

# Toward Construction of Spoken Dialogue System that Evokes Users' Spontaneous Backchannels

Teruhisa Misu, Etsuo Mizukami, Yoshinori Shiga, Shinichi Kawamoto,\*  
Hisashi Kawai and Satoshi Nakamura

National Institute of Information and Communications Technology (NICT), Kyoto, Japan.  
teruhisa.misu@nict.go.jp

## Abstract

This paper addresses a first step toward a spoken dialogue system that evokes user's spontaneous backchannels. We construct an HMM-based dialogue-style text-to-speech (TTS) system that generates human-like cues that evoke users' backchannels. A spoken dialogue system for information navigation was implemented and the TTS was evaluated in terms of evoked user backchannels. We conducted user experiments and demonstrated that the user backchannels evoked by our TTS are more informative for the system in detecting users' feelings than those by conventional reading-style TTS.

## 1 Introduction

One of the most enduring problems in spoken dialogue systems research is realizing a natural dialogue in a human-human form. One direction researchers have been utilizing spontaneous nonverbal and paralinguistic information. For example,

This paper focuses on backchannels, one of the most common forms of para-linguistic information in human-human dialogue. In particular, we focus on users' verbal feedback, such as “*uh-huh*” (called *Aizuchi* in Japanese), and non-verbal feedback in the form of nods. Such backchannels are very common phenomena, and considered to be used to facilitate smooth human-human communications. In this regard, Maynard (Maynard, 1986) indicated that such backchannels are listener's signals to let the speaker continue speaking (continuer), to indicate that the listener understands and consents. It was also hypothesized that humans detect feelings expressed via backchannels, and the correlation between backchannel patterns and user interests was examined (Kawahara et al., 2008). These studies indicate that detection of spontaneous user backchan-

nels can benefit spoken dialogue systems by providing informative cues that reflect the user's situation. For instance, if a spoken dialogue system can detect user's backchannels, it can facilitate smooth turn-taking. The system can also detect user's feelings and judge if it should continue the current topic or change it.

Despite these previous studies and decades of analysis on backchannels, few practical dialogue systems have made use of them. This is probably due to the fact that users do not react as spontaneously to dialogue systems as they do to other humans. We presume one of the reasons for this is the unnatural intonation of synthesized speech. That is, conventional speech synthesizers do not provide users with signs to elicit backchannels; an appropriate set of lexical, acoustic and prosodic cues (or backchannel-inviting cues (A. Gravano and J. Hirschberg, 2009)), which tends to precede the listener's backchannels in human-human communication. Though recorded human speech can provide such cues, it is costly to re-record system's speech every time system scripts are updated. In this work, we therefore tackle the challenge of constructing dialogue-style text-to-speech (TTS) system that inspires users to make spontaneous backchannels under the hypothesis of:

*People will give more spontaneous backchannels to a spoken dialogue system that makes more spontaneous backchannel-inviting cues than a spoken dialogue system that makes less spontaneous ones.*

which is derived from the Media Equation (Reeves and Nass, 1996).

## 2 Related Works

A number of studies have aimed at improving the naturalness of TTS. Though most of these have focused on means of realizing a clear and easy-to-listen-to reading-style speech, some attempts have been made at spontaneous conversational speech. Andersson (Andersson et al., 2010) and Marge (Marge et al., 2010) focused on lexi-

\* currently with Japan Advanced Institute of Science and Technology (JAIST)

cal phenomena such as lexical filler and acknowledgments in spontaneous speech, and showed that inserting them improves the naturalness of human-computer dialogues. In this work, we tackle constructing a natural dialogue-style TTS system focusing on prosodic phenomena such as intonation and phoneme duration.

In the field of conversation analysis, many studies analyzed backchannels in human-human dialogue focusing on lexical and non-verbal cues (Koiso et al., 1998; Ward and Tsukahara, 2000; A. Gravano and J. Hirschberg, 2009). For instance these cues were examined in preceding utterances, such as in part-of-speech tags, length of pause, power contour pattern, and  $F_0$  contour pattern around the end of the Inter-Pausal Units (IPUs). (A. Gravano and J. Hirschberg, 2009) showed that when several of the above cues occur simultaneously, the likelihood of occurrence of a backchannel will increase.

Several studies also utilized the above findings for spoken dialogue systems. Okato (Okato et al., 1996) and Fujie (Fujie et al., 2005) trained models to predict backchannels, and implemented spoken dialogue systems that make backchannels. Our goal differs in that it is to inspire users to give backchannels.

### 3 Construction of Spoken Dialogue TTS

#### 3.1 Spoken Dialogue Data collection for TTS

In order to make spontaneous dialogue-style TTS that can evoke backchannels, we construct a spontaneous dialogue-style speech corpus that contains backchannel-inviting cues, and then train an HMM acoustic model for synthesis.

We collected our training data by dubbing a script of our Kyoto Sightseeing Guidance Spoken Dialogue Corpus (Misu et al., 2009), a set of itinerary-planning dialogues in Japanese. In the dialogue task, the expert guide has made recommendations on sightseeing spots and restaurants until has decided on a plan for the day. With the guide’s recommendations, many users give spontaneous backchannels. We made a set of dialogue scripts from the corpus, and asked voice actors to act them out.

When preparing the dialogue script for dubbing, we first removed fillers and backchannels from the transcripts of the dialogue corpus. We then annotated the guide’s end of the IPUs, where the user made backchannels, with #. A sample dialogue script is shown in Figure 6. We asked two professional voice actresses to duplicate the spoken dia-

logue of the script, with playing the role of the tour guide, and the other as the tourist, sitting face-to-face. During the recording, we asked the tour guide role to read the scenario with intonation so that the tourist role would spontaneously make backchannels at the points marked with #. The tourist was allowed to make backchannels at will at any pause segments the guide made. We recorded 12 dialogue sessions in total. The speech data was manually labeled, and 239.3 minutes of tour guide utterances, which are used to train our HMM for the TTS system, were collected. The training data is complemented by the ATR 503 phonetically balanced sentence set (Abe et al., 1990), so as to cover deficiencies in the phoneme sequence. The sentence set is collected from news articles, and data consists of 43.1 minutes of reading-style speech.

#### 3.2 Analysis of Collected Speech Data

Before training the HMM, we analyzed the collected spoken dialogue data to confirm if the recorded dialogue speech data contained backchannel-inviting prosodic cues. We compared prosodic features of the dialogue speech data with those of the reading-style speech data (phonetically balanced sentences that we collected). Following the findings of a previous study (Koiso et al., 1998), we investigated the duration,  $F_0$  contour pattern and power contour pattern of the final phoneme of the IPUs<sup>1</sup>.

In conversation analysis of Japanese, the  $F_0$  contour pattern label of the final phoneme is often used. While the contour pattern is usually manually labeled, we roughly determined the patterns based on the following procedure. We first normalized the log  $F_0$  scale using all utterances so that it has zero mean and one standard deviation (z-score:  $z = (x - \mu) / \sigma$ ). We then divided each final phoneme of the IPU into former and latter parts, and calculated the  $F_0$  slope of each segment by linear regression. By combination of following three patterns, we defined nine  $F_0$  contour patterns for the final phonemes of the IPUs. The pattern of the segment was judged as *rise* if the slope was larger than a threshold  $\theta$ . If the slope was less than the threshold  $-\theta$ , the pattern was judged as *fall*. Otherwise, it was judged as *flat*. Here,  $\theta$  was empirically set to 5.0. The power contour patterns of the IPUs were estimated by a similar procedure.

We analyzed 3,311 IPUs that were not followed

<sup>1</sup>For this study, we define an IPU as a maximal sequence of words surrounded by silence longer than 200 ms. This unit usually coincides with one Japanese phrasal unit.

Table 1: Prosodic analysis of final phonemes of IPU (dialogue script vs. newsarticle script)

	dialogue		newsarticle	
dur. phoneme [msec]	177.1 ( $\pm$ 83.6)		119.4 ( $\pm$ 31.3)	
	average ( $\pm$ standard deviation)			
	$F_0$		power	
pattern	dialogue	news	dialogue	news
rise-rise	3.7 %	10.4 %	0.0 %	0.0 %
rise-flat	2.6 %	2.1 %	0.0 %	0.0 %
rise-fall	18.8 %	3.2 %	0.0 %	0.0 %
flat-rise	4.8 %	11.5 %	0.0 %	0.0 %
flat-flat	3.5 %	1.8 %	0.0 %	9.2 %
flat-fall	12.6 %	2.7 %	13.6 %	0.1 %
fall-rise	29.2 %	47.0 %	0.0 %	0.0 %
fall-flat	7.7 %	9.0 %	86.0 %	90.7 %
fall-fall	17.1 %	12.3 %	0.0 %	0.0 %

by a turn-switch in the dialogue-style speech data and 645 non-sentence-end IPUs in the reading-style speech data. The prosodic features of final phonemes of these IPUs are listed in Table 1.

According to a study (Koiso et al., 1998), in which prosodic features of IPUs followed by a turn-hold with backchannel, without backchannel and turn-switch were compared, a long duration in the final phoneme is a speaker’s typical sign to keep floor. The same study also reported that the *flat-fall* and *rise-fall* pattern of  $F_0$  and power are more likely to be followed by a backchannel than a turn-hold without a backchannel and turn-switch. In our collected speech corpus, there were actually significant ( $p < 0.01$ ) differences in the duration of the final phoneme between that in the dialogue-style speech and in reading-style speech. There was also significant ( $p < 0.01$ ) difference in the occurrence probability of the above two prosodic patterns between dialogue-style speech and reading-style speech data. These figures indicate that as a whole the collected dialogue-style data contains more backchannel-inviting cues than collected reading-style speech data.

We trained HMM for our TTS system Ximera using the HMM-based Speech Synthesis System (HTS) (Zen et al., 2007). We adopted mel log spectrum approximation (MLSA) filter-based vocoding (SPTK, 2011), a quint-phone-based phoneme set and five state HMM-based acoustic modeling. All training data including reading-style speech data were used for model training.

## 4 User Experiment

### 4.1 Dialogue System used for Experiment

To evaluate our TTS system based on users’ reactions, a sightseeing guidance spoken dialogue sys-



Figure 1: Screen shot of the dialogue system

tem that assist users in making decision was implemented. The system can explain six sightseeing spots in Kyoto. The system provides responses to user requests for explanation about a certain spot. Each descriptive text on a sightseeing spot consists of 500 ( $\pm 1\%$ ) characters, 30 phrases. The text is synthesized using section 3 TTS<sup>2</sup>. We set the speech rate of our TTS as nine phoneme per second.

A display is used to present photos of the target sightseeing spot and an animated 3D desktop avatar named Hanna. Figure 1 shows the GUI the user sees. The avatar can express its status through several motions. For example, when the user begins speaking, it can express the state of listening using the listener’s motion, as shown in the figure. A sample dialogue with the system is shown in Table 7. A video (with English subtitles) of an sample dialogue with a user can be seen at <http://mastarpj.nict.go.jp/~xtmisu/video/TTS.wmv>.

To compare the effectiveness of our TTS in evoking users’ spontaneous backchannels, we constructed a comparison system that adopts a conventional reading-style TTS system. An HMM model was trained using 10-hour reading-style speech by another professional female narrator. Other settings, such as the descriptive text and avatar agent, were the same as those of the base system.

### 4.2 Comparison of Prosodic Features of the Synthesized Speech

Prior to the experiments, we investigated the prosodic features of the final phoneme of IPUs in the synthesized explanations on six spots to confirm if they contain backchannel-inviting cues. The results are given in Table 2.

Tendencies in the duration of the final phoneme and prosody pattern distribution of the synthesized

<sup>2</sup>The descriptive texts are not included in the training data.

Table 2: Prosodic analysis of final phonemes of IPUs (dialogue-style TTS vs. reading-style TTS)

	dialogue synth.		reading synth.	
dur. phoneme [msec]	172.9 ( $\pm$ 29.6)		126.1 ( $\pm$ 19.1)	
average ( $\pm$ standard deviation)				
	$F_0$		power	
pattern	dialogue	reading	dialogue	reading
rise-rise	5.4 %	0.0 %	0.0 %	0.0 %
rise-flat	2.0 %	0.0 %	1.7 %	0.0 %
rise-fall	23.5 %	0.0 %	46.3 %	5.3 %
flat-rise	5.0 %	0.0 %	0.0 %	0.0 %
flat-flat	1.7 %	0.0 %	4.0 %	9.2 %
flat-fall	15.8 %	0.0 %	22.8 %	18.1 %
fall-rise	15.8 %	0.0 %	0.7 %	0.0 %
fall-flat	3.4 %	0.0 %	7.0 %	0.0 %
fall-fall	27.5 %	100.0 %	17.4 %	76.5 %

speech by the dialogue-style TTS system were similar to that of recorded dialogue speech, suggests that the constructed dialogue-style TTS system can duplicate the backchannel-inviting cues of the recorded original speech. The synthesized dialogue-style speech also contained much more *rise-fall* and *flat-fall* patterns in  $F_0$  and power than that generated by the reading-style TTS system. The average duration of the final phoneme was also longer. Considering the fact that the speech data was generated from the same script, this indicates that the synthesized speech by the dialogue-style TTS system contains more backchannel-inviting features than that by the reading-style TTS system.

### 4.3 Experimental Setup

We evaluated the TTS systems using 30 subjects who had not previously used spoken dialogue systems. Subjects were asked to use the dialogue system in two settings; dialogue-style TTS system and reading-style TTS system. The experiment was conducted in a small (about 2  $m^2$ ) soundproof room with no one else present.

We instructed the subjects to speak with the avatar agent Hanna (not with the system). We also told them that the avatar agent was listening to their speech at all times using the microphone, and was observing their reactions using the camera above the display<sup>3</sup>. Subjects were given the task of acquiring information about three candidate sightseeing spots in Kyoto shown on the display and then selecting one that they liked. An example dialogue with the system is shown in Table 7. A video (with English subtitles) showing a real user dialogue can be seen at <http://mastarpj.nict.go.jp/~xtmisu/video/exp.avi>.

<sup>3</sup>The system did not actually sense the subjects’ reactions.

Table 3: Questionnaire items

1. Overall, which speech was better?
2. Which speech had easier-to-understand explanations?
3. For which speech did you feel compelled to give backchannels?
4. Which speech was more appropriate for this system?
5. Which speech had more human-like explanation?

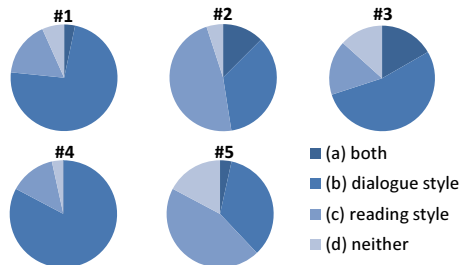


Figure 2: Questionnaire results

After the subject selected from candidate spots, we changed the TTS system settings and instructed the user to have another dialogue session selecting one of another three spots. Considering the effects of the order, the subjects were divided into four groups; the first group (Group 1) used the system in the order of “Spot list A with dialogue-style speech Spot list B with reading-style speech,” the second group (Group 2) worked in reverse order. Groups 3 and 4 used a system alternating the order of the spot sets.

## 5 Experimental Results

### 5.1 Questionnaire Results

After the experiments, subjects were asked to fill in a questionnaire about the system. Table 3 shows the questionnaire items. The subjects selected (a) both are good, (b) dialogue-style speech was better, (c) reading-style speech was better, or (d) neither were good. Figure 2 shows the results.

The dialogue-style speech generally earned higher ratings, but reading-style was slightly higher in items #2 and #5. This tendency is likely attributable to the fact that the dialogue-style speech had worse clarity and naturalness than reading-style. The mean opinion score (MOS), which is often used to measure clarity and naturalness of TTS, of the dialogue-style TTS was in fact 2.79, worse than 3.74 for the reading-style.

### 5.2 Analysis of Frequency of Backchannels

We analyzed the number of backchannels that users made during the dialogue session. We manually annotated subjects’ verbal feedbacks, such as “uh-huh” and nodding of the head using the recorded video. Out of 30 subjects, 26 gave some form of

Table 4: Percentages and average number of users who made backchannels

	TTS	% users made BCs	# average BCs taken
Group 1: (Dialogue → Reading) (Spot list A → Spot list B)	Dialogue-style	100.0% (50.0%, 100.0%)	30.4 (1.8, 28.6)
	Reading-style	100.0% (50.0%, 87.5%)	26.1 (3.1, 23.0)
Group 2: (Reading → Dialogue) (Spot list A → Spot list B)	Dialogue-style	75.0% (25.0%, 62.5%)	12.7 (0.5, 12.2)
	Reading-style	75.0% (25.0%, 62.5%)	12.9 (1.3, 11.6)
Group 3: (Dialogue → Reading) (Spot list B → Spot list A)	Dialogue-style	100.0% (28.6%, 100.0%)	14.0 (0.4, 13.6)
	Reading-style	100.0% (0%, 100.0%)	19.3 (0, 19.3)
Group 4: (Reading → Dialogue) (Spot list B → Spot list A)	Dialogue-style	87.5% (42.9%, 87.5%)	28.2 (4.7, 23.5)
	Reading-style	100.0% (71.4%, 87.5%)	24.8 (6.5, 18.3)
All:	Dialogue-style	86.7% (36.7%, 86.7%)	21.1 (1.7, 19.4)
	Reading-style	90.0% (40.0%, 83.3%)	20.6 (2.4, 18.2)

Total backchannel (verbal feedback [Aizuchi], nodding)

backchannel to the system. Table 4 shows the percentages and average number of times subjects gave backchannels. Many users made more backchannels using the dialogue-style TTS system. Despite the significant difference in questionnaire item #3, there were no significant differences in the average number of users’ backchannels.

### 5.3 Informativeness of Backchannels

We then evaluated the TTS in terms of the informativeness of evoked backchannels. The spontaneous prosodic pattern of the backchannels is expected to suggest positive/negative feelings on regarding the recommended candidate. One promising use of backchannels in our application is for detecting users’ feelings about the currently focused on spot, and choosing to continue the explanation on the current topic if the user seems interested, or otherwise change the topic. We therefore label backchannels made during the systems explanation of the spot that the user finally selected as “positive” and those made during the explanations of the other two spots as “negative” and consider distinguishing between them. In human-human dialogues, it was confirmed that when a user responds promptly, the majority of responses are positive, and more backchannels also suggest positive responses (Kawahara et al., 2008).

We investigated the informativeness of the backchannels based on their classification rate, or whether the system can distinguish positive and negative backchannels, using 10-fold cross-validation. That is, the backchannels evoked by the dialogue-style TTS system were divided into 10 groups and nine were used for training and the other for classification tests. We trained decision trees using J4.8 algorithm using timing, frequency, total frequency throughout the session and type of backchannel (verbal feedback or nod) as the feature set. The classification error cost of the positive sample was set to (# negative samples / # positive samples) considering

the difference in the number of positive and negative samples. Ten trials were conducted by changing the test set and the average classification rate was calculated. The classification rate of backchannels evoked by the system with dialogue-style TTS was 71.4%. The confusion matrix of the classification is shown below. We obtained precisions of 62.8% in the classification of the positive backchannels, and 73.2% in that of the negative backchannels. The rates are significantly higher than chance rates of 33.5% and 66.5%. This result indicates the backchannels evoked by the dialogue-style TTS were informative for the system.

Table 5: Confusion matrix of classification

→ classified as ↓ label	positive	negative
positive	76	141
negative	45	386

The classification rate of the reading-style TTS system was calculated in the same way. The average classification rate of backchannels evoked by reading-style TTS was a significantly lower 47.4%, meaning they were not informative at all.

These results suggest that our dialogue-style TTS system can evoke more spontaneous and informative backchannels that reflects users’ intentions than the conventional reading-style one. This classification rate is not completely satisfactory, but we expect that users’ feeling can be detected after observing several backchannels. We also believe that we can estimate users’ interest more precisely by combining verbal information of dialogue acts (Misu et al., 2010).

## 6 Conclusions

This paper presented our first steps toward a spoken dialogue system that evokes users’ spontaneous listener’s reactions. We constructed a dialogue-style TTS and confirmed that by generating human-like backchannel-inviting cues, the system can evoke user’s spontaneous backchannels, which are informative for the system.

## References

- A. Gravano and J. Hirschberg. 2009. Backchannel-inviting cues in task-oriented dialogue. In *Proc. Interspeech*, pages 1019–1022.
- M. Abe, Y. Sagisaka, T. Umeda, and H. Kuwabara. 1990. *Speech Database User's Manual*. ATR Technical Report TR-I-0166.
- S. Andersson, K. Georgila, D. Traum, and R. Clark M. Aylett. 2010. Prediction and Realisation of Conversational Characteristics by Utilising Spontaneous Speech for Unit Selection. In *Proc. Speech Prosody*.
- S. Fujie, K. Fukushima, and T. Kobayashi. 2005. Backchannel feedback generation using linguistic and non-linguistic information and its application to spoken dialogue system. In *Proc. Interspeech*, pages 889–892.
- T. Kawahara, M. Toyokura, T. Misu, and C. Hori. 2008. Detection of Feeling Through Back-Channels in Spoken Dialogue. In *Proc. Interspeech*, pages 1696–1696.
- H. Koiso, Y. Horiuchi, S. Tutiya, A. Ichikawa, and Y. Den. 1998. An Analysis of Turn-Taking and Backchannels based on Prosodic and Syntactic Features in Japanese Map Task Dialogue. *Language and Speech*, 41(3-4):295–322.
- M. Marge, J. Miranda, A. Black, and A. I. Rudnicky. 2010. Towards Improving the Naturalness of Social Conversations with Dialogue Systems. In *Proc. SIGDIAL*, pages 91–94.
- S. Maynard. 1986. On back-channel behavior in japanese and english casual conversation. *Linguistics*, 24(6):1079–1108.
- T. Misu, K. Ohtake, C. Hori, H. Kashioka, and S. Nakamura. 2009. Annotating Communicative Function and Semantic Content in Dialogue Act for Construction of Consulting Dialogue Systems. In *Proc. Interspeech*.
- Teruhisa Misu, Komei Sugiura, Kiyonori Ohtake, Chiori Hori, Hideki Kashioka, Hisashi Kawai, and Satoshi Nakamura. 2010. Dialogue Strategy Optimization to Assist User's Decision for Spoken Consulting Dialogue Systems. In *Proc. IEEE-SLT*, pages 342–347.
- Y. Okato, K. Kato, M. Yamamoto, and S. Itahashi. 1996. Insertion of interjectory response based on prosodic information. In *Proc. of IEEE Workshop Interactive Voice Technology for Telecommunication Applications*, pages 85–88.
- B. Reeves and C. Nass. 1996. *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. Cambridge University Press.
- SPTK, 2011. *Speech Signal Processing Toolkit (SPTK)*. <http://sp-tk.sourceforge.net/>.
- N. Ward and W. Tsukahara. 2000. Prosodic features which cue backchannel responses in English and Japanese. *Journal of Pragmatics*, 32(8):1177–1207.
- H. Zen, T. Nose, J. Yamagishi, S. Sako, T. Masuko, A.W. Black, and K. Tokuda. 2007. The HMM-based speech synthesis system version 2.0. In *Proc. ISCA SSW6*.

## A Appendix

Table 6: Example dialogue scenario used for recording (translation of Japanese)

---

Guide	And Kyomizu-dera.
Guide	Yeah, it is very famous for.
Guide	There is this like wooden balcony#, and you can't see the whole city#, but you can see a good part of the city from there. And it is a beautiful view.
User	It is nice. You have been there?
Guide	Yes, many a times.
Guide	Unfortunately, in winter it is not very green#, and there are not any flowers.
User	So if there is snow will they close the stairs or will they - we are just going to be careful walking the steps?
Guide	You have to be careful going up.
Guide	Basically, it's a slope it's a slope#, so if it's very icy - if there is, you know, like black ice#, you might want to rethink hiking up that slope. But usually it is okay# in the winter. I mean, Kyomizu-dera and this whole area, like every tourist eventually makes their way here#, whether they are Japanese or foreign.
Guide	So, I am sure the paths should be clear.
Guide	Yeah, that is a good spot too.

---

Table 7: Example dialogue with the system (translation of Japanese)

---

System	Hello, my name is Hanna, a tour guide for sightseeing in Kyoto. Nice to meet you. Please say "Okay" when you have decided which spot you would like to know about first.
User	Okay.
System	Which spot would you like to know about?
User	Tell me about Shinnyo-do temple.
System	Sure, I'll explain Shinnyo-do temple.
System	Shinnyo-do is a Buddhist Tendai Temple in Sakyo-ku, Kyoto City and the formal name is Shinsho Gokuraku-ji. You may take the bus directly from Kyoto Station for about thirty minutes, and the fee is two hundred twenty yen. The admission fee for the temple is five hundred yen. The temple was established in nine hundred eighty-four A.D., when Amida Nyorai-zo; a principle image from Mt. Hiei was placed in the Imperial villa of this area.
System	Shall I continue?
User	Sure.
System	The temple is famous, especially for its coloring of autumn leaves; Red Maple stand on both sides of the stone-paved approach, and it appears as though the temple is placed inside a maple forest. ...
...	
System	That's all about this spot.
System	Which spot would you like to know about next?
User	Daitoku-ji please.
...	

---