
phonation tended to be shorter. We found larger Fmax(joy)/
F0(joy) values for the participants with children than for
those without children, and the correlations were 0.48 with
attentiveness and 0.45 with years of experience. This finding
suggests that having younger siblings, high attentiveness
scores, and longer years of service tends to be manifested as
louder and higher voices when expressing joy.

Now turning to sadness, the Fmax(sadness)/Fmax(neutral)
values were smaller for the participants with greater
attentiveness (show in Table 1). From this result, we
observed a tendency of expressing sadness, which were the
participants who exhibited attentiveness using with lower
voices when speaking in a loud voice.

For the emotion of anger, the participants with more
experience had larger MaxdB(anger)-MaxdB(neutral)
values in Table1(show it in Figure 4), thus confirming that
those with more experience speak raise their voices when
expressing this emotion. The participants with younger
siblings had higher Fmax(anger)/F0(anger) values than
the participants without younger siblings, suggesting that
women at least with younger siblings tend to speak in a
higher voice when angered than women without younger
siblings. We found a positive significant correlation of
0.49 between years of experience and MaxdB(anger)-
MaxdB(neutral), thus revealing that more years of
experience correlates with a louder voice when expressing
this emotion.

For the emotion surprise, we found significantly large
changes in the values Fmax(surprise)/Fmax(neutral) for the
younger participants in Table 1, indicating high frequency
when their voices reached the loudest level.

Finally, for the neutral emotion, the Fmax(neutral)/
F0(neutral) values in Table1, indicated that the older
participants were speaking in higher voices when they raised
their voices.

4. Discussion

In previous research involving the analysis of speech and
emotion, many studies have used actual words or phrases
as speech samples, many studies have used expressions
of emotion created by professional actors and actresses as
data 1)-4), and other studies have analyzed the characteristic
attributes of the speech of patients and military personnel. However, these studies mainly focus on connections with depression and stress\textsuperscript{11)-12)}, and no one had investigated the effects of occupations and attribute features on the expression of emotion. In this study we have explored for the first time how attentiveness toward others is related to emotional expression.

Assuming that the way people express emotion varies depending on their roles in society and life experiences and the associative implications of real words can affect emotional expression, we conducted this study using a nonsensical made-up word as the speech sample, and used this sample to draw comparisons between different emotions and measure variations from a neutral emotion state. Specifically, we analyzed the features of emotional expression for a particular occupation, that of nursing.

Furthermore, the findings of previous research on speech and emotion have mainly focused on how the voice changes in expressing different emotions such as showing how the fundamental frequency F0 is affected by changes in emotion\textsuperscript{5)}, or highlighting the differences in fundamental frequency F0 and sound pressure between different emotions\textsuperscript{6).} But very few studies have examined variations from a neutral state of emotion, or attempted to compare the effects on emotional expression of different attributes. With this study we offer a new approach and a number of interesting insights.

Among the four emotions studied, we found especially in the speech parameters for expressing joy that there was a range of differences in degree of attentiveness. This we think is because a person’s age and attentiveness related parameters are more readily manifested when expressing joy. Particularly in the parameters relating to nursing experience, there was considerable Fmax variation for the emotion joy manifested by the voice becoming higher as it got louder. When attentiveness scores were high, the intonation of the voices tended to become clearer than for the set of 0 attentiveness scores, suggesting that the participants could control their speech even as they raised their voices. When expressing anger, for example, the participants with attentiveness showed a tendency for their voices to raise or become louder. In the study we did not observe a clear connection between the emotions of sadness and surprise and nursing experience parameters. We surmise this is because there are fewer occasions to express these emotions compared to joy and anger at hospitals and other places where nurses work.

The participants ranged in age from 22 to 40, so we could see a close connection to age, number of years work experience, whether the participants had children or not, and other parameters. However we infer that the items conceived here as attentiveness-related parameters are all closely interrelated except for the question of whether the participants had younger siblings. This should be borne in mind when considering the interrelations between parameters.

Putting it all together, it is thought that the speech-related characteristics found in joy would imply the sympathy for the others. With voice, behavior and facial expression, a nurse would cheer up for patients as same as an elder sister would encourage her younger siblings. In such cases, a nurse or an elder sister express her joy and that brings patients or siblings can-do spirits. Patients and children are sensitive to pretended emotion so that nurse or elders would be trained their expression to be faithful. In case of anger, it is thought to be contrary. A nurse would often forbid a patient smoking or drinking, sometimes sweets for the treatment, and sometimes the patient would be disobedient. In such
case, the nurse wouldn't be angered actually against the patient. She would pretend being angry as a professional. Namely, the anger expressed by a nurse or elder sister seems including some pretence. These would be the reason that there found relationship between attentive-related parameter and speech-related parameters in joy and anger.

Continuing this approach, we plan to conduct a study to see if the speech parameters derived from the occupational characteristics of nurses are different for other occupations. For example, this approach would be very useful for comparing teachers, doctors, councilors, and other occupations filled by women and for assessing the effects that having children and/or younger siblings might have on women in other occupations.

Now that we have fairly well identified the items associated with the speech-related parameters of emotional expression, we would like to explore skills-related parameters in a future study.

5. Conclusions

Starting from the premise that expressions of emotion are affected by life experiences and one's role in society, here we analyzed and compared how the expression of emotion is affected by being a nurse (who must be a good communicator to communicate with patients) and by life experiences of having younger siblings and children, and how emotional expression varies from a neutral state. Characteristic relation was observed between the attentiveness-related parameters and speech-related parameters in joy and anger. The relation found in the study and the approach used would be valuable and significant to further research in this field.
References


13) Bowles N., Mackintosh C. & Torn A. Nurses’ communication skills: an evaluation of the impact of solution-focused communication training, Journal of


感情表出の特性に関連する音声関連パラメータの抽出
-看護実務経験などの気配り関連パラメータを考えて-

高亜罕*1 大野 ゆう子*1 岡田 志麻*2

*1 大阪大学大学院医学系研究科数理保健学研究室 *2 立命館大学理工学部ロボティクス学科

看護師は看護業務において、さまざまな性格や職種、健康状態の人との会話が大きな比重を占めるため、高いコミュニケーション能力が必要とされる。一方感情の表出には看護経験以外の影響因子も存在する。本研究では患者、弟妹、子供など他者への気配りに注目し、これら気配り関連パラメータと音声関連パラメータとの関係を検討した。まず、「喜んでいる」「悲しんでいる」「怒っている」「驚いている」「平静」という5つの感情について感情を込めた発声実験を行い、感情別音声解析を行い、基本周波数F0(Hz)、音圧最大時周波数Fmax(Hz)、最大音圧MaxdB(dB)、音圧平均AvedB(dB)、発声持続時間(s)を測定し、感情ごとの変動について統計的分析を行った。被験者数は女性22人（年齢：29.6±6.4歳）、うち、看護経験者12人（年齢：32.4±6.0歳）、看護未経験者10人（年齢：26.1±5.1歳）であった。その結果、感情別で音声の表出の特性が異なり、気配り関連パラメータとの関連が示唆され、5つの感情の内特に「喜んでいる」「怒っている」が関連性を示した。

和文キーワード
音声パラメータの解析、感情、感情表出、職業特徴、看護、気配り

2009年11月26日受付；2010年6月25日受理
連絡先: 〒565-0871 大阪府吹田市山田丘1-7
大阪大学大学院医学系研究科数理保健学研究室
Tel: 06-6879-2526
大阪大学大学院医学系研究科数理保健学研究室
高亜罕
E-mail: gaoyahan@sahs.med.osaka-u.ac.jp

ITヘルスケア 第5巻2号, 2010：114-123