

Monologue Summarization: Extraction of Important Sentences for TV News Commentary Programs

Takahiro Ito[†], Kenji Matsumoto[†], Yasuo Tanida[‡], Hideki Kashioka[†], Hideki Tanaka[†]

[†]ATR Spoken Language Translation Research Laboratory

[‡]Advanced Telecommunications Research Institute International

2-2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0288, JAPAN

{takahiro.ito, kenji.matsumoto, yasuo.tanida, hideki.kashioka, hideki.tanaka}@atr.co.jp

Abstract

The extraction of important sentences is a key technique for automatic summarization. Whereas most research in this area has targeted written language, we are conducting research on spoken language monologues such as presentations and TV news commentary programs. We collected 50 TV news commentary programs, and experimented with the extraction of important sentences from transcriptions. We used two extraction methods. The first one uses word statistics, and the second one uses the surface features of the sentences. In order to use the latter method, we analyzed the transcriptions and obtained surface features related to the importance of the sentences. The experiments showed that the latter method was better than the former one especially when extracting small sets of sentences. We also mention the ambiguity of judgment by individuals and the contribution of each surface feature to the importance of the sentences.

1 Introduction

With rapid advances in communications technology, the importance of grasping important points in a large number of documents is gaining ever-increasing attention. In recent years, an overwhelming amount of information is being supplied not only in text form but also in media such as graphics and sounds. Therefore, there is a growing need for a technique to summarize data in non-textual media.

With this in mind, we took up monologue as our summarization target. The goal of our research is to establish a suitable method to generate monologue summaries. Some techniques already exist for automatic summarization, such as extracting

important sentences, eliminating unimportant parts in sentences, and connecting important words or phrases into sentences. We employed the extraction of important sentences since it is the most commonly used and therefore most important technique in automatic summarization. Traditionally, research on the extraction of important sentences has targeted written language. We thus used the framework of the technique developed for written language. In order to apply the framework to monologues, we analyzed the transcriptions of TV news commentary programs and obtained monologue features.

In section 2, we refer to some conventional methods for extracting important sentences, and describe the motivation for our research. In section 3, we describe the target data of the experiments. In section 4, we present an experiment in the manual extraction of important sentences. Three people took part in the extraction. The extracted sentences are used as correct answer sets for experiments in automatic extraction. In section 5, we describe an experiment using word statistics. This method can be applied independent of the type of document. Through this experiment, we examined the quality of the extraction with a widely used method. In section 6, we examine the surface features of TV news commentary programs, and conduct an experiment using them. We use a decision tree learning method and set the examined surface features as the attributes for the learning. In section 7, we discuss the result of the experiments.

2 Conventional sentence extraction methods

In conventional research, the importance of sentences has been evaluated by the following three types of information.

- (1) Statistical information such as word frequency.
- (2) Surface features such as sentence position in a document or clue phrases.

(3) Document structure derived from rhetorical analysis or a text segmentation technique.

Examples of type (1) include a method proposed by Luhn (1958) and one proposed by Zechner (1996). Both methods measure the importance of the words in the documents by using word statistics, and calculate the importance of sentences from the number of important words in the sentences. This approach does not depend on the type of document and therefore can be applied to monologues directly. But it has a demerit in that it cannot reflect the features found in the object data on the judgment of the importance of sentences.

An example of type (2) is proposed by Yamamoto *et al.* (1995). Their target is newspaper editorials. They use clue words, the location of sentences in the documents or paragraphs, and the types of sentences (e.g., sentences that describe facts, sentences that describe opinions) extracted from the surface patterns of sentences to judge the importance of the sentences. Their algorithm cannot be applied to a monologue directly because it assumes the rhetorical structure that is characteristic of editorials and uses paragraphs and surface patterns of sentences that are exclusive to written language. Watanabe (1996) and Nomoto *et al.* (1997) use learning methods to extract important sentences. Watanabe uses multiple-regression analysis and Nomoto uses a decision tree. Both of them manually extracted the surface features of sentences from newspaper articles and used these features as attributes for the learning. In order to apply these methods to a monologue, it is necessary to obtain the characteristic features of the monologue.

Examples of type (3) are proposed by Sumita *et al.* (1995) and Nakao (1999). Sumita targeted technical papers. He used the tree structure of sentences derived from rhetorical analysis to judge the importance of sentences. Nakao targeted white papers. He judged the range of topics from the distribution of words in the documents and extracted the important sentences from each topic.

In our research, we applied approaches (1) and (2) to monologues. We did not try approach (3) this time, because it requires a high-quality rhetorical analysis technique before the extraction of important sentences.

3 Overview of the data

We used transcriptions of a TV news commentary program called "ASU-WO-YOMU" for our research. It is a 10-minute program and the size of the transcription is about 3,000 Japanese characters

(about 1,135 words in English) per program. The number of transcriptions for the experiments is 50, and the number of sentences in each transcription is 60.2 on average. We found the following characteristics in the data.

(1) Questions and calling expressions (e.g., Let's) are commonly observed. They are used to draw the attention of the viewers.

(2) Conjunctive expressions and adverbial expressions appear frequently. According to our investigation, the rates of conjunctions and adverbs in "ASU-WO-YOMU" (from the 50 transcriptions used in our experiments) are 1.22% and 1.75%, respectively. Meanwhile, their rates in newspaper articles (Nihon Keizai Shinbun, 1990-1995, about 920,000 articles) are 0.29% and 0.76%, respectively.

4 Manual extraction of important sentences

Three people manually extracted important sentences. The purposes of this extraction were to make correct answer sets and to examine the ambiguity of judgments among individuals.

4.1 Extraction procedure

The extraction of important sentences proceeded in the following way.

(1) The importance decision was left to each individual. No importance criteria were given.

(2) Two types of extractions were made: 5-sentence extraction (EX5) and 20-sentence extraction (EX20). If extraction with the exact number was difficult, a range of ± 2 sentences for the 5-sentence extraction and ± 3 sentences for the 20-sentence extraction was allowed.

(3) The set of sentences extracted in the 20-sentence extraction did not have to include the sentences extracted in the 5-sentence extraction.

Table 1 shows the result of the experiment.

Table 1. Manual Extraction Result

	EX5	EX20
By one person on average.	5.2	20.5
By three people	1.5	9.0
By at least one person	9.9	32.6
Average agreement rate	46.7	61.9

The figures in the first three rows are the average numbers of extracted sentences for the 50

transcriptions in this experiment. The "average agreement rate" means the average value of precision and recall between all combinations of two people. We use this term because the values of precision and recall are the same if they are calculated for all combinations of two people.

The result shows that even in judgments among people, the agreement rate was no more than 46.7% for the 5-sentence extraction and 61.9% for the 20-sentence extraction. These can be the upper limit for the automatic extraction experiments.

4.2 Result evaluation with kappa

We calculated the kappa coefficient to evaluate the degree of agreement among people. The kappa coefficient is a value to calculate the proportion of agreement considering agreement occurring by chance. It can be calculated by the following formula.

$$K = \frac{Pa - Pe}{1 - Pe}$$

Pa is the proportion of agreement between two people, and Pe is the proportion of agreement by chance.

We express the set of sentences judged as important or not important by persons A and B as follows.

	Important	Not important
A	α	β
B	γ	δ

In our experiment, we defined Pa and Pe as follows.

$$Pa = \frac{N(\alpha \cap \gamma)}{Ns} + \frac{N(\beta \cap \delta)}{Ns}$$

$$Pe = \frac{N(\alpha)}{Ns} \times \frac{N(\gamma)}{Ns} + \frac{N(\beta)}{Ns} \times \frac{N(\delta)}{Ns}$$

where the notation N(X) means the number of elements in set X.

Table 2 shows the values of the kappa coefficient between each pair of the three people, A, B, and C, who took part in this experiment. They are all average values of the 50 transcriptions in this experiment. Table 3 shows the interpretation of the kappa coefficient (Carletta *et al.*, 1997).

Table 2. Kappa Coefficient between Each Pair of People

	EX5	EX20
A - B	0.58	0.55
B - C	0.35	0.33
C - A	0.33	0.39
Average	0.42	0.42

Table 3. Interpretation of Kappa Coefficient

Kappa coefficient	Degree of agreement
< 0	poor
0 - 0.20	slight
0.21 - 0.40	fair
0.41 - 0.60	moderate
0.61 - 0.80	substantial
0.81 - 1	near perfect

As seen in Table 2, the average kappa coefficient shows the same value of 0.42 for both the 5-sentence extraction and the 20-sentence extraction. The degree of agreement in Table 3 is between fair and moderate. From this result, it can be said that the task of extracting important sentences is rather difficult because the standard of importance is ambiguous even for people. The fact that the kappa coefficient of the 5-sentence extraction and the 20-sentence extraction is the same value can be interpreted as meaning that the two tasks are equally difficult.

As for the tendency for agreement among individuals, the pairs that show high agreement in the 5-sentence extraction tend to show high agreement in the 20-sentence extraction.

4.3 Baseline agreement ratio

We defined the baseline agreement ratio as recall and precision when the extraction was done randomly. They are calculated by the following formulas.

$$recall = \frac{\frac{Np}{Ns} \times Nc}{Np} = \frac{Nc}{Ns}$$

$$precision = \frac{\frac{Np}{Ns} \times Nc}{Nc} = \frac{Np}{Ns}$$

where Np: the number of sentences extracted by a person.

Nc: the number of important sentences extracted by a computer.

Ns: the total number of sentences in a transcription.

In the above formulas, (Np / Ns) is the probability that a sentence which is randomly output by a computer corresponds to one of the sentences extracted by a person. So, (Np / Ns) * Nc means the expected number of sentences extracted by a person that appear in the sentences extracted by a computer. Table 4 shows the values of recall and precision calculated by the above formulas assuming that Ns is 60.2, Np is 5.2 for the 5-sentence extraction and 20.5 for the 20-sentence extraction, and Nc is 5 for

the 5-sentence extraction and 20 for the 20-sentence extraction. These values can be the baseline for the automatic extraction evaluation.

Table 4. Baseline of the Experiments

EX5		EX20	
Rec.	Prec.	Rec.	Prec.
8.3	8.6	33.2	34.1

5 Sentence extraction using word statistics

We tried the method proposed by Luhn (1958) and the one proposed by Zechner (1996) as typical methods using word statistics. These methods can be applied to TV news commentary programs directly.

5.1 Experiment procedure

We conducted the experiments with the following procedure.

(1) Remove function words from the result of the morphological analysis of the transcriptions. We used only content words for word statistics.

(a) Luhn's method

(a-2) Define important words as the words ranking within 10th in frequency for each transcription.

(a-3) Calculate the importance of each sentence by the following formula.

$$imp(S) = \frac{Nwimp^2}{Nwcnt}$$

where Nwimp: the number of important words in the sentence.

Nwcnt: the number of content words in the sentence.

(b) Zechner's method

(b-2) Calculate the tf*idf value of each content word by the following formula. We call this value the "word tf*idf" in the rest of this paper.

$$f(w) * \log \frac{Nd}{d(w)}$$

where f(w): the frequency of the word w in the transcription

Nd : the total number of transcriptions in our experiment (i.e., 50)

d(w): the number of transcriptions that the word w appears in.

(b-3) Calculate the importance of each sentence by summing up the "word tf*idf" in the sentence. We call this value the "sentence tf*idf" in the rest of this paper.

(4) Extract 5 and 20 sentences by their order of importance.

5.2 Experiment result

Table 5 shows the result of the experiments. In the table, "Lu" is Luhn's method, and "Ze" is Zechner's method. The scores are average values when each set is extracted by three people as the correct answer.

Table 5. Result Using Word Statistics

	EX5		EX20	
	Rec.	Prec.	Rec.	Prec.
Lu	14.7	15.1	42.1	43.1
Ze	17.4	18.1	45.0	45.9

In both the 5-sentence extraction and the 20-sentence extraction, the scores are higher than the baseline (see Table 4), but considerably lower than the average agreement rate between people (see Table 1). Especially in the 5-sentence extraction, the scores show a large drop from 46.7% to less than 20%.

Comparing the two methods, Zechner's method performed better than Luhn's method. One of the reasons is that general words, such as "case" or "person" are selected as "important words" in Luhn's method. Another reason is that in Luhn's method, the importance of the words is indicated by 2 values of "important" / "not important", but in Zechner's method, the importance is expressed by consecutive values.

6 Sentence extraction using surface features

We analyzed the transcriptions of TV news commentary programs and extracted the features that are considered to be related to the importance of sentences. We took a decision tree learning method (Quinlan, 1993), and used the features as the attributes for the learning. We also examined the contribution of each feature to the importance of the sentences.

The output of the decision tree learning is not only the class showing whether the sentence is important or not, but also the probability that the sentence belongs to each class, which can be calculated by the following formula.

$$P(C_i) = \frac{|C_i| + 1}{N + 2}$$

where N: the number of sentences in the leaf node of the decision tree.

C_i : the class of the learning. In our experiment, C_0 is the class for important sentences, and C_1 is the class for not important sentences.

$|C_i|$: the number of sentences that belong to class C_i in the leaf node of the decision tree.

6.1 Attributes for the learning

From our analysis, we extracted the following features as the attributes of the decision tree learning and set the values of the attributes to each sentence in the following way.

In former research (Yamamoto *et al.*, 1995; Watanabe, 1996; Nomoto *et al.*, 1997), (2) and (6) in the following are not used, and in (4), sentences are roughly classified into only 2 or 3 types (i.e., sentences to describe facts, opinions, and guesses).

(1) The appearance of conjunctions (CNJ)

In Japanese, conjunctions such as "shikashi" (but) or "sate" (by the way) express the semantic relation between sentences. So, in this research, they are considered to be related to the importance of the sentence. We extracted all conjunctions in the transcriptions, and set them as attributes for learning. We assigned the value 1 to a sentence if a conjunction appears in the sentence and 0 if not.

(2) The appearance of adverbs (ADV)

In Japanese, some adverbs such as "yousuruni" (in short) or "tatoeba" (for example) modify the whole sentence. So, in this research, they are considered to be related to the importance of the sentence. We extracted all adverbs in the transcriptions and set them as attributes for learning. We assigned the value 1 to a sentence if an adverb appears in the sentence and 0 if not.

(3) The location of sentences (LS)

The beginning of the transcription is the part that introduces the theme of the program, and the end of the transcription is the part that concludes the program. They are both considered to be related to the importance of the sentences in this research. We assigned values from 10 to 1 for the first sentence to the 10th sentence, and values from 10 to 1 for the last sentence to the 10th sentence from the last. We assigned the value 0 to the rest of the sentences.

(4) Types of sentences (TS)

The roles of the sentences, such as to describe facts or to describe opinions, are considered to be related

to the importance of the sentence in this research. Even in sentences that have the same role, the importance might be different according to the manner of speech. We analyzed the surface patterns of the sentences that are characteristic of TV news commentary programs, and classified the sentences by type automatically.

The following are the types of sentences, the number of times and the rate that each type appears in the transcriptions, and examples of expressions in Japanese and English. The whole patterns in Japanese are attached in appendix A.

We assigned "type A" to "type O" to each sentence as the value. The values are mutually exclusive. The total number of sentences is 3,010.

- (a) Conclude (136, 4.52%)
...wake-desu. (To sum up, ...)
- (b) Explain (23, 0.76%)
...kara-desu. (Because ...)
- (c) Emphasize (14, 0.47%)
...to-iu-koto-nano-de-ari-masu. (It is ... that ...)
- (d) Deduce (85, 2.82%)
...to-iu-koto-ni-nari-masu. (This leads to ...)
- (e) Express general view (11, 0.37%)
...to-kangae-rare-masu. (It seems that...)
- (f) Guess (29, 0.96%)
...de-shou. (This may ...)
- (g) Express greeting (71, 2.36%)
konbanha. (Good evening.)
- (h) Order (7, 0.23%)
goran-kudasai. (Look at ...)
- (i) Introduce the theme of the program (34, 1.13%)
...ni-tsuite-otsutae-shimasu.
(Today I will talk about ...)
- (j) Call to viewers (74, 2.46%)
...te-mi-mashou. (Let's think about...)
- (k) Express opinion directly (179, 5.95%)
...to-omoi-masu. (I think ...)
- (l) Express opinion by rhetorical question (25, 0.83%)
...deha-nai-deshou-ka? (Isn't that ...?)
- (m) Bring up subjects (64, 2.13%)
naze...de-shou-ka? (Why ...?)
- (n) Express question (17, 0.56%)
...de-yoi-no-deshou-ka? (Is it OK with...?)
- (o) Other expressions (2,241, 74.45%)

This consists of sentences that are not classified in the above types. Most of them are sentences describing facts.

(5) Tense (TNS)

The difference between sentences expressed in present tense and past tense is considered to be related to a difference in importance in this research. The sentences whose last letter is "ta" are set to be "past", and others are set to be "present".

(6) Enumerating expressions (EE)

Enumerating expressions such as "First .." "Second ..." can be related to the importance of sentences in this research, because they express the points of content. We assigned the value 1 to a sentence if enumerating expressions appear in the sentence and 0 if not.

6.2 Experiment procedure

We tried 10 folded cross validation experiments for 150 data that were made by three people for 50 transcriptions (see section 4). Because three correct answer sets extracted by three people exist for one transcription, we made the data for the same transcription not appear in both the training data and evaluating data at the same time (Figure 1).

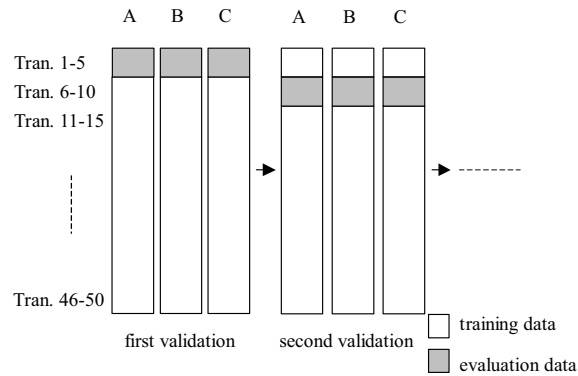


Figure 1. Cross Validation

The output of this experiment is not a class showing whether the sentence is important or not, but the 5 and 20 sentences with a higher order of probability of being important sentences. This was for comparison with other experiments.

6.3 Experiment result

Table 6 shows the result of the experiment. The scores in the table are average values for each set being extracted by three people as the correct answer. In addition to the experiment with the attributes described in section 6.1, we tried an experiment with the values of the "sentence tf*idf" added to the attributes. We considered the "sentence tf*idf" as being an attribute related to the content of the sentence, and examined the effect of adding the attribute to the six surface features (i.e., not related to the content).

Table 6. Result Using Surface Features

	EX5		EX20	
	Rec.	Prec.	Rec.	Prec.
Surf	29.6	30.1	48.4	49.3
Surf+tf*idf	30.8	30.8	47.4	48.3

The result of these experiments was better than the result of the methods using word statistics (see Table 5). The improvement is great especially in the 5-sentence extraction. But we did not observe any improvement by adding the values of the sentence tf*idf to the attributes for learning.

Table 7 shows the result of the experiments of decision tree learning when they were conducted with each unique feature.

Table 7. Result Using Unique Features

	EX5		EX20	
	Rec.	Prec.	Rec.	Prec.
CNJ	13.7	13.9	36.4	37.1
ADV	13.1	13.1	39.5	40.3
LS	38.8	38.7	45.5	46.2
TS	21.3	21.6	41.1	41.9
TNS	20.7	20.4	33.0	33.7
EE	21.8	22.4	38.2	39.1

The order in which each feature contributes to the judgment of the important sentences is as follows.

5-sentence extraction

LS > EE > TS > TNS > CNJ > ADV

20-sentence extraction

LS > TS > ADV > EE > CNJ > TNS

The contribution of the location of the sentences is the biggest. In the 5-sentence extraction, it exceeds the result of using all features.

7 Discussion

The results of the experiments described in sections 5 and 6 show that the method that uses surface features performs better than the method that uses word statistics. This tendency is more conspicuous in 5-sentence extractions.

This means that the importance of sentences can be judged more accurately from surface features without information of the content. In monologues, the viewers or audience cannot rehear the parts that they missed, so it is believed that speakers use emphatic expressions explicitly to show that it is an important part.

Next, we investigate the difference in tendencies between human extraction and machine extraction. Table 8 shows the recall and precision of human extraction and Table 9 shows those of machine extraction (Luhn's method, Zechner's method, and the decision tree learning method with six features in section 6.1) for extractions by each person, A, B, and C, as correct answer sets.

Table 8. Scores between Each Pair of People in 20-sentence Extraction

	A		B		C	
	Rec.	Prec.	Rec.	Prec.	Rec.	Prec.
A			65.3	74.3	59.3	61.8
B	74.3	65.3			57.9	52.6
C	61.8	59.3	52.6	57.9		

Table 9. Scores between Each Person and Each Method in 20-sentence Extraction

	Lu		Ze		Surf	
	Rec.	Prec.	Rec.	Prec.	Rec.	Prec.
A	41.6	45.0	43.8	47.1	48.1	52.0
B	42.2	39.7	45.4	42.9	52.8	49.7
C	42.7	44.5	45.7	47.7	49.7	46.1

Comparing Tables 8 and 9, the difference in judgment between each pair of people is larger than that of each person and each method. The same tendency is observed in 5-sentence extraction. As these experiments show, the recall and precision of important sentences extracted by automatic extractions are almost the same regardless of the answer sets extracted by people. In other words, the machine extraction methods get sentences that do not depend on the individuality of people.

Finally, we think about the reason that the result using only the attribute of the location of sentences is better than the result using all of the attributes in the 5-sentence extraction (see Table 6, scores of "Surf" in EX5 and Table 7, scores of "LS" in EX5). This seems to be an effect of overfitting. We set the gain ratio of entropy as 0, for the stop condition in making the decision tree. So the constructed decision tree overfit the training data. When we set the gain ratio of entropy as 0.03, the values of recall and precision increased 29.6 and 30.1 (see Table 6) to 37.5 and 37.6, respectively, in the 5-sentence extraction.

8 Conclusion

In this paper, we described experiments in extracting important sentences from monologues. We used TV news commentary programs as an example of monologues, and conducted experiments on the transcription. We found that a method using decision tree learning with attributes of surface features performed better than methods using statistical information of words, and this tendency was more conspicuous in 5-sentence extractions.

In these experiments, we calculated the importance of each sentence independently, and did not consider the coherence of the extracted

sentences. In the future, we plan to generate a coherent summary that reflects the structure of the transcriptions.

References

- Carletta J., Isard A., Isard S., Kowtko J. C., Doherty-Sneddon G., and Anderson A. H. 1997. The Reliability of a Dialogue Structure Coding Scheme. *Computational Linguistics*, 23(1): 13-31.
- Luhn, H. P. 1958. The Automatic Creation of Literature Abstracts. *IBM Journal of Research and Development*, 2(2): 159-165.
- Nakao Y. 1999. An Algorithm for Text Summarization base on Thematic Hierarchy Detection: How to Generate a One-page Summary of a Long Text. (In Japanese). *Information Processing Society of Japan SIG Notes*, 99(62): 49-56.
- Nomoto T., and Matsumoto Y. 1997. The Reliability of Human Coding and Effects on Automatic Abstracting. (In Japanese). *Information Processing Society of Japan SIG Notes*, 97(69): 71-76.
- Quinlan, J. R. 1993. C4.5 Programs for Machine Learning. *Morgan Kaufmann Publishers, Inc.*
- Sumita K., Chino T., Ono K., and Miike S. 1995. Automatic Abstract Generation Based on Document Structure Analysis and Its Evaluation as a Document Retrieval Presentation Function. (In Japanese) *The Transactions of the Institute of Electronics, Information and Communication Engineers D-II*, J78-D-II(3): 511-519.
- Watanabe H. 1996. A Method for Abstracting Newspaper Articles by Using Surface Clues. In *Proc. of the 16th International Conference on Computational Linguistics*. 974-979.
- Yamamoto K., Masuyama S., and Naito S. 1995. Green: An Experimental System Generating Summary of Japanese Editorials by Combining Multiple Discourse Characteristics. (In Japanese) *Journal of Natural Language Processing*, 2(1): 39-55.
- Zechner K. 1996. Fast Generation of Abstracts from General Domain Text Corpora by Extracting Relevant Sentences. In *Proc. of the 16th International Conference on Computational Linguistics*. 986-989.

Appendix A. Types of sentences

The following are types of sentences classified by the surface patterns of the sentences. The patterns are expressed in regular expression. The symbols used here are as follows.

- . : Symbol that matches any character.
- ^ : Symbol that matches the head of the sentence.
- | : Symbol that expresses OR
- * : Symbol that expresses the repetition more than 0 times
- ? : Symbol that expresses the repetition either 0 or 1 times.

If the sentence matches more than one pattern, and can be classified into more than one type, the former type in the following list is preferred. (i.e., (a) is preferred to (b))

- (a) Conclude
(わけ|訳)(なん|なの)?です(ね)?。
(わけ|訳)であります。
- (b) Explain
(から|ため|為)(なん|なの)?です(ね)?。
(から|ため|為)であります。
- (c) Emphasize
ということ|事|もの|物)(なん|なの)であります。
ということ|事|もの|物)(なん|なの)です(ね)?。
- (d) Deduce
(こと|事)になります。
(こと|事)になるんです(ね)?。
- (e) Express general view
思われます。
と考えられます。
望まれます。
(心配|予想|懸念)されます。
- (f) Guess
でしょう。
でありましょう。
- (g) Express greeting
(こんばん|今晚)(わ|は)。
明けましておめでとうございます。
失礼(を)?(いた|致)?します。
^それではこの辺で
^それでは。
- (h) Order
(下|くだ)さい。
- (i) Introduce the theme of the program
を(取|と)(り)?(あ|上)(げ)?ます。
についてお伝えします。

- について考えてまいります。
についてお話をしようと思っております。
(今日は|今晩は|今夜は).*?(たい|よう)と(いう
ふうに)?思います。
- (j) Call to viewers
て(み|おき)ましょう。
(致|いた)しましょう。
(たい|よう)と思う(の|ん)です(ね|が)?。
(たい|よう)と(いう|ふうに)?思います。
- (k) Express opinion directly
思います。
思う(の|ん)です(ね|が)?。
思っ(て|おり)?ます。
べきです。
たいものです。
気が(致|いた)?します。
と(私|は)?見(て|おり|い)ます。
- (l) Express opinion by rhetorical question
(では|じゃ)ない(の|ん)?でしょうか。
- (m) Bring up subjects
(どう|どの|どれ|どんな|どこ|なぜ|何|なん|な
に).*か。
- (n) Express question
でしょうか。