Introduction to the Real-time Articulatory Movement Database - Version 2

Kikuo Maekawa
(National Institute for Japanese Language and Linguistics)
2024.04.01

Table of Contents Page

1. Introduction

We are pleased to announce the release of the second version of the "Real-time MRI Articulatory Motion Database. This is a database of moving images of the midsagittal section of the vocal tract during the production of Japanese utterances, recorded at a rate of 14 or 27 frames per second by using a medical MRI system with special operating settings. This data has realized the dream of articulatory phoneticians to visualize articulatory movements and may be widely used for critical review of the existing articulatory phonetic descriptions, education of Japanese pronunciation, or speech synthesis from vocal tract shape.

The data has been collected by the author's research group since 2017. In April 2021 the trial version containing data of about 13,000 utterances by 10 speakers was released. In April 2022, Version 1 of the database (abb. rtMRIDB_v1) containing more than 26,000 utterances by 22 speakers was released. This time, Version 2 of the database (abb. rtMRIDB_v2) is publicly available. The new version includes the following enhancements:

(1) With 3 new speakers, the total number of utterances by 25 speakers is now about 30,000.
(2) New movies with superposed dentition image are provided for all utterances of all speakers.
(3) New audio noise reduction technique is applied for all audio data.
(4) Minor improvements in the search system

For the contents of the database, please refer to sections 2 to 6 of this document. If you want to know more about MRI imaging conditions, please refer to the appendix of Maekawa et al (2020). When using this database, please be aware that the database is released under the Creative Commons CC-BY-NC-SA license, and, the Principal Investigator and the National Institute for Japanese Language and Linguistics are not liable for any problems arising from the use of this database.

Lastly, an information for the users using the desktop version of the rtMRIDB_v1: the rtMRIDB_v2 is provided only as a web version. The desktop version provided by the National Institute of Informatics Spoken Language Resources Consortium (NII-SRC) will not be updated to version 2 this time.

2. Fields of the rtMRIDB_v2 data

The rtMRIDB_v2 covers 29,984 utterance samples. All samples are video movie in the MP4 format. To make them searchable, all the samples are given the information in Table 1. In the rest of this document, we will call these the fields of a database record.

Table 1: The 20 fields that make up a record of the rtMRIDB_v2

	FIELD NAME	DESCRIPTION
1	file	Name of the MP4 video file (not including the file extension)
2	start	Time at the beginning of the sample (time from the beginning of the file)
3	end	Time at the end of sample (time from the beginning of the file)
4	date	Date of data recording
5	fps	Number of frames per second (14 or 27)
6	text	A character string that uniquely identifies the content of the utterance as instructed to the speaker
7	jtext	Text in Japanese character displayed for the speaker in recording
8	phoneme	A phoneme sequence that constitutes an utterance.
9	tag	Sample-specific events
10	class	Class of speech content (either 'MU', 'MB', 'MP', 'TT', 'SR', 'PL', or 'NS')
11	slide	ID of the slide presented to the speaker (without the repetition index)
12	slide2	ID of the slide presented to the speaker (with the repeat index)
13	ser	Location of the text in a slide
14	session	Session number of the MRI recording (the serial order in the recordings sessions performed in a single day)
15	subject	Speaker ID (begins with the letter 's' like s1, s2 etc.)
16	subjectID	Integer part of the speaker ID
17	gender	Speaker's gender (either 'F' or 'M')
18	birthYear	Year of birth of the speaker (A.D.)
19	birthPlace	Birthplace of the speaker (name of prefecture or city)
20	dialect	Speaker's dialect (Standard or Kinki)

2．1　The ‘file’ field

This field stands for the file name of the MP4 video file in which the utterance in question is recorded. The file name is a string such as “s5_40_mb26b”. This field consists of the “Subject”, “Session2”, and “Slide2” fields in Table 1, which are concatenated, from left to right, with underscores. In some files, a suffix is added after the file name like “s7_11_mp1_add” or “s7_22_mu5a_27_add”. In these examples, "_add" and "_27_add" are the suffixes. The suffix "add" indicates that it is an additionally recorded sample, i.e., the second recording from the same speaker who has already made recording of that utterance. The suffix "27" (sometimes "28") indicates the frame rate of 27 fps. Users are requested to note that the number 27 and 28 also appear in the session information, so the meaning of “27” (or “28”) differs depending on where it appears in a “File” field.

2．2　The ‘start’ field

The sequences of still image data captured by the MRI machine are converted to a MP4 video for each session (see 2.14), and, in most cases, a single MP4 file contains multiple utterances. The “Start” field contains the starting time of a given sample as measured from the beginning of the file. The unit is seconds [sec].

2．3　The ‘end’ field

This field contains the end time of a sample corresponding to the start field.

2．4　The ‘date’ field

This field indicates the date of data recording using an 8-digit integer such as 20171105. It is not possible for the same speaker to record twice on the same day, but there are cases where data was recorded on different occasions (different dates). It is important to notice that the head position in the MRI system is different for each date, and the midsagittal plane is not exactly the same. These issues should be taken into consideration when analyzing the data.

2．5　The ‘fps’ field

This field contains the frame rate of the sample. The frame rate is the number of frames per second. This field takes values of either 14 or 27, but these are not exact values. They are approximations to 13.79 ... and 27.17... These values are calculated by the following formula with the number of still images in the session being 512.

Number of still images in the session / Length of time in the session [sec]

Samples recorded at 27 fps contain the number "27" or "28" in their ‘file’ field. The difference of “27” and “28” does not mean anything, it simply reflects that “28” was used in early days of recording and replaced by “27”. The recording rate is always 27.17... fps.

Table 2：List of homophonic pairs in the utterances

Text	Jtext	Class	Slide	N
ae_1	アエ	MU	mu4	35
ae_2	会え	MP	mp4	36
ai_1	アイ	MU	mu4	35
ai_2	愛	MP	mp2	33
aki_1	秋	MP	mp3	35
aki_2	空き	MP	mp4	36
ama-imo	甘芋	MP	mp3	35
amai-mo	甘い藻	MP	mp4	36
ao_1	アオ	MU	mu4	34
ao_2	青	MP	mp3	34
au_1	アウ	MU	mu4	34
au_2	会う	MP	mp4	36
eHgo_1	英語	MP	mp1	55
eHgo_2	Ａ５	MP	mp4	36
ha-isja	歯医者	MP	mp4	36
hai-sja	敗者	MP	mp3	35
kaeru_1	帰る	MP	mp3	35
kaeru_2	飼える	MP	mp4	35
kai_1	貝	MP	mp3	35
kai_2	下位	MP	mp4	36
kehe	ケヘ	MB	mb5	27
kehe_1	ケヘ	MP	mp2	36
koHni	高二	MP	mp4	36
ko-oni	小鬼	MP	mp4	36
sato-oja	里親	MP	mp4	36
satoH-ja	砂糖屋	MP	mp4	34
sjoHka-izjoH	消化異常	MP	mp4	35
sjoHkai-zjoH	紹介状	MP	mp3	35

２．６　The ‘text’ field

The information in this field is alpha-numeric characters that uniquely identify the text shown to the speaker in the recording. For example, if the speaker recites the Chinese character string "新案", the “Text” field contains an alphabetic string of “siNaN”. The phoneme sequence of the item is used as the basis of the “Text” field, but in the case of homophones like “英語” (/eHgo/ ‘English’) and “A5” (paper size) or “帰る” (/kaeru/ ‘return’) and “飼える” (‘be able to keep’), the suffixes consisting of underscore and integer are used: “eHgo_1” and “eHgo_2” in the case of former and “kaeru_1” and “kaeru_2” in the case of the latter.

There are also cases where a hyphen is used to distinguish homonyms by indicating morphological boundaries. An example is “amaimo” (‘sweet potato’) and “amai-mo” (‘sweet seaweed’), or “haisja” (‘loser’) and “ha-isja” (‘dentist’). Note that in these cases, the hyphen is not attached to all morpheme boundaries, but only to those that are minimally necessary for the distinction.

Finally, some ‘text’ fields begin with an underscore, such as “_neutral”, “_suspicion”, “_focus0”, “_takegaki”, and “_northwind”. These are exceptional entries where the string represents information other than phonemes. For example, “_suspicion” represents the specification of so-called paralinguistic information in the PL class (see below), and “_norhwind” represents the reading of the "Northwind and the Sun" fable in the NS class (see below). Table 2 shows the list of homophonic pairs in the utterance list, and Table 3 shows the list of utterance items whose ‘text’ field begin with underscore. In both tables, N indicates the frequency of occurrence of the item in rtMRIDB_v2.

Table 3 : List of utterances whose ‘text’ field starts with an underscore

text	jtext	class	slide	N
_northwind	『北風と太陽』	NS	ns1	22
_space	『膨張する宇宙』	NS	ns2	20
_kikuguri	『菊栗』	TT	tt1	25
_takegaki	『竹垣』	TT	tt1	23
_tora	『虎』	TT	tt1	16
_neutral	『中立』	PL	pl1	34
_suspicion	『反問』	PL	pl1	33
_admiration	『感心』	PL	pl1	27
_disappointment	『落胆』	PL	pl1	27
_focus0	『0』	PL	pl2	34
_focus1	『1』	PL	pl2	29
_focus2	『2』	PL	pl2	26
_focus3	『3』	PL	pl2	25

2．7 The ‘jtext’ field

This is a string of the Japanese characters shown to the speaker at the time of data recording. Basically, it is a Japanese string that corresponds to the phonetic information in the ‘text’ or ‘phoneme’ fields. Some kanji strings have furigana (pronunciation keys written in Japanese syllabic characters) to prevent misreading, but they are omitted in the ‘jtext’ field. In the recording of the utterances in the PL, TT, and NS classes (see below), the specification of the semantic content of the utterance is given instead of the phonetic information. These specifications are given enclosed by the Japanese ”『』” parentheses as in Table 3.

2．8　The ‘phoneme’ field

This field indicates the segmental phoneme representation of an utterance. It is identical to the string obtained by excluding suffixes and hyphens from the string in the “text’ field in most cases. If, however, the ‘text’ field starts with an underscore, the ‘phoneme’ field is blank (See sections 2.6 and 2.7 above). Also, the ‘phoneme’ fields for "スィ" and "ズィ" in slide MU4 are "s_i" and "z_i", which contain underscores. These stand for the IPA phonetic symbols of "si" [si] and "zi" [zi]. Note ordinary “シ” [ɕi] and “ジ” [ʑi] are represented as "si" and "zi" respectively in the ‘phoneme’ field.

2．9　The ‘tag’ field

This field contains various tags attached to samples to represent sample-specific events. The following 12 tags are prepared for various, mostly irregular, articulatory events. Table 4 shows the total frequency of each tag (Total) and the breakdowns by utterance classes (MU-SR). Utterance classes will be explained in section 3.

Table 4 : Frequency of tags and the breakdowns by the utterance classes

Tag	Total	MU	MB	MP	SR
[noRPf]	284	87	184	13	0
[pz]	246	0	246	0	0
[noRPp]	227	24	167	35	1
[?]	81	2	73	6	0
[d]	69	0	67	2	0
[noRPp][noRPf]	23	6	17	0	0
[head_mv]	15	0	11	4	0
[lipLick]	5	2	0	2	1
[pz][d]	5	0	5	0	0
[pz][noRPf]	5	0	5	0	0
[pz][noRPp]	4	0	4	0	0
[fp][noRPp]	2	0	1	1	0
[pz][?]	1	0	1	0	0

[noRPf] tag: This tag is applied when the vocal tract has not returned to the so-called resting position at the end of an utterance. There are three main criteria for recognizing a vocal tract in its resting position: (1) lowered uvular and opening of the nasal passage, (2) no contact between the tongue and the palate, and (3) lowered larynx. In a typical resting state, all of these are satisfied. However, there are utterances that are lacking a typical resting position either at the beginning or the end. The [noRPf] tag is given when it is difficult to identify the resting position after the end of the utterance. In the same way, [noRPp] tag is applied when the

resting position immediately before the start of an utterance cannot be confirmed. See section 5 for the criteria of recognizing an utterance.

[pz] tag: As described below, in the utterances of the MB class, the target bimora is located in the carrier sentence "korega___gata" (this is ~type). Although speakers are instructed not to insert a silent pause after "korega" (this is), it is not uncommon for a perceptible pause to still be inserted between the carrier sentence and the target bimora. In such cases, this tag is applied. The majority of pauses are inserted immediately after "korega", but there are a few rare cases where a pause is inserted immediately after the target (just before "gata").

[noRPp] tag: This tag is applied when the vocal tract in its resting position cannot be observed in the time interval preceding the articulatory movement of the speech item. See section 5 for more details on this issue.

[d] tag: This tag applies to the utterances of the class MB in which there is some kind of disfluency in the target bimora. A typical disfluency is related to the length of the consonantal closure, when, for example, "korega kaka gata" sounds like "korega kakka gata".

[noRPp][noRPf] tag: This composite tag is applied to utterances where the vocal tract in its resting position cannot be observed both before and after the utterance.

[?] tag: This tag is applied to utterances where it is difficult to confirm whether the utterance has been realized correctly either from an articulatory or perceptual point of view. For example, when listening to an utterance of an MB class whose 'jtext' is "samo", one may not be able to clearly decide whether it is "korega samo gata" or "korega sama gata". In such a case, the [?] tag is applied if it is still difficult to make the decision after referring to the recorded MRI image of the articulation. In a case where articulatory movement suggests that the speaker pronounced a different phoneme mistakenly, then [err] tag (see below) is applied instead of the [?].

[head_mv] tag: This tag is applied when a positional change in the speaker's head is clearly observed during an utterance (in most cases, the change is a nodding-like movement of the head). In utterances of the MB class, it is not uncommon to observe a slight upward or downward movement at the end of an utterance, but any positional changes that occur outside of the target bimora are ignored.

[lipLick] tag: This tag applies to the utterances in which the speaker licks his lips with the tip of his tongue just before or just after the start of the utterance. See section 5 for details.

[pz][d] tag: This composite tag is applied when the [pz] and [d] described above cooccur in the same utterance.

[pz][noRPf] tag: This composite tag is applied when the [pz] and [noRPf] cooccur in the same utterance.

[pz][noRPp] tag: This composite tag is applied when the [pz] and [noRPp] cooccur in the same utterance.

[fp][noRPp] tag: Sometimes a filled pause such as "eh" is uttered before the articulation of a speech item. This tag is applied to those utterances where it is impossible to separate the filled pause from the target utterance. Note that the articulation of the filled pause is included in the records with this tag.

[err] tag: This tag is applied to utterances in which the speaker articulated a phoneme sequence that is completely different from the text printed on the slide. In the original data of the database, there are 28 occurrences of the [err] tag, but all those utterances are removed in the process of database compilation. So, you can't find this tag in the database.

2．10 The 'class' field

An utterance of rtMRIDB_v2 belongs to one of the following utterance classes: MU, MB, MP, SR, PL, TT, and NS. See section 3 for details of each class.

2．11 The 'slide' field

In the recording of the rtMRI data, the speaker reads one slide per session (see 2.14), and around 50 slides are presented in a single recording. These slides are identified by the 'slide' field, which is recorded in the form of a 'class' name followed by a serial number, such as MU1, MU2, ..., MB1, MB2, ..., MP1, MP2, ...

2．12 The 'slide2' field

A speaker reads one slide per session (see 2.14). However, sometimes it happens that the speaker may not be able to read through all items printed in a slide within the time of one session (about 37 and 19 seconds in the case of 14 and 27 fps recordings respectively). In such a case, recording of the same slide will take place again. When the same slide is recorded more than once on the same recording date, each recording session needs to be distinguished. For this purpose, we made it a rule to add alphabetical suffixes (i.e., 'a', 'b', 'c' …) to the end of the 'slide' field, which makes the 'slide2' field. The suffix 'a' indicates the first recording, 'b' the

second, ‘c’ the third, and ‘d’ the fourth. No suffix is given to utterances with only one recording. In most slides, the texts to be pronounced are printed in four or five lines (see Figure 1). When rereading a slide, the speaker is instructed to start at the beginning of the third line and read to the end, then go back to the first line and read to the end of the second line.

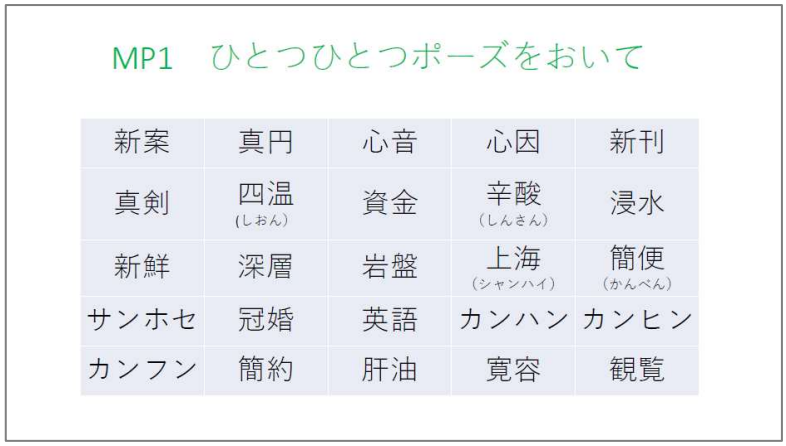


Figure 1 : Example of slide (the first slide in the MP class)
Note some items have ‘furigana’ (pronunciation key).

２．１３ The ‘ser’ field

As can be seen in Figure 1, a slide usually contains multiple items (usually words), and the speaker starts reading with the top-left item and continues his way through the items. The ‘ser’ field indicates the serial location of an item on a slide. In the case of Figure 1, the ‘ser’ values for “新案”, "心音" and "浸水" are 1, 3 and 10, respectively.

The recordings of the rtMRIDB_v2 data were conducted from 2017 to 2023. During this period, the utterance list was expanded from time to time. However, although we added utterances, we do not delete utterances or change the serial order of utterances in a slide. For this reason, the ‘ser’ value is constant regardless of the time of data collection.

But there is one exception. In the MP1 slide, there were two items, "完備" and "幹部", immediately after "上海" (whose ‘ser’ being 14) in early days of recording. These two items were deleted in the recordings later than the 20180233 (including the recording of this date). For the sake of raw data management, these two items are also assigned their ‘ser’ values, so in the MP1 slide, the ‘ser’ values of items after "簡便" are given a ‘ser’ value that is two larger than it appears. For example, the ‘ser’ values of "簡便" and “観覧” are 17 and 27 respectively. This irregular manipulation is concerned only with the MP1 slide. No other slide has such a problem.

２．１４　The ‘session’ field

In the recording, an rtMRI data is captured continuously for about 37 seconds (14 fps) or 19 seconds (27 fps) at a time. We call it a "session". Usually, 50-60 sessions are involved in a single recording that usually lasts 60-70 minutes. The MRI image data is managed by referring to the

session. The ‘session’ field is redundant because, as explained in 2.1, the ‘file” field contains the session information, but we prepared this field for the convenience of data search and analysis.

２．１５ The ‘subject’ field

This field uniquely identifies speakers. It consists of a letter ‘s’ followed by an integer in the range of 1-30, like ‘s1’ or ‘s24’. The number of speakers of rtMRIDB_v2 is 25, so there are missing values in the integer part of the field. This is because some of the subjects are not native speakers Japanese.

２．１６ The ‘subjectID’ field

This field corresponds to the integer part of the ‘subject’ field. Needless to say, this is redundant information, but it is provided for the convenience of data search.

２．１７ The ‘gender’ field

This field indicates the gender of the speaker. It takes the value ‘F’ for females and ‘M’ for males.

２．１８ The ‘birthYear’ field

This field contains the birthyear of the speaker. By comparing this information with the first four digits of the ‘date’ field, it is possible to obtain an approximation of the speaker's age at the time of data acquisition.

２．１９ The ‘birthPlace’ field

This field contains the birthplace (mostly at the level of the prefecture) of the speaker.

２．２０ The ‘dialect’ field

This field takes the value of ‘Standard’ if the speaker is speaking Standard (Tokyo-style) Japanese, or ‘Kinki’ if the speaker is speaking a Kansai dialect like Kyoto, Osaka, or Kobe.

３．Utterance class

As mentioned earlier in section 2.10, rtMRIDB_v2 has seven different utterance classes. The following sub-sections describe the characteristics of each class. The list of utterance items belonging to each class is shown in Appendix I. The number of utterances in each class for each speaker is shown in Table 5 below.

３．１ The ‘MU’ class

MU stands for “mora unigram”. It is a collection of separated Japanese morae, such as "ka", "su", "sja", "me", "rju", etc. As with the other classes, the items belonging to the MU class have been

expanded from time to time by including so-called peripheral mora such as "sje", "wo", "fe", "s_i" (See 2.8), and so on. Currently, 145 morae belong to the MU class.

In addition to the above, there are two types of additional items in the MU class. First, the sequences of two vowels like "ai", "ao", "ae", etc. are included in the MU class. Second, in the data recorded after (including) the date of 20180720, the vowel sequences separated by an /r/ like “ara”, “ari”, “ora”, “ori” etc. are also included in this class.

The MU class data was recorded at 14 fps from all speakers (although the number of items varied depending on the recording date), and also at 27 fps from some speakers. Items in the MU class provide the most basic information about the articulation of the Japanese language. The 14-fps data are collected from all 25 subjects.

3．2 The ‘MP’ class

MP stands for “mora phoneme”. Items in this class include words (either meaningful or meaningless) containing Japanese special mora, i.e., the moraic nasal (/N/), geminate (/Q/), long vowel (/H/), and diphthong (/J/, but see below). The number of speech items in the MP class has been expanded, from 97 in 2017 to 151 in 2022. This is because the analysis of special morae, especially the moraic nasal, has been carried out in parallel with the data recording. The MP class data was recorded at 14 fps from all speakers (although the number of items varied depending on the recording dates), and at 27 fps from some speakers. The latter, however, is mainly limited to the MP1 slides.

It is mentioned above that Japanese special morae contain diphthongs. But the phonemic distinction is not made between, for example, the diphthongal /aJ/ and non-diphthongal /ai/ in this database, because the realization of the diphthong is highly variable in Japanese. All items are treated as having the /ai/ phoneme strings. Other vowel sequences that can be realized as diphthong are treated alike.

3．3 The ‘MB’ class

MB stands for “mora bigram”, which is a combination of two morae pronounced in a carrier sentence. The target bigram consists of the combination of 26 morae, viz., ”ka, ki, ku, ke, ko, kja, kju, kjo, sa, si, su, se, so, sja, sju, sjo, ha, hi, hu, he, ho, ma, mi, mu, me, mo", resulting in 26 x 26 = 676 two-mora meaningless words. They are uttered in a carrier sentence "korega ___ gata” (Gloss: “this is the ~ type").

The MB class makes up the bulk of rtMRIDB_v1 quantitatively. The main purpose of this class is to provide material for quantitative analysis of coarticulation in Japanese segmental phonemes. The suffix morpheme /kata/ (pronounced as [gata]) used in the carrier sentence is known as a deaccenting morpheme that deletes the lexical accent in the immediately preceding morpheme. As a result of this, all target bigrams are realized as unaccented words (both in the

Standard and Kinki Japanese). In addition, the same "ga" mora is placed immediately before and after the target bigram to provide a uniform phonemic context.

The number of items in this class is 676, regardless of the recording date; data in the MB class were recorded at 14 fps from all speakers, and at 27 fps from some speakers. The latter recording was done only for the MB1 slide.

3．4　The 'SR' class

SR stands for "speaking rate". In this class, we asked speakers who finished the recording of MU, MP, and MB classes to read aloud two slides of the MB class with intentionally decreased speaking rates. The data of the SR class was collected from only ten speakers in 2017 and 2018 (s1, s2, s5, s7, s8, s9, s10, s11, s12, and s14). Recording of the SR class was ceased in late 2018 when stable imaging at 27 fps became possible.

3．5　The 'PL' class

PL stands for "para-language". This class aims to examine the effects of paralinguistic information on articulatory movements. Among the various types of paralinguistic information, rtMRIDB_v1 focuses on two types. One of them is the speaker's intentions or attitudes, where speakers are asked to read aloud the same text with four types of intentions: "neutral," "suspicion," "admiration," and "disappointment. We used semantically neutral sentence "Yamada-san ga" (Gloss, Mr. Yamada-AGENT) as the text. The acoustic effects of these intentions are explained in chapter 3 of 森・前川・粕谷(2014).

The other type of paralinguistic information is the contrastive focus. In this task, speakers are asked to utter a sentence repeatedly placing prosodic emphasis on different parts of the text. We used the text "yachin-no takai manshon-ni haitta (Gloss, "I entered an apartment with high rent") as the text and specified four types of foci: type 0 (no focus), type 1 (focus on "yachin-no"), type 2 (focus on "takai"), and type 3 (focus on "manshon-ni"). The "Phoneme" fields of these speech items are blank, as described in 2.8.

The PL class data was collected from 14 speakers between 2017 and 2019 (s1, s2, s7, s8, s9, s10, s11, s12, s16, s17, s18, s19, s20, s21). All samples were captured at 14 fps. After 2020, the recording was ceased because the recording time became tight due to increase in the number of items in other classes.

3．6　The 'TT' class

TT stands for "tongue twister". Three phrases were used as texts: "kikuguri kikuguri mikikuguri awasete kikuguri mukikuguri", "kono takegakini take tatekaketanowa take tatekaketakattakara take tatekaketanoda", and, "torao torunara torao toruyori torio tori torio otorini torao tore". Data of the TT class were collected from 14 speakers between 2017 and 2019 (s1, s2, s5, s7, s8, s9, s10, s11, s12, s14, s18, s19, s20, s21). All were captured at 14 fps. After 2020, the recording

was ceased because the recording time became tight due to increase in the number of items in other classes. Another reason for the cease was that many of the speakers had practiced well beforehand, so their speech was much smoother than we had initially expected.

3．7　The ‘NS’ class

NS stands for “narrative speaking”. This class aims to compare the word-level material (those in the MU, MP, MB classes) with so-called continuous speech with a coherent story. The texts used were "The North Wind and the Sun", which is used in the IPA (International Phonetic Alphabet) handbook (IPA 1999), and "The expanding universe", which is one of the read speech texts in the Corpus of Spontaneous Japanese (Maekawa 2003). The former text consists mainly of Japanese native words ("wago") , while the latter contains many Sino-Japanese words including palatalized phonemes (so-called "yoo-on").

Data collection of the NS class started in 2019, and 15 speakers participated in the recording so far (s1, s2, s4, s5, s7, s8, s12, s14, s19, s20, s21, s24, s25, s26, s27).

4．Differences of data by the speaker

Although the rtMRIDB_v2 contains the data of 25 speakers, the number of recorded utterances differs depending on speakers. Table 5 shows, from left to right, the properties (‘dialect‘, ‘gender‘, ‘year of birth‘) of the 25 speakers in the rtmRIDB_v2, the number of times they were recorded on different dates (‘N dates‘), the number of utterances by class (‘MU‘-‘NS‘), the total number of utterances (‘Total’), and the number of utterances captured at 27 fps out of the total (‘27 fps‘).

As you see, the number of utterances is not identical across speakers. There are three reasons for this variation in the number of utterances. The first reason is the expansion of the utterance list. As explained in the previous section, the utterance list used in the rtMRI recording is constantly being expanded. As a result, for MU and MP classes, the number of utterances tends to increase as the recording period becomes later. To correct this problem, for speakers who were recorded earlier, additional items that were missing in the early utterance lists were recorded later. As of the time of writing, nine speakers (s1, s2, s4, s5, s7, s8, s12, s14, s19) have been additionally recorded.

The second reason is errors and repetitions. Occasionally, a speaker makes an utterance that is different from the one indicated on the slide and proceeds to the next utterance without noticing the error. This does not happen frequently because, during the recording, an experimenter monitors the speaker's speech in real-time. If an error is detected by the experimenter, the session is re-recorded. However, on a rare occasion, an error is overlooked both by the speaker and experimenter (partly because the MRI machine makes quite a lot of noise when it is in operation). In that case, the data for that item will be tagged with [err] and become a missing value.

On the other hand, repetition of utterance occurs when a single slide is read aloud multiple times. There are several reasons why this can happen. One is when an error is found in the speech, as mentioned immediately above. Repetition can also occur when a single slide is not fully read in a single session, as described in 2.12.

The final source of variation in the number of utterances is the additional recording described at the beginning of this section. Repetition occurs because, in the additional recordings of the MU and MP classes, speakers read not only the newly added items but also all other items printed in the same slide.

In rtMRIDB_v2, the same item recorded on the same date can be identified by the suffixes 'a', 'b', 'c' in the "Slide2" field (see 2.12). Multiple readings resulting from additional recordings can be identified by the value of the "Date" field (see 2.4).

Table 5: Variation in the number of utterances by speaker and class

Subject	Dialect	Gender	BirthYear	N date	MU	MB	MP	SR	PL	TT	NS	Total	27fps
s1	Standard	M	1956	2	177	797	272	35	15	4	2	1302	68
s2	Standard	M	1970	2	197	865	364	31	20	4	2	1483	145
s3	Kinki	F	1952	1	137	716	140	0	0	0	0	993	65
s4	Standard	M	1969	2	244	775	315	0	0	0	2	1336	131
s5	Standard	M	1958	2	195	746	305	41	0	3	4	1294	127
s7	Standard	M	1955	2	210	800	320	38	12	10	3	1393	77
s8	Kinki	F	1968	2	175	811	299	31	16	4	2	1338	66
s9	Standard	M	1990	1	109	846	129	41	8	3	0	1136	51
s10	Standard	F	1967	1	108	744	147	31	18	3	0	1051	58
s11	Standard	F	1971	1	109	741	150	31	24	3	0	1058	48
s12	Kinki	M	1961	2	228	716	310	41	10	5	2	1312	148
s14	Standard	M	1950	2	212	855	319	31	0	5	2	1424	117
s16	Standard	F	1969	1	175	752	217	0	20	0	0	1164	52
s17	Standard	F	1969	1	167	752	201	0	16	0	0	1136	55
s18	Standard	M	1958	1	140	789	208	0	18	3	0	1158	59
s19	Standard	M	1961	2	230	774	408	0	18	8	2	1440	128
s20	Standard	M	1994	1	142	734	149	0	19	4	1	1049	0
s21	Standard	M	1991	1	142	782	171	0	21	5	1	1122	0
s24	Standard	F	1956	1	192	787	196	0	0	0	4	1179	0
s25	Standard	M	1969	1	145	676	149	0	0	0	2	972	0
s26	Kinki	M	1964	1	205	745	174	0	0	0	2	1126	0
s27	Kinki	F	1970	1	144	731	191	0	0	0	2	1068	0
s28	Kinki	M	1950	1	143	786	220	0	0	0	4	1153	0
s29	Kinki	F	1978	1	145	730	218	0	0	0	3	1096	0
s30	Standard	F	1982	1	145	817	220	0	0	0	2	1184	0

5. Segmentation of utterance from a point of view of articulatory movement

Among the tags described in 2.9, [noRPp] and [noRPf] are closely related to the segmentation of utterance, i.e., the determination of the "start" and "end" values. Compared to the segmentation of utterance by audio signals, articulatory segmentation is much more complex. In this section, some basic issues will be explained.

The first thing to recognize is that articulatory movements begin in advance of the acoustic signal. For example, in producing the mora "pa", the perceivable speech signal is produced only at the burst stage of the bilabial closure of [p], but the various speech organs begin their articulatory movements before the timing of the burst; for example, the approximation of two lips (mainly by raising the lower lip) and tight closure of lips are completed before the timing of the burst. In the segmentation of audio signals, a silent pause is an important criterion for the segmentation, while in articulatory segmentation, it is the vocal tract in its resting position that plays a crucial role. Figure 2 shows a typical example of the vocal tract in its resting position. The soft palate is lowered, the larynx is in a slightly lowered position, and there is no closure anywhere in the oral cavity of the vocal tract. When such a vocal tract in its resting position is observed before or after an utterance, the “start” tag is assigned to the pause preceding the utterance, and the “end “tag to the pause immediately following the utterance.

Most utterances can be segmented by these criteria, but there are some deviant cases to which these criteria cannot be applied. For example, at the end of an utterance, articulatory organs may preempt the articulation of the consonant or vowel at the beginning of the following utterance. For example, in the case where the current utterance is "bada" and the following utterance is "mahha", the lip closure for the [m] in “mahha” can be formed immediately after the end of “bada” without any intervening resting position. Note that this kind of articulatory anticipation cannot be interpreted as a simple coarticulation, since articulatory anticipation can be observed even in cases where two adjacent utterances are separated by a long acoustic pause.

Another commonly occurring deviation is related to the movement of the soft palate. Normally, the soft palate descends between adjacent utterances (not necessarily for breathing), but sometimes subsequent utterances are initiated without this descent being observed. A similar problem can be noted with the vertical movement of the larynx. The larynx usually rises slightly before the start of a speech and falls at the end of the speech, but sometimes the larynx remains elevated over two successive utterances.

Figure 2：Example of vocal tract in the resting position

As explained in 2.9, the [noRPp] and/or [noRPf] tags are assigned when these "deviant" articulatory movements cause uncertainty in the articulatory segmentation. Note, however, that these tags are assigned when uncertainty of segmentation remains even after referring to all criteria discussed above. For example, no uncertainty tag is applied to the example above ("bada" followed by "mahha" with preempted lip closure) if the end of the utterance can be confirmed by a pronounced lowering of the soft palate.

The three types of deviations described above are relatively common. However, other problems can also arise. One of them is the act of licking the lips. In the MRI recording, the speaker often lick the upper lip with the tongue tip to moisten the lips. Most speakers do this between sessions, but some lick their lips between adjacent utterances within a session. In such cases, the lip-licking behavior is recorded at the beginning or end of utterance. It is such utterances that are given the [LipLick] tag.

Finally, filled pauses can also cause problems. Some speakers utter filled pauses like “eh” or “ah” before the target utterance. If the utterance and the filled pause were separable in terms of articulatory motion (and in acoustic signal), we treated it as a normal utterance. In some cases, however, it was difficult to separate them. In such cases, the [fp] tag was applied together with the [noRPp] tag.

6. The making of MP4 data

6. 1 Movie without the dentition information

An MRI machine (MAGNETOM Prisma fit 3T, Siemens) settled at ATR-Promotions Brain Activity Imaging Center (ATR-BAIC), Inc. was used to record data for the rtMRIDB. This machine can capture 512 frames of MRI images (still images with a resolution of 256 x 256 pixels and a slice width of 10 mm) at an imaging speed at 14 fps for about 37 seconds and at 27 fps for about 19 seconds. The frame number is inserted at the bottom left of each frame during the process of movie compilation from the still images captured by the MRI. The frame number starts from zero, so it takes a value from 0 to 511. For more specialized information on the imaging conditions of the MRI movies in this database, please refer to the appendix of 前川ほか (2020).

Still image data captured by the MRI system is saved in DCM format, the standard format for MRI data. The speech signal uttered by the speaker is recorded separately using a DAT (digital tape recorder) with a sampling frequency of 44.1 kHz and a quantization precision of 16 bits.

The rtMRIDB video was created by dubbing the DAT audio onto an MP4 video consisting of JPG data converted from DCM data. The beginning of the operation noise of the MRI machine recorded in the DAT was used to synchronize the image and audio. However, since the frame rate of the MRI movie and the sampling frequency of the audio are significantly different, the possibility cannot be excluded that there is a slight discrepancy in the synchronization. The large operation noise of the MRI machine overlaid onto the speech signal was digitally reduced. In the case of rtMRIDB_v1, the noise-reduced audio had a slightly metallic sound. This problem was resolved in the rtMRIDB_v2, resulting in a highly natural speech sound. Finally, this database only publishes the MP4 video data, and does not include the DCM image data and DAT audio data.

6．2　Movie with the dentition information

RtMRIDB_v2 now supports the retrieval of movie data with dentition information. Unlike X-ray image, the image of bones does not appear in the MRI image. This is due to the basic principle of MRI imaging: materials that do not contain hydrogen atoms, such as bone, cannot be imaged by MRI. Basically, this is a major advantage of MRI data over the X-ray films because X-ray films clearly show the skull and dentition, and the researcher has to find out the contour of the tongue through these images, causing great difficulty in observation. However, the images of the incisors are sometimes necessary for phonetic research because some consonants

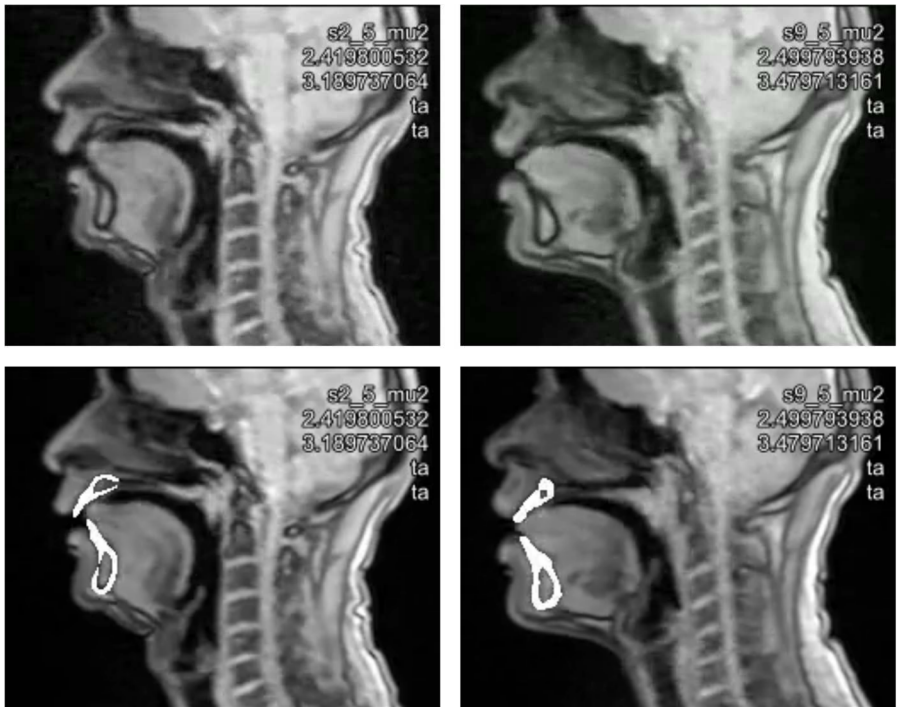


Figure 3: Example of the usefulness of dentition image

are classified in terms of their places of articulation with reference to the contact with the incisors (like dental and inter-dental consonants). We have developed a technique to insert the contour images of incisors onto rtMRI images through image processing (常盤ほか 2023).

Figure 3 compares the /ta/ (in the MU class) pronounced by speakers s2 and s9. The upper part of the figure shows a video clip from a previously published movie without dentition information, with the data for s2 on the left and s9 on the right. On the other hand, the lower part of the figure shows the video image from the movie on which the dentition information was inserted.

Observing only the video in the upper part of the figure, we can imagine that there is an articulatory difference between s2 and s9 /t/, which may be related to the presence or absence of the tongue contact with the upper incisors, but we cannot be sure if such articulation is actually taking place. On the other hand, comparing the images in the lower part of the figure, we can see that in the case of s2, the tongue apex is not in contact with the insicor but with the

gums, whereas in the case of s9, the contact between the tongue apex and the teeth forms a closure of the vocal tract. Using the terminology of articulatory phonetics, /t/ by s2 is a alveolar stop and /t/ by s9 is a dental stop, and it would be reasonable to transcribe the consonant by s9 as [t̪] if expressed using IPA's auxiliary symbol.

Of course, the articulatory difference in /t/ between s2 and s9 is not phonologically distinctive. However, this example suggests that it is dangerous to simply describe Japanese /t/ as an alveolar consonant. The fact that a certain number of native speakers of standard Japanese also articulate /t/ as a dental consonant as s9 is a finding that should be further explored in contrastive phonetics of Japanese and English.

7. Searching the rtMRIDB_v2

While rtMRIDB_v1 was released on the web as well as a Windows desktop application distributed by the National Institute of Informatics Speech Resource Consortium (NII-SRC http://research.nii.ac.jp/SRc/), rtMRIDB_v2 will be available only on the web for the time being. What follows is an instruction of how to use the search system available on the web. Basically, it is the same as the search system of rtMRIDB_v1, but there are some improvements.

7．1 Access to the database

The rtMRIDB_v2 contains 29,984 utterances by 25 speakers. Users can search this data by specifying various criteria and can see the retrieved data as MP4 video. Users can also download the video.

To search rtMRIDB_v2, access https://rtmridb.ninjal.ac.jp using a browser. Chrome is recommended as the browser, but other browsers can also be used. After successful access, the page shown in Figure 4 will appear. At first, an English page is displayed, but you can switch to Japanese by using the switch in the upper right corner of the screen. The following explanation is given using the English page.

7．2 Specifying search conditions

The menu for searching is displayed in the lower half of the screen in Figure 4. Although, usually, a search consists of multiple searching conditions, a single search condition is specified in three steps: "select attributes," "select operators," and "enter search string.

Let us now execute a simple search. Figure 5 specifies a condition to search for samples beginning with the phoneme sequence /sak/ in the Phoneme field. The user determines the search condition by specifying three elements: "attribute" (left menu, here ‘Phoneme’), "operator" (middle menu, here ‘begins with’), and "search string" (right menu, here ‘sak’). After completing the specifications shown in Figure 5, select the button just below the menu to select whether to search for samples at 14 fps or 27 fps. You can search for both, but if you choose both 14 and 27 fps, the search results cannot be converted to a video. Here we choose 14 fps. Then, click on the light blue magnifying glass icon just below the 14-fps button, and the search

results will immediately appear at the bottom of the screen. The search results show the values for all 20 fields for each sample.

Executing a search using the conditions in Figure 5 yields 259 hits for samples with Phoneme containing /sak/, but this is a little too many, so let's narrow down the search by adding additional search conditions. Click the green + icon to the right of the search condition to add a new condition.

Figure 4: Initial page of the search system on the web (English version on the top and Japanese version in the bottom)

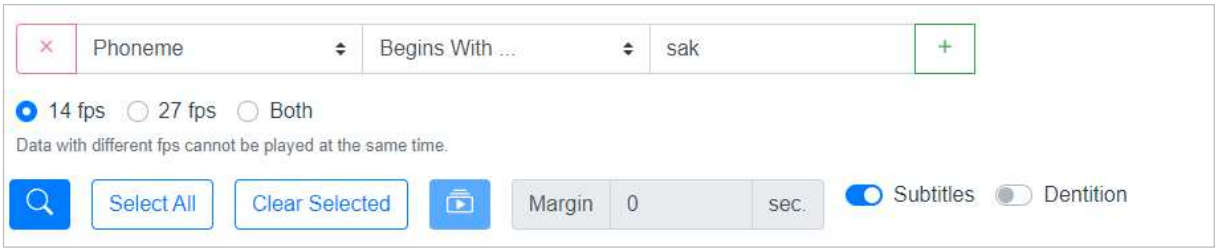


Figure 5: Specification of search (only one condition)

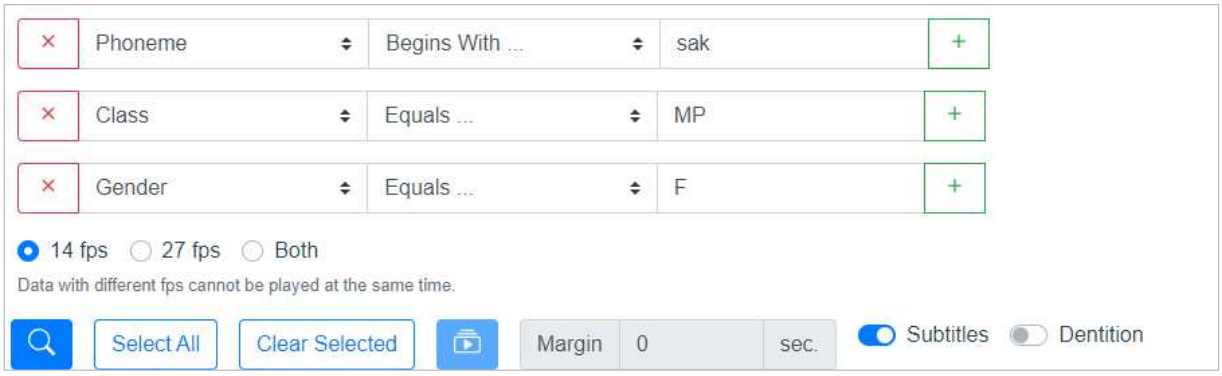


Figure 6: Specification of search (three conditions on different fields)

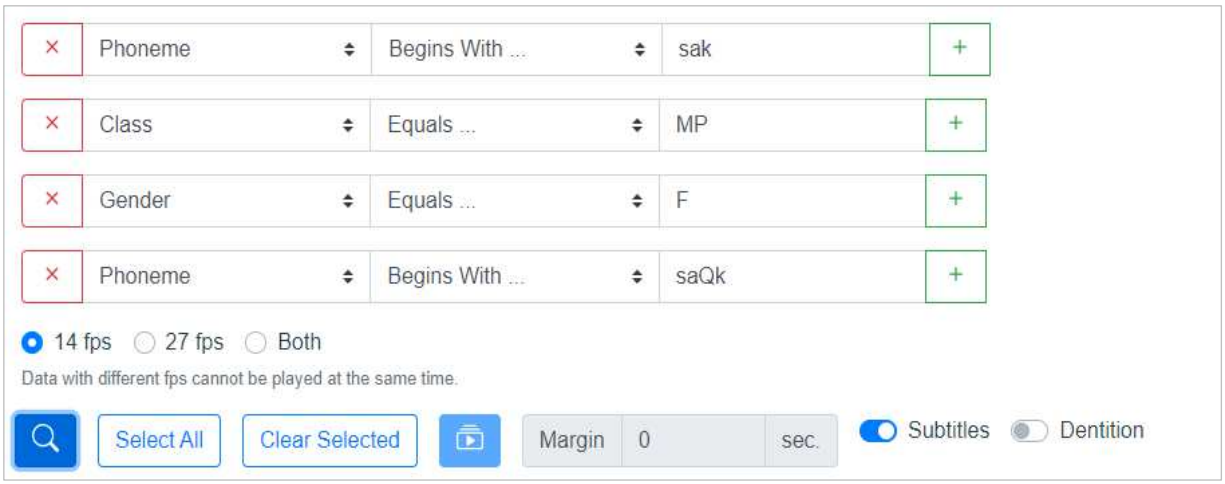
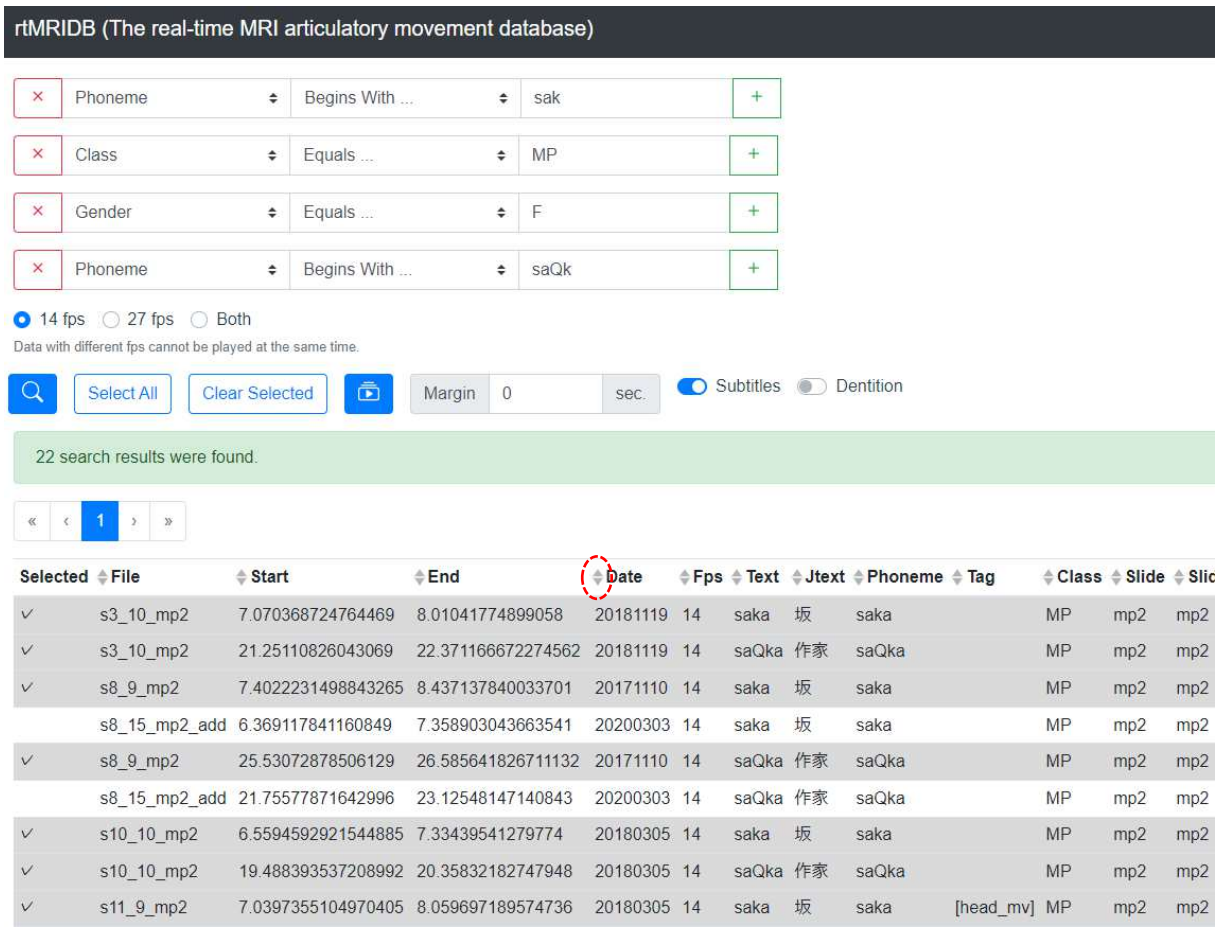


Figure 7 ：Specification of search (multiple conditions for the same attribute)

Figure 6 shows the search screen with the conditions "Class is equal to 'MP'" and "Gender is equal to 'F'" added to the conditions in Figure 5. When multiple conditions are specified in this way, the conditions concerning different fields are interpreted as AND (logical product) relationships, while the conditions for the same field are interpreted as OR (logical sum) relationships. Conditions shown in Figure 6 can be read "Phoneme begins with 'sak' AND Class is equal to 'MP' AND Gender is equal to 'F'". When the search is executed, the number of hit samples is 22.

Finally, Figure 7 shows an example of a search in which multiple criteria are specified for the same attribute (in this case, ‘Phoneme’). In Figure 7, "Phoneme begins with saQk" is added to the search condition. As mentioned earlier, multiple conditions for one field are interpreted as logical ORs. Accordingly, the logical operation can be expressed as (("Phoneme begins with ‘sak’" OR "Phoneme begins with ‘saQk’") AND "Class is equal to ‘MP’" AND "Gender is equal to ‘F’"), where parentheses are used to show the range of logical operators. Executing this search will result in 22 samples containing the two words /saka/ "slope" and /saQka/ "writer".

If you wish to remove a search condition, you can delete it by clicking on the red X mark to the left of the search condition. For example, if you delete the Gender criteria in Figure 7, the search result will increase to 71 samples.



Selected	File	Start	End	Date	Fps	Text	Jtext	Phoneme	Tag	Class	Slide	Slic
✓	s3_10_mp2	7.070368724764469	8.01041774899058	20181119	14	saka	坂	saka		MP	mp2	mp2
✓	s3_10_mp2	21.25110826043069	22.371166672274562	20181119	14	saQka	作家	saQka		MP	mp2	mp2
✓	s8_9_mp2	7.4022231498843265	8.437137840033701	20171110	14	saka	坂	saka		MP	mp2	mp2
	s8_15_mp2_add	6.369117841160849	7.358903043663541	20200303	14	saka	坂	saka		MP	mp2	mp2
✓	s8_9_mp2	25.53072878506129	26.585641826711132	20171110	14	saQka	作家	saQka		MP	mp2	mp2
	s8_15_mp2_add	21.75577871642996	23.12548147140843	20200303	14	saQka	作家	saQka		MP	mp2	mp2
✓	s10_10_mp2	6.5594592921544885	7.33439541279774	20180305	14	saka	坂	saka		MP	mp2	mp2
✓	s10_10_mp2	19.488393537208992	20.35832182747948	20180305	14	saQka	作家	saQka		MP	mp2	mp2
✓	s11_9_mp2	7.0397355104970405	8.059697189574736	20180305	14	saka	坂	saka	[head_mv]	MP	mp2	mp2

Figure 8 ：Selection of the search results (part of the screen displayed)

7．3 Editing, playing, and downloading the search results

Now, let's edit select the samples you need from the retrieved samples into a video file and play it back. Since editing takes time when the number of samples is large, let's work with 18 samples found by searching with the condition shown in Figure 7 (including "Gender equals to F").

When you have finished your search, click the "Select All" button to the right of the magnifying glass icon, and all samples will be highlighted in gray with a ✓ mark to the left of

them, indicating that they are selected. You can deselect individual samples by clicking on the ✓ mark in this state. Figure 8 shows the screen after all samples have been selected, with the fourth and sixth samples deselected.

In this state, clicking the light blue ▶ icon between "Clear selected" and "Margin" will start editing the video file, combining the selected samples in the order they appear into a single MP4 movie file. When editing is complete, the movie will be played back (Figure 9). The subtitle superimposed on the upper right corner of the movie indicates the 'File', 'Start', 'End', 'Text', and 'Phoneme' values of the sample being played back (red dashed rectangle in the figure; Also see 7.4).

The video can be repeatedly played and paused using the ▶ and = icons above the left end of the slider bar. You can select any part of the movie using the slider bar and start playing the movie. You can also click on the icons ⏮ and ⏭ placed in the white frame below the video to move frame by frame, allowing you to observe the articulatory movement in detail.

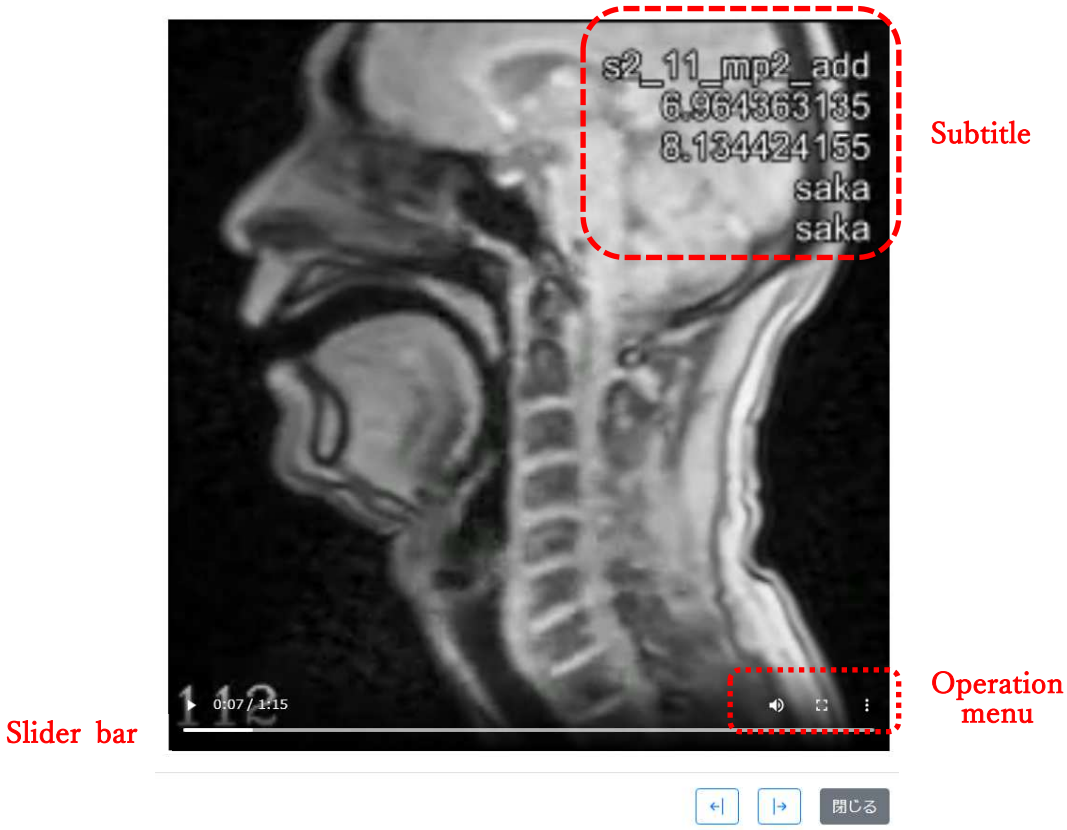


Figure 9 ：Playback screen of videos generated from search results

The three small icons in the lower right corner of the movie (enclosed by a red dotted rectangle in the figure) are, from left to right, the volume control, full screen, and file operation menus. The file operation menu has two options: "Download" and "Picture-in-picture. By clicking "Download," you can download the edited movie. Clicking on "Picture-in-picture" will

open a small window outside the window in which the search system is operating. The size of this window can be freely changed, but it does not have a frame-by-frame function.

7．4 Margins, Subtitles, and Dentition buttons

The search screen (see Figure 4) has buttons for "Margin", "Subtitle" and "Dentition". The "Margin" button expands the start and end time (Start, End) of the sample displayed on the search screen forward and backward. The maximum time range can be expanded up to 0.5 seconds in both directions. Note that for some samples, if the time range is expanded too much, there may be overlap between the sample immediately before and after. Normally, leave the value at 0.

“Subtitles" is turned on by default, and information on ‘File’, ‘Start’, ‘End’, ‘Text’, and ‘Phoneme’ is displayed in the upper right corner of the video. The frame number in the lower left corner of the video is not a subtitle, so it will not disappear even if you turn subtitles off.

Finally, "Dentition" is a switch to change the search target to the movie data with dentition information newly released in the rtMRIDB_v2. When this is turned on, the search will be performed only on movies with dentition information. Note it is impossible to search for movies with and without the dentition information at the same time.

7．5 Sorting function

A simple sorting function is implemented in the rtMRIDB search system. Look at Figure 8 again. At the top of the search results, the field names are displayed, and to the left of each field name is a small gray ▲▼ icon (red oval in the figure). By clicking on these, you can change the sort order (ascending/descending) of the samples. However, this sorting function is simple and does not support complex sorting over multiple columns. Therefore, it cannot be used for purposes such as rearranging samples into an arbitrary order for presentation purposes.

To solve this problem, a new function that can change the ranking of samples arbitrarily is available in rtMRIDB_v2. Although not shown in Figure 8, if you scroll the search result display screen to the right edge, you will see ↑ and ↓ symbols surrounded by circles in the right-ends of samples. By clicking on these symbols, you can move the position of the relevant sample up or down one by one. Although it is a tedious operation, this function allows you to sort the searched samples in any order.

7．6 Suggest function

As is clear from the description so far, knowledge of the contents of the database is required to use the search function of rtMRIDB. Knowing the possible values of the fields like 'Text', 'Jtext', 'Phoneme' is essential. A list of the values recorded in these fields is provided in Appendix 1 of this document. In addition, a suggestion function is implemented in the search system to reduce the burden of users.

This function predicts and displays subsequent candidates at the time when users enter a part of a character string. Suppose, for example, "Phoneme" and "Contains" are selected as an

attribute and operator. If two letters 'aQ' are entered as the initial part of the search string, the system will show a drop-dwon menu of candidate values of "Phoneme" including 'aQpa', 'aQsaku', 'baQda', 'baQku', 'baQda', 'baQnjari', 'baQta', 'gaQ', and more., and users can select the desired value from this menu.

7．7　Limitations of the system

The interface of the search system described so far was developed to provide an opportunity for users with limited computer experience to get in touch with the data of speech production. It does not provide a complete search function.

For example, as noted in 7.2, it is not possible to set multiple search criteria in a logical OR relationship across different fields. Also, the ability to further narrow down the results of a given search is not implemented. Therefore, it may not be possible to perfectly narrow down the sample desired by the user. If this is the case, please consider deselecting the unnecessary samples from an over-selected set of samples.

A problem can be pointed out regarding the playback of the search result movie. When playing back the movie in this system, there may be a delay of the video image to the audio signal, especially in the latter half of a movie of relatively large size. This problem is regarded to be related to several causes and may also depend on the performance of the computer being used. When this problem occurs, users may be able to avoid the problem in several ways. When you have a large number of search results, try to mitigate the problem by splitting the search results into multiple videos of modest size instead of editing one large video. Second, in case you playback a video of large size, you may be able to avoid the problem by jumping to the part of the video you wish to observe by use of the slider bar at the bottom of the screen before you playback the video. Lastly, you may be able to avoid the problem by downloading the movie and playing it back with a common video playback tool (such as Media player).

7．8　Data Viewing and Analysis Environment

For users who want to analyze the MP4 movies of the search results for purposes such as speech production studies, a data viewing and analysis environment specifically designed for the rtMRIDB data is available (“MRI Vuewer Ver. 2.0”. https://kikuchiken-waseda.github.io/mri-vuewer.ver2/). This is a browser-based application developed by Takuya Asai of Kikuchi Laboratory in the School of Human Sciences, Waseda University (See 浅井・菊池・前川 2021, 前川ほか 2021, Maekawa in press). It allows users to perform annotation along the time axis ("time series transcription") and annotation for a specific single frame ("frame transcription"). The former allows users to work with MRI images along with speech signals and sound spectrograms, while the latter allows users to perform various measurements on vocal tract information. Annotated results are recorded in the browser database, but it can also be exported in JSON, xlsx, and other data formats. Incidentally, the Start, End, Text, and Tag information

recorded in rtMRIDB have been annotated using this system.

Figure 10 shows an image of annotation work using the MRI Vuewer Ver.2. Below the three consecutive MRI images in the upper left corner are shown sound spectrogram and audio signals. The central MRI image corresponds to the time cursor on the spectrogram and audio signal, and the images to its left- and right-side are the frames that immediately preceding and following the central frame. A frame transcription window is shown in the foreground, in which the tongue contours are marked with measurement dots. MRI Vuewer Ver.2 and its documentation are available online at https://github.com/kikuchiken-waseda/mri-vuewer.ver2/wiki/Usage. Chrome is recommended as browser.

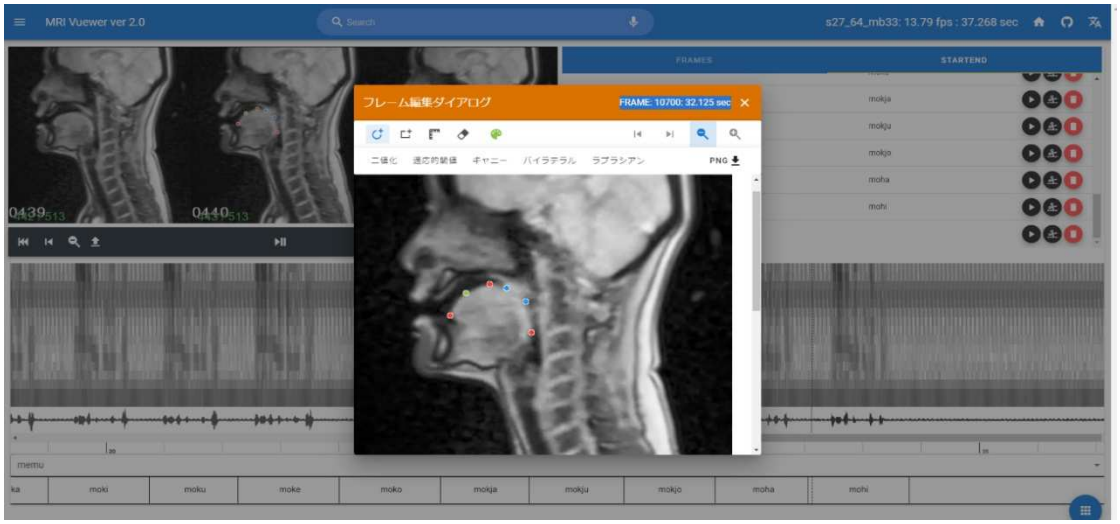


Figure 10: Working image of the MRI Vuewer Ver. 2

8. Restrictions on this database

This database (rtmridb_v2) is released under the Creative Commons License CC BY-NC-SA (Attribution-NonCommercial-ShareAlike). Please note that there is a possibility that this database will be substantially upgraded in the future.

Acknowledgments: We are deeply grateful to the 25 speakers who agreed to make their data publicly available. The development of this database was supported by two Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (17H02339, 20H01265), Discretion Grant from the Director of the National Institutes for the Humanities for the years 2022-2023, and budget from the Corpus Development Center, National Institute for Japanese Language and Linguistics. Data recording and processing/editing were supported by Yukiko Nota, Takuya Asai, Yuichi Ishimoto, Koichi Katsurada, Ryo Tanji, and the members of Professor Hironori Takemoto's laboratory at Chiba Institute of Technology. Real-time MRI video data were recorded at the Brain Activity Imaging Center (BAIC), ATR-

Promotions, Inc. with assistance from Drs. Shinobu Masaki, Ikuhiro Shimada, Nobukazu Nishikido, and Nobuyoshi Tanki. The database search environment was developed under contract to PicoLab Inc.

(31 March 2024)

For inquiries about this database, please contact Kikuo Maekawa at kikuoATninjal.ac.jp (Please replace 'AT' with an appropriate character.)

RERERENCES

Maekawa, K. (To appear). "Real-time MRI articulatory movement database and its application to articulatory phonetics". *Acoustical Science and Technology*, 2024.

浅井拓也・菊池英明・前川喜久雄(2021)「リアルタイム MRI 調音動画データの閲覧および解析環境の開発」言語資源活用ワークショップ 2021 発表論文集, 国立国語研究所コーパス開発センター.

常盤朔也・一岡昴輝・大谷幸聖・竹本浩典・前川喜久雄「頭頸部の正中矢状面における rtMRI 動画への歯列補填」日本音響学会第 149 回(2023 年春季)研究発表会, 2023.

前川喜久雄・西川賢哉・浅井拓也・能田由紀子・正木信夫・島田育廣・竹本浩典・北村達也・斎藤純男・籠宮隆之・石本祐一・菊池英明・藤本雅子・八木豊「リアルタイム MRI 動画日本語調音運動データベースの設計」(2020). 言語資源活用ワークショップ 2020 発表論文集, 国立国語研究所コーパス開発センター.

前川喜久雄(2021).「リアルタイム MRI 動画データベースプロジェクトの概要と調音音声学への応用」言語資源活用ワークショップ 2021 発表論文集, 国立国語研究所コーパス開発センター.

Appendix 1: List of utterance items in rtMRIDB_v2

This is a list of the utterance items recorded in the rtMRIDB_v2. Items are listed according to the 'Ser' values (See 2.13) for each slide (See 2.11). Each line shows the 'Jtext', 'text', and 'phoneme' of the utterance. The 'AddDate' at the end of each utterance indicates the date the item was added or deleted from the list in the format yyyymmdd. If this field is blank, it means that the item was included in the list from the beginning. A + at the end of 'AddDate' field indicates the addition of the item, while a - indicates the deletion. Only two items, ser=15 and 16 in the MP1 slide, were deleted. Note that the three classes MU, MP, and MB contain data from all speakers, while the other classes (NS, PL, SR, TT) may or may not contain data depending on the speaker. See Appendix 2 below for the differences in recording due to speakers.

SLIDE	SER	JTEXT	TEXT	PHONEME	ADDDATE
MU1	1	ア	a	a	
MU1	2	イ	i	i	
MU1	3	ウ	u	u	
MU1	4	エ	e	e	
MU1	5	オ	o	o	
MU1	6	ヤ	ja	ja	
MU1	7	ユ	ju	ju	
MU1	8	ヨ	jo	jo	
MU1	9	イェ	je	je	
MU1	10	カ	ka	ka	
MU1	11	キ	ki	ki	
MU1	12	ク	ku	ku	
MU1	13	ケ	ke	ke	
MU1	14	コ	ko	ko	
MU1	15	キャ	kja	kja	
MU1	16	キュ	kju	kju	
MU1	17	キョ	kjo	kjo	
MU1	18	キェ	kje	kje	
MU1	19	サ	sa	sa	
MU1	20	シ	si	si	
MU1	21	ス	su	su	
MU1	22	セ	se	se	
MU1	23	ソ	so	so	

MU1	24	シャ	sja	sja
MU1	25	シュ	sju	sju
MU1	26	ショ	sjo	sjo
MU1	27	シェ	sje	sje
MU2	1	タ	ta	ta
MU2	2	チ	ci	ci
MU2	3	ツ	cu	cu
MU2	4	テ	te	te
MU2	5	ト	to	to
MU2	6	チャ	cja	cja
MU2	7	チュ	cju	cju
MU2	8	チョ	cjo	cjo
MU2	9	チェ	cje	cje
MU2	10	ナ	na	na
MU2	11	ニ	ni	ni
MU2	12	ヌ	nu	nu
MU2	13	ネ	ne	ne
MU2	14	ノ	no	no
MU2	15	ニャ	nja	nja
MU2	16	ニュ	nju	nju
MU2	17	ニョ	njo	njo
MU2	18	ニェ	nje	nje
MU2	19	ハ	ha	ha
MU2	20	ヒ	hi	hi
MU2	21	フ	hu	hu
MU2	22	ヘ	he	he
MU2	23	ホ	ho	ho
MU2	24	ヒャ	hja	hja
MU2	25	ヒュ	hju	hju
MU2	26	ヒョ	hjo	hjo
MU2	27	ヒェ	hje	hje
MU3	1	パ	pa	pa
MU3	2	ピ	pi	pi
MU3	3	プ	pu	pu
MU3	4	ペ	pe	pe
MU3	5	ポ	po	po

MU3	6	ピャ	pja	pja
MU3	7	ピュ	pju	pju
MU3	8	ピョ	pjo	pjo
MU3	9	ピェ	pje	pje
MU3	10	ザ	za	za
MU3	11	ジ	zi	zi
MU3	12	ズ	zu	zu
MU3	13	ゼ	ze	ze
MU3	14	ゾ	zo	zo
MU3	15	ジャ	zja	zja
MU3	16	ジュ	zju	zju
MU3	17	ジョ	zjo	zjo
MU3	18	ジェ	zje	zje
MU3	19	マ	ma	ma
MU3	20	ミ	mi	mi
MU3	21	ム	MU	MU
MU3	22	メ	me	me
MU3	23	モ	mo	mo
MU3	24	ミャ	mja	mja
MU3	25	ミュ	mju	mju
MU3	26	ミョ	mjo	mjo
MU3	27	ミェ	mje	mje
MU4	1	ラ	ra	ra
MU4	2	リ	ri	ri
MU4	3	ル	ru	ru
MU4	4	レ	re	re
MU4	5	ロ	ro	ro
MU4	6	リャ	rja	rja
MU4	7	リュ	rju	rju
MU4	8	リョ	rjo	rjo
MU4	9	リェ	rje	rje
MU4	10	ワ	wa	wa
MU4	11	ウィ	wi	wi
MU4	12	ウェ	we	we
MU4	13	ウォ	wo	wo
MU4	14	ン	N	N

MU4	15	ファ	fa	fa	
MU4	16	フィ	fi	fi	
MU4	17	フェ	fe	fe	
MU4	18	フォ	fo	fo	
MU4	19	スィ	s_i	s_i	
MU4	20	ティ	ti	ti	
MU4	21	トゥ	tu	tu	
MU4	22	ズィ	z_i	z_i	
MU4	23	ディ	di	di	
MU4	24	ドゥ	du	du	
MU4	25	アイ	ai_1	ai	
MU4	26	アオ	ao_1	ao	
MU4	27	アエ	ae_1	ae	
MU4	28	アウ	au_1	au	
MU4	29	イア	ia	ia	
MU4	30	イウ	iu	iu	
MU4	31	イエ	ie	ie	
MU4	32	イオ	io	io	
MU5	1	アラ	ara	ara	20180720+
MU5	2	アリ	ari	ari	20180720+
MU5	3	アル	aru	aru	20180720+
MU5	4	アレ	are	are	20180720+
MU5	5	アロ	aro	aro	20180720+
MU5	6	イラ	ira	ira	20180720+
MU5	7	イリ	iri	iri	20180720+
MU5	8	イル	iru	iru	20180720+
MU5	9	イレ	ire	ire	20180720+
MU5	10	イロ	iro	iro	20180720+
MU5	11	オラ	ora	ora	20180720+
MU5	12	オリ	ori	ori	20180720+
MU5	13	オル	oru	oru	20180720+
MU5	14	オレ	ore	ore	20180720+
MU5	15	オロ	oro	oro	20180720+
MU5	16	オア	oa	oa	20180720+
MU5	17	オイ	oi	oi	20180720+
MU5	18	オウ	ou	ou	20180720+

MU5	19	オエ	oe	oe	20180720+
MU5	20	アー	aH	aH	20180720+
MU5	21	イー	iH	iH	20180720+
MU5	22	ウー	uH	uH	20180720+
MU5	23	エー	eH	eH	20180720+
MU5	24	オー	oH	oH	20180720+
MU5	25	バ	ba	ba	20190111+
MU5	26	ベ	be	be	20190111+
MU5	27	ボ	bo	bo	20190111+
MU5	28	フュ	fju	fju	20190527+
MU5	29	デュ	dju	dju	20190527+
MU5	30	テャ	tja	tja	20201001+
MU5	31	テュ	tju	tju	20201001+
MU5	32	テョ	tjo	tjo	20201001+
MP1	1	新案	siNaN	siNaN	
MP1	2	真円	siNeN	siNeN	
MP1	3	心音	siNoN	siNoN	
MP1	4	心因	siNiN	siNiN	
MP1	5	新刊	siNkaN	siNkaN	
MP1	6	真剣	siNkeN	siNkeN	
MP1	7	四温（しおん）	sioN	sioN	
MP1	8	資金	sikiN	sikiN	
MP1	9	辛酸（しんさん）	siNsaN	siNsaN	
MP1	10	浸水	siNsui	siNsui	
MP1	11	新鮮	siNseN	siNseN	
MP1	12	深層	siNsoH	siNsoH	
MP1	13	岩盤	gaNbaN	gaNbaN	
MP1	14	上海（シャンハイ）	sjaNhai	sjaNhai	20180223+
MP1	15	完備	kaNbi	kaNbi	20180223-
MP1	16	幹部	kaNbu	kaNbu	20180223-
MP1	17	簡便	kaNbeN	kaNbeN	
MP1	18	サンホセ	saNhose	saNhose	20180223+
MP1	19	冠婚	kaNkoN	kaNkoN	
MP1	20	英語	eHgo_1	eHgo	
MP1	21	カンハン	kaNhaN	kaNhaN	
MP1	22	カンヒン	kaNhiN	kaNhiN	

MP1	23	カンフン	kaNhuN	kaNhuN
MP1	24	簡約	kaNjaku	kaNjaku
MP1	25	肝油	kaNju	kaNju
MP1	26	寛容	kaNjoH	kaNjoH
MP1	27	観覧	kaNraN	kaNraN
MP2	1	カンヘン	kaNheN	kaNheN
MP2	2	カンホン	kaNhoN	kaNhoN
MP2	3	蜜柑（みかん）	mikaN	mikaN
MP2	4	麒麟（きりん）	kiriN	kiriN
MP2	5	坂	saka	saka
MP2	6	肩	kata	kata
MP2	7	浅く	asaku	asaku
MP2	8	アパ	apa	apa
MP2	9	マハ	maha	maha
MP2	10	升席（ますせき）	masu-seki	masuseki
MP2	11	西小（にししょう）	nisi-sjoH	nisisjoH
MP2	12	ガス栓	gasu-seN	gasuseN
MP2	13	ケヘ	kehe_1	kehe
MP2	14	ウェブ	webu	webu
MP2	15	バタ	bata	bata
MP2	16	バダ	bada	bada
MP2	17	マッハ	maQha	maQha
MP2	18	作家	saQka	saQka
MP2	19	買った	kaQta	kaQta
MP2	20	圧搾（あっさく）	aQsaku	aQsaku
MP2	21	愛	ai_2	ai
MP2	22	簡単	kaNtaN	kaNtaN
MP2	23	広東（カントン）	kaNtoN	kaNtoN
MP2	24	緩和（かんわ）	kaNwa	kaNwa
MP2	25	ガッ	gaQ	gaQ
MP3	1	アッパ	aQpa	aQpa
MP3	2	ケッヘ	keQhe	keQhe
MP3	3	バック	baQku	baQku
MP3	4	別途	beQto	beQto
MP3	5	末席	maQseki	maQseki
MP3	6	日照	niQsjoH	niQsjoH

MP3	7	合戦（かっせん）	kaQseN	kaQseN
MP3	8	バッタ	baQta	baQta
MP3	9	バッダ	baQda	baQda
MP3	10	グッズ	guQzu	guQzu
MP3	11	ウェッブ	weQbu	weQbu
MP3	12	バッグ	baQgu	baQgu
MP3	13	ベッド	beQdo	beQdo
MP3	14	甘芋（あまいも）	ama-imo	amaimo
MP3	15	秋	aki_1	aki
MP3	16	敗者	hai-sja	haisja
MP3	17	貝	kai_1	kai
MP3	18	紹介状	sjoHkai-zjoH	sjoHkaizjoH
MP3	19	帰る	kaeru_1	kaeru
MP3	20	孤児	kozi	kozi
MP3	21	青	ao_2	ao
MP3	22	パンチ	paNci	paNci
MP3	23	パンツ	paNcu	paNcu
MP3	24	失敗	siQpai	siQpai
MP3	25	サッ	saQ	saQ
MP4	1	里親	sato-oja	satooja
MP4	2	小鬼	ko-oni	kooni
MP4	3	甘い藻	amai-mo	amaimo
MP4	4	空き	aki_2	aki
MP4	5	歯医者	ha-isja	haisja
MP4	6	下位	kai_2	kai
MP4	7	消化異常	sjoHka-izjoH	sjoHkaizjoH
MP4	8	飼える	kaeru_2	kaeru
MP4	9	工事	koHzi	koHzi
MP4	10	砂糖屋	satoH-ja	satoHja
MP4	11	高二	koHni	koHni
MP4	12	新聞	siNbuN	siNbuN
MP4	13	風船	huHseN	huHseN
MP4	14	ゴッホ	goQho	goQho
MP4	15	安全	aNzeN	aNzeN
MP4	16	系統	keHtoH	keHtoH
MP4	17	毛糸	ke-ito	ke-ito

MP4	18	夜景	jakeH	jakeH	
MP4	19	焼板（やけいた）	jakeita	jakeita	
MP4	20	Ａ５	eHgo_2	eHgo	
MP4	21	会う	au_2	au	
MP4	22	会え	ae_2	ae	
MP4	23	案内	aNnai	aNnai	
MP4	24	欠品	keQpiN	keQpiN	
MP4	25	パッ	paQ	paQ	
MP5	1	カラン	karaN	karaN	20181106+
MP5	2	カリン	kariN	kariN	20181106+
MP5	3	イラン	iraN	iraN	20181106+
MP5	4	ピッ	piQ	piQ	20181106+
MP5	5	カッ	kaQ	kaQ	20181106+
MP5	6	キッ	kiQ	kiQ	20181106+
MP5	7	ペッ	peQ	peQ	20181106+
MP5	8	ポッ	poQ	poQ	20181106+
MP5	9	シッ	siQ	siQ	20181106+
MP5	10	プッ	puQ	puQ	20181106+
MP5	11	ソッ	soQ	soQ	20181106+
MP5	12	スッ	suQ	suQ	20181106+
MP5	13	コッ	koQ	koQ	20181106+
MP5	14	ケッ	keQ	keQ	20181106+
MP5	15	セッ	seQ	seQ	20181106+
MP5	16	クッ	kuQ	kuQ	20181106+
MP5	17	近隣（きんりん）	kiNriN	kiNriN	20181106+
MP5	18	官林（かんりん）	kaNriN	kaNriN	20181106+
MP5	19	淫乱（いんらん）	iNraN	iNraN	20181106+
MP5	20	完全	kaNzeN	kaNzeN	20181106+
MP5	21	唖然（あぜん）	azeN	azeN	20181106+
MP5	22	果然（かぜん）	kazeN	kazeN	20181106+
MP5	23	イ段	idaN	idaN	20181106+
MP5	24	花壇	kadaN	kadaN	20181106+
MP5	25	アンフェア	aNfea	aNfea	20190111+
MP6	1	カンフー	kaNhuH	kaNhuH	20190111+
MP6	2	インフォーム	iNfoHmu	iNfoHmu	20190111+
MP6	3	アンペア	aNpea	aNpea	20190111+

MP6	4	画商	gasjoH	gasjoH	20190111+
MP6	5	完封	kaNpuH	kaNpuH	20190111+
MP6	6	決勝	keQsjoH	keQsjoH	20190111+
MP6	7	鑑賞	kaNsjoH	kaNsjoH	20190111+
MP6	8	検証	keNsjoH	keNsjoH	20190111+
MP6	9	合掌	gaQsjoH	gaQsjoH	20190111+
MP6	10	化粧	kesjoH	kesjoH	20190111+
MP6	11	山陰	saNiN	saNiN	20190111+
MP6	12	婚姻	koNiN	koNiN	20190111+
MP6	13	牽引	keNiN	keNiN	20190111+
MP6	14	雰囲気	huNiki	huNiki	20190111+
MP6	15	金運	kiNuN	kiNuN	20190111+
MP6	16	甲板	kaNpaN	kaNpaN	20190111+
MP6	17	カラカラ	karakara	karakara	20190111+
MP6	18	カンラカラ	kaNrakara	kaNrakara	20190111+
MP6	19	カッラカラ	kaQrakara	kaQrakara	20190111+
MP6	20	バニャリ	banjari	banjari	20190208+
MP6	21	バンニャリ	baNnjari	baNnjari	20190208+
MP6	22	バッニャリ	baQnjari	baQnjari	20190208+
MP6	23	フュージョン	fjuHzjoN	fjuHzjoN	20190527+
MP6	24	デューク	djuHku	djuHku	20190527+
MP7	1	艱難	kaNnaN	kaNnaN	20201228+
MP7	2	堪忍	kaNniN	kaNniN	20201228+
MP7	3	カンヌ	kaNnu	kaNnu	20201228+
MP7	4	三年	saNneN	saNneN	20201228+
MP7	5	観音	kaNnoN	kaNnoN	20201228+
MP7	6	散漫	saNmaN	saNmaN	20201228+
MP7	7	官民	kaNmiN	kaNmiN	20201228+
MP7	8	任務	niNmu	niNmu	20201228+
MP7	9	三面	saNmeN	saNmeN	20201228+
MP7	10	審問	siNmoN	siNmoN	20201228+
MP7	11	心眼	siNgaN	siNgaN	20201228+
MP7	12	震源	siNgeN	siNgeN	20201228+
MP7	13	呻吟	siNgiN	siNgiN	20201228+
MP7	14	真言	siNgoN	siNgoN	20201228+
MP7	15	進軍	siNguN	siNguN	20201228+

MP7	16	三本	saNboN	saNboN	20201228+
MP7	17	信任	siNniN	siNniN	20201228+
MP7	18	信念	siNneN	siNneN	20201228+
MP7	19	金満	kiNmaN	kiNmaN	20201228+
MP7	20	山門	saNmoN	saNmoN	20201228+
MP7	21	宦官	kaNgaN	kaNgaN	20201228+
MP7	22	勧銀	kaNgiN	kaNgiN	20201228+
MP7	23	サンゴ	saNgo	saNgo	20201228+
MP7	24	臣民	siNmiN	siNmiN	20201228+
MP7	25	新免	siNmeN	siNmeN	20201228+
MP8	1	三絃	saNgeN	saNgeN	20201228+
MP8	2	三軍	saNguN	saNguN	20201228+
MP8	3	割烹	kaQpoH	kaQpoH	20201228+
MP8	4	サッフォー	saQfoH	saQfoH	20201228+
MP8	5	ン	N	N	MU4 と同一
MP8	6	カップ	kaQpu	kaQpu	20201228+
MP8	7	スタッフ	sutaQfu	sutaQfu	20201228+
MP8	8	他人（たにん）	taniN	taniN	20201228+
MP8	9	カヌー	kanuH	kanuH	20201228+
MP8	10	カノン	kanoN	kanoN	20201228+
MP8	11	我慢	gamaN	gamaN	20201228+
MP8	12	仮眠	kamiN	kamiN	20201228+
MP8	13	指紋	simoN	simoN	20201228+
MP8	14	志願	sigaN	sigaN	20201228+
MP8	15	市民	simiN	simiN	20201228+
MP8	16	資源	sigeN	sigeN	20201228+
MP8	17	詩吟	sigiN	sigiN	20201228+
MP8	18	泡（あわ）	awa	awa	202110928+
MP8	19	永久（とわ）	towa	Towa	20210928+
MP8	20	桑（くわ）	kuwa	kuwa	20210928+
MP8	21	庭（にわ）	niwa	niwa	20210928+
MP8	22	岩（いわ）	Iwa	iwa	20210928+
MP8	23	ティーカップ	tiHkaQpu	tiHkaQpu	20210928+
NS1	1	『北風と太陽』	_northwind		20190527+
NS2	1	『膨張する宇宙』	_space		20191210+
MB1	1	カカ	kaka	kaka	

MB1	2	カキ	kaki	kaki
MB1	3	カク	kaku	kaku
MB1	4	カケ	kake	kake
MB1	5	カコ	kako	kako
MB1	6	カキャ	kakja	kakja
MB1	7	カキュ	kakju	kakju
MB1	8	カキョ	kakjo	kakjo
MB1	9	カハ	kaha	kaha
MB1	10	カヒ	kahi	kahi
MB1	11	カフ	kahu	kahu
MB1	12	カヘ	kahe	kahe
MB1	13	カホ	kaho	kaho
MB1	14	カサ	kasa	kasa
MB1	15	カシ	kasi	kasi
MB1	16	カス	kasu	kasu
MB1	17	カセ	kase	kase
MB1	18	カソ	kaso	kaso
MB1	19	カシャ	kasja	kasja
MB1	20	カシュ	kasju	kasju
MB2	1	カショ	kasjo	kasjo
MB2	2	カマ	kama	kama
MB2	3	カミ	kami	kami
MB2	4	カム	kamu	kamu
MB2	5	カメ	kame	kame
MB2	6	カモ	kamo	kamo
MB2	7	キカ	kika	kika
MB2	8	キキ	kiki	kiki
MB2	9	キク	kiku	kiku
MB2	10	キケ	kike	kike
MB2	11	キコ	kiko	kiko
MB2	12	キキャ	kikja	kikja
MB2	13	キキュ	kikju	kikju
MB2	14	キキョ	kikjo	kikjo
MB2	15	キハ	kiha	kiha
MB2	16	キヒ	kihi	kihi
MB2	17	キフ	kihu	kihu

MB2	18	キヘ	kihe	kihe
MB2	19	キホ	kiho	kiho
MB2	20	キサ	kisa	kisa
MB3	1	キシ	kisi	kisi
MB3	2	キス	kisu	kisu
MB3	3	キセ	kise	kise
MB3	4	キソ	kiso	kiso
MB3	5	キシャ	kisja	kisja
MB3	6	キシュ	kisju	kisju
MB3	7	キショ	kisjo	kisjo
MB3	8	キマ	kima	kima
MB3	9	キミ	kimi	kimi
MB3	10	キム	kimu	kimu
MB3	11	キメ	kime	kime
MB3	12	キモ	kimo	kimo
MB3	13	クカ	kuka	kuka
MB3	14	クキ	kuki	kuki
MB3	15	ククク	kuku	kuku
MB3	16	クケ	kuke	kuke
MB3	17	クコ	kuko	kuko
MB3	18	クキャ	kukja	kukja
MB3	19	クキュ	kukju	kukju
MB3	20	クキョ	kukjo	kukjo
MB4	1	クハ	kuha	kuha
MB4	2	クヒ	kuhi	kuhi
MB4	3	クフ	kuhu	kuhu
MB4	4	クヘ	kuhe	kuhe
MB4	5	クホ	kuho	kuho
MB4	6	クサ	kusa	kusa
MB4	7	クシ	kusi	kusi
MB4	8	クス	kusu	kusu
MB4	9	クセ	kuse	kuse
MB4	10	クソ	kuso	kuso
MB4	11	クシャ	kusja	kusja
MB4	12	クシュ	kusju	kusju
MB4	13	クショ	kusjo	kusjo

MB4	14	クマ	kuma	kuma
MB4	15	クミ	kumi	kumi
MB4	16	クム	kumu	kumu
MB4	17	クメ	kume	kume
MB4	18	クモ	kumo	kumo
MB4	19	ケカ	keka	keka
MB4	20	ケキ	keki	keki
MB5	1	ケク	keku	keku
MB5	2	ケケ	keke	keke
MB5	3	ケコ	keko	keko
MB5	4	ケキャ	kekja	kekja
MB5	5	ケキュ	kekju	kekju
MB5	6	ケキョ	kekjo	kekjo
MB5	7	ケハ	keha	keha
MB5	8	ケヒ	kehi	kehi
MB5	9	ケフ	kehu	kehu
MB5	10	ケヘ	kehe	kehe
MB5	11	ケホ	keho	keho
MB5	12	ケサ	kesa	kesa
MB5	13	ケシ	kesi	kesi
MB5	14	ケス	kesu	kesu
MB5	15	ケセ	kese	kese
MB5	16	ケソ	keso	keso
MB5	17	ケシャ	kesja	kesja
MB5	18	ケシュ	kesju	kesju
MB5	19	ケショ	kesjo	kesjo
MB5	20	ケマ	kema	kema
MB6	1	ケミ	kemi	kemi
MB6	2	ケム	kemu	kemu
MB6	3	ケメ	keme	keme
MB6	4	ケモ	kemo	kemo
MB6	5	コカ	koka	koka
MB6	6	コキ	koki	koki
MB6	7	コク	koku	koku
MB6	8	コケ	koke	koke
MB6	9	ココ	koko	koko

MB6	10	コキャ	kokja	kokja
MB6	11	コキュ	kokju	kokju
MB6	12	コキョ	kokjo	kokjo
MB6	13	コハ	koha	koha
MB6	14	コヒ	kohi	kohi
MB6	15	コフ	kohu	kohu
MB6	16	コヘ	kohe	kohe
MB6	17	コホ	koho	koho
MB6	18	コサ	kosa	kosa
MB6	19	コシ	kosi	kosi
MB6	20	コス	kosu	kosu
MB7	1	コセ	kose	kose
MB7	2	コソ	koso	koso
MB7	3	コシャ	kosja	kosja
MB7	4	コシュ	kosju	kosju
MB7	5	コショ	kosjo	kosjo
MB7	6	コマ	koma	koma
MB7	7	コミ	komi	komi
MB7	8	コム	komu	komu
MB7	9	コメ	kome	kome
MB7	10	コモ	komo	komo
MB7	11	キャカ	kjaka	kjaka
MB7	12	キャキ	kjaki	kjaki
MB7	13	キャク	kjaku	kjaku
MB7	14	キャケ	kjake	kjake
MB7	15	キャコ	kjako	kjako
MB7	16	キャキャ	kjakja	kjakja
MB7	17	キャキュ	kjakju	kjakju
MB7	18	キャキョ	kjakjo	kjakjo
MB7	19	キャハ	kjaha	kjaha
MB7	20	キャヒ	kjahi	kjahi
MB8	1	キャフ	kjahu	kjahu
MB8	2	キャヘ	kjahe	kjahe
MB8	3	キャホ	kjaho	kjaho
MB8	4	キャサ	kjasa	kjasa
MB8	5	キャシ	kjasi	kjasi

MB8	6	キャス	kjasu	kjasu
MB8	7	キャセ	kjase	kjase
MB8	8	キャソ	kjaso	kjaso
MB8	9	キャシャ	kjasja	kjasja
MB8	10	キャシュ	kjasju	kjasju
MB8	11	キャショ	kjasjo	kjasjo
MB8	12	キャマ	kjama	kjama
MB8	13	キャミ	kjami	kjami
MB8	14	キャム	kjamu	kjamu
MB8	15	キャメ	kjame	kjame
MB8	16	キャモ	kjamo	kjamo
MB8	17	キュカ	kjuka	kjuka
MB8	18	キュキ	kjuki	kjuki
MB8	19	キュク	kjuku	kjuku
MB8	20	キュケ	kjuke	kjuke
MB9	1	キュコ	kjuko	kjuko
MB9	2	キュキャ	kjukja	kjukja
MB9	3	キュキュ	kjukju	kjukju
MB9	4	キュキョ	kjukjo	kjukjo
MB9	5	キュハ	kjuha	kjuha
MB9	6	キュヒ	kjuhi	kjuhi
MB9	7	キュフ	kjuhu	kjuhu
MB9	8	キュヘ	kjuhe	kjuhe
MB9	9	キュホ	kjuho	kjuho
MB9	10	キュサ	kjusa	kjusa
MB9	11	キュシ	kjusi	kjusi
MB9	12	キュス	kjusu	kjusu
MB9	13	キュセ	kjuse	kjuse
MB9	14	キュソ	kjuso	kjuso
MB9	15	キュシャ	kjusja	kjusja
MB9	16	キュシュ	kjusju	kjusju
MB9	17	キュショ	kjusjo	kjusjo
MB9	18	キュマ	kjuma	kjuma
MB9	19	キュミ	kjumi	kjumi
MB9	20	キュム	kjumu	kjumu
MB10	1	キュメ	kjume	kjume

MB10	2	キュモ	kjumo	kjumo
MB10	3	キョカ	kjoka	kjoka
MB10	4	キョキ	kjoki	kjoki
MB10	5	キョク	kjoku	kjoku
MB10	6	キョケ	kjoke	kjoke
MB10	7	キョコ	kjoko	kjoko
MB10	8	キョキャ	kjokja	kjokja
MB10	9	キョキュ	kjokju	kjokju
MB10	10	キョキョ	kjokjo	kjokjo
MB10	11	キョハ	kjoha	kjoha
MB10	12	キョヒ	kjohi	kjohi
MB10	13	キョフ	kjohu	kjohu
MB10	14	キョヘ	kjohe	kjohe
MB10	15	キョホ	kjoho	kjoho
MB10	16	キョサ	kjosa	kjosa
MB10	17	キョシ	kjosi	kjosi
MB10	18	キョス	kjosu	kjosu
MB10	19	キョセ	kjose	kjose
MB10	20	キョソ	kjoso	kjoso
MB11	1	キョシャ	kjosja	kjosja
MB11	2	キョシュ	kjosju	kjosju
MB11	3	キョショ	kjosjo	kjosjo
MB11	4	キョマ	kjoma	kjoma
MB11	5	キョミ	kjomi	kjomi
MB11	6	キョム	kjomu	kjomu
MB11	7	キョメ	kjome	kjome
MB11	8	キョモ	kjomo	kjomo
MB11	9	ハカ	haka	haka
MB11	10	ハキ	haki	haki
MB11	11	ハク	haku	haku
MB11	12	ハケ	hake	hake
MB11	13	ハコ	hako	hako
MB11	14	ハキャ	hakja	hakja
MB11	15	ハキュ	hakju	hakju
MB11	16	ハキョ	hakjo	hakjo
MB11	17	ハハ	haha	haha

MB11	18	ハヒ	hahi	hahi
MB11	19	ハフ	hahu	hahu
MB11	20	ハヘ	hahe	hahe
MB12	1	ハホ	haho	haho
MB12	2	ハサ	hasa	hasa
MB12	3	ハシ	hasi	hasi
MB12	4	ハス	hasu	hasu
MB12	5	ハセ	hase	hase
MB12	6	ハソ	haso	haso
MB12	7	ハシャ	hasja	hasja
MB12	8	ハシュ	hasju	hasju
MB12	9	ハショ	hasjo	hasjo
MB12	10	ハマ	hama	hama
MB12	11	ハミ	hami	hami
MB12	12	ハム	hamu	hamu
MB12	13	ハメ	hame	hame
MB12	14	ハモ	hamo	hamo
MB12	15	ヒカ	hika	hika
MB12	16	ヒキ	hiki	hiki
MB12	17	ヒク	hiku	hiku
MB12	18	ヒケ	hike	hike
MB12	19	ヒコ	hiko	hiko
MB12	20	ヒキャ	hikja	hikja
MB13	1	ヒキュ	hikju	hikju
MB13	2	ヒキョ	hikjo	hikjo
MB13	3	ヒハ	hiha	hiha
MB13	4	ヒヒ	hihi	hihi
MB13	5	ヒフ	hihu	hihu
MB13	6	ヒヘ	hihe	hihe
MB13	7	ヒホ	hiho	hiho
MB13	8	ヒサ	hisa	hisa
MB13	9	ヒシ	hisi	hisi
MB13	10	ヒス	hisu	hisu
MB13	11	ヒセ	hise	hise
MB13	12	ヒソ	hiso	hiso
MB13	13	ヒシャ	hisja	hisja

MB13	14	ヒシュ	hisju	hisju
MB13	15	ヒショ	hisjo	hisjo
MB13	16	ヒマ	hima	hima
MB13	17	ヒミ	himi	himi
MB13	18	ヒム	himu	himu
MB13	19	ヒメ	hime	hime
MB13	20	ヒモ	himo	himo
MB14	1	フカ	huka	huka
MB14	2	フキ	huki	huki
MB14	3	フク	huku	huku
MB14	4	フケ	huke	huke
MB14	5	フコ	huko	huko
MB14	6	フキャ	hukja	hukja
MB14	7	フキュ	hukju	hukju
MB14	8	フキョ	hukjo	hukjo
MB14	9	フハ	huha	huha
MB14	10	フヒ	huhi	huhi
MB14	11	フフ	huhu	huhu
MB14	12	フヘ	huhe	huhe
MB14	13	フホ	huho	huho
MB14	14	フサ	husa	husa
MB14	15	フシ	husi	husi
MB14	16	フス	husu	husu
MB14	17	フセ	huse	huse
MB14	18	フソ	huso	huso
MB14	19	フシャ	husja	husja
MB14	20	フシュ	husju	husju
MB15	1	フショ	husjo	husjo
MB15	2	フマ	huma	huma
MB15	3	フミ	humi	humi
MB15	4	フム	humu	humu
MB15	5	フメ	hume	hume
MB15	6	フモ	humo	humo
MB15	7	ヘカ	heka	heka
MB15	8	ヘキ	heki	heki
MB15	9	ヘク	heku	heku

MB15	10	ヘケ	heke	heke
MB15	11	ヘコ	heko	heko
MB15	12	ヘキャ	hekja	hekja
MB15	13	ヘキュ	hekju	hekju
MB15	14	ヘキョ	hekjo	hekjo
MB15	15	ヘハ	heha	heha
MB15	16	ヘヒ	hehi	hehi
MB15	17	ヘフ	hehu	hehu
MB15	18	ヘヘ	hehe	hehe
MB15	19	ヘホ	heho	heho
MB15	20	ヘサ	hesa	hesa
MB16	1	ヘシ	hesi	hesi
MB16	2	ヘス	hesu	hesu
MB16	3	ヘセ	hese	hese
MB16	4	ヘソ	heso	heso
MB16	5	ヘシャ	hesja	hesja
MB16	6	ヘシュ	hesju	hesju
MB16	7	ヘショ	hesjo	hesjo
MB16	8	ヘマ	hema	hema
MB16	9	ヘミ	hemi	hemi
MB16	10	ヘム	hemu	hemu
MB16	11	ヘメ	heme	heme
MB16	12	ヘモ	hemo	hemo
MB16	13	ホカ	hoka	hoka
MB16	14	ホキ	hoki	hoki
MB16	15	ホク	hoku	hoku
MB16	16	ホケ	hoke	hoke
MB16	17	ホコ	hoko	hoko
MB16	18	ホキャ	hokja	hokja
MB16	19	ホキュ	hokju	hokju
MB16	20	ホキョ	hokjo	hokjo
MB17	1	ホハ	hoha	hoha
MB17	2	ホヒ	hohi	hohi
MB17	3	ホフ	hohu	hohu
MB17	4	ホヘ	hohe	hohe
MB17	5	ホホ	hoho	hoho

MB17	6	ホサ	hosa	hosa
MB17	7	ホシ	hosi	hosi
MB17	8	ホス	hosu	hosu
MB17	9	ホセ	hose	hose
MB17	10	ホソ	hoso	hoso
MB17	11	ホシャ	hosja	hosja
MB17	12	ホシュ	hosju	hosju
MB17	13	ホショ	hosjo	hosjo
MB17	14	ホマ	homa	homa
MB17	15	ホミ	homi	homi
MB17	16	ホム	homu	homu
MB17	17	ホメ	home	home
MB17	18	ホモ	homo	homo
MB17	19	サカ	saka	saka
MB17	20	サキ	saki	saki
MB18	1	サク	saku	saku
MB18	2	サケ	sake	sake
MB18	3	サコ	sako	sako
MB18	4	サキャ	sakja	sakja
MB18	5	サキュ	sakju	sakju
MB18	6	サキョ	sakjo	sakjo
MB18	7	サハ	saha	saha
MB18	8	サヒ	sahi	sahi
MB18	9	サフ	sahu	sahu
MB18	10	サヘ	sahe	sahe
MB18	11	サホ	saho	saho
MB18	12	ササ	sasa	sasa
MB18	13	サシ	sasi	sasi
MB18	14	サス	sasu	sasu
MB18	15	サセ	sase	sase
MB18	16	サソ	saso	saso
MB18	17	サシャ	sasja	sasja
MB18	18	サシュ	sasju	sasju
MB18	19	サショ	sasjo	sasjo
MB18	20	サマ	sama	sama
MB19	1	サミ	sami	sami

MB19	2	サム	samu	samu
MB19	3	サメ	same	same
MB19	4	サモ	samo	samo
MB19	5	シカ	sika	sika
MB19	6	シキ	siki	siki
MB19	7	シク	siku	siku
MB19	8	シケ	sike	sike
MB19	9	シコ	siko	siko
MB19	10	シキャ	sikja	sikja
MB19	11	シキュ	sikju	sikju
MB19	12	シキョ	sikjo	sikjo
MB19	13	シハ	siha	siha
MB19	14	シヒ	sihi	sihi
MB19	15	シフ	sihu	sihu
MB19	16	シヘ	sihe	sihe
MB19	17	シホ	siho	siho
MB19	18	シサ	sisa	sisa
MB19	19	シシ	sisi	sisi
MB19	20	シス	sisu	sisu
MB20	1	シセ	sise	sise
MB20	2	シソ	siso	siso
MB20	3	シシャ	sisja	sisja
MB20	4	シシュ	sisju	sisju
MB20	5	シショ	sisjo	sisjo
MB20	6	シマ	sima	sima
MB20	7	シミ	simi	simi
MB20	8	シム	simu	simu
MB20	9	シメ	sime	sime
MB20	10	シモ	simo	simo
MB20	11	スカ	suka	suka
MB20	12	スキ	suki	suki
MB20	13	スク	suku	suku
MB20	14	スケ	suke	suke
MB20	15	スコ	suko	suko
MB20	16	スキャ	sukja	sukja
MB20	17	スキュ	sukju	sukju

MB20	18	スキョ	sukjo	sukjo
MB20	19	スハ	suha	suha
MB20	20	スヒ	suhi	suhi
MB21	1	スフ	suhu	suhu
MB21	2	スヘ	suhe	suhe
MB21	3	スホ	suho	suho
MB21	4	スサ	susa	susa
MB21	5	スシ	susi	susi
MB21	6	ススス	susu	susu
MB21	7	スセ	suse	suse
MB21	8	スソ	suso	suso
MB21	9	スシャ	susja	susja
MB21	10	スシュ	susju	susju
MB21	11	スショ	susjo	susjo
MB21	12	スマ	suma	suma
MB21	13	スミ	sumi	sumi
MB21	14	スム	sumu	sumu
MB21	15	スメ	sume	sume
MB21	16	スモ	sumo	sumo
MB21	17	セカ	seka	seka
MB21	18	セキ	seki	seki
MB21	19	セク	seku	seku
MB21	20	セケ	seke	seke
MB22	1	セコ	seko	seko
MB22	2	セキャ	sekja	sekja
MB22	3	セキュ	sekju	sekju
MB22	4	セキョ	sekjo	sekjo
MB22	5	セハ	seha	seha
MB22	6	セヒ	sehi	sehi
MB22	7	セフ	sehu	sehu
MB22	8	セヘ	sehe	sehe
MB22	9	セホ	seho	seho
MB22	10	セサ	sesa	sesa
MB22	11	セシ	sesi	sesi
MB22	12	セス	sesu	sesu
MB22	13	セセ	sese	sese

MB22	14	セソ	seso	seso
MB22	15	セシャ	sesja	sesja
MB22	16	セシュ	sesju	sesju
MB22	17	セショ	sesjo	sesjo
MB22	18	セマ	sema	sema
MB22	19	セミ	semi	semi
MB22	20	セム	semu	semu
MB23	1	セメ	seme	seme
MB23	2	セモ	semo	semo
MB23	3	ソカ	soka	soka
MB23	4	ソキ	soki	soki
MB23	5	ソク	soku	soku
MB23	6	ソケ	soke	soke
MB23	7	ソコ	soko	soko
MB23	8	ソキャ	sokja	sokja
MB23	9	ソキュ	sokju	sokju
MB23	10	ソキョ	sokjo	sokjo
MB23	11	ソハ	soha	soha
MB23	12	ソヒ	sohi	sohi
MB23	13	ソフ	sohu	sohu
MB23	14	ソヘ	sohe	sohe
MB23	15	ソホ	soho	soho
MB23	16	ソサ	sosa	sosa
MB23	17	ソシ	sosi	sosi
MB23	18	ソス	sosu	sosu
MB23	19	ソセ	sose	sose
MB23	20	ソソ	soso	soso
MB24	1	ソシャ	sosja	sosja
MB24	2	ソシュ	sosju	sosju
MB24	3	ソショ	sosjo	sosjo
MB24	4	ソマ	soma	soma
MB24	5	ソミ	somi	somi
MB24	6	ソム	somu	somu
MB24	7	ソメ	some	some
MB24	8	ソモ	somo	somo
MB24	9	シャカ	sjaka	sjaka

MB24	10	シャキ	sjaki	sjaki
MB24	11	シャク	sjaku	sjaku
MB24	12	シャケ	sjake	sjake
MB24	13	シャコ	sjako	sjako
MB24	14	シャキャ	sjakja	sjakja
MB24	15	シャキュ	sjakju	sjakju
MB24	16	シャキョ	sjakjo	sjakjo
MB24	17	シャハ	sjaha	sjaha
MB24	18	シャヒ	sjahi	sjahi
MB24	19	シャフ	sjahu	sjahu
MB24	20	シャヘ	sjahe	sjahe
MB25	1	シャホ	sjaho	sjaho
MB25	2	シャサ	sjasa	sjasa
MB25	3	シャシ	sjasi	sjasi
MB25	4	シャス	sjasu	sjasu
MB25	5	シャセ	sjase	sjase
MB25	6	シャソ	sjaso	sjaso
MB25	7	シャシャ	sjasja	sjasja
MB25	8	シャシュ	sjasju	sjasju
MB25	9	シャショ	sjasjo	sjasjo
MB25	10	シャマ	sjama	sjama
MB25	11	シャミ	sjami	sjami
MB25	12	シャム	sjamu	sjamu
MB25	13	シャメ	sjame	sjame
MB25	14	シャモ	sjamo	sjamo
MB25	15	シュカ	sjuka	sjuka
MB25	16	シュキ	sjuki	sjuki
MB25	17	シュク	sjuku	sjuku
MB25	18	シュケ	sjuke	sjuke
MB25	19	シュコ	sjuko	sjuko
MB25	20	シュキャ	sjukja	sjukja
MB26	1	シュキュ	sjukju	sjukju
MB26	2	シュキョ	sjukjo	sjukjo
MB26	3	シュハ	sjuha	sjuha
MB26	4	シュヒ	sjuhi	sjuhi
MB26	5	シュフ	sjuhu	sjuhu

MB26	6	シュヘ	sjuhe	sjuhe
MB26	7	シュホ	sjuho	sjuho
MB26	8	シュサ	sjusa	sjusa
MB26	9	シュシ	sjusi	sjusi
MB26	10	シュス	sjusu	sjusu
MB26	11	シュセ	sjuse	sjuse
MB26	12	シュソ	sjuso	sjuso
MB26	13	シュシャ	sjusja	sjusja
MB26	14	シュシュ	sjusju	sjusju
MB26	15	シュショ	sjusjo	sjusjo
MB26	16	シュマ	sjuma	sjuma
MB26	17	シュミ	sjumi	sjumi
MB26	18	シュム	sjumu	sjumu
MB26	19	シュメ	sjume	sjume
MB26	20	シュモ	sjumo	sjumo
MB27	1	ショカ	sjoka	sjoka
MB27	2	ショキ	sjoki	sjoki
MB27	3	ショク	sjoku	sjoku
MB27	4	ショケ	sjoke	sjoke
MB27	5	ショコ	sjoko	sjoko
MB27	6	ショキャ	sjokja	sjokja
MB27	7	ショキュ	sjokju	sjokju
MB27	8	ショキョ	sjokjo	sjokjo
MB27	9	ショハ	sjoha	sjoha
MB27	10	ショヒ	sjohi	sjohi
MB27	11	ショフ	sjohu	sjohu
MB27	12	ショヘ	sjohe	sjohe
MB27	13	ショホ	sjoho	sjoho
MB27	14	ショサ	sjosa	sjosa
MB27	15	ショシ	sjosi	sjosi
MB27	16	ショス	sjosu	sjosu
MB27	17	ショセ	sjose	sjose
MB27	18	ショソ	sjoso	sjoso
MB27	19	ショシャ	sjosja	sjosja
MB27	20	ショシュ	sjosju	sjosju
MB28	1	ショショ	sjosjo	sjosjo

MB28	2	ショマ	sjoma	sjoma
MB28	3	ショミ	sjomi	sjomi
MB28	4	ショム	sjomu	sjomu
MB28	5	ショメ	sjome	sjome
MB28	6	ショモ	sjomo	sjomo
MB28	7	マカ	maka	maka
MB28	8	マキ	maki	maki
MB28	9	マク	maku	maku
MB28	10	マケ	make	make
MB28	11	マコ	mako	mako
MB28	12	マキャ	makja	makja
MB28	13	マキュ	makju	makju
MB28	14	マキョ	makjo	makjo
MB28	15	マハ	maha	maha
MB28	16	マヒ	mahi	mahi
MB28	17	マフ	mahu	mahu
MB28	18	マヘ	mahe	mahe
MB28	19	マホ	maho	maho
MB28	20	マサ	masa	masa
MB29	1	マシ	masi	masi
MB29	2	マス	masu	masu
MB29	3	マセ	mase	mase
MB29	4	マソ	maso	maso
MB29	5	マシャ	masja	masja
MB29	6	マシュ	masju	masju
MB29	7	マショ	masjo	masjo
MB29	8	ママ	mama	mama
MB29	9	マミ	mami	mami
MB29	10	マム	mamu	mamu
MB29	11	マメ	mame	mame
MB29	12	マモ	mamo	mamo
MB29	13	ミカ	mika	mika
MB29	14	ミキ	miki	miki
MB29	15	ミク	miku	miku
MB29	16	ミケ	mike	mike
MB29	17	ミコ	miko	miko

MB29	18	ミキャ	mikja	mikja
MB29	19	ミキュ	mikju	mikju
MB29	20	ミキョ	mikjo	mikjo
MB30	1	ミハ	miha	miha
MB30	2	ミヒ	mihi	mihi
MB30	3	ミフ	mihu	mihu
MB30	4	ミヘ	mihe	mihe
MB30	5	ミホ	miho	miho
MB30	6	ミサ	misa	misa
MB30	7	ミシ	misi	misi
MB30	8	ミス	misu	misu
MB30	9	ミセ	mise	mise
MB30	10	ミソ	miso	miso
MB30	11	ミシャ	misja	misja
MB30	12	ミシュ	misju	misju
MB30	13	ミショ	misjo	misjo
MB30	14	ミマ	mima	mima
MB30	15	ミミ	mimi	mimi
MB30	16	ミム	mimu	mimu
MB30	17	ミメ	mime	mime
MB30	18	ミモ	mimo	mimo
MB30	19	ムカ	muka	muka
MB30	20	ムキ	muki	muki
MB31	1	ムク	muku	muku
MB31	2	ムケ	muke	muke
MB31	3	ムコ	muko	muko
MB31	4	ムキャ	mukja	mukja
MB31	5	ムキュ	mukju	mukju
MB31	6	ムキョ	mukjo	mukjo
MB31	7	ムハ	muha	muha
MB31	8	ムヒ	muhi	muhi
MB31	9	ムフ	muhu	muhu
MB31	10	ムヘ	muhe	muhe
MB31	11	ムホ	muho	muho
MB31	12	ムサ	musa	musa
MB31	13	ムシ	musi	musi

MB31	14	ムス	musu	musu
MB31	15	ムセ	muse	muse
MB31	16	ムソ	muso	muso
MB31	17	ムシャ	musja	musja
MB31	18	ムシュ	musju	musju
MB31	19	ムショ	musjo	musjo
MB31	20	ムマ	muma	muma
MB32	1	ムミ	mumi	mumi
MB32	2	ムム	mumu	mumu
MB32	3	ムメ	mume	mume
MB32	4	ムモ	mumo	mumo
MB32	5	メカ	meka	meka
MB32	6	メキ	meki	meki
MB32	7	メク	meku	meku
MB32	8	メケ	meke	meke
MB32	9	メコ	meko	meko
MB32	10	メキャ	mekja	mekja
MB32	11	メキュ	mekju	mekju
MB32	12	メキョ	mekjo	mekjo
MB32	13	メハ	meha	meha
MB32	14	メヒ	mehi	mehi
MB32	15	メフ	mehu	mehu
MB32	16	メヘ	mehe	mehe
MB32	17	メホ	meho	meho
MB32	18	メサ	mesa	mesa
MB32	19	メシ	mesi	mesi
MB32	20	メス	mesu	mesu
MB33	1	メセ	mese	mese
MB33	2	メソ	meso	meso
MB33	3	メシャ	mesja	mesja
MB33	4	メシュ	mesju	mesju
MB33	5	メショ	mesjo	mesjo
MB33	6	メマ	mema	mema
MB33	7	メミ	memi	memi
MB33	8	メム	memu	memu
MB33	9	メメ	meme	meme

MB33	10	メモ	memo	memo
MB33	11	モカ	moka	moka
MB33	12	モキ	moki	moki
MB33	13	モク	moku	moku
MB33	14	モケ	moke	moke
MB33	15	モコ	moko	moko
MB33	16	モキャ	mokja	mokja
MB33	17	モキュ	mokju	mokju
MB33	18	モキョ	mokjo	mokjo
MB33	19	モハ	moha	moha
MB33	20	モヒ	mohi	mohi
MB34	1	モフ	mohu	mohu
MB34	2	モヘ	mohe	mohe
MB34	3	モホ	moho	moho
MB34	4	モサ	mosa	mosa
MB34	5	モシ	mosi	mosi
MB34	6	モス	mosu	mosu
MB34	7	モセ	mose	mose
MB34	8	モソ	moso	moso
MB34	9	モシャ	mosja	mosja
MB34	10	モシュ	mosju	mosju
MB34	11	モショ	mosjo	mosjo
MB34	12	モマ	moma	moma
MB34	13	モミ	momi	momi
MB34	14	モム	momu	momu
MB34	15	モメ	mome	mome
MB34	16	モモ	momo	momo
SR1	1	カカ	kaka	kaka
SR1	2	カキ	kaki	kaki
SR1	3	カク	kaku	kaku
SR1	4	カケ	kake	kake
SR1	5	カコ	kako	kako
SR1	6	カキャ	kakja	kakja
SR1	7	カキュ	kakju	kakju
SR1	8	カキョ	kakjo	kakjo
SR1	9	カハ	kaha	kaha

SR1	10	カヒ	kahi	kahi
SR2	1	カフ	kahu	kahu
SR2	2	カヘ	kahe	kahe
SR2	3	カホ	kaho	kaho
SR2	4	カサ	kasa	kasa
SR2	5	カシ	kasi	kasi
SR2	6	カス	kasu	kasu
SR2	7	カセ	kase	kase
SR2	8	カソ	kaso	kaso
SR2	9	カシャ	kasja	kasja
SR2	10	カシュ	kasju	kasju
SR3	1	カショ	kasjo	kasjo
SR3	2	カマ	kama	kama
SR3	3	カミ	kami	kami
SR3	4	カム	kamu	kamu
SR3	5	カメ	kame	kame
SR3	6	カモ	kamo	kamo
SR3	7	キカ	kika	kika
SR3	8	キキ	kiki	kiki
SR3	9	キク	kiku	kiku
SR3	10	キケ	kike	kike
SR3	11	キコ	kiko	kiko
TT1	1	『菊栗』	_kikuguri	
TT1	2	『竹垣』	_takegaki	
TT1	3	『虎』	_tora	
PL1	1	『中立』	_neutral	
PL1	2	『反問』	_suspicion	
PL1	3	『感心』	_admiration	
PL1	4	『落胆』	_disappointment	
PL2	1	『0』	_focus0	
PL2	2	『1』	_focus1	
PL2	3	『2』	_focus2	
PL2	4	『3』	_focus3	

Appendix 2: Number of recorded utterances per speaker, FPS, and recording classes

For each speaker and FPS, the total number of recorded utterances and its break-down for utterance classes are indicated in this table.

SUBJECT	FPS	MU	MP	MB	SR	PL	TT	NS	TOTAL
s1	14	109	272	797	35	15	4	2	1234
s1	27	68	0	0	0	0	0	0	68
s2	14	109	336	836	31	20	4	2	1338
s2	27	88	28	29	0	0	0	0	145
s3	14	137	97	694	0	0	0	0	928
s3	27	0	43	22	0	0	0	0	65
s4	14	162	287	754	0	0	0	2	1205
s4	27	82	28	21	0	0	0	0	131
s5	14	134	273	712	41	0	3	4	1167
s5	27	61	32	34	0	0	0	0	127
s7	14	133	320	800	38	12	10	3	1316
s7	27	77	0	0	0	0	0	0	77
s8	14	109	299	811	31	16	4	2	1272
s8	27	66	0	0	0	0	0	0	66
s9	14	109	100	824	41	8	3	0	1085
s9	27	0	29	22	0	0	0	0	51
s10	14	108	100	733	31	18	3	0	993
s10	27	0	47	11	0	0	0	0	58
s11	14	109	123	720	31	24	3	0	1010
s11	27	0	27	21	0	0	0	0	48
s12	14	137	274	695	41	10	5	2	1164
s12	27	91	36	21	0	0	0	0	148
s14	14	140	292	837	31	0	5	2	1307
s14	27	72	27	18	0	0	0	0	117
s16	14	175	187	730	0	20	0	0	1112
s16	27	0	30	22	0	0	0	0	52
s17	14	167	168	730	0	16	0	0	1081
s17	27	0	33	22	0	0	0	0	55
s18	14	140	168	770	0	18	3	0	1099
s18	27	0	40	19	0	0	0	0	59
s19	14	167	365	752	0	18	8	2	1312

s19	27	63	43	22	0	0	0	0	128
s20	14	142	149	734	0	19	4	1	1049
s21	14	142	171	782	0	21	5	1	1122
s24	14	192	196	787	0	0	0	4	1179
s25	14	145	149	676	0	0	0	2	972
s26	14	205	174	745	0	0	0	2	1126
s27	14	144	191	731	0	0	0	2	1068
s28	14	145	220	790	0	0	0	4	1159
s29	14	145	219	732	0	0	0	3	1009
s30	14	145	221	824	0	0	0	2	1192