

An Analysis of the Phonetic and Acoustic Phenomenon on Four-Letter words

中英文髒話語音及聲學現象之研究

Nien-Chen Lee (李念貞)¹ Li-mei Chen (陳麗美)²

¹國立成功大學外國語文學系

²國立成功大學外文系(所)助理教授

住址：701 台南市大學路1號。成功大學外文系

電話：(06) 275-7575 Ext 52231 傳真：(06) 200-4319

Email：leemay@mail.ncku.edu.tw

摘要

本研究探討在中、英文最常使用的髒話中的特殊語音及聲學現象及其造成的心理影響，其主要目的是為人類語言的普遍性尋求一些具體的科學證據，期可作為未來相關研究之參考。研究結果顯示：1)語音類別分布有特殊共同性-字首塞音及擦音，2)中英文髒話皆較一般用語的響度來的高 3)中文髒話聲調多屬高降調(第四聲)，英文則傾向下降語調，且其組成母音多為中央元音(schwa)，4)重音落於第一音節。

Abstract

This paper, based on the phonetic and psychoacoustic theories, is an analysis of the phonetic phenomenon of the most frequently used taboo words in both English and Chinese, and their psychoacoustic impacts made in production and perception. The experimental results showed four similarities in FLW (four-letter-words) of these two languages: 1) the shared phonetic features of plosive and fricative-initials, 2) greater intensity than the neutral, 3) high-falling tone in Chinese and falling intonation in English, and 4) the stress in first syllable. It was hypothesized that dirty words, functioning as a speech act, were used to express particular emotions, whose specific phonetic features had certain psychoacoustic effects upon both the speakers and listeners.

關鍵字： 禁忌用語、心理聲學影響、語音象徵

Keywords: taboo words, psychoacoustic effect, sound symbolism

I. INTRODUCTION

Taboo words are regarded by the majority as “a bad language”, most of which denote certain taboo subjects, including bodily effluvia, sex, and private body parts. However, despite its pejorative disposition, there are swearwords in most languages worldwide. It is also a linguistic behavior performed commonly by the average people, either out of some emotional arousals or merely personal habits. From the perspectives of linguistics, the high frequency and wide distribution of dirty words deserve particular investigation, for they may indicate a possible universality. Moreover, from the natural language processing viewpoint, since a speech signal includes both linguistic information and the speaker’s emotions, it might make some contribution to future works on emotion recognition technology by examining the association between the two.

The following sections introduce the recent linguistic study of dirty words and the researches from other disciplines on emotion recognition in TTS (Text-to-Speech) technology, both of which dealt with human speech sounds with emotions through different methodologies. Also the theory of sound symbolism adopted to support the experimental results in this paper is elucidated here.

LINGUISTIC LITERATURE

The word “taboo” was originally from Togan, meaning “prohibited behavior” (Allan & Burridge, 1991), yet many works devoted to the research of dirty words indicated that it seemed to be a very widespread linguistic phenomenon. The comment that only Japaneses, Amerindian and Polynesian language lacked swearwords made by both Montagu (1967, p.55) and Bryson (1900, p.214) was also shown commonly in the literature in this area (cited in Kidman, 1993, p.2). This assumption implied the possibility that swearing might be a universal linguistic behavior.

Jay (1992) in *Cursing in America* has done numerous field researches and laboratory experiments on swearing in American English. His study has covered a wide range from the acquisition theory, the pragmatic factors, to the human psychological mechanism, all demonstrated with apt scientific evidences to support the hypothesis of swearing as a shared mechanic function in human linguistic system, which was necessary for human to express emotions, especially the suppressed anger or depression, as a relief mechanism (Averill, 1983; Bao & Chen, 2003). As for phonology, McCarthy

(1982) has made research to analyze the prosodic structure and the formation of infixation of dirty words, such as “fuck” in “abso-fuckin-lutely”.

Also in the field of TESL (Teaching English as a Second Language), Register (1996) proposed that the wide exposure, high frequency of usage, and the simple monosyllabic structure of dirty words made it easier for ESL students to acquire English swearwords than the non-obscene colloquial expression. As for Chinese swearwords, Huang and Tian (1990) have conducted a research on linguistic taboo in Chinese from the sociolinguistic perspectives.

EMOTION RECOGNITION REVIEW

The recognition of emotion in human speech has drawn increasing attention recently due to its application in natural language processing. Because of the monotonousness of the speech synthesized by computers, many researches have been done to extract particular acoustic features from human sounds with emotional contour so as to promote the human-likeness in voice quality (Boa & Chan, 2003; Kwon et al, 2003). However, performance in literatures was hard to compare due to the lack of common database. The database for testing emotion recognition was mainly retrieved from three methods, first, recording of utterances with target emotions simulated by actors, second, recording of the driven response of the actor in emotion triggered context, and third, which is hard to obtain, the actual real-world recording of utterances that express emotions (Yacoub et al, 2003).

SOUND SYMBOLISM (PHONAESTHESIA)

Phonaesthesia, also called sound symbolism, is a theory that states there are links between particular phonological patterns and certain meanings (Crystal, 1997). Klerk and Bosch (1997) indicated that some recognizable phonological preferences have been found in the English lexicon, of which certain phonemes and stress patterns may bear positive or negative connotations.

The sound-meaning associations were also observed in Chan's (1996) investigation, which exhibited the supportive evidences from both English and Chinese for his postulated typology of the universal synthetic sound symbolism. Chan has provided the linguistic data of the /i/ sound for the concept of “small” and /a/ for that of “large”, and of words that validated the association of labialization with “roundness” concept in both English and Chinese. This result led him to assert that this sound-meaning

correspondence was perception-based, reflecting universal tendencies, and hence the acoustic consideration should be accounted for.

The cross-language discovery of sound-meaning correspondences implied that there indeed were certain psychoacoustic effects a sound might impose on human cognitive systems.

II. METHODOLOGY

Two approaches, including the phonetic analysis and the acoustic experiment, were adopted in this study.

PHONETIC APPROACH

The data of English dirty words were extracted from Jay's field research, in which the most frequently used English FLW were gathered by surveying both the teenage and adult speakers in the United States. For Chinese swearwords, the data were collected through surveying the native Chinese speakers of the age from 18 to 25 on campus. The questionnaire was designed with several questions for the targeted subjects, including the possible social context in which they may hear or utter dirty words, the psychological impact in the perception and production of FLW, and the most frequently heard and spoken swearwords in their daily life. The retrieved preferred FLW in both languages were categorized according to their phonetic attributes.

ACOUSTIC APPROACH

The acoustic investigation mainly focused on Chinese dirty words due to the difficulty of recording the actual utterance by native English speakers in Taiwan. In order to keep the natural quality of the database, the utterance of FLW were gathered by recording the real-world conversation, from which the dirty words were extracted. There were totally four record files, three by males and one by females, each of which contained a conversation of about 45 minutes, and carried out by three to four speakers. Each recording was divided into two parts, the neutral utterances and the FLW by using the speech analysis software CoolEdit. The average values of frequency (Hz) and intensity (dB) of the two parts in each recording were analyzed and compared so as to extract the major factors in discriminating FLW from the neutral.

III. FINDINGS AND DISCUSSION

From the analysis of retrieved linguistic data and recording, certain phonetic and acoustic phenomenon was discovered in the preferred English and Chinese FLW. Both the most frequently used Chinese FLW and the English ones obtained from Jay's study could be classified into at least three categories according to their phonetic features in word initial positions, which were the plosive-initial, fricative-initial, and others. Plosives are consonants made when a complete closure in the vocal tract is suddenly released, causing the air pressure which had built up behind the closure rush out with an explosive sound. Fricatives are sometimes called "spirant" or hissing sound, made when two articulators come so close to each other, causing air moving between them to generate audible friction (Giegerich, 2000).

From Table 1 we can see the salient distribution of plosive and fricative-initial in the most frequently used dirty words of the two languages. In addition to the phonetic similarity in word-initial, most of the preferred Chinese FLW are marked with high-falling tone, while the English one with descending intonation, accented in the first syllable, most of which have the schwa (reduced vowel) as their peaks. The sound spectrum of the top three most frequently used English dirty words displays similar patterns.

From the results above, the hypothesis that such particular phonetic features among FLW across these two languages may function as phonoaesthesia, a sound symbolism that linked phonological patterns to certain meanings, was therefore proposed. As what mentioned in the background introduction, the phonetic characteristics in FLW may have some psychological influence on both the speakers and hearers, which resulted in their tendency in choosing such words to express certain emotions.

Besides the phonetic examination, in order to find psychoacoustic evidences to support the hypothesis, the acoustic attributes of dirty words were investigated as well. Table 2 shows the average value of the pitch and energy contour of the neutral and FLW in each recording file. However, the experimental results in this part did not demonstrate a strong backup for the postulated hypothesis, for the average value of frequency in FLW showed no drastic difference from the neutral, and the intensity was merely slightly larger. It might be the interruption of background noises in real-world conversations that biased the consequence.

Table 1: Phonetic distribution in most frequently used FLW

	Chinese	English
Plosive-initial	9	18
Fricative-initial	7	10
Others	3	6
Total	19	34

Table 2: Average frequencies and amplitude in neutral utterances and four-letter-words

N (neutral utterances), F (four-letter-words)

Recording		Frequency(Hz)	Intensity(dB)
A	N	195.6	49.6
	F	183.6	53.2
B	N	222.8	42.2
	F	191	49.2
C	N	208.6	51.9
	F	212.7	57.1
D	N	234.3	46.3
	F	214.7	51.5
Average	N	215.3	47.5
	F	200.5	52.8

IV. CONCLUSION

The result of this research on the phonetic and acoustic phenomenon of both English and Chinese dirty words demonstrated the shared phonetic qualities in the two languages, which were the plosive and fricative-initial that may serve as an expressive speech act to release one's emotions. The discovery not only echoed with the theory of the universal synthetic sound symbolism (Chan, 1996), but also solidified the hypothesis that the phonological patterns might influence human speech perception. The purpose of this study was to find some psychoacoustic evidences of sound-meaning correspondence in the universal linguistic phenomenon of taboo words, which may be applied to advance the modeling of emotional contours in speech recognition and synthesis systems, and also contributed to further researches on human cognition of sounds in future works.

V. ACKNOWLEDGEMENT

本文撰寫期間承蒙成大資訊工程研究所及所有參與錄音工作人員之熱心協助，特此申致謝忱。

VI. REFERENCES

- (1) Allan, K. and Burrige, K., Euphemism and dysphemism: Language used as shield and weapon, Oxford University Press, New York (1991).
- (2) Averill, J. R., Study on anger and aggression: Implications for theories of emotion. *American Psychology*, 38(1), 1145-1160 (1983).
- (3) Bao, T. L. and Chen, Y. D., Emotion recognition in Mandarin speech, Diss. National Taiwan Normal U (2003).
- (4) Chan, M. K. M., Some thought on the typology of the sound symbolism and the Chinese language, in C. C. Cheng et al. (Ed.), *Proceeding of the English North American Conference on Chinese linguistics*, Vol. 2. pp.1-15, GSIL, Los Angeles(1996).
- (5) Crystal, D, *A Dictionary of Linguistics and Phonetics*, Blackwell, Massachusetts (1997).
- (6) Giegerich, H. J, *English phonology: An introduction*, Cambridge University Press, Cambridge (2000).
- (7) Huang, H. and Tian, G., A sociolinguistic view of linguistic taboo in Chinese, *International Journal of the Sociology of Language*, 81, pp.63-85(1990).
- (8) Jay, T., *Cursing in America*, John Benjamins, Philadelphia (1992).
- (9) Kidman, A., *How to do things with four-letter-words: A study of the semantics of swearing in Australia*, Unpublished Dissertation, University of New England(1993)
- (10) Klerk V. D. and Bosch, B., The sound patterns of English nicknames, *Language Sciences*, 19(4), pp.289-301 (1997)
- (11) McCarthy, J. J., Prosodic structure and expletive infixation, *Laguage*, 58(3), pp.574-590 (1982).
- (12) O-W. Kwon, K.-L. Chan, J. Hao, and T-W. Lee, Emotion recognition by speech signals, *Eurospeech 2003*, pp.125 – 128 (2003, September).
- (13) Register, N. A., Second language learners and taboo words in American English, *English Today*, 12.3(47), pp.44-49 (1996, July).
- (14) Yacoub, S. et al, *Recognition of Emotions in Interactive Voice Response Systems*, Hewlett-Packard Laboratories (2003).