Vulgarities are fucking funny, or at least make things a little bit funnier

Jonas Sjöbergh
KTH CSC
SE-100 44 Stockholm, Sweden
jsh@nada.kth.se

Abstract

We present a program that generates puns in Japanese. The new idea of this program is that by using vulgar or taboo words the generated jokes become a little bit more funny. The evaluation shows that while the generated jokes are not very funny, they are funnier than jokes generated using non-vulgar words.

1 Introduction

This paper presents a system that generates puns in Japanese. There have been some attempts at this earlier (Yokogawa, 2001). In (Binsted and Takizawa, 1998) a system for generating punning riddles in Japanese is described. We developed a very similar system to try out a new hypothesis. This hypothesis is that while computer generated jokes may not be very funny, if you use "bad words" they might become a little bit funnier.

Two different kinds of jokes were generated, riddles using puns and proverbs changed in a pun-like way to new expressions.

2 Preliminary test: Punning riddles

A very simple program for generating riddles was created. When generating punning riddles, three connected things are searched for and then inserted into a fixed template. This template is "An X is an X but what kind of X is Y? Z!". Here X and Z are two words that have similar pronunciation. Y is a description that matches the meaning of Z. To make things funnier, Z is always chosen from a list of vulgar or taboo words. This list was constructed for this program, with words collected from different sources.

Sound	Similarity
i, e	0.7
u, o	0.7
a, other vowel	0.55
V_1, V_2	0.5
n, m	0.9
g, b, d	0.9
k, p, t	0.9
ki$V$, kiy$V$	0.95
ki$V$, ky$V$	0.9
kiy$V$, ky$V$	0.9
V, yV	0.7
shi, hi	0.95
voiced/unvoiced (z-s, d-t, g-k, . . .)	0.7

Table 1: Pronunciation similarity scores used. V indicates a vowel sound.

An example joke (freely translated into English) generated by the program is: "Sisters *(shimai)* are sisters, but what kind of sisters are untidy? Sloppy bitches *(darashinai)*."

X and Z are found by looking through a dictionary of words, checking all word pairs to see how similar the pronunciation is. Similarity is based on a few simple rules for which sounds are similar in Japanese. Similarity scores used are shown in table 1.

When Z has been selected, the description Y is generated by looking in a dictionary of Japanese, with descriptions of the words also written in Japanese. The Sanseido online dictionary was used for this. If the first sentence in the description is short, the whole sentence is used as Y. Otherwise the first word of the sentence is used.

Since the joke is not funny if X and Z are synonyms, a check is also done to see if the meanings are too similar. This is done by checking the word overlap of the English de-

scriptions of X and Z in a Japanese-English dictionary, for which the EDICT (Breen, 1995) was used.

A very small evaluation of this program was done by letting four Japanese readers read jokes and decide how funny they were on a scale from 1 (not funny) to 5 (very funny). It was also possible to select "I don't understand" if for some reason it was impossible to tell if the joke was funny or not; for instance if it contained difficult words that the reader did not understand.

The evaluation contained 5 examples of similar (though not vulgar) jokes created by humans, found on the Internet. There were also 10 non-jokes, created by selecting X, Y and Z by randomly drawing words from the dictionary. Finally there were 15 jokes from the program described above and another 15 jokes generated not using the vulgar words (i.e. Z was selected as a similar sounding word from the normal word list).

The reason that the evaluation was very small was that the jokes were considered almost completely unfunny by most readers. All readers did give vulgar jokes higher scores than non-vulgar jokes, though, and most readers gave human generated jokes the highest scores.

The comments from these preliminary evaluations indicated that this kind of punning riddles were not considered funny, even if they were cleverly created puns by humans. Thus it was decided to change the generated joke type to something with more humor potential and discontinue the evaluation of the riddles.

3 Proverb punning

The second type of jokes that were generated was based on Japanese proverbs and idiomatic expressions. A joke is generated by presenting a proverb and then the same proverb with one word changed to another similar sounding word, changing the meaning of the phrase. The similar sounding word is always chosen from the list of vulgar or taboo words.

An example joke from the program (with approximate English translations) is: "*isogaba maware* (more haste, less speed) – *kusobaba maware* (turn away, you old hag)".

The evaluation was done in a similar way as in the previous section, by having six Japanese

Type	Score	Too hard
Human	2.6	22%
Vulgar	2.2	22%
Normal	1.8	36%
Random	1.5	40%

Table 2: Evaluation of generated proverb jokes. Mean value of funniness (from 1 to 5) and the percentage of the jokes that were not understood by the readers.

readers read jokes and selecting from 1 to 5 how funny they were, or indicating that the joke was not understood.

Five non-jokes were generated by changing a random word from the proverb to a randomly selected word from a dictionary. 10 jokes were taken from a web site with proverbs and a changed form of the proverb. These human generated jokes were usually more sophisticated than the computer generated jokes, for instance changing more than one word. Many of them also turned out to be very very vulgar. Then there were 17 jokes generated by the program above and 15 jokes generated in the same way but using normal words from the dictionary instead of the list of vulgar words.

Table 2 shows the results of the evaluation. As expected, human produced jokes are considered funnier than the computer generated jokes, though still not very funny. Vulgar jokes are considered funnier than jokes generated in the same way but using normal words. Least funny are the randomly changed proverbs which are also the ones that are hardest to find any meaning in. Vulgar jokes are more easily understood than non-vulgar jokes.

4 Discussion

The generated jokes were not considered very funny, most jokes scored the lowest possible value of one. Jokes by humans also scored very low. The scores varied quite a lot between different readers, though, with one reader assigning a mean score of 4.1 and another 1.2.

Many jokes were hard to understand. This was caused by among other things using proverbs that the reader was not familiar with, using slang words or difficult words that the reader did not know, writing otherwise known words using difficult kanji that reader did not

understand and similar things. There were also some mistakes in the automatic assignment of pronunciation to some words written with kanji, which was confusing.

Many jokes were also incomprehensible for the simple reason that the new word did not make any sense in the changed proverb, so it was impossible to construe a reasonable meaning for the new phrase. This was less of a problem with the vulgar words than other words, likely because the vulgar words have many meanings or can be used in many ways.

The readers also had the possibility to write any comment they liked about the jokes. One comment was that some jokes were cleverly created but too vulgar or offensive, so the total funniness was low, at least for this reader. This usually referred to the human generated jokes.

Other comments included things like stating that a certain joke was not funny, but with a very small change to another word too, it would be much funnier. This indicates that it is probably a good idea to change more than one word, selecting several new words that are related to each other.

5   Conclusions and future work

The theory that bad words are funny seems to hold for the generated jokes. When using bad words, it also seems to be easier to find a reasonable interpretation of the new generated proverb. While the generated jokes were not considered very funny in general, neither were the jokes generated by humans. There were some automatically generated jokes that were considered quite funny by most readers.

Future possibilities include changing the proverbs in more sophisticated ways, by for instance selecting words that have a meaning related to the remaining words in the proverb, by changing more than one word in each proverb, by changing words to words without similar pronunciation but with related meanings (such as antonyms) etc. Another possibility is to generate puns on other types of texts, such as titles of famous movies or books.

Generating jokes that are funny regardless of context, such as these jokes have to be since they have no context, is quite hard. Another possibility is to generate jokes in a certain con-text, which then can be used. A simple example would be a dialogue system, were the previous sentences can be used as a context to base jokes on. While in general it is easier to produce jokes that are funny in a certain context than jokes that are always funny, understanding how to use the context to make a funny joke is of course a quite hard problem.

References

Kim Binsted and Osamu Takizawa. 1998. BOKE: A Japanese punning riddle generator. *Journal of the Japanese Society for Artificial Intelligence*, 13(6):920–927.

Jim Breen. 1995. Building an electronic Japanese-English dictionary. In *Japanese Studies Association of Australia Conference*, Brisbane, Australia.

Toshihiko Yokogawa. 2001. Generation of Japanese puns based on similarity of articulation. In *Proceedings of IFSA/NAFIPS 2001*, Vancouver, Canada.