

GRIPS Discussion Paper 16-22

# Linking Parliamentary Minutes and Videos in the Japanese Diet

Mikitaka Masuyama

November 2016



**GRIPS**

NATIONAL GRADUATE INSTITUTE  
FOR POLICY STUDIES

National Graduate Institute for Policy Studies  
7-22-1 Roppongi, Minato-ku,  
Tokyo, Japan 106-8677

## **Linking Parliamentary Minutes and Videos in the Japanese Diet**

Mikitaka Masuyama

National Graduate Institute for Policy Studies

7-22-1 Roppongi

Minato-ku, Tokyo 106-8677

Japan

mmasuyama\_at\_grips.ac.jp (replace \_at\_ with @)

### Summary

This paper offers an overview of the video retrieval system we have developed for the Japanese Diet. With our video retrieval system one can directly retrieve the video feed segment of interest, gain a visual understanding of the flow of parliamentary debate, and check the facial expressions and body language of the speaker. In this paper, we demonstrate how one can retrieve video streaming on user terminals that do not support Japanese language input, and suggest a variety of ways in which our video retrieval system can be utilized. Also, we report a preliminary analysis on the correspondence between the official minutes and the results of speech recognition of recordings of parliamentary meetings. We believe that our system encourages research on the utilization of visual information in policy-making and marks a step toward the provision of universal access to policy information.

This work is supported by JSPS Kakenhi Grant Number 15H05727 and the Diet Archives Project funded by the GRIPS Policy Research Center. An earlier version (Masuyama 2016b) was presented at the 2016 General Conference of the European Consortium for Political Research, Charles University, Prague, Czech Republic, September 7-10, 2016.

## Introduction

Seeking innovative ways to provide universal access to policy information, we have launched an internet video retrieval system for the Japanese Diet.<sup>1</sup> Although many parliaments offer online video streaming of speech and debate, those video library systems merely provide a search engine for retrieval of videos by date and by parliamentary meeting theme. However, even if we successfully retrieve the parliamentary video we want, we need to watch the video from the beginning to the point in the speech or debate we are particularly interested in. Using the latest sound recognition techniques to create timestamp data to match parliamentary video feeds and the minutes of proceedings, we have developed an internet video retrieval system for Diet deliberations in which one can pinpoint and play the parliamentary video clips corresponding to the minutes of proceedings by means of keyword search.

With our video retrieval system one can directly retrieve the video feed segment one is particularly interested in, gain a visual understanding of the flow of parliamentary debate, and check the facial expressions and body language of the speaker. Our system captions the videos, so it offers an alternative means of accessing parliamentary deliberation clips for those who have visual and hearing impairments. Since it is easy to share the URL identifying a moment in a video feed via SNS, our system has great potential to boost the usage of Diet deliberation videos by researchers and ordinary citizens as well.

In this paper, we offer an overview of the video retrieval system we have developed, and demonstrate how one can retrieve video streaming on user terminals that do not support Japanese language input. We also suggest a variety of ways in which our video retrieval system can be utilized, and report a preliminary analysis of the correspondence between the official minutes and the results of speech recognition of recordings of parliamentary meetings.

## Video Retrieval System for Diet Deliberations

The Japanese Constitution stipulates that each house of the Diet shall keep a record of proceedings and put it into general circulation. The Diet Library currently provides via internet the digitized minutes of parliamentary meetings since the opening of the Imperial Diet in 1890. Although those are not considered as “official” records, they are amenable to keyword searching. On the other hand, we can watch the online live streaming of proceedings at the secretariat website of each house. We can also search the video library and watch video streaming of parliamentary proceedings. Both houses originally made video streaming of plenary and committee meetings available for only one year, but the House of Representatives alone changed its policy so that the videos of proceedings since 2010 are currently available for viewing at the following locations.

<http://www.shugiintv.go.jp/index.php>

<http://www.webtv.sangiin.go.jp/webtv/index.php>

Diet deliberation videos can be searched by meeting date, meeting title, subject and speaker, although only the first two search options are offered in the English interface. However, even if we successfully retrieve the desired Diet deliberation video, we need to watch the video streaming from the beginning to the speech or debate segment we are particularly interested in. It is not uncommon for a committee meeting

---

<sup>1</sup> Masuyama and Takeda (2014, 2015).

to last more than 7 hours. While the breakdown of the video by questioner is available in the Japanese interface, streamed video segments are usually 30 to 60 minutes long. No such breakdown is available in the English interface. Moreover, replies to parliamentary questions are included in the video, arranged by questioner. Thus, we have no way to search the Diet deliberation videos of prime ministers and cabinet ministers answering parliamentary questions.

By linking the Diet Library's proceedings database and the Diet secretariats' deliberation video libraries, our "Video Retrieval System for Diet Deliberations" makes it possible to retrieve the deliberation video clips corresponding to the minutes of the proceedings through keyword searching:

<http://gclip1.grips.ac.jp/video/>

For instance, our system allows us to pinpoint and play the related video clip when we see an item such as "MP X made remark Y during parliamentary debate." With our system we can directly retrieve the portion of the video feed we are particularly interested in, gain a visual understanding of the flow of parliamentary debate and check the facial expressions and body language of the speaker, all of which are not possible from a simple reading of the minutes of parliamentary proceedings.

Unlike the Diet secretariats' websites, our system creates and adds subtitles to the Diet deliberation videos, and thus offers those suffering from hearing impairment a means of accessing the deliberation video library. At present the Diet Library's website is not compatible with the screen readers used to assist blind and visually impaired users. Even if we can successfully search the proceedings database and reach the speech or debate of particular interest, we need to use text vocalizing software that produces synthetic voices which do not resemble that of the original speaker. In contrast, our video retrieval system allows us to see instantly video of what was actually spoken in the Diet.

Technically speaking, our video retrieval system consists of two sub-systems. As illustrated in Figure 1, one of the sub-systems uses the latest sound recognition techniques to create timestamp data to match the Diet Library's proceedings database (Minute DB) and the Diet secretariats' deliberation video databases (Video DB). The second sub-system uses the timestamp data to search the Diet proceedings and retrieve the Diet deliberation videos corresponding to the minutes of the proceedings by means of keyword search (Web-based Search Interface). The results of keyword searches are deliberation video links, and the portion of video we are particularly interested in can be played partially by clicking the URL link for the deliberation video stored in the Diet secretariats' databases.

"Video Retrieval System for Diet Deliberations" has been in operation and publicly available since November 2012. It is possible to keyword search all the plenary and committee meetings in the House of Representatives since January 2010 and those in the House of Councillors since December 2012. Below, we briefly describe how our video retrieval system works. Figure 2 shows the top page of our web-based search interface that allows us to search for deliberation video segments by typing keywords. The English interface will appear when the user clicks on "English" in the upper right hand corner.

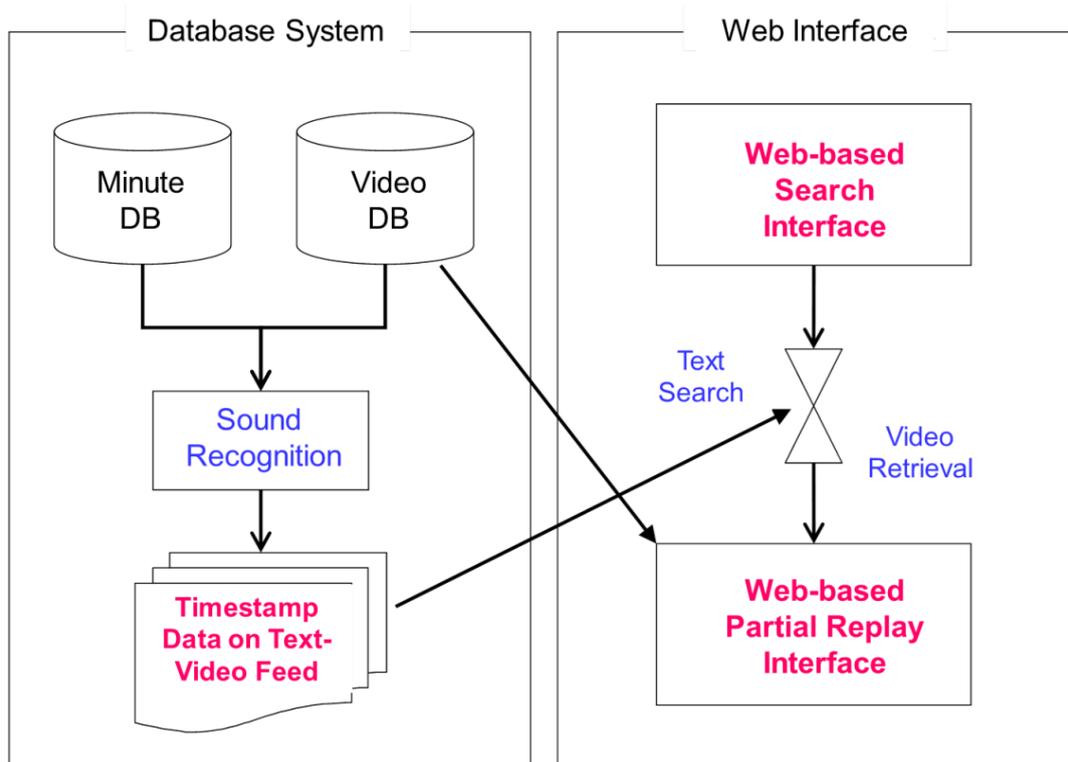


Figure 1: Process for Linking Diet Proceedings and Deliberation Videos



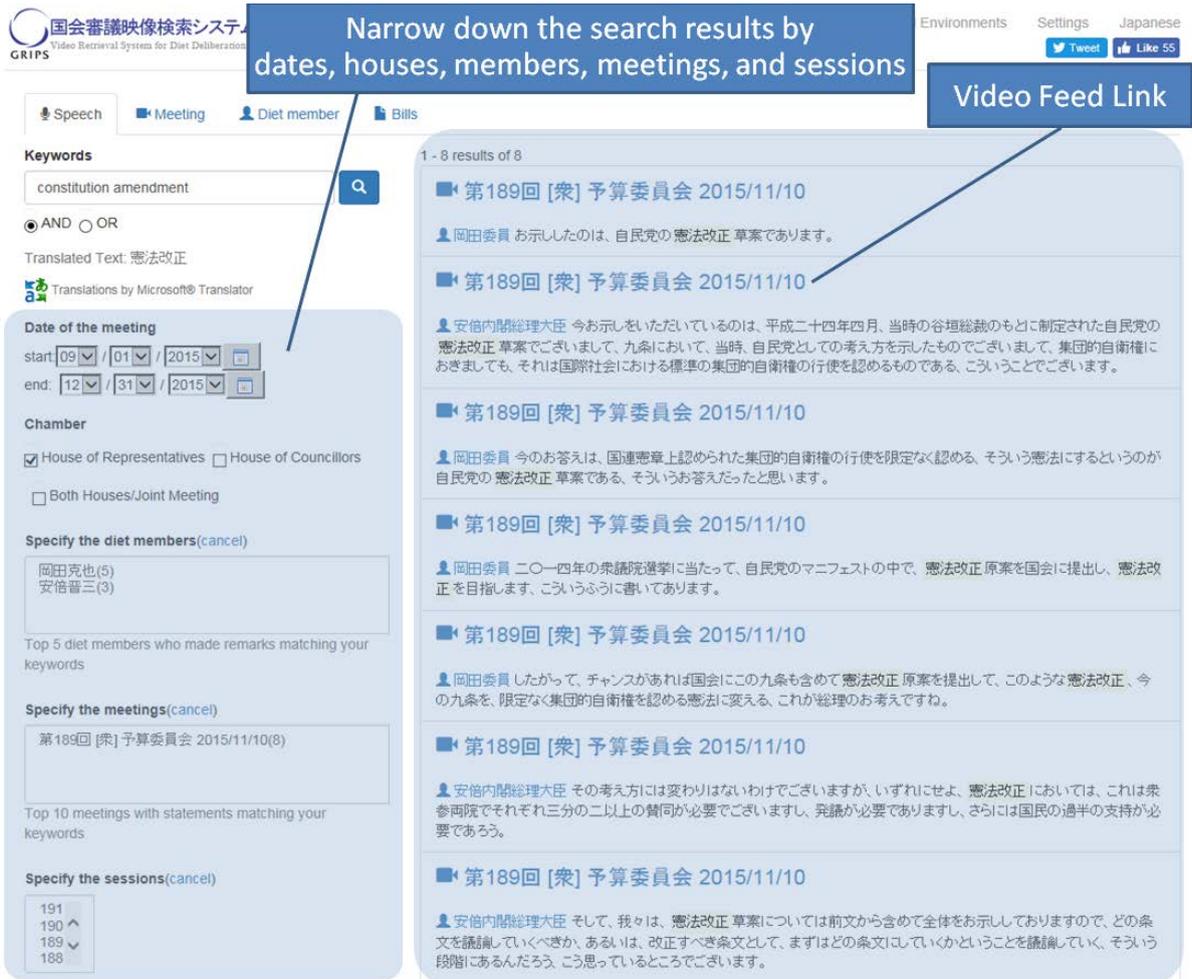
Figure 2: Keyword Search Interface in Japanese

As shown in Figure 3, one can type English keywords separated by spaces in the search field, and they will be translated automatically into Japanese and used in keyword searching. Any combination of keywords can be used. For instance, one could type “Shinzo Abe” (the name of the current Prime Minister of Japan) and “constitution amendment” in the search field. Alternatively, one could first select one of the four categories (speech, meeting, Diet member, and bill) from “search option” and type keywords in the search field. If one selects “Diet member” from “search option” and types “Shinzo Abe constitution amendment” in the search field, each of the keywords will be first considered as the name of a person and then as either a word in a speech, the title of a meeting, or the agenda of a meeting, giving priority to search results where Shinzo Abe uttered ‘constitution’ and ‘amendment’ in his speech, over results where some Diet members uttered ‘Shinzo’ ‘Abe’ ‘constitution’ and ‘amendment’ in their speeches.



**Figure 3: Keyword Search Interface in English**

The search results are listed as URL links. Once we hit the search button a screen similar to Figure 4 is displayed, listing the search results in order of date (keywords “constitution amendment” in the House of Representatives for the period September through December 2015). In addition to specifying AND/OR search options by selecting one of the boxes under the search field, we can filter the search results by date and limit the search to one or both of the two houses, or to joint meetings of the two houses. As shown in Figure 4, additional search options include speakers (top 5) and meetings (top 10) with words with utterances matching the search text, and allow us to further narrow the search results by selecting one of the meetings and one of the speakers.



**Figure 4: Video Feed Link**

If we click one of the video feed links the screen content will be similar to that in Figure 5. Subtitles are shown under the video. The speech list is shown on the right side, and speech currently playing is highlighted. The video will play for either one minute or for three speeches. Alternatively, we can keep playing the video by clicking the play button in the toolbar under the video. By double-clicking any speech in the list on the right side, we can instantly watch the video stream of the speeches before and after the speech found by keyword search. Once the user has moved on to another speech, the original speech found by keyword search remains highlighted in yellow.

Further, the URL for the corresponding segment of video streaming is shown below the video, and we can easily share the URL via SNS by clicking the tweet button above the URL bar while the video stream is playing. The text of the speech and the URL will immediately appear in the tweet box after the tweet button is clicked. At the bottom of the page the profile of the speaker is provided, followed by a list of agendas and a list of the Diet members attending the meeting (not shown in Figure 5)

To assist keyword searching, “Keyword Ranking” on the right side of the top page lists 15 words uttered in the Diet proceedings, in descending order of frequency, placing more weight on frequency in later parliamentary meetings than earlier and more weight on single meetings than multiple meetings. The font-size of the words increases with the number of searches, reflecting the attention given to Diet deliberation videos. An English translation pops up when the mouse hovers over any word.

第189回 [衆] 予算委員会 2015/11/10

The interface is divided into several key sections:

- Video Replay Area:** A video player showing a speaker at a podium. Below the video is a **Caption Area** with Japanese text.
- Speech List:** A scrollable list of speech segments with timestamps and speaker names.
- Speech Replaying:** A section for selecting and replaying specific speech segments.
- Play Control:** A control bar with play, pause, and volume buttons.
- SNS Link:** A field for the URL of the current speech, with a Twitter icon.
- Speaker's Profile:** A profile card for the speaker, including a photo, name, and biographical information.

Figure 5: Video Replay Interface

### Usage beyond Keyword Search

Our video retrieval system can be utilized in a variety of ways. For instance, we can create a list of search results with the query “constitution amendment” and the name of the speaker, “Abe Shinzo.” By clicking one of the video links in the list we can instantly retrieve video of the speeches by Prime Minister Shinzo Abe, with his own voice, facial expressions and body language, where he mentioned constitutional amendment.

Diet members are increasingly posting information about their activities on the websites. Some use their websites to display the minutes of parliamentary proceedings, and some even edit and upload deliberation videos on their websites. In contrast, our video retrieval system allows us to obtain the URL for a moment of video streaming and to create a list of video links without downloading and editing the video files.

Furthermore, we can use Twitter to create a list of parliamentary speeches. For instance, prime ministers customarily begin answering questions in plenary meetings by saying “There is a question regarding X.” Thus we can narrow the above-mentioned search results to speeches beginning with “There is a question regarding constitutional amendment,” and tweet the speeches and their video links to create a list of Prime Minister Abe’s plenary speeches on the issue of constitutional amendment.

Another way of utilizing the interfaces for keyword searching and partial replay is to post deliberation video links to internet news.

*On February 13, 2014, Shinzo Abe told a Lower House Budget Committee that he is the “ultimate arbiter” of affairs concerning the constitutional interpretation.*

For instance, if we find a report like the above in an internet newspaper, featuring the remarks made by Prime Minister Abe in the Diet, we can enhance the internet news visually by using our video retrieval system and inserting the video link (below) for the moment of video streaming in question. Clicking on the link will result in the instant playing of the video of the moment of Prime Minister Abe’s controversial remark (Figure 6).

[http://gclip1.grips.ac.jp/video/video/1117/speech\\_id/832311#main](http://gclip1.grips.ac.jp/video/video/1117/speech_id/832311#main)



**Figure 6: Shinzo Abe on Constitutional Interpretation**

The minutes of the proceedings are an important source of the content of discussion in the Diet, but they are not the whole story. For instance, supplementary materials are often used in parliamentary meetings, and discussions refer extensively to graphic materials such as figures and tables. Such supplementary materials are not included as part of the minutes unless a Diet member attending the meeting requests that they be included. Although the secretariats and the library of the Diet keep the supplementary materials used in parliamentary meetings, the existence of those materials is not widely known and they are hardly used.

As a preliminary analysis of the committee deliberation videos, we extracted video clips involving such supplementary materials by retrieving the deliberation videos corresponding to the minutes of the proceedings through keyword searching. For the period January 2010 to June 2016 (the 174<sup>th</sup> session to the 190<sup>th</sup> session), we found 34 speeches with the query “panel present” in the House of Representatives committee meetings, and played each of the deliberation videos by clicking the URL links listed in the search results. For each of the 34 speeches, a video clip focusing on a supplementary

material is displayed within 10 seconds or so. Figures 7 and 8 show the supplementary materials found by means of keyword searching.<sup>2</sup>

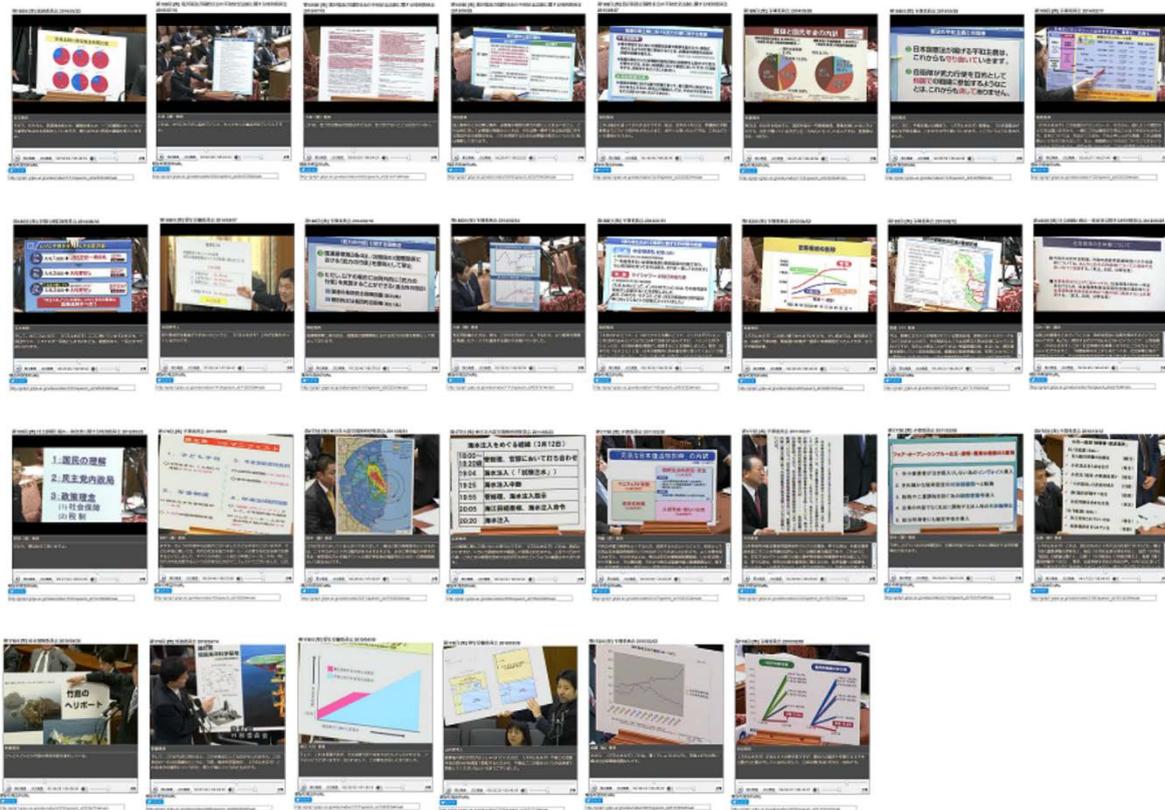


Figure 7: Supplementary Materials Found by Keyword Searching

Moreover, the minutes are silent with regard to non-verbal communication. For instance, a committee meeting may be suspended due to disorder or irregularities. The committee chair and senior members usually conduct under-the-table negotiations to deal with the situation, and there may be no written records of the suspension. Visual information may be useful in this regard. By checking deliberation videos, we may be able to get a sense of how long it took to resume and how the situation was dealt with during the suspension.<sup>3</sup>

By taking advantage of our system's combining of the interfaces for keyword searching and partial replay, we may be able to conduct experimental studies to determine how visual information affects the understanding of Diet deliberations.<sup>4</sup> We can think of the informatics of Diet deliberations as a multidisciplinary scientific approach to parliamentary deliberations, involving analysis not only of textual

---

<sup>2</sup> We are working on combining sound and pattern recognition techniques to distinguish between the portions of videos that do or do not focus on the speaker and creating a list of thumbnails for the videos including the moments focusing on supplementary materials used in committee meetings.

<sup>3</sup> Matsuura (2015).

<sup>4</sup> Kinoshita (2015).

information, but also of audio-visual information.<sup>5</sup>

第186回 [衆] 決算行政監視委員会 2014/06/16

The screenshot shows a video player interface. The main content is a presentation slide with a blue background and white text. The slide is titled 'ムリに予算を使わないとする厚労省' (Ministry of Health, Labour and Welfare aims to avoid overusing the budget). It lists three dates and corresponding bid statuses: 3月3日 (March 3rd) '入札1回目 → J.E.E.Dが一者応札 1アウト' (1st bid → J.E.E.D. one bid, 1 out); 4月8日 (April 8th) '入札2回目 → 入札者なし 2アウト!' (2nd bid → no bidder, 2 out!); 5月29日 (May 29th) '入札を分割しても 全国6地区のうち5地区で入札3回目 → 入札者なし 3アウト!!' (Even if bids are split, in 5 of 6 regions, 3rd bid → no bidder, 3 out!!). A red banner at the bottom of the slide reads '「不正入札」「ゾンビ復活」いわく付きの基金は 国庫返納すべき!!' (For 'unfair bids' and 'zombie revival', the fund with a bad reputation should be returned to the national treasury!!). Below the slide, the speaker's name '玉木委員' (Tanimaki Member) is shown, followed by a transcript snippet: 'そして、出てこないのです、(パネルを示す)ここに書いていますけれども、一回目やって、J.E.E.Dが一回落としますけれども、疑惑が出て、一回これでためになります。' (And, it doesn't come out, (pointing to the panel) it's written here, but after the first round, J.E.E.D. falls once, but there's a suspicion, and it's done for the first time, so it's a relief). The video player controls at the bottom show a progress bar at 02:23:26 / 04:55:04, a volume icon, and a '字幕' (Subtitles) button. Below the player, there is a '現在の発言のURL' (Current speech URL) section with a 'ツイート' (Tweet) button and the URL 'http://gclip1.grips.ac.jp/video/video/1518/speech\_id/1205890#main'.

Figure 8: An Example of Supplementary Materials

A unique aspect of our video retrieval system is that we use sound recognition techniques to create timestamp data to match Diet proceedings and deliberation videos. In other words, we deal with two types of text information related to parliamentary meetings. Both are written versions of speeches made in the Diet. The information derived from speech recognition is “correct” in the sense that it captures 100 percent of what was actually spoken in the Diet, although it may contain irrelevant filler and words wrongly recognized due to both individual speaker factors such as intonation and pronunciation and technical and environmental factors such as recording quality and noise. On the other hand, the minutes of the Diet proceedings become “official” after transcription to eliminate filler, correct inappropriate wording, and add commas and periods so that the speech in the Diet can make sense as a written language.<sup>6</sup>

<sup>5</sup> Our video retrieval system offers a particularly useful tool for scholars interested in the differences between spoken and written language. Matsuda (2016) used our system to check whether Diet members say “Ba’ai” or “Bayai,” which are written with the same Chinese characters. The minutes of parliamentary proceedings do not enable us to check if the two words are pronounced differently.

<sup>6</sup> The House of Representatives Secretariat in 2011 started using speech recognition to create a draft for meeting transcripts. A statistical model showed that the content of the minutes of the House of Representatives meetings matches approximately 87 percent of actual utterance. A language model was constructed to predict the actual utterance from minutes consisting of approximately 200 million words over a 10-year period, and was

Table 1: Correspondence Rates for the 190<sup>th</sup> Diet Session (Jan. 4 – Jun. 1, 2016)

House	Committee	Mean	St. Dev	Min	Max	N
Total		84.50	5.57	64.60	95.80	238
Representatives (R)	Total	84.91	5.58	64.60	95.80	125
	Plenary	85.49	5.03	73.10	93.40	37
	Budget	82.65	5.40	69.80	94.20	21
	Finance	81.72	6.31	64.60	94.90	20
	Justice	86.13	5.50	72.00	95.80	20
	Interior	86.42	4.58	76.40	95.80	18
	Agriculture	89.21	3.79	82.50	95.10	9
Councillors (C)	Total	84.05	5.56	68.50	93.90	113
	Plenary	84.52	5.83	68.50	93.60	32
	Budget	78.55	4.50	68.80	85.40	23
	Finance	84.04	4.99	75.80	93.90	16
	Justice	87.63	3.64	81.00	92.20	18
	Interior	86.67	3.48	79.70	91.40	15
	Agriculture	84.88	4.29	77.60	89.80	9
Difference (R-C)	Total	0.87	0.02	-3.90	1.90	12
	Plenary	0.97	-0.80	4.60	-0.20	5
	Budget	4.10	0.91	1.00	8.80	-2
	Finance	-2.33	1.32	-11.20	1.00	4
	Justice	-1.51	1.86	-9.00	3.60	2
	Interior	-0.25	1.10	-3.30	4.40	3
	Agriculture	4.33	-0.51	4.90	5.30	0

As a preliminary analysis, we collected data on the correspondence between the official minutes and the results of speech recognition of Diet proceedings videos. Table 1 reports statistics summarizing the correspondence rates for the plenary and committee meetings of both houses in the Diet during the 190<sup>th</sup> session.<sup>7</sup> To see if there is any systematic difference between the two houses or among policy areas, we selected the meetings of four standing committees in addition to the budget committee and plenary meetings. For a total of 238 meetings, meeting time ranges from 5 minutes to over 7 hours. To reduce computational demand, we calculated the rates of correspondence between the official minutes and speech recognition results for the first 10 minutes of each meeting and for the entire meeting time of those lasting less than 10 minutes.<sup>8</sup>

---

applied to more than 500 hours of recordings of parliamentary meetings to create an acoustic model. In the process of speech recognition, these models are used and updated in a semi-automatic manner. See Kawahara (2012) and Akita and Kawahara (2013) for the details. The House of Councillors Secretariat still relies solely on stenographic transcripts.

<sup>7</sup> There are three types of Diet sessions. An ordinary session is convened once per year for a term of 150 days by law, subject to extension by Diet resolution or to early termination due to parliamentary dissolution. Important legislative and budgetary matters are deliberated in ordinary sessions. The 190<sup>th</sup> Diet session is an ordinary session, held during the period January 4 to June 1, 2016.

<sup>8</sup> We are developing a web-based program to calculate correspondence rates as part of the standard procedure for creating timestamp data to match proceedings and deliberation videos. This will enable us to make the videos searchable in our video retrieval system so that we can systematically analyze the correspondence between the official minutes and

For the 190<sup>th</sup> Diet session, the matching rate ranges from 64.6 to 95.8 percent, with an average of 84.5 percent (see also Figure 9). The correspondence between the official minutes and speech recognition results is slightly higher for the House of Representatives. For the House of Representatives, the average matching rates of budget and finance committees are relatively low, and the average increases in order for plenary, justice, interior, and agriculture. The bottom part of Table 1 shows the differences between the two houses. The differences suggest that the matching rates are generally higher in budget and agriculture committees in the House of Representatives, but higher in finance and justice committees in the House of Councillors. The average matching rate for House of Councillors meetings is lowest for budget and increases, increasing in order for finance, plenary, agriculture, interior, and justice.

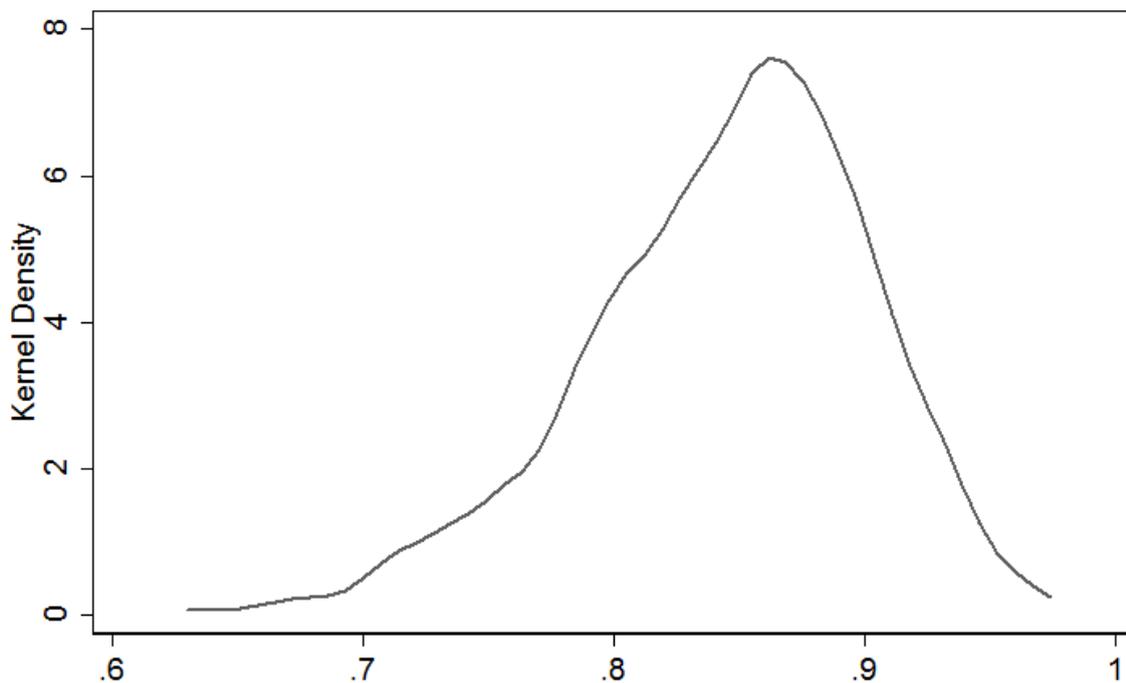


Figure 9: Distribution of Correspondence Rates

Although it is understandable that the correspondence rate is lower for the budget committees, the relatively low correspondence for the finance committee in the House of Representatives is somewhat unexpected and constitutes a potentially important finding, which may lead us to move from simply reading the official minutes towards a multifaceted understanding of parliamentary discussion. The matching rate is the lowest, 64.6 percent, for the Finance Committee meeting on February 29, in which expert witnesses made statements on the issue of earthquake recovery. Similarly, the matching rate for the Justice Committee is low, 72 percent for the May 10 meeting, when witnesses were requested to attend. The matching rate is 73.1 percent for the plenary meeting on January 22, when Prime Minister Abe and his cabinet ministers

---

speech recognition results by meeting, speaker, etc. for all meetings and for the entire duration of each meeting.

were hooted and jeered through much of their policy speeches. However, there is nothing out of the ordinary in the way parliamentary business was carried out in the House of Councillors for the budget committee and the plenary meetings on March 25 and 30 respectively, but the correspondence rates are below 70 percent.

The results of regression analysis for the rates of correspondence between the official minutes and speech recognition results are presented in Table 2. We use a generalized linear model (GLM) with a logit link and the binomial family to analyze the correspondence rates with values between zero and one. To take into account technical advancements in speech recognition technology, we incorporate April and August in the estimation as dummy variables, distinguishing the dates of speech recognition before April 1 and after August 15. As Table 1 suggests, there are some organizational factors that may systematically change the nature of parliamentary meetings. Four dummy variables included in the estimation model are HC, Plenary, Budget, and Finance, each assigned a value of one for meetings in the House of Councillors, plenary meetings, Budget Committee meetings, and Finance Committee meetings, and zero otherwise. As can be seen in Table 2, Plenary, Budget and Finance are estimated to have a negative impact on the correspondence rate between the minutes and the speech recognition results. April, August, and HC are estimated to be statistically insignificant.

Table 2: GLM Estimation of Correspondence Rates

	B	Robust SE	P
April	0.168	0.111	0.130
August	0.111	0.113	0.328
HC	-0.003	0.054	0.951
Plenary	-0.439	0.130	0.001
Budget	-0.538	0.068	0.000
Finance	-0.326	0.075	0.000
Log Words	0.049	0.035	0.160
Log Freq	-0.120	0.068	0.078
Address	-0.199	0.146	0.172
Witness	-0.378	0.137	0.006
Constant	2.051	0.338	0.000

N238. Log pseudolikelihood: -72.119.

Residual Deviance: 4.237. Pearson Deviance: 4.299.

Parliamentary meetings range in duration from 5 minutes to over 7 hours. The longer the meeting, the more complex the discussion may become. However, meeting length, measured as the total number of words in the minutes of the meeting in a logarithmic scale (Log Words), is estimated to have a positive but statistically insignificant impact on the correspondence rate. Also, speech recognition success may depend on whether or not the discussion is focused on specific issues. The logarithmic frequency of the word most frequently uttered in each of the parliamentary videos under study, Log Freq, is estimated to have a negative coefficient with modest statistical significance, implying that the more often a word is repeated in a meeting the wider the gap becomes between the minutes and the speech recognition results.

A dummy variable for the word in meeting agendas is included in the estimation model. Address indicates whether or not cabinet members made a policy address in the meeting; the estimate for Address is negative but not statistically significant. Finally, Witness is a dummy variable equal to one if either a government or an expert witness is

requested to attend the meeting. The estimate for Witness suggests that speech recognition may not perform well for meetings in which witnesses make statements.

We are interested in how these factors affect the rate of correspondence between the minutes and the speech recognition results. It is more intuitive to interpret the predicted values based on the logit coefficients. Assuming that both Log Words and Log Freq are held at their average values and all the dummy variables are set at zero, the predicted baseline correspondence rate would be 89.8 percent. For plenary meetings in the two houses, the correspondence rate decreases to 85.1 percent, with all other factors remaining the same (4.7 percentage points down from the baseline). Similarly, the correspondence rates are predicted as 83.8 and 86.5 percent, respectively, for meetings of budget and finance committees (6.0 and 3.3 percentage points down).

The correspondence rate decreases to 85.8 percent when witnesses are requested to make statements in a meeting (4.0 percentage points down). The average frequency of the word most frequently uttered is approximately 24 and the standard deviation is around 10. The predicted correspondence rate decreases from 91.6 percent to 90.3, 89.7, 89.3, and 89.0 percent, while the frequency of the word most frequently uttered increases by 10 (one standard deviation) over the range from 4 to 44 (two standard deviations down and up from the average), assuming that Log Words is held at its average value and all the dummy variables are set at zero.<sup>9</sup>

### Conclusion

Although our video retrieval system is still at the preliminary stage, we believe that our system offers alternative means of access to parliamentary video streaming for those who have visual and hearing impairments. Also, by incorporating automatic translation into the web-based keyword search interface, we have made our system available from computers without Japanese text capability. We found substantial variation in the rate of correspondence between the official minutes and the results of speech recognition of videos of Diet proceedings, which may be systematically related to the nature of parliamentary meetings by committee and policy area.

Our video retrieval system has great potential to boost the usage of parliamentary videos in relation to policy-making. The sound recognition techniques for creating timestamp data for matching video and text information are applicable to a wide range of proceedings, including those of local assemblies and administrative councils,<sup>10</sup> as well as other types of videos such as TV news clips.<sup>11</sup> Our system is

---

<sup>9</sup> Variance inflation factors range from 1.21 to 7.48 with a mean of 2.76. The observed and predicted dependent variables had a correlation of 0.490. See Appendix for summary statistics. In addition to using a logit link, we used a GLM with a complementary log log link to account for the skewed distribution of correspondence rates (see Figure 9). The estimation results are not substantially different from those reported in Table 2.

<sup>10</sup> We have developed a similar video retrieval system for local assemblies, and a sample of deliberation videos are searchable by keywords for 10 local assemblies of Kitami (Hokkaido), Miyako (Iwate), Chikusei (Ibaraki), Nerima (Tokyo), Tsubame (Niigata), Hashima (Gifu), Kameoka (Kyoto), Setouchi (Okayama), Naruto (Tokushima), and Kitakyushu (Fukuoka).

<http://gclip1.grips.ac.jp/local-assembly/>

The system is not currently publicly available. Please contact us if you are interested in trial use.

expected to enhance discussion thanks to a better understanding of policy making in the Japanese Diet, and to indicate a direction for the development of parliamentary video streaming systems in other countries. Much remains to be done in this regard, but we believe that our system encourages research on the use of visual information in policy-making and marks a step towards the provision of universal access to policy information.

### *References*

- Akita, Yuya and Tatsuya Kawahara. 2013. “Onsei Ninshiki wo Mochiita Onrain Jido Jimaku Sakusei Henshu Shisutemu” *Nihon Onkyo Gakkai Shuki Kenkyu Happyokai Koen Ronbunshu*. 2-8-4.
- Kawahara, Tatsuya. 2012. “Gikai no Kaigiroku Sakusei no tameno Onsei Ninshiki: Shugiin no Shisutemu no Gaiyo” *Joho Shori Gakkai Kenkyu Hokoku*. SLP—93—5.
- Kinoshita, Ken. 2015. “Kokkai Shingi no Eizo Joho to Moji Joho no Ninchiteki Sai” *Revaiasan*. 56: 117-138.
- Masuyama, Mikitaka. 2016a. “Kokkai Shingi Eizo Kensaku Shisutemu to sono Nyusu Doga eno Oyo Kanosei” Mimeo, National Graduate Institute for Policy Studies.
- Masuyama, Mikitaka. 2016 b. “Reducing the Space and Time between Citizens and Parliament: Video Retrieval System for Diet Deliberations” *Prepared for delivery at the 2016 General Conference of the European Consortium for Political Research*, Charles University, Prague, Czech Republic.
- Masuyama, Mikitaka and Kaori Takeda. 2014. “Instant Parliamentary Deliberations Are in Our Reach” *Presented at the 2014 Annual Meeting of the American Political Science Association*, Washington, D.C., USA, August 28-31, 2014.
- Masuyama, Mikitaka and Kaori Takeda. 2015. “Ikani Mitai Kokkai Shingi Eizo ni Totatsu suruka?” *Revaiasan*. 56: 54-79.
- Matsuda, Kenjiro. 2016. “Taisho Showa Senzenki no SPban Enzetsu Rekodo niokeru ‘Baa’ no Yomi nitsuite” *Kokuritsu Kokugo Kenkyusho Ronshu*. 11: 63-81.
- Matsuura, Junsuke. 2015. “Higashi Nihon Daishinsai no Hassei to Nihon no Kokkai Seiji” *Revaiasan*. 56: 102-116.

---

<sup>11</sup> We collected data on the correspondence between news transcripts and the results of speech recognition of internet videos dealing with the Diet for the period December 15, 2015 to March 31, 2016. Video duration is around 1 minute for most of the 271 news videos analyzed. Since the speech recognition procedures are designed for parliamentary meetings, it is not surprising that the average matching rate of 77.53 for political news, which may devote coverage to gaffes, blunders and scandals, is slightly lower than that reported in Table 1. The correspondence rate ranges from 15.9 to 98.2, with a standard deviation of 12.7. See Masuyama (2016a) for the details.

## Appendix

### 1. Summary Statistics

	Mean	St. Dev.	Min	Max
Correspondence Rate	0.845	0.056	0.646	0.958
April	0.046	0.210	0	1
August	0.176	0.382	0	1
HC	0.475	0.500	0	1
Plenary	0.290	0.455	0	1
Budget	0.185	0.389	0	1
Finance	0.151	0.359	0	1
Address	0.050	0.219	0	1
Witness	0.563	0.497	0	1
Words	55,247	46,438	278	171,565
Freq	23.62	10.09	4	69
Log Words	10.245	1.500	5.628	12.053
Log Freq	3.074	0.431	1.386	4.234

### 2. Deviance

$$\begin{aligned} \text{Residual Deviance} & \quad \sum_j d_j^2 & \quad d_j^2 = 2 \left( y_j \log \frac{y_j}{\mu_j} + (1 - y_j) \log \frac{1 - y_j}{1 - \mu_j} \right) \\ \text{Pearson Deviance} & \quad \sum_j r_j^2 & \quad r_j^2 = \frac{(y_j - \mu_j)^2}{\mu_j(1 - \mu_j)} \end{aligned}$$

### 3. GLM Estimation: Cloglog Link

	B	Robust SE	P
April	0.076	0.053	0.149
August	0.048	0.051	0.347
Log Words	0.023	0.016	0.145
Log Freq	-0.053	0.031	0.090
HC	-0.002	0.024	0.927
Plenary	-0.196	0.057	0.001
Budget	-0.247	0.032	0.000
Finance	-0.146	0.035	0.000
Address	-0.091	0.067	0.175
Witness	-0.173	0.061	0.004
Constant	0.769	0.152	0.000

N238. Log pseudolikelihood: -72.118.

Residual Deviance: 4.234. Pearson Deviance: 4.298.