

TR-IT-0301

Japanese Speech Recognition Lexicon
Revision

Eric M. Visser

March 1999

c ATR Interpreting Telecommunications
Research Laboratories

c ATR 音声翻訳通信研究所

Contents

1	Introduction	1
2	Tuning the lexicon	1
3	Revisions	1
3.1	Doubles (code #H)	2
3.2	Names (code #J)	2
3.3	Homonyms (code #M)	2
3.4	Combinations (codes #CH/#C/#D and #FH/#F/#G)	3
3.5	Concatenations (code #R)	4
3.6	Errors (codes #X/#Y, #X2)	5
3.7	Reparanda (code #S)	5
4	Lexicon Statistics	6
5	Methods	7
6	Evaluation	7
7	Conclusions	9
	Appendix: Lexicon Revisions	10

1 Introduction

In the course of a detailed examination of the recognition results (lattices) output by ATR-SPREC, I have come across a large number of phenomena that were detrimental to recognition quality, yet seemed unnecessary. As a consequence I resolved to clean up the recognition lexicon; this Technical Report details my labours, and presents an evaluation. Cleaning up the lexicon involved removing “superfluous” entries (I will explain below which entries I considered superfluous, and why) and correcting some errors. One important reason why I did this, admittedly boring, work is that the low quality of the lexicon seriously hinders rule-based post-processing. The result of my work is a master lexicon which contains all the information in the original lexicon plus all the corrections, and a number of scripts to create training and/or recognition lexicons and abstract various kinds of data from the master lexicon. To evaluate the lexicon I modified the answer files used for training and testing so that they reflected the changes in the lexicon, re-trained the language model, and did recognition experiments.

2 Tuning the lexicon

As mentioned above, the revisions consist mainly of removing superfluous entries from the lexicon. However, these entries have not been physically removed: rather, they have been “commented out”. That is to say, they have been prefixed with the symbol #, which makes SPREC ignore them as if they were comments. This is just as good as physically removing them, but it can be undone and therein lies the revised lexicon’s tunability.

Not all entries are removed for the same reason. Some of them are removed because they differ from each other only in surface realisation (i.e., the character string output by SPREC); others are removed because they can easily be combined; others again are removed for strategic reasons. What I have done is to categorise the superfluous entries, and to comment them out with a different string for each category. For instance, one of the things I have done is to remove all entries with Part-of-Speech “sonota” (reparanda): for this I have used the string “#S”. Entries that are simply doubles are commented out with “#H”, and include a reference to the ID number of the entry they double with. There are 8 categories of revisions, more if you include combinations of categories. In order to tune the lexicon to specific requirements, (for instance, produce a lexicon that leaves the reparanda in, but removes the doubles) it suffices to execute a script (see below). With the right settings, this script will even produce the original lexicon.

3 Revisions

The following kinds of lexicon revisions have been implemented. They are discussed in more detail below.

- I’ve removed doubles, entries that refer to the same word but are merely written differently.
- I’ve also removed phonetically similar names with different writings.
- I’ve removed homonyms, different words that are pronounced similarly.
- I’ve combined pronunciation variants of the same word into one entry. I have distinguished between interjections (PoS “kantoushi”) and others.
- I’ve removed entries that could be seen as being made up of concatenated simpler entries.
- I’ve corrected errors.
- I’ve removed all reparanda (PoS “sonota”) from the lexicon.

3.1 Doubles (code #H)

There are many morphemes in the lexicon which are for all practical purposes nothing but doubles: they refer to the same word, but are just written differently. Examples of these include for instance interjections that are written once in hiragana and once in katakana; words that differ only in the number of "okurigana"; and words that are written using different kanji which have the same meaning. Such differences probably stem from tagging errors.

Note that when two entries doubled with each other, I had no specific criteria to decide which one to leave in and which one to throw out. Mostly I just chose the one whose surface form I liked best. Also there are some cases where the #H tag combines with the #C and #S tags.

ATR

Cleanup 1: Doubles

Doubles:

- same phoneme string (or subset),
- same category,
- same word

Examples:

```
10207% [てお] t {e o|o {o|}} # てお|テオ|てお|助動詞|五段力|語幹||
#H 10207 10411% [と] t o # と|ト|と|助動詞|五段力|語幹||

#H 17756 13703% [ジェーティービー] zh e e t i i b i i {|-} # ...|固有名詞|
17756% [ジェイティービー] zh e e t i i b i i {|-} # ...|固有名詞|

17202% [振り替え] h u r i k a e {|-} # ...|振り替え|本動詞|一段|語幹||
#H 17202 17207% [振替] h u r i k a e {|-} # ...|振替|本動詞|一段|語幹||
```

116 times

Eric M. Visser - ATR Interpreting Telecommunications Research Laboratories - 1998

3.2 Names (code #J)

Names of people, both family names and given names, have the Part-of-Speech "jinmei" in the SR Lexicon. In Japanese there may be several ways to write the same name: for instance, the lexicon contains no less than 4 versions of "Akiko". However, it is fundamentally impossible for Speech Recognition to distinguish between different writings of the same name. (In fact this is true not only for class-based language models but also for word-based language models, with the possible exception of fixed combinations but such cases are close to cheating anyway.) It may be noted that this revision can be considered a special case of either the doubles or the homonyms, depending on your point of view.

(Example in the next paragraph)

3.3 Homonyms (code #M)

Words that have the same pronunciation and the exact same grammatical attributes cannot be distinguished by a Speech Recognition system that uses only these grammatical attributes, even if they are really different words. I've found a fair number of such cases, and marked them with the symbol #M. However, it should be noted that some of these morphemes are split off into separate classes when using the composite N-gram method. In such cases, the ambiguity becomes (at least in theory) resolvable. I

have looked up such entries for the case of a composite N-gram with 1000 classes split off, and marked them doubly with the tag #V1000.

Note: I have generally retained the word that I thought to be the most general, though my intuitions may be wrong.

Another note: while examples such as the one below, where the kanji are different, are invariably treated as different words, this is not the case for words of foreign origin. The entry “basu”, for instance, is used to indicate both “bath” and “bus”. This is obviously wrong, but hasn’t been corrected because it would have to be done manually.

ATR

Cleanup 2/3: Names/Homonyms

Names: people's names that are pronounced the same but written with different kanji.

```
#J 15177 14058% [貴島] k i z h i m a # 貴島|キジマ|貴島|人名| || |
15177% [木島] k i z h i m a # 木島|キジマ|木島|人名| || |
```

40 times

Homonyms: same grammar, same pronunciation, different word.

```
#M 16479 15918% [放送] h o o s o o {-} # ...|放送|サ変名詞| || |
16479% [包装] h o o s o o {-} # ...|包装|サ変名詞| || |
```

```
10201% [早] h a j a # 早|ハヤ|早|形容詞|形容詞|語幹| |
#M 10201 10740% [速] h a j a # 速|ハヤ|速|形容詞|形容詞|語幹| |
```

** note: some homonyms are split off as a separate class in the composite N-gram method.*

116 times (- 26 times)

Eric M. Visser - ATR Interpreting Telecommunications Research Laboratories - 1998

3.4 Combinations (codes #CH/#C/#D and #FH/#F/#G)

To a certain extent, the lexicon developers have taken to combining pronunciation variants of a word into a single entry in the lexicon. For instance, in the lexicon you will find the pronunciation variant “beya” combined with the original “heya” into a phoneme string [{h—b} e j a]. The Interjection “e” has been combined with its lengthened variant “ee” into [e {—e}]. However, this has by no means been done consistently: there are still lots of other morphemes which can be combined. A painful example is the aforementioned interjection “e”: in all there are 6 possible pronunciation variants (long or short, followed by a pause, a glottal stop or nothing: 2 times 3 is 6). However, in the lexicon, there are 7 entries incorporating these variants a total of 17 times. It goes without saying that this is wasteful; in addition, I find it hard to believe that tagging is (or even: can be?) done precisely enough to distinguish between such entries.

The revision is implemented using three tags: #CH marks an entry which is combined with another and removed (needs to specify the ID of the entry with which is is combined); #C also marks an entry which is combined with another, but marks the one whose ID number is retained; and #D marks the new entry, which has a modified phoneme string which combines those of the two aforementioned morphemes. Also, interjections have been tagged separately using #FH, #F and #G instead of the tags mentioned above.

Combinations: same word, slightly different pronunciations

```
#C 15638% [アスレチック] a s u r e c h i q k u {|-} # ...|普通名詞|||
#CH 15638 16731% [アスレティック] a s u r e t i q k u {|-} # ...|普通名詞|||
#D 15638% [アスレチック] a s u r e {ch|t} i q k u {|-} # ...|普通名詞|||

#C 10544% [コンピューター] k o n g p j u u t a a {|-} # ...|普通名詞|||
#CH 10544 14999% [コンピュータ] k o n g p j u u t a {|-} # ...|普通名詞|||
#D 10544% [コンピューター] k o n g p j u u t a {a|} {|-} # ...|普通名詞|||
```

97 groups, 203 entries

(+ the same thing for Interjections)

```
#F 10017% [えー] e {e|} {|-} # えー|エー|えー|間投詞|||
#FH 10017 10203% [え] e {|e} {|-} # え|エ|え|間投詞|||
#FH 10017 10917% [えっ] e {q|-} # えっ|エツ|えっ|間投詞|||
#FH 10017 11802% [ええ] e e {|-} # ええ|エエ|ええ|間投詞|||
#FH 10017 14603% [エー] e e {|-} # エー|エー|エー|間投詞|||
#FH 10017 14772% [えーっ] e e {q|-} # えーっ|エーツ|えーっ|間投詞|||
#FH 10017 17936% [ええっ] e e q # ええっ|エエツ|ええっ|間投詞|||

#G 10017% [えー] e {e|} {|q|-} # えー|エー|えー|間投詞|||
```

49 groups, 168 entries*Eric M. Visser - ATR Interpreting Telecommunications Research Laboratories - 1998***3.5 Concatenations (code #R)**

The lexicon contains a good number of morphemes which can be seen as being made up of smaller morphemes. This phenomenon is especially marked in the category of interjections or fillers ("kantoushi"). I have taken a quick look at this category, and found that about a third of them were in reality nothing but concatenations of shorter interjections. I have removed these on the assumption that they could just as well be recognised as two consecutive interjections. Note that the same applies for an enormous number of other (non-interjection) entries (I count at least 1,500!), but I haven't tagged those.

Concatenations: one entry consisting of two or more parts which are also in the lexicon

```
#R 10017 10054 #/ 15101% [えーあの一] e e a n o o {|-} # ...|...|...|間投詞|...|
#R 10017 10054 #/ 15241% [えーあの一] e e a n o o {|-} # ...|...|...|間投詞|...|

#R 10005 10263 #/ 14755% [あそうですね] a s o o d e s u n e {|-} # ...|...|...|
|間投詞|...|

#R 10017 10005 10498 #/ 19611% [えっあっえーと] e q a q e e t o {|-} # ...|...|
|...|間投詞|...|
```

(note: only Interjections)

92 times

Eric M. Visser - ATR Interpreting Telecommunications Research Laboratories - 1998

3.6 Errors (codes #X/#Y, #X2)

The lexicon contains a number of errors, some of which have an influence on recognition per se and some of which might make trouble for post-processing. These errors have been corrected. There are two types of error: ones that can be solved simply by removing a morpheme, and ones that require the incorrect morpheme to be replaced. The former type has been marked with the tag #X2; of the latter type, the incorrect morphemes have the tag #X and the correct ones that replace them have the tag #Y.

Note: to be completely honest I've also changed a couple of unfelicitous surface notations using this tag, e.g. "mon" for "mono", "n" for "ni", "hoidewa" for "soredewa". Since they are only surface notations they are not important, but they make the recognition results difficult to read (for humans that is). To be precise, there were 11 of those.

(Examples in the next paragraph)

3.7 Reparanda (code #S)

Reparanda (mispronunciations and such) have in the ATR corpus been treated just like regular words: they have been given their own Part-of-Speech ("sonota"), added to the lexicon, and processed as if they were regular words. Which they are not, of course. There are some very strong arguments for removing reparanda from the lexicon:

1. Only about 1 % of the utterances in ITL's Speech and Language Data Base (SLDB) contains the PoS "sonota". (234 out of 21,769) (In other words, the recognition rate of 99 % of the utterances is unlikely to be adversely affected.)
2. However, fully 8 % of the morphemes in the speech recognition dictionary have PoS "sonota". (586 out of 7,219) (Clearly, many "sonota" morphemes are used only once. This is of course not surprising.)

3. Moreover, when running the speech recognition system, close to 20 % of the morphemes that will be considered in the lattice are "sonota" morphemes. Even if they were not in the utterance and do not come out in the actual nbest list.
4. And finally the coup de grace: the recognition rate for "sonota" morphemes is very low. I ran a little test with 4 utterances from the SLDB (which all contained a "sonota" morpheme), and in 3 of the 4 cases the right morpheme was nowhere in the 10-best list, in one case it was in the 5th-best candidate. This is not surprising if you think of it: they occur very rarely (in the DB anyway) and their use is irregular, so they will have a low probability both as a category and as an entity. Reparanda are principally ill-suited to statistical language models.

ATR

Cleanup 7/8: Errors, "sonota"

Errors: in surface form, PoS, phoneme string

```
#X 19894% [インターナショナル] ... {-} # ...|...|サ変名詞|
#Y 19894% [インターナショナル] ... {-} # ...|...|形容名詞|

#X 16886% [京都運輸] k j o o t o o n g j u {-} # ...|...|普通名詞|
#Y 16886% [京都運輸] k j o o t o {-} u n g j u {-} # ...|普通名詞|

#X2 18414 14839% [そこら] s o k o r a {-} # ...|副詞|
18414% [そこら] s o k o r a {-} # ...|代名詞|
# a corpus check reveals these are the exact same usage
```

70 times (-11)

Reparanda: mispronunciations, partial words etc.

```
#S 14215% [にひゃ] n i h j a {-} # にひゃ|ニヒャ|にひゃ|その他|
#S 14060% [さんびゃくろくじゅうえん] ... {-} # ...|...|その他|
```

585 times

Eric M. Visser - ATR Interpreting Telecommunications Research Laboratories - 1998

4 Lexicon Statistics

The following table shows the number of entries per PoS for the old lexicon and for two revised versions: one that does not implement the removal of homonyms (the second column) and one that does (the rightmost one; I would consider the former one to be the more realistic revised version). Although the lexicon has been reduced by about 13 %, it should be noted that the lion's share of the reduction (about 60 %) has been achieved merely by removing the reparanda.

ATR		Total	合計	7219	6246 (-13 %)	6157 (-15 %)
Comparison:	Nominals	名詞類	4503	4382	4329	
	- Regular Nouns	- 普通名詞	2383	2358	2315	
	- Verbal Nouns	- サ変名詞	549	535	527	
	- Adjectival Nouns	- 形容名詞	209	210	209	
	- Pronouns	- 代名詞	39	37	37	
	- Times and Dates	- 日時	115	108	108	
	- Proper Nouns	- 固有名詞	673	640	639	
	- Names	- 人名	477	435	435	
	- Addresses	- 住所名	58	58	58	
	Verbals	動詞類	812	788	769	
	- Regular Verbs	- 本動詞	617	606	587	
	- Auxiliary Verbs	- 補助動詞	39	38	38	
	- Affixed Verbs	- 助動詞	156	144	144	
	Adjectives	形容詞	102	101	100	
	Endings	語尾	77	77	77	
	Adverbs	副詞	305	296	296	
	Prefixes	接頭辞	25	25	22	
	Suffixes	接尾辞	153	141	130	
	Numerals	数詞	36	34	34	
	Conjunctions	接続詞	93	88	88	
	Adnominals	連体詞	28	28	28	
Particles	助詞	136	129	129		
Interjections	間投詞	265	76	76		
Idioms	感動詞	98	82	82		
Reparanda	その他	586	0	0		

lexical categories before and after revision

column 1: original lexicon
column 2: Exp. 4 (didn't remove homonyms)
column 3: Exp. 5 (removed homonyms that are not split off)

Eric M. Visser - ATR Interpreting Telecommunications Research Laboratories - 1998

5 Methods

While it would admittedly be nice if there were a script or something you could run to achieve this kind of cleanup, the reality is that I did it all by hand. I expanded all the morpheme lists, marked entries that had the same morpheme list, double-marked entries that also had the same category (up to this point I used emacs' search-and-replace function with regular expressions), and from there I've checked each case individually. I ran a special hand check of the interjections, because this category obviously had a lot of room for improvement (as the results show). In addition there are also quite a few revisions I merely chanced upon, and it is extremely unlikely that the lexicon is error-free even now.

6 Evaluation

In order to evaluate the new lexicon, I took the following steps.

- Produced a training lexicon from the Master Lexicon which was discussed above, using a python script. The script requires a set of "reform conditions", which state for each of the revisions (discussed above) whether it is to be implemented or not.
- Adapted the answer files used for training, again using a python script. I used the same set of answer files as presented in [1]. I also adapted the answer files of the test set.

	dialogues	utterances	words
Training	2,514	30,112	565,935
Testing	64	1,331	19,927

- Produced a class definition file from the lexicon, using yet another python script.
- Trained a language model and produced the recognition lexicon using the regular ATR-SPREC scripts. The language model is a composite N-gram model with backoff smoothing and 1000 classes split off.

- Ran recognition tests. The recognition tests used `atrlattres`, and the settings were copied from [1]. The test set was as indicated above. At the time the current version of ATR-SPREC was `r05r05`, which was used in these experiments both for training and for recognition. The acoustic model was `AM.MF.bin` Version 1, a gender-independent model. The parameters were as follows.

lmscale	8.0,20.0
beam	90,90
work area	3000,200

The experiments were carried out for 5 different lexicons; the results are shown in the table below. Exp04 (experiment 4) uses the lexicon that's considered to be the "most realistic" revised version. As you can see, the recognition results show a slight (probably not significant) improvement, while the recognition efficiency (as expressed by the number of nodes, arcs and words in the lattice as well as the real-time rate) goes up significantly. In fact, on individual utterances, I have observed speed gains of up to 500 %. However, again the majority of the improvement is attained merely by removing the reparanda from the lexicon.

ATR V1

	Exp00	Exp01	Exp02	Exp04	Exp05
(1331) total utterances	1330	1330	1330	1329	1329
# perfect	472	471	498	517	518
% accuracy	35.49	35.41	37.44	38.90	38.98
# net perfect	656	656	654	668	671
% accuracy	49.32	49.32	49.17	50.26	50.49
total words	17232	17232	17232	17179	17179
T # ins / del / sub	732/ 762/ 2475	728/ 759/ 2510	703/ 782/ 2372	729/ 754/ 2225	731/ 755/ 2220
O % accuracy	(74.66) 76.97	(76.15) 76.80	77.62	78.42	78.43
P % correct	81.22	81.03	81.70	82.66	82.68
acc. # ins / del / sub	405/ 415/ 1541	409/ 415/ 1570	403/ 427/ 1536	425/ 425/ 1446	423/ 423/ 1441
N % accuracy	(80.37) 86.30	(81.78) 86.11	86.27	86.63	86.69
E correct # del / sub	472 / 1386	471 / 1413	496 / 1376	489 / 1302	492 / 1290
T % correct	89.22	89.07	89.14	89.57	89.63
node / arc	1160 / 16135	987 / 12150	974 / 11681	994 / 11777	996 / 11801
word	167	146	148	151	150
rate	(485) 605	(438) 476	463	460	455

Exp00 = original Lex/LM (01 = remove "sonota"); Exp02 = combine Interjections, remove "sonota"
Exp04 = all revisions except homonyms; Exp05 = all revisions

Eric M. Visser - ATR Interpreting Telecommunications Research Laboratories - 1998

A closer look at the test results, concentrating on experiments 1 and 4, shows that about two-thirds of the test set is unaffected by the changes in the lexicon. (Note that the discussion here is restricted to the effect of the changes in the lexicon excluding the removal of reparanda.) About 20 but half of this gain is offset by the fact that close to 10 % of the recognition results actually deteriorate.

	Exp01	Exp04
(1331) total utterances	1330	1329
# perfect	471	517
% accuracy	35.41	38.90
# net perfect	656	668
% accuracy	49.32	50.26
total words	17232	17179
T # ins / del / sub	728 / 759 / 2510	729 / 754 / 2225
O % accuracy	76.80	78.42
P % correct	81.03	82.66
acc. # ins / del / sub	409 / 415 / 1570	425 / 425 / 1446
N % accuracy	86.11	86.63
E correct # del / sub	471 / 1413	489 / 1302
T % correct	89.07	89.57
node / arc	987 / 12150	994 / 11777
word	146	151
rate	476	460

Individual utterances

Training set

- number of items: 565,935
 modified: 14,002 (2.47 %)

Test set

- number of items: 19,927
 modified: 536 (2.69 %)

per-utterance comparison

seikai changed	396
both perfect	454
Exp01 is better	124
Exp04 is better	268
same error	431
same score	54
total	1331

Eric M. Visser - ATR Interpreting Telecommunications Research Laboratories - 1998

7 Conclusions

I started this work with the goal of working on a (more or less) intelligent post-processing scheme for Speech Recognition, which would take the lattice as its basic input. From that point of view I saw the enormous amount of junk in the lattice as a major impediment to post-processing, and I felt I needed to work on lightening it. This I think has been achieved, to a point: the test results show a sizable reduction in the number of nodes, arcs and words in the lattice. It is clear though that the lattice still allows a great number of linguistically impossible options. I feel it is important to strengthen the linguistic relevance of the language model processing in order to further reduce the size of the lattice, but it is obvious that this cannot be achieved by merely cleaning up the lexicon.

A final note: the master lexicon, scripts and various other files associated with this project can be found in the directory

/dept4/work3/eric/cleanup/

References

- [1] Hirokazu Masataki, Kazuhiro Takahashi, Hajime Tsukada, Kouichi Tanigaki, Hirofumi Yamamoto, Atsushi Nishino and Atsushi Nakamura: *Baseline Language Model for Continuous Speech Recognition*. ATR Technical Report, TR-IT-0235, 1997. (in Japanese)

Appendix: Speech Recognition Lexicon Revisions

#####

The following is a list of revisions to the speech recognition lexicon of ATR-SPREC (SLDB format).
The revisions are grouped by type.

#####

DOUBLES

doubles of reparanda

#S #H 15120 10376% [は] w a {|-} # ..|その他|||
#S #H 14848 10395% [ス] s u {|-} # ..|その他|||
#S #H 14848 10545% [す] s u {|-} # ..|その他|||
#S #H 14646 11054% [ツ] t s u {|-} # ..|その他|||
#S #H 14769 11115% [し] s h i {|-} # ..|その他|||
#S #H 10993 11534% [を] o {|-} # ..|その他|||
#S #H 15986 11564% [ちよ] c h j o {|-} # ..|その他|||
#S #H 14904 11974% [も] m o {|-} # ..|その他|||
#S #H 13970 12366% [け] k e {|-} # ..|その他|||
#S #H 14769 14041% [し] s h i {|-} # ..|その他|||
#S #H 14186 14079% [とお] t o o {|-} # ..|その他|||
#S #H 15420 14116% [さ] s a {|-} # ..|その他|||
#S #H 17222 14571% [場] b a {|-} # ..|その他|||
#S #H 15291 14572% [合] a i {|-} # ..|その他|||
#S #H 12020 14660% [ホ] h o {|-} # ..|その他|||
#S #H 15081 14745% [じゅーは] z h j u u h a {|-} # ..|その他|||
#S #H 14858 14760% [ぶ] p u {|-} # ..|その他|||
#S #H 17406 14885% [おー] o o {|-} # ..|その他|||
#S #H 15222 14941% [うけ] u k e {|-} # ..|その他|||
#S #H 10661 15021% [何] n a n g {|-} # ..|その他|||
#S #H 13513 15022% [時] z h i {|-} # ..|その他|||
#S #H 14080 15176% [シン] s h i n g {|-} # ..|その他|||
#S #H 19931 15240% [すー] s u u {|-} # ..|その他|||
#S #H 19547 15286% [泊] t o m a {|-} # ..|その他|||
#S #H 15302 15303% [日] n i c h i {|-} # ..|その他|||
#S #H 15891 15305% [ル] r u {|-} # ..|その他|||
#S #H 11507 15315% [取] t o {|-} # ..|その他|||
#S #H 20199 15386% [小会] s h j o o k a i {|-} # ..|その他|||
#S #H 11507 15510% [ト] t o {|-} # ..|その他|||
#S #H 19745 15720% [スイ] s u i {|-} # ..|その他|||
#S #H 19745 15745% [スイ] s u i {|-} # ..|その他|||
#S #H 17406 16266% [おう] o o {|-} # ..|その他|||

#S #H 14042 16744% [風] hu {-} # ..|その他|||
 #S #H 14419 16745% [呂] ro {-} # ..|その他|||
 #S #H 10893 16758% [ハ] ha {-} # ..|その他|||
 #S #H 15144 16918% [質] sh i ts u {-} # ..|その他|||
 #S #H 19707 19678% [イン] i ng {-} # ..|その他|||
 #S #H 13991 19943% [万] m a ng {-} # ..|その他|||
 #S #H 14744 20207% [チ] ch i {-} # ..|その他|||

just doubles

#H 10207 10411% [と] to # ..|助動詞|五段カ|語幹||
 #H 17593 11175% [ブロードウェイ] bu ro o do we {e|i} {-} # ..|固有名詞|||
 #H 11728 11371% [エー] ee {-} # ..|普通名詞|||
 #H 12539 11380% [ケー] ke {e|i} {-} # ..|普通名詞|||
 #H 11748 11749% [十日] to o k a {-} # ..|日時|||
 #H 12195 11757% [エクストラベッド] ek us u to ra be q do {-} # ..|普通名詞|||
 #H 11767 11779% [ニューヨークシティーホテル] n j u u j o o k u sh i t i i ho te ru {-} # ..|固有名詞|||
 #H 11184 12028% [七時] hi ch i zh i {-} # ..|日時|||
 #H 12030 12047% [てらっしゃ] te ra q sh ja # ..|助動詞|特殊ラ|語幹||
 #H 12195 12188% [エキストラベット] ek is u to ra be q to {-} # ..|普通名詞|||
 #H 18839 12340% [撮] to # ..|本動詞|五段ラ|語幹||
 #H 10126 12706% [らっしゃ] ra q sh ja # ..|補助動詞|特殊ラ|語幹||
 #H 16702 12886% [グランビュウ] gu ran b j u u {-} # ..|固有名詞|||
 #H 18665 12931% [間違え] ma chi ga e {-} # ..|普通名詞|||
 #H 13050 13007% [パーティー] pa a t i i {-} # ..|普通名詞|||
 #H 17756 13703% [ジューティービー] zhe e t i i b i i {-} # ..|固有名詞|||
 #H 12823 14275% [治] na o # ..|本動詞|五段サ|語幹||
 #H 14516 14482% [エアメール] ea me e ru {-} # ..|普通名詞|||
 #H 11892 14556% [ベット] be q {t|d} o {-} # ..|普通名詞|||
 #H 10152 14652% [ありがとう] a ri ga to {-} # ..|感動詞|||
 #H 13886 14675% [治] na o # ..|本動詞|五段ラ|語幹||
 #H 14097 15380% [ルームメーク] ru um e e ku {-} # ..|普通名詞|||
 #H 11106 15903% [こたえ] ko ta e {-} # ..|本動詞|一段|語幹||
 #H 15916 15923% [清算書] se e sa ng sh jo {-} # ..|普通名詞|||
 #H 16210 16213% [ショウケース] sh jo o ke e su {-} # ..|普通名詞|||
 #H 17183 17158% [エルエーエックス] er ue e e q ku su {-} # ..|固有名詞|||
 #H 17202 17207% [振替] hu ri ka e {-} # ..|本動詞|一段|語幹||
 #H 17290 17263% [ソルトレイクシティー] so ru to re e ku sh i t i {-} # ..|固有名詞|||
 #H 16590 17557% [さあ] sa a {-} # ..|間投詞|||
 #H 18839 17622% [採] to # ..|本動詞|五段ラ|語幹||
 #H 15600 17686% [わあ] wa a {-} # ..|間投詞|||
 #H 17290 17709% [ソルトレイクシティー] so ru to re e ku sh i t i i {-} # ..|固有名詞|||

#H 17880 17892% [フォームゲーム] h o o m u g e e m u {|-} # ..|普通名詞|||
#H 20103 17941% [お薦め品] o s u s u m e h i n g {|-} # ..|普通名詞|||
#H 19884 18385% [べんけい] b e n g k e e {|-} # ..|固有名詞|||
#H 18806 18754% [と] t o # ..|助動詞|五段ラ|語幹||
#H 13769 19317% [上] n o b o # ..|本動詞|五段ラ|語幹||
#H 14003 19517% [ダブルベット] d a b u r u b e t t o {|-} # ..|普通名詞|||
#H 17175 19737% [ジェーエフケー] z h e e e h u k e e {|-} # ..|固有名詞|||
#H 19643 19838% [大通公園] o o d o o r i k o o e n g {|-} # ..|固有名詞|||
#H 19766 19991% [張] h a # ..|本動詞|五段ラ|語幹||
#H 15834 20079% [預り票] a z u k a r i h j o o {|-} # ..|普通名詞|||
#H 11598 20204% [十] t o o {|-} # ..|数詞|||
#H 17358 20244% [コロンバス・サークル駅] k o r o n g b a s u s a a k u r u e k i {|-} # ..|固有名詞|||
#H 13656 20246% [サウスストリート・シーポート] s a u s u s u t o r i i t o s h i i p o o t o {|-} # ..|固有名詞|||
#H 17250 20367% [こんにちわ] k o n g n i c h i w a {|-} # ..|感動詞|||

COMBINATIONS (not interjections)

combination groups including errors

#CH 19059 #H 19059 11940% [それじゃあ] s o r e z h j a a {|-} # ..|接続詞|||
#CH 19059 10770% [それじゃ] s o r e z h j a {|-} # ..|接続詞|||
#C #X 19059% [はいでは] s o r e {dew|zh j a} a {|-} # ..|接続詞|||
#D #X 19059% [はいでは] s o r e {dew|zh j {|a}} a {|-} # ..|接続詞|||
#C #Y 19059% [それでは] s o r e {dew|zh j a} a {|-} # ..|接続詞|||
#D #Y 19059% [それでは] s o r e {dew|zh j {|a}} a {|-} # ..|接続詞|||

#CH 11505 #X 16856% [二十七日] n i z h j u u h i c h i n i c h i {|-} # ..|日時|||
#CH 11505 #Y 16856% [二十七日] n i z h j u u h i c h i {n i|ng} c h i {|-} # ..|日時|||
#C #X 11505% [二十七日] n i z h j u u s h i c h i n i c h i {|-} # ..|日時|||
#C #Y 11505% [二十七日] n i z h j u u s h i c h i {n i|ng} c h i {|-} # ..|日時|||
#D #X 11505% [二十七日] n i z h j u u {s h|h} i c h i n i c h i {|-} # ..|日時|||
#D #Y 11505% [二十七日] n i z h j u u {s h|h} i c h i {n i|ng} c h i {|-} # ..|日時|||

#CH 13171 #X 13734% [十七日] z h j u u h i c h i n i c h i {|-} # ..|日時|||
#CH 13171 #Y 13734% [十七日] z h j u u h i c h i {n i|ng} c h i {|-} # ..|日時|||
#C #X 13171% [十七日] z h j u u s h i c h i n i c h i {|-} # ..|日時|||
#C #Y 13171% [十七日] z h j u u s h i c h i {n i|ng} c h i {|-} # ..|日時|||
#D #X 13171% [十七日] z h j u u {s h|h} i c h i n i c h i {|-} # ..|日時|||
#D #Y 13171% [十七日] z h j u u {s h|h} i c h i {n i|ng} c h i {|-} # ..|日時|||

#C #X 18511% [て] {t {o|e}|q t e} {|-} # ..|引用助詞|||

#C #Y 18511% [と] {t {o|e}|q t e} {-} # ..|引用助詞|||
#D #X 18511% [て] {|q} t {o|e} {-} # ..|引用助詞|||
#D #Y 18511% [と] {|q} t {o|e} {-} # ..|引用助詞|||

combination groups including doubles

#CH 10128 #H 10128 15559% [えー] e e {-} # ..|感動詞|||
#C 10128% [ええ] e e {-} # ..|感動詞|||
#CH 10128 11053% [え] e {-} # ..|感動詞|||
#D 10128% [ええ] e {e|} {-} # ..|感動詞|||

#CH 14271 #H 14271 20001% [グラン・セントラル駅] g u r a n g s e n g t o r a r u e k i {-} # ..|固有名詞|||
#C 14271% [グランセントラル駅] g u r a n g s e n g t o r a r u e k i {-} # ..|固有名詞|||
#CH 14271 14287% [グランドセントラル駅] g u r a n g d o s e n g t o r a r u e k i {-} # ..|固有名詞|||
#D 14271% [グランセントラル駅] g u r a n g {|d o} s e n g t o r a r u e k i {-} # ..|固有名詞|||

simple combination groups

#C 10016% [けど] k e d o {|o} {-} # ..|接続助詞|||
#CH 10016 10202% [けれど] k e r e d o {-} # ..|接続助詞|||
#CH 10016 10292% [けれども] k e r e d o m o {-} # ..|接続助詞|||
#CH 10016 10324% [けども] k e d o m o {-} # ..|接続助詞|||
#D 10016% [けど] k e {|r e} d o {|o|m o} {-} # ..|接続助詞|||

#C 10039% [ね] n e {-} # ..|終助詞|||
#CH 10039 11870% [ねえ] n e e {-} # ..|終助詞|||
#D 10039% [ね] n e {|e} {-} # ..|終助詞|||

#C 10040% [はい] {h|} a i {-} # ..|感動詞|||
#CH 10040 14633% [はいつ] h a i {q|-} # ..|感動詞|||
#D 10040% [はい] {h|} a i {|q|-} # ..|感動詞|||

#C 10055% [わたし] w a t a s h i {-} # ..|代名詞|||
#CH 10055 15550% [あたし] a t a s h i {-} # ..|代名詞|||
#D 10055% [わたし] {w|} a t a s h i {-} # ..|代名詞|||

#C 10086% [じゃ] z h j a {-} # ..|接続詞|||
#CH 10086 10393% [じゃあ] z h j a a {-} # ..|接続詞|||
#CH 10086 14700% [じゃっ] z h j a {q|-} # ..|接続詞|||
#D 10086% [じゃ] z h j a {|a} {-}|q} # ..|接続詞|||

#C 10099% [そうですね] s o o d e s u n e {-} # ..|感動詞|||

#CH 10099 12917% [そうですねえ] s o o d e s u n e e {|-} # ..|感動詞|||
#D 10099% [そうですね] s o o d e s u n e {e} {|-} # ..|感動詞|||

#C 10137% [わたくし] w a t a k u s h i {|-} # ..|代名詞|||
#CH 10137 13727% [あたくし] a t a k u s h i {|-} # ..|代名詞|||
#D 10137% [わたくし] {w} a t a k u s h i {|-} # ..|代名詞|||

#C 10260% [百] h j a k u {|-} # ..|接尾辞|||
#CH 10260 11157% [百] b j a k u {|-} # ..|接尾辞|||
#CH 10260 11214% [百] p j a k u {|-} # ..|接尾辞|||
#D 10260% [百] {h|b|p} j a k u {|-} # ..|接尾辞|||

#C 10280% [とても] t o t e m o {|-} # ..|副詞|||
#CH 10280 10908% [とつても] t o t t e m o {|-} # ..|副詞|||
#D 10280% [とても] t o {q} t e m o {|-} # ..|副詞|||

#C 10313% [十] z h j u u {|-} # ..|接尾辞|||
#CH 10313 10107% [十] z h j u q # ..|接尾辞|||
#D 10313% [十] z h j u {u {|-}|q} # ..|接尾辞|||

#C 10427% [さようなら] s a j o o n a r a {|-} # ..|感動詞|||
#CH 10427 11191% [さよなら] s a j o n a r a {|-} # ..|感動詞|||
#D 10427% [さようなら] s a j o {o} n a r a {|-} # ..|感動詞|||

#C 10429% [六] r o k u {|-} # ..|数詞|||
#D 10429% [六] r o {k u|q} {|-} # ..|数詞|||

#C 10491% [申し訳ないのですが] m o o s h i w a k e n a i n o d e s u g a {|-} # ..|副詞|||
#CH 10491 12778% [申し訳ないんですが] m o o s h i w a k e n a i n g d e s u g a {|-} # ..|副詞|||
#D 10491% [申し訳ないのですが] m o o s h i w a k e n a i {n o|ng} d e s u g a {|-} # ..|副詞|||

#C 10496% [本] b o n g {|-} # ..|接尾辞|||
#CH 10496 10500% [本] h o n g {|-} # ..|接尾辞|||
#CH 10496 11664% [本] p o n g {|-} # ..|接尾辞|||
#D 10496% [本] {h|b|p} o n g {|-} # ..|接尾辞|||

#C 10502% [千] s e n g {|-} # ..|接尾辞|||
#CH 10502 11636% [千] z e n g {|-} # ..|接尾辞|||
#D 10502% [千] {s|z} e n g {|-} # ..|接尾辞|||

#C 10520% [十] z h j u u {|-} # ..|数詞|||
#CH 10520 10508% [十] z h j u q # ..|数詞|||

#D 10520% [十] zh j u {u {-}|q} # ..|数詞|||

#C 10529% [いいえ] i i e {-} # ..|感動詞|||

#CH 10529 10537% [いえ] i e {-} # ..|感動詞|||

#CH 10529 14647% [いえつ] i e {q|-} # ..|感動詞|||

#D 10529% [いいえ] {i} i e {q|-} # ..|感動詞|||

#C 10544% [コンピューター] k o n g p j u u t a a {-} # ..|普通名詞|||

#CH 10544 14999% [コンピュータ] k o n g p j u u t a {-} # ..|普通名詞|||

#D 10544% [コンピューター] k o n g p j u u t a {a} {-} # ..|普通名詞|||

#C 10551% [クレジットカード] k u r e z h i q t o k a a d o {-} # ..|普通名詞|||

#CH 10551 15456% [クレジットカード] k u r e d i q t o k a a d o {-} # ..|普通名詞|||

#D 10551% [クレジットカード] k u r e {zh|d} i q t o k a a d o {-} # ..|普通名詞|||

#C 10585% [ロサンゼルス] r o s a n g z e r u s u {-} # ..|固有名詞|||

#CH 10585 11797% [ロサンゼルス] r o s a n g z h e r u s u {-} # ..|固有名詞|||

#CH 10585 17160% [ロスアンゼルス] r o s u a n g z h e r u s u {-} # ..|固有名詞|||

#D 10585% [ロサンゼルス] r o s {u} a n g {z|zh} e r u s u {-} # ..|固有名詞|||

#C 10616% [セント] s e n g t o {-} # ..|接尾辞|||

#CH 10616 17532% [センツ] s e n g t s u {-} # ..|接尾辞|||

#D 10616% [セント] s e n g {t o|t s u} {-} # ..|接尾辞|||

#C 10625% [左側] h i d a r i {g|q k} a w a {-} # ..|普通名詞|||

#CH 10625 17454% [左っかわ] h i d a r i q k a w a {-} # ..|普通名詞|||

#D 10625% [左側] h i d a r i g a w a {-} # ..|普通名詞|||

#C 10668% [分] h u n g {-} # ..|接尾辞|||

#CH 10668 10108% [分] p u n g {-} # ..|接尾辞|||

#D 10668% [分] {h|p} u n g {-} # ..|接尾辞|||

#C 10984% [階] k a i {-} # ..|接尾辞|||

#CH 10984 12689% [階] g a i {-} # ..|接尾辞|||

#D 10984% [階] {k|g} a i {-} # ..|接尾辞|||

#C 11031% [それは] s o r e w a {-} # ..|感動詞|||

#CH 11031 14266% [そりゃ] s o r j a {-} # ..|感動詞|||

#D 11031% [それは] s o r {e w|j} a {-} # ..|感動詞|||

#C 11059% [ても] t e m o {-} # ..|接続助詞|||

#CH 11059 17439% [つても] q t e m o {-} # ..|接続助詞|||

#D 11059% [ても] { |q} t e m o { | - } # .. | 接続助詞 | | | |

#C 11370% [エイチ] e i c h i { | - } # .. | 普通名詞 | | | |

#CH 11370 15563% [エッチ] e q c h i { | - } # .. | 普通名詞 | | | |

#D 11370% [エイチ] e { i | q } c h i { | - } # .. | 普通名詞 | | | |

#C 11429% [な] n a { | - } # .. | 終助詞 | | | |

#CH 11429 12698% [なあ] n a a { | - } # .. | 終助詞 | | | |

#D 11429% [な] n a { | a } { | - } # .. | 終助詞 | | | |

#C 11510% [アドリアン] a d o r i a n g { | - } # .. | 固有名詞 | | | |

#CH 11510 16929% [エイドリアン] e e d o r i a n g { | - } # .. | 固有名詞 | | | |

#D 11510% [アドリアン] { a | e e } d o r i a n g { | - } # .. | 固有名詞 | | | |

#C 11515% [泊] h a k u { | - } # .. | 接尾辞 | | | |

#CH 11515 10036% [泊] p a k u { | - } # .. | 接尾辞 | | | |

#D 11515% [泊] { h | p } a k u { | - } # .. | 接尾辞 | | | |

#C 11554% [なくてはならぬ] n a k u t e w a n a r a n a # .. | 助動詞 | 形容詞 | 語幹 | |

#CH 11554 15410% [なきやなんな] n a k j a n a n g n a # .. | 助動詞 | 形容詞 | 語幹 | |

#CH 11554 15494% [なくちゃならぬ] n a k u c h j a n a r a n a # .. | 助動詞 | 形容詞 | 語幹 | |

#D 11554% [なくてはならぬ] n a k { u t e w a n a r a | u c h j a n a r a | j a n a n g } n a # .. | 助動詞 | 形容詞 | 語幹 | |

#C 11594% [申し訳ないんですけれども] m o o s h i w a k e n a i n g d e s u k e r e d o m o { | - } # .. | 副詞 | | | |

#CH 11594 15224% [申し訳ないんですけれども] m o o s h i w a k e n a i d e s u k e r e d o m o { | - } # .. | 副詞 | | | |

#D 11594% [申し訳ないんですけれども] m o o s h i w a k e n a i { n g } d e s u k e r e d o m o { | - } # .. | 副詞 | | | |

#C 11652% [品] p i n g { | - } # .. | 接尾辞 | | | |

#CH 11652 17917% [品] h i n g { | - } # .. | 接尾辞 | | | |

#D 11652% [品] { p | h } i n g { | - } # .. | 接尾辞 | | | |

#C 11853% [いや] i j a { | - } # .. | 感動詞 | | | |

#CH 11853 20193% [や] j a { | - } # .. | 感動詞 | | | |

#D 11853% [いや] { i } j a { | - } # .. | 感動詞 | | | |

#C 12061% [ずうっと] z u u q t o { | - } # .. | 副詞 | | | |

#CH 12061 19155% [ずうと] z u q t o { | - } # .. | 副詞 | | | |

#D 12061% [ずうっと] z u { u } q t o { | - } # .. | 副詞 | | | |

#C 12441% [コンシェルジュ] k o n g s h e r u z h e { | - } # .. | 普通名詞 | | | |

#CH 12441 16136% [コンシェールジュ] k o n g s h e e r u z h j u { | - } # .. | 普通名詞 | | | |

#D 12441% [コンシェルジュ] k o n g s h e { | e } r u z h { e | j u } { | - } # .. | 普通名詞 | | | |

#C 12450% [軒] k e ng {|-} # ..|接尾辞|||
 #CH 12450 19916% [軒] g e ng {|-} # ..|接尾辞|||
 #D 12450% [軒] {k|g} e ng {|-} # ..|接尾辞|||

#C 12473% [そしたら] s o sh i t a r a {|-} # ..|接続詞|||
 #CH 12473 15409% [そうしたら] s o o sh i t a r a {|-} # ..|接続詞|||
 #D 12473% [そしたら] s o {|o} sh i t a r a {|-} # ..|接続詞|||

#C 12895% [でい] d e i # ..|助動詞|一段|語幹||
 #CH 12895 11092% [で] d e {|-} # ..|助動詞|一段|語幹||
 #D 12895% [でい] d e {i|} # ..|助動詞|一段|語幹||

#C 12929% [どうも] d o o m o {|-} # ..|感動詞|||
 #CH 12929 14701% [ども] d o m o {|-} # ..|感動詞|||
 #D 12929% [どうも] d o {o|} m o {|-} # ..|感動詞|||

#C 12968% [ランドリーバッグ] r a n g d o r i i b a q g u {|-} # ..|普通名詞|||
 #CH 12968 14613% [ランドリーバッグ] r a n g d o r i i b a q k u {|-} # ..|普通名詞|||
 #D 12968% [ランドリーバッグ] r a n g d o r i i b a q {g|k} u {|-} # ..|普通名詞|||

#C 13084% [なるほど] n a r u h o d o {|-} # ..|感動詞|||
 #CH 13084 17803% [なるほどお] n a r u h o d o o {|-} # ..|感動詞|||
 #D 13084% [なるほど] n a r u h o d o {o|} {|-} # ..|感動詞|||

#C 13224% [ユナイテッド] j u n a i t e q d o {|-} # ..|固有名詞|||
 #CH 13224 14988% [ユナイテイト] j u n a i d i q t o {|-} # ..|固有名詞|||
 #D 13224% [ユナイテッド] j u n a i {t e q d|d i g t} o {|-} # ..|固有名詞|||

#C 13600% [まあ] m a a {|-} # ..|副詞|||
 #CH 13600 13999% [ま] m a {|-} # ..|副詞|||
 #D 13600% [まあ] m a {a|} {|-} # ..|副詞|||

#C 13646% [七月] s h i c h i g a t s u {|-} # ..|日時|||
 #CH 13646 11842% [七月] h i c h i g a t s u {|-} # ..|日時|||
 #D 13646% [七月] {sh|h} i c h i g a t s u {|-} # ..|日時|||

#C 13803% [万が一] m a n g g a i c h i {|-} # ..|副詞|||
 #CH 13803 15622% [万一] m a n g i c h i {|-} # ..|副詞|||
 #D 13803% [万が一] m a n g {g a|} i c h i {|-} # ..|副詞|||

#C 14010% [ワシントンシェラトンホテル] w a s h i n g t o n g s h e r a t o n g h o t e r u {|-} # ..|固有名詞|||

#CH 14010 14016% [ワシントンシュラトンホテル] w a s h i n g t o n g s h j u r a t o n g h o t e r u {|-} # ..|固有名詞|||
 #D 14010% [ワシントンシュラトンホテル] w a s h i n g t o n g s h {e|j u} r a t o n g h o t e r u {|-} # ..|固有名詞|||

#C 14593% [コンチネンタルブレックファースト] k o n g c h i n e n g t a r u b u r e q k u h a a s u t o {|-} # ..|普通名詞|||
 #CH 14593 14606% [コンチネンタルブレックファースト] k o n g c h i n e n g t a r u b u r e q k u h a s u t o {|-} # ..|普通名詞|||
 #D 14593% [コンチネンタルブレックファースト] k o n g c h i n e n g t a r u b u r e q k u h a {a|} s u t o {|-} # ..|普通名詞|||

#C 14599% [イングリッシュスタイル] i n g g u r i s h j u s u t a i r u {|-} # ..|普通名詞|||
 #CH 14599 14601% [イングリッシュスタイル] i n g g u r i q s h j u s u t a i r u {|-} # ..|普通名詞|||
 #D 14599% [イングリッシュスタイル] i n g g u r i {|q} s h j u s u t a i r u {|-} # ..|普通名詞|||

#C 14619% [ブルックリン美術館] b u r u q k u r i n g b i z h j u t s u k a n g {|-} # ..|固有名詞|||
 #CH 14619 14620% [ブリュックリン美術館] b u r j u q k u r i n g b i z h j u t s u k a n g {|-} # ..|固有名詞|||
 #D 14619% [ブルックリン美術館] b u r {|j} u q k u r i n g b i z h j u t s u k a n g {|-} # ..|固有名詞|||

#C 14716% [ユナイテッドエアライン] j u n a i t e q d o e a r a i n g {|-} # ..|固有名詞|||
 #CH 14716 14834% [ユナイテッドエアライン] j u n a i q t e d o e a r a i n g {|-} # ..|固有名詞|||
 #D 14716% [ユナイテッドエアライン] j u n a i {t e q|q t e} d o e a r a i n g {|-} # ..|固有名詞|||

#C 14827% [ておけ] t e o k e {|-} # ..|助動詞|一段|語幹||
 #CH 14827 14809% [とけ] t o k e {|-} # ..|助動詞|一段|語幹||
 #D 14827% [ておけ] t {e|} o k e {|-} # ..|助動詞|一段|語幹||

#C 14914% [すいませんけれども] s u {m|} i m a s e n g k e r e d o m o {|-} # ..|副詞|||
 #CH 14914 20055% [すいませんけれども] s u m i m a s e n g k e d o m o {|-} # ..|副詞|||
 #D 14914% [すいませんけれども] s u {m|} i m a s e n g k e {r e|} d o m o {|-} # ..|副詞|||

#C 15119% [しかたがな] s h i k a t a g a n a # ..|形容詞|形容詞|語幹||
 #CH 15119 10692% [しかたな] s h i k a t a n a # ..|形容詞|形容詞|語幹||
 #D 15119% [しかたがな] s h i k a t a {g a|} n a # ..|形容詞|形容詞|語幹||

#C 15514% [コテージ] k o t e e z h i {|-} # ..|普通名詞|||
 #CH 15514 15657% [コテージ] k o q t e e z h i {|-} # ..|普通名詞|||
 #D 15514% [コテージ] k o {|q} t e e z h i {|-} # ..|普通名詞|||

#C 15638% [アスレチック] a s u r e c h i q k u {|-} # ..|普通名詞|||
 #CH 15638 16731% [アスレティック] a s u r e t i q k u {|-} # ..|普通名詞|||
 #D 15638% [アスレチック] a s u r e {c h|t} i q k u {|-} # ..|普通名詞|||

#C 15658% [プリンスコテージ] p u r i n g s u k o q t e e z h i {|-} # ..|固有名詞|||
 #CH 15658 15654% [プリンスコテージ] p u r i n g s u k o t e e z h i {|-} # ..|固有名詞|||
 #D 15658% [プリンスコテージ] p u r i n g s u k o {q|} t e e z h i {|-} # ..|固有名詞|||

#C 15666% [神山コテージ] k a m i j a m a k o q t e e z h i {|-} # ..|固有名詞|||
 #CH 15666 15665% [神山コテージ] k a m i j a m a k o t e e z h i {|-} # ..|固有名詞|||
 #D 15666% [神山コテージ] k a m i j a m a k o {q|} t e e z h i {|-} # ..|固有名詞|||

 #C 15763% [マジソンホテル] m a z h i s o n g h o t e r u {|-} # ..|固有名詞|||
 #CH 15763 15769% [マディソンホテル] m a d i s o n g h o t e r u {|-} # ..|固有名詞|||
 #D 15763% [マジソンホテル] m a {zh|d} i s o n g h o t e r u {|-} # ..|固有名詞|||

 #C 15838% [バッジ] b a q z h i {|-} # ..|普通名詞|||
 #CH 15838 15843% [バッヂ] b a q c h i {|-} # ..|普通名詞|||
 #D 15838% [バッジ] b a q {zh|ch} i {|-} # ..|普通名詞|||

 #C 16112% [シャンパン] s h j a n g p a n g {|-} # ..|普通名詞|||
 #CH 16112 15724% [シャンペン] s h j a n g p e n g {|-} # ..|普通名詞|||
 #CH 16112 15746% [シャンペーン] s h j a n g p e e n g {|-} # ..|普通名詞|||
 #D 16112% [シャンパン] s h j a n g p {a|e|e} n g {|-} # ..|普通名詞|||

 #C 16380% [ダイレクトコール] d a i r e k u t o k o o r u {|-} # ..|普通名詞|||
 #CH 16380 16388% [ダイレクトコール] d i r e k u t o k o o r u {|-} # ..|普通名詞|||
 #D 16380% [ダイレクトコール] d {a|} i r e k u t o k o o r u {|-} # ..|普通名詞|||

 #C 16460% [なくてはいけな] n a k u t e w a i k e n a # ..|助動詞|形容詞|語幹||
 #CH 16460 11604% [なきゃいけな] n a k j a i k e n a # ..|助動詞|形容詞|語幹||
 #CH 16460 15496% [なくちゃいけな] n a k u c h j a i k e n a # ..|助動詞|形容詞|語幹||
 #D 16460% [なくてはいけな] n a k {u t e w|u c h j|j} a i k e n a # ..|助動詞|形容詞|語幹||

 #C 16759% [ハッピーバースデー] h a q p i i b a a s u d e e {|-} # ..|感動詞|||
 #CH 16759 16798% [ハッピーバースデー] h a q p i b a a s u d e e {|-} # ..|感動詞|||
 #D 16759% [ハッピーバースデー] h a q p i {i|} b a a s u d e e {|-} # ..|感動詞|||

 #C 16764% [ロングヘア] r o n g g u h e a {|-} # ..|普通名詞|||
 #CH 16764 16804% [ロングヘア-] r o n g g u h e a a {|-} # ..|普通名詞|||
 #D 16764% [ロングヘア] r o n g g u h e a {a} {|-} # ..|普通名詞|||

 #C 16826% [十一分] z h j u u i q p u n g {|-} # ..|日時|||
 #CH 16826 16825% [十一分] z h j u u i c h i h u n g {|-} # ..|日時|||
 #D 16826% [十一分] z h j u u i {q p|c h i h} u n g {|-} # ..|日時|||

 #C 16828% [三十八分] s a n g z h j u u h a c h i h u n g {|-} # ..|日時|||
 #CH 16828 16853% [三十八分] s a n g z h j u u h a q p u n g {|-} # ..|日時|||
 #D 16828% [三十八分] s a n g z h j u u h a {c h i h|q p} u n g {|-} # ..|日時|||

#C 16866% [なくてはいけません] nakutewa ikemaseng {-} # ..|助動詞|特殊|終止||
 #CH 16866 17336% [なくちゃいけません] nakuch ja ikemaseng {-} # ..|助動詞|特殊|終止||
 #D 16866% [なくてはいけません] naku {tew|ch j} a ikemaseng {-} # ..|助動詞|特殊|終止||

 #C 17016% [エクササイズマシン] ekusasaizumashining {-} # ..|普通名詞|||
 #CH 17016 17027% [エクササイズマシン] ekusasaizumashining {-} # ..|普通名詞|||
 #D 17016% [エクササイズマシン] ekusasaizumashi {i} ng {-} # ..|普通名詞|||

 #C 17061% [ロサンジェルス支店] rosang zherusushiteng {-} # ..|固有名詞|||
 #CH 17061 17036% [ロサンゼルス支店] rosang zerusushiteng {-} # ..|固有名詞|||
 #D 17061% [ロサンジェルス支店] rosang {zh|z} erusushiteng {-} # ..|固有名詞|||

 #C 17099% [桔梗の間] kikjoonoma {-} # ..|固有名詞|||
 #CH 17099 16540% [桔梗間] kikjooma {-} # ..|固有名詞|||
 #D 17099% [桔梗の間] kikjoonoma {-} # ..|固有名詞|||

 #C 17156% [エアーパシフィック] eaapashihiku {-} # ..|固有名詞|||
 #CH 17156 17159% [エアパシフィック] eapashihiku {-} # ..|固有名詞|||
 #D 17156% [エアーパシフィック] ea {a} pashihiku {-} # ..|固有名詞|||

 #C 17360% [リバティー島] ribatitoo {-} # ..|固有名詞|||
 #CH 17360 17374% [リバティ島] ribatitoo {-} # ..|固有名詞|||
 #D 17360% [リバティー島] ribati {i} too {-} # ..|固有名詞|||

 #C 17443% [向こう側] mukoo gawa {-} # ..|普通名詞|||
 #CH 17443 17455% [向こうっかわ] mukooqkawa {-} # ..|普通名詞|||
 #D 17443% [向こう側] mukoo {g|q k} awa {-} # ..|普通名詞|||

 #C 17453% [向かいっかわ] mukaiqkawa {-} # ..|普通名詞|||
 #CH 17453 18706% [向かえっ側] mukai gawa {-} # ..|普通名詞|||
 #D 17453% [向かいっかわ] mukai {q k|g} awa {-} # ..|普通名詞|||

 #C 17467% [フリーマイリッジ] hurimaiirizhi {-} # ..|普通名詞|||
 #CH 17467 17468% [フリーマイレッジ] hurimairerizhi {-} # ..|普通名詞|||
 #D 17467% [フリーマイリッジ] hurimaiir {i|e} qzhi {-} # ..|普通名詞|||

 #C 17513% [オートマチック] ootomachiku {-} # ..|普通名詞|||
 #CH 17513 17522% [オートマテック] ootomatiku {-} # ..|普通名詞|||
 #D 17513% [オートマチック] ootoma {ch|t} iku {-} # ..|普通名詞|||

 #C 17537% [オートマテック車] ootomatiku shja {-} # ..|普通名詞|||
 #CH 17537 17564% [オートマチック車] ootomachiku shja {-} # ..|普通名詞|||

#D 17537% [オートマチック車] o o t o m a {t|ch} i q k u s h j a {|-} # ..|普通名詞|||

#C 17576% [ウインドーショッピング] u i n g d o o s h j o q p i n g g u {|-} # ..|普通名詞|||

#CH 17576 17590% [ウインドショッピング] u i n g d o s h j o q p i n g g u {|-} # ..|普通名詞|||

#D 17576% [ウインドーショッピング] u i n g d o {o|} s h j o q p i n g g u {|-} # ..|普通名詞|||

#C 17718% [グレートソルトレイク] g u r e e t o s o r u t o r e e k u {|-} # ..|固有名詞|||

#CH 17718 17710% [グレードソルトレイク] g u r e e d o s o r u t o r e e k u {|-} # ..|固有名詞|||

#D 17718% [グレートソルトレイク] g u r e e {t|d} o s o r u t o r e e k u {|-} # ..|固有名詞|||

#C 17773% [日本ヘリコプター] n i h o n g h e r i k o p u t a a {|-} # ..|固有名詞|||

#CH 17773 17786% [日本ヘリコプター] n i q p o n g h e r i k o p u t a a {|-} # ..|固有名詞|||

#D 17773% [日本ヘリコプター] n i {h|q p} o n g h e r i k o p u t a a {|-} # ..|固有名詞|||

#C 17876% [マジステイック劇場] m a z h e s u t i q k u g e k i z h j o o {|-} # ..|固有名詞|||

#CH 17876 17877% [マジステイック劇場] m a z h i s u t i q k u g e k i z h j o o {|-} # ..|固有名詞|||

#D 17876% [マジステイック劇場] m a z h {e|i} s u t i q k u g e k i z h j o o {|-} # ..|固有名詞|||

#C 17957% [もの] m o n o {|-} # ..|終助詞|||

#CH 17957 15419% [もん] m o n g {|-} # ..|終助詞|||

#D 17957% [もの] m o {n o|ng} {|-} # ..|終助詞|||

#C 18486% [デビッド] d e b i q d o {|-} # ..|人名|||

#CH 18486 18535% [デービット] d e e b i q t o {|-} # ..|人名|||

#CH 18486 19541% [デビット] d e b i q t o {|-} # ..|人名|||

#D 18486% [デビッド] d e {e|} b i q {d|t} o {|-} # ..|人名|||

#C 19163% [余り] a m a r i {|-} # ..|副詞|||

#CH 19163 11005% [あんまり] a n g m a r i {|-} # ..|副詞|||

#D 19163% [余り] a {ng} m a r i {|-} # ..|副詞|||

#C 19684% [ニューワーク] n j u u w a a k u {|-} # ..|固有名詞|||

#CH 19684 17168% [ニューアーク] n j u u a a k u {|-} # ..|固有名詞|||

#D 19684% [ニューワーク] n j u u {w|} a a k u {|-} # ..|固有名詞|||

#C 20191% [あのね] a n o n e {|-} # ..|感動詞|||

#CH 20191 20222% [あのねえ] a n o n e e {|-} # ..|感動詞|||

#D 20191% [あのね] a n o n e {e} {|-} # ..|感動詞|||

#C 20296% [ジョッキークラブ] z h j o q k i i k u r a b u {|-} # ..|固有名詞|||

#CH 20296 20292% [ジョッキークラブ] z h j o q k i k u r a b u {|-} # ..|固有名詞|||

#D 20296% [ジョッキークラブ] z h j o q k i {i|} k u r a b u {|-} # ..|固有名詞|||

#####

COMBINATIONS (interjections)

combination groups including errors

#FH 11280 #X #H 11280 13651% [へえー] h e e {-} # ..|間投詞|||
#FH 11280 #Y 13651% [へえー] h e e e {-} # ..|間投詞|||
#F 11280% [へー] h e e {-} # ..|間投詞|||
#G 11280% [へー] h e e {e} {-} # ..|間投詞|||

combination groups including doubles

#FH 10005 #H 10005 10098% [あー] a a {-} # ..|間投詞|||
#FH 10005 #H 10005 10648% [ああ] a a {-} # ..|間投詞|||
#F 10005% [あ] a {a} {-} # ..|間投詞|||
#FH 10005 10526% [あっ] a {q|-} # ..|間投詞|||
#G 10005% [あ] a {a} {q|-} # ..|間投詞|||

#FH 10017 #H 10017 10203% [え] e {e} {-} # ..|間投詞|||
#FH 10017 #H 10017 11802% [ええ] e e {-} # ..|間投詞|||
#FH 10017 #H 10017 14603% [エー] e e {-} # ..|間投詞|||
#FH 10017 #H 14772 17936% [ええっ] e e q # ..|間投詞|||
#F 10017% [えー] e {e} {-} # ..|間投詞|||
#FH 10017 10917% [えっ] e {q|-} # ..|間投詞|||
#FH 10017 14772% [えーっ] e e {q|-} # ..|間投詞|||
#FH 10017 15094% [えーえ] e e e {-} # ..|間投詞|||
#FH 10017 19800% [えーえっ] e e e q # ..|間投詞|||
#G 10017% [えー] e {e|e e} {q|-} # ..|間投詞|||

#FH 10934 #H 10934 11019% [んー] ng {-} # ..|間投詞|||
#FH 10934 #H 11291 11826% [うん] u ng {-} # ..|間投詞|||
#FH 10934 #H 15776 15592% [ふうーん] h u u ng {-} # ..|間投詞|||
#F 10934% [ん] ng {-} # ..|間投詞|||
#FH 10934 11291% [うーん] u {u} ng {-} # ..|間投詞|||
#FH 10934 12054% [う] u {-} # ..|間投詞|||
#FH 10934 13811% [ふん] h u ng {-} # ..|間投詞|||
#FH 10934 15586% [うー] u u {-} # ..|間投詞|||
#FH 10934 15776% [ふーん] h u u ng {-} # ..|間投詞|||
#FH 10934 16538% [ふう] h u u {-} # ..|間投詞|||
#FH 10934 19556% [んっ] ng q # ..|間投詞|||
#G 10934% [ん] {(h u u|u u|u)}{|h u} {|u|u u} ng) {q|-} # ..|間投詞|||

#FH 10644 #H 10644 11060% [まあ] m a a {|-} # ..|間投詞|||
 #F 10644% [まー] m a a {|-} # ..|間投詞|||
 #FH 10644 11993% [ま] m a {|-} # ..|間投詞|||
 #FH 10644 14804% [まつ] m a q # ..|間投詞|||
 #G 10644% [まー] m a {a} {|q|-} # ..|間投詞|||

#FH 10263 #H 10365 11875% [そうですねえ] s o o d e s u n e e {|-} # ..|間投詞|||
 #FH 10263 #H 10263 15283% [そーですね] s o o d e s u n e {|-} # ..|間投詞|||
 #F 10263% [そうですね] s o o d e s u n e {|-} # ..|間投詞|||
 #FH 10263 10365% [そうですねー] s o o d e s u n e e {|-} # ..|間投詞|||
 #FH 10263 15244% [そうですねえー] s o o d e s u n e e e {|-} # ..|間投詞|||
 #G 10263% [そうですね] s o o d e s u n e {|e|e|e} {|-} # ..|間投詞|||

#FH 10054 #H 10227 12488% [あのう] a n o o {|-} # ..|間投詞|||
 #F 10054% [あの] a n o {|u} {|-} # ..|間投詞|||
 #FH 10054 10227% [あの一] a n o o {|-} # ..|間投詞|||
 #FH 10054 14402% [あの一の一] a a n o o {|-} # ..|間投詞|||
 #G 10054% [あの] a {|a} n o {|o|u} {|-} # ..|間投詞|||

#FH 14506 #H 14506 14508% [はあ] h a a {|-} # ..|間投詞|||
 #F 14506% [はー] h a a {|-} # ..|間投詞|||
 #FH 14506 13272% [は] h a {|-} # ..|間投詞|||
 #FH 14506 15246% [はっ] h a {q|-} # ..|間投詞|||
 #G 14506% [はー] h a {a} {|q|-} # ..|間投詞|||

#FH 12870 #H 12870 15753% [でえ] d e e {|-} # ..|間投詞|||
 #F 12870% [でー] d e e {|-} # ..|間投詞|||
 #FH 12870 11836% [で] d e {|-} # ..|間投詞|||
 #G 12870% [でー] d e {|e} {|-} # ..|間投詞|||

#FH 12359 #H 12359 16081% [んーと] n g t o {|-} # ..|間投詞|||
 #FH 12359 #H 15784 16088% [んーとー] n g t o o {|-} # ..|間投詞|||
 #F 12359% [んと] n g t o {|-} # ..|間投詞|||
 #FH 12359 11986% [うーんと] u u n g t o {|-} # ..|間投詞|||
 #FH 12359 14923% [んつと] n g q t o {|-} # ..|間投詞|||
 #FH 12359 15768% [うんと] u n g t o {|-} # ..|間投詞|||
 #FH 12359 15784% [んとー] n g t o o {|-} # ..|間投詞|||
 #FH 12359 16154% [うーんとー] u u n g t o o {|-} # ..|間投詞|||
 #FH 12359 17907% [うんとー] u n g t o o {|-} # ..|間投詞|||
 #G 12359% [んと] {|u|u u} n g {|q} t o {|o} {|-} # ..|間投詞|||

#FH 16441 #H 14784 16487% [そうでございますねえ] s o o d e g o z a i m a s u n e e {|-} # ..|間投詞|||

#F 16441% [そうでございますね] soodegozaimasune {-} # ..|間投詞|||
 #FH 16441 14784% [そうでございますねー] soodegozaimasunee {-} # ..|間投詞|||
 #G 16441% [そうでございますね] soodegozaimasune {e} {-} # ..|間投詞|||

 #FH 10498 #H 11962 17166% [えつーとー] eqtoo {-} # ..|間投詞|||
 #FH 10498 #H 10498 17373% [ええと] eeto {-} # ..|間投詞|||
 #FH 10498 #H 11183 18064% [ええつと] eeeto {-} # ..|間投詞|||
 #FH 10498 11462% [えと] eto {-} # ..|間投詞|||
 #F 10498% [えーと] ee {-} to {-} # ..|間投詞|||
 #FH 10498 11033% [えつと] eqto {-} # ..|間投詞|||
 #FH 10498 11183% [えーつと] ee{q|-} to {-} # ..|間投詞|||
 #FH 10498 11962% [えつとー] eqtoo {-} # ..|間投詞|||
 #FH 10498 13657% [えーとー] eeto {-} # ..|間投詞|||
 #FH 10498 14189% [えとー] etoo {-} # ..|間投詞|||
 #FH 10498 14806% [えーつとー] eeeto {-} # ..|間投詞|||
 #FH 10498 17541% [ええーと] eeeto {-} # ..|間投詞|||
 #G 10498% [えーと] e{e|e} {q|-} to {o} {-} # ..|間投詞|||

 #FH 11819 #H 11819 17281% [おう] oo {-} # ..|間投詞|||
 #F 11819% [おー] oo {-} # ..|間投詞|||
 #FH 11819 10408% [お] o {-} # ..|間投詞|||
 #G 11819% [おー] o{o} {-} # ..|間投詞|||

 #FH 10950 #H 10950 20205% [やあ] jaa {-} # ..|間投詞|||
 #F 10950% [いや] ija {-} # ..|間投詞|||
 #FH 10950 11297% [いやー] ija {-} # ..|間投詞|||
 #FH 10950 18030% [やー] jaa {-} # ..|間投詞|||
 #FH 10950 19667% [や] ja {-} # ..|間投詞|||
 #G 10950% [いや] {i} ja {a} {-} # ..|間投詞|||

 #FH 14203 #R 10005 10017 #/ 19553% [あーえー] aae {-} # ..|間投詞|||
 #FH 14203 #R 10005 10017 #/ 19576% [あつえー] aqee {-} # ..|間投詞|||
 #FH 14203 #R 10005 10017 #/ 19758% [あーえ] aae {-} # ..|間投詞|||
 #F #R 10005 10017 #/ 14203% [あえー] aee {-} # ..|間投詞|||
 #G #R 10005 10017 #/ 14203% [あえー] a{q|a} e{e} {-} # ..|間投詞|||

 #FH 14270 #R 10017 12870 #/ 14774% [えでー] edee {-} # ..|間投詞|||
 #F #R 10017 12870 #/ 14270% [えーで] eede {-} # ..|間投詞|||
 #G #R 10017 12870 #/ 14270% [えーで] e{e} de{e} {-} # ..|間投詞|||

 #FH 14331 #R 10005 10498 #/ 19652% [あつえーとー] aqeeto {-} # ..|間投詞|||
 #FH 14331 #R 10005 10498 #/ 19936% [あつえーと] aqeto {-} # ..|間投詞|||

#F #R 10005 10498 #/ 14331% [あえーと] a e e t o {-} # ..|間投詞|||
 #G #R 10005 10498 #/ 14331% [あえーと] a {q} e e t o {o} {-} # ..|間投詞|||

 #FH 14665 #R 10054 10017 #/ 19944% [あのえー] a n o e e {-} # ..|間投詞|||
 #F #R 10054 10017 #/ 14665% [あのえー] a n o o {-} e e {-} # ..|間投詞|||
 #G #R 10054 10017 #/ 14665% [あのえー] a n o {o} {-} e e {-} # ..|間投詞|||

 #FH 14748 #R 13107 10017 #/ 19820% [とえー] t o e e {-} # ..|間投詞|||
 #F #R 13107 10017 #/ 14748% [ととえ] {q|-} t o e {-} # ..|間投詞|||
 #G #R 13107 10017 #/ 14748% [ととえ] {q|-} t o e {e} {-} # ..|間投詞|||

 #FH 14755 #R 10005 10263 #/ 19802% [あっそうですねえー] a q s o o d e s u n e e e {-} # ..|間投詞|||
 #F #R 10005 10263 #/ 14755% [あっそうですね] a s o o d e s u n e {-} # ..|間投詞|||
 #G #R 10005 10263 #/ 14755% [あっそうですね] a {q} s o o d e s u n e {e} {-} # ..|間投詞|||

 #FH 14913 #R 10054 10005 #/ 19905% [あのーあつ] a n o o a q # ..|間投詞|||
 #F #R 10054 10005 #/ 14913% [あのーあ] a n o o a {-} # ..|間投詞|||
 #G #R 10054 10005 #/ 14913% [あのーあ] a n o o a {q|-} # ..|間投詞|||

 #FH 15121 #R 10017 10263 #/ 15199% [えーそうですねー] e e s o o d e s u n e e {-} # ..|間投詞|||
 #FH 15121 #R 10017 10263 #/ 15293% [えーそうですね] e s o o d e s u n e {-} # ..|間投詞|||
 #FH 15121 #R 10017 10263 #/ 19799% [えーそうですねえー] e e s o o d e s u n e e e {-} # ..|間投詞|||
 #FH 15121 #R 10017 10263 #/ 19805% [ええーそうですね] e e e s o o d e s u n e {-} # ..|間投詞|||
 #FH 15121 #R 10017 10263 #/ 20214% [えーっそうですね] e e q s o o d e s u n e {-} # ..|間投詞|||
 #F #R 10017 10263 #/ 15121% [えーそうですね] e e s o o d e s u n e {-} # ..|間投詞|||
 #G #R 10017 10263 #/ 15121% [えーそうですね] e {e|e|e} {q} s o o d e s u n e {e|e|e} {-} # ..|間投詞|||

 #FH 15148 #R 10498 10263 #/ 15147% [えーつとそうですねー] e e q t o s o o d e s u n e e {-} # ..|間投詞|||
 #FH 15148 #R 10498 10263 #/ 19679% [えーとーそうですね] e e t o o s o o d e s u n e {-} # ..|間投詞|||
 #F #R 10498 10263 #/ 15148% [えーつとそうですね] e e q t o s o o d e s u n e {-} # ..|間投詞|||
 #G #R 10498 10263 #/ 15148% [えーつとそうですね] e e {q} t o {o} s o o d e s u n e {e} {-} # ..|間投詞|||

 #FH 15241 #R 10017 10054 #/ 14184% [えあのー] e a n o o {-} # ..|間投詞|||
 #FH 15241 #R 10017 10054 #/ 15101% [えーあの一] e e a n o o {-} # ..|間投詞|||
 #F #R 10017 10054 #/ 15241% [えーあの一] e e {-} a n o {-} # ..|間投詞|||
 #G #R 10017 10054 #/ 15241% [えーあの一] e {e} {-} a n o {o} {-} # ..|間投詞|||

 #FH 15284 #R 10498 10054 #/ 20120% [えつとあの一] e q t o a n o {-} # ..|間投詞|||
 #F #R 10498 10054 #/ 15284% [えーとあの一] e e t o a n o {-} # ..|間投詞|||
 #G #R 10498 10054 #/ 15284% [えーとあの一] e {e|q} t o a n o {-} # ..|間投詞|||

 #FH 15320 #R 10005 13227 #/ 19714% [あっえーとですな] a q e e t o d e s u n e {-} # ..|間投詞|||

#F #R 10005 13227 #/ 15320% [あーえーとですねー] a a e e t o d e s u n e e {|-} # ..|間投詞|||
 #G #R 10005 13227 #/ 15320% [あーえーとですねー] a {a|q} e e t o d e s u n e {|e} {|-} # ..|間投詞|||
 #FH 17196 #R 10005 11823 #/ 20223% [あーそー] a a s o o {|-} # ..|間投詞|||
 #F #R 10005 11823 #/ 17196% [あそう] a s o o {|-} # ..|間投詞|||
 #G #R 10005 11823 #/ 17196% [あそう] a {|a} s o o {|-} # ..|間投詞|||
 #FH 19666 #R 11819 10017 #/ 20179% [おえー] o e e {|-} # ..|間投詞|||
 #F #R 11819 10017 #/ 19666% [おえー] o o e e {|-} # ..|間投詞|||
 #G #R 11819 10017 #/ 19666% [おえー] o {|o} e e {|-} # ..|間投詞|||
 #FH 19677 #R 13227 10054 #/ 19706% [えーとですねあのー] e e t o d e s u n e a n o o {|-} # ..|間投詞|||
 #F #R 13227 10054 #/ 19677% [えーとですねあの] e e t o d e s u n e a n o {|-} # ..|間投詞|||
 #G #R 13227 10054 #/ 19677% [えーとですねあの] e e t o d e s u n e a n o {|o} {|-} # ..|間投詞|||
 #FH 19930 #R 10005 10663 #/ 19926% [あじゃ] a z h j a {|-} # ..|間投詞|||
 #FH 19930 #R 10005 10663 #/ 19927% [あーじゃあ] a a z h j a a {|-} # ..|間投詞|||
 #F #R 10005 10663 #/ 19930% [あーじゃ] a a z h j a {|-} # ..|間投詞|||
 #G #R 10005 10663 #/ 19930% [あーじゃ] a {|a} z h j a {|a} {|-} # ..|間投詞|||
 #FH 20226 #R 10950 10054 #/ 20225% [やーあの] j a a a n o {|-} # ..|間投詞|||
 #F #R 10950 10054 #/ 20226% [いやあの] i j a a n o {|-} # ..|間投詞|||
 #G #R 10950 10054 #/ 20226% [いやあの] {|i} j a {|a} a n o {|-} # ..|間投詞|||
 #F 10435% [いー] i i {|-} # ..|間投詞|||
 #FH 10435 10941% [い] i {|-} # ..|間投詞|||
 #G 10435% [いー] i {|i} {|-} # ..|間投詞|||
 #F 10663% [じゃ] z h j a {|-} # ..|間投詞|||
 #FH 10663 15083% [じゃあ] z h j a a {|-} # ..|間投詞|||
 #G 10663% [じゃ] z h j a {|a} {|-} # ..|間投詞|||
 #F 11855% [もう] m o o {|-} # ..|間投詞|||
 #FH 11855 15125% [も] m o {|-} # ..|間投詞|||
 #G 11855% [もう] m o {|o} {|-} # ..|間投詞|||
 #F 12518% [その] s o n o {|-} # ..|間投詞|||
 #FH 12518 19301% [そのー] s o n o o {|-} # ..|間投詞|||
 #G 12518% [その] s o n o {|o} {|-} # ..|間投詞|||
 #F 12633% [あーと] a a t o {|-} # ..|間投詞|||
 #FH 12633 14173% [あーつと] a a q t o {|-} # ..|間投詞|||

#G 12633% [あーと] a a { |q } t o { | - } # .. | 間投詞 | | | |

#F 12838% [さようでございますね] s a j o o d e g o z a i m a s u n e { | - } # .. | 間投詞 | | | |

#FH 12838 12816% [さようでございますねえ] s a j o o d e g o z a i m a s u n e e { | - } # .. | 間投詞 | | | |

#G 12838% [さようでございますね] s a j o o d e g o z a i m a s u n e { | e } { | - } # .. | 間投詞 | | | |

#F 12855% [そうですか] s o o d e s u k a { | - } # .. | 間投詞 | | | |

#FH 12855 15209% [そうですかー] s o o d e s u k a a { | - } # .. | 間投詞 | | | |

#G 12855% [そうですか] s o o d e s u k a { | a } { | - } # .. | 間投詞 | | | |

#F 13107% [つと] { q | - } t o { | - } # .. | 間投詞 | | | |

#FH 13107 11626% [と] t o { | - } # .. | 間投詞 | | | |

#FH 13107 14118% [とー] t o o { | - } # .. | 間投詞 | | | |

#FH 13107 15469% [つとー] { q | - } t o o { | - } # .. | 間投詞 | | | |

#G 13107% [つと] { | q | - } t o { | o } { | - } # .. | 間投詞 | | | |

#F 13227% [えーとですね] e e t o d e s u n e { | - } # .. | 間投詞 | | | |

#FH 13227 14141% [えーとですねー] e e t o d e s u n e e { | - } # .. | 間投詞 | | | |

#FH 13227 14160% [えーつとですね] e e q t o d e s u n e { | - } # .. | 間投詞 | | | |

#FH 13227 14830% [えつとですね] e q t o d e s u n e { | - } # .. | 間投詞 | | | |

#FH 13227 15078% [えつとですねー] e q t o d e s u n e e { | - } # .. | 間投詞 | | | |

#FH 13227 19240% [えーつとですねー] e e q t o d e s u n e e { | - } # .. | 間投詞 | | | |

#FH 13227 19811% [えーとですねえー] e e t o d e s u n e e e { | - } # .. | 間投詞 | | | |

#G 13227% [えーとですね] e { e { | q } | q } t o d e s u n e { | e | e e } { | - } # .. | 間投詞 | | | |

#F 13429% [す] s u { | - } # .. | 間投詞 | | | |

#FH 13429 19554% [すー] s u u { | - } # .. | 間投詞 | | | |

#FH 13429 20164% [すつ] s u q # .. | 間投詞 | | | |

#G 13429% [す] s u { | u } { | q | - } # .. | 間投詞 | | | |

#F 13806% [あのですね] a n o d e s u n e { | - } # .. | 間投詞 | | | |

#FH 13806 10930% [あのですねー] a n o d e s u n e e { | - } # .. | 間投詞 | | | |

#FH 13806 14011% [あのーですね] a n o o d e s u n e { | - } # .. | 間投詞 | | | |

#G 13806% [あのですね] a n o { | o } d e s u n e { | e } { | - } # .. | 間投詞 | | | |

#F 14158% [あとー] a t o o { | - } # .. | 間投詞 | | | |

#FH 14158 14867% [あと] a t o { | - } # .. | 間投詞 | | | |

#G 14158% [あとー] a t o { | o } { | - } # .. | 間投詞 | | | |

#F 14903% [あう] a u { | - } # .. | 間投詞 | | | |

#FH 14903 16512% [あうー] a u u { | - } # .. | 間投詞 | | | |

#G 14903% [あう] a u { | u } { | - } # .. | 間投詞 | | | |

```

#F 14947% [の] n o {|-} # ..|問投詞|||
#FH 14947 19179% [の一] n o o {|-} # ..|問投詞|||
#G 14947% [の] n o {o} {|-} # ..|問投詞|||

#F 15450% [えーとね] e e q t o n e {|-} # ..|問投詞|||
#FH 15450 16301% [えーとね] e e t o n e {|-} # ..|問投詞|||
#G 15450% [えーとね] e e {q} t o n e {|-} # ..|問投詞|||

#F 16368% [ひー] h i i {|-} # ..|問投詞|||
#FH 16368 15252% [ひ] h i {|-} # ..|問投詞|||
#G 16368% [ひー] h i {i} {|-} # ..|問投詞|||

#F 16636% [こうー] k o {o o|u u} {|-} # ..|問投詞|||
#FH 16636 17589% [こう] k o o {|-} # ..|問投詞|||
#G 16636% [こうー] k o {o|o o|u u} {|-} # ..|問投詞|||

#F 19524% [えの一] e n o o {|-} # ..|問投詞|||
#FH 19524 14868% [えの] e n o {|-} # ..|問投詞|||
#G 19524% [えの一] e n o {o} {|-} # ..|問投詞|||

```

```

#####
### NAMES

```

```

#J 11223 16635% [ヒリップス] h i r i q p u s u {|-} # ..|人名|||
#J 14015 15273% [純] z h j u n g {|-} # ..|人名|||
#J 14253 14967% [令子] r e e k o {|-} # ..|人名|||
#J 14293 14289% [祐一] j u u i c h i {|-} # ..|人名|||
#J 14394 14488% [博之] h i r o j u k i {|-} # ..|人名|||
#J 14854 13960% [孝子] t a k a k o {|-} # ..|人名|||
#J 14856 19563% [皆子] m i n a k o {|-} # ..|人名|||
#J 14911 19508% [友広] t o m o h i r o {|-} # ..|人名|||
#J 15122 14106% [明美] a k e m i {|-} # ..|人名|||
#J 15158 20313% [櫻井] s a k u r a i {|-} # ..|人名|||
#J 15177 14058% [貴島] k i z h i m a {|-} # ..|人名|||
#J 15184 15142% [紀子] n o r i k o {|-} # ..|人名|||
#J 15287 20005% [麻衣子] m a i k o {|-} # ..|人名|||
#J 15390 15092% [野澤] n o z a w a {|-} # ..|人名|||
#J 15424 19837% [浩子] h i r o k o {|-} # ..|人名|||
#J 15426 11488% [健史] t a k e s h i {|-} # ..|人名|||
#J 15558 11574% [メアリー] m e a r i i {|-} # ..|人名|||
#J 16757 16797% [裕二] j u u z h i {|-} # ..|人名|||

```

#J 16874 20009% [正弘] m a s a h i r o {|-} # ..|人名|||
 #J 19485 19907% [雅子] m a s a k o {|-} # ..|人名|||
 #J 19488 19632% [良子] j o s h i k o {|-} # ..|人名|||
 #J 19647 15026% [裕] j u t a k a {|-} # ..|人名|||
 #J 19761 14250% [利恵] r i e {|-} # ..|人名|||
 #J 19815 18321% [金野] k o n g n o {|-} # ..|人名|||
 #J 20050 14008% [昌範] m a s a n o r i {|-} # ..|人名|||
 #J 20056 13979% [幸] m i j u k i {|-} # ..|人名|||

names that have more than 2 different writings

#J 14066 14075% [揺子] j o o k o {|-} # ..|人名|||
 #J 14066 19819% [陽子] j o o k o {|-} # ..|人名|||
 #J 15100 14148% [朋子] t o m o k o {|-} # ..|人名|||
 #J 15100 14990% [知子] t o m o k o {|-} # ..|人名|||
 #J 15100 19769% [とも子] t o m o k o {|-} # ..|人名|||
 #J 15561 14083% [晶子] a k i k o {|-} # ..|人名|||
 #J 15561 14970% [明子] a k i k o {|-} # ..|人名|||
 #J 15561 20014% [亜希子] a k i k o {|-} # ..|人名|||
 #J 19597 15131% [義美] j o s h i m i {|-} # ..|人名|||
 #J 19597 19521% [好美] j o s h i m i {|-} # ..|人名|||
 #J 19779 19664% [洋] h i r o s h i {|-} # ..|人名|||
 #J 19779 19875% [裕] h i r o s h i {|-} # ..|人名|||
 #J 19817 14033% [聡] s a t o s h i {|-} # ..|人名|||
 #J 19817 19491% [智] s a t o s h i {|-} # ..|人名|||

#####

HOMONYMS

#M 20196 10193% [書] k a # ..|本動詞|五段カ|語幹||
 #M 11145 10556% [濟] s u # ..|本動詞|五段マ|語幹||
 #M 15208 10765% [分かれ] w a k a r e {|-} # ..|本動詞|一段|語幹||
 #M 11243 10890% [端] h a s h i {|-} # ..|普通名詞|||
 #M 12640 10980% [よ] j o # ..|本動詞|五段ラ|語幹||
 #M 14085 11076% [秋] a k i {|-} # ..|普通名詞|||
 #M 14091 11164% [何回] n a n g k a i {|-} # ..|普通名詞|||
 #M 15831 11441% [正装] s e e s o o {|-} # ..|サ変名詞|||

#M 15739 11638% [懷石] k a i s e k i {|-} # ..|普通名詞|||
 #M 12758 11736% [変わり] k a w a r i {|-} # ..|普通名詞|||
 #M 12450 12002% [県] k e n g {|-} # ..|接尾辞|||
 #M 11782 12223% [訪ね] t a z u n e {|-} # ..|本動詞|一段|語幹||
 #M 12549 12283% [干菓子] h i g a s h i {|-} # ..|普通名詞|||
 #M 13118 12552% [制] s e e {|-} # ..|接尾辞|||
 #M 14953 12572% [傷害] s h j o o g a i {|-} # ..|普通名詞|||
 #M 10796 12617% [件] k e n g {|-} # ..|普通名詞|||
 #M 14098 12625% [発] t a # ..|本動詞|五段夕|語幹||
 #M 11136 12667% [合] a # ..|本動詞|五段ワ|語幹||
 #M 13794 12795% [機械] k i k a i {|-} # ..|普通名詞|||
 #M 14098 12846% [立] t a # ..|本動詞|五段夕|語幹||
 #M 19863 12954% [のぞ] n o z o # ..|本動詞|五段カ|語幹||
 #M 16051 12974% [洗濯] s e n g t a k u {|-} # ..|サ変名詞|||
 #M 14067 13022% [用] j o o {|-} # ..|普通名詞|||
 #M 12012 13061% [髪] k a m i {|-} # ..|普通名詞|||
 #M 12661 13104% [大] d a i {|-} # ..|接頭辞|||
 #M 16371 13275% [長] c h j o o {|-} # ..|接頭辞|||
 #M 16140 13698% [話せ] h a n a s e {|-} # ..|本動詞|一段|語幹||
 #M 10222 13833% [古都] k o t o {|-} # ..|普通名詞|||
 #M 10447 13868% [清算] s e e s a n g {|-} # ..|サ変名詞|||
 #M 17235 13961% [州] s h j u u {|-} # ..|普通名詞|||
 #M 12139 14068% [字] z h i {|-} # ..|普通名詞|||
 #M 14067 14070% [洋] j o o {|-} # ..|普通名詞|||
 #M 13012 14087% [晶] s h j o o {|-} # ..|普通名詞|||
 #M 16911 14128% [全階] z e n g k a i {|-} # ..|普通名詞|||
 #M 14067 14490% [様] j o o {|-} # ..|普通名詞|||
 #M 14509 14492% [文書] b u n g s h j o {|-} # ..|普通名詞|||
 #M 13048 14706% [搭乘] t o o z h j o o {|-} # ..|サ変名詞|||
 #M 10862 14762% [市外] s h i g a i {|-} # ..|普通名詞|||
 #M 12539 15089% [計] k e e {|-} # ..|普通名詞|||
 #M 16184 15172% [自身] z h i s h i n g {|-} # ..|普通名詞|||
 #M 15441 15573% [差] s a {|-} # ..|普通名詞|||
 #M 10845 15611% [待ち] m a c h i {|-} # ..|普通名詞|||
 #M 12973 15651% [浴え] s o e {|-} # ..|本動詞|一段|語幹||
 #M 15213 15683% [乗り] n o r i {|-} # ..|普通名詞|||
 #M 15095 15760% [停め] t o m e {|-} # ..|本動詞|一段|語幹||
 #M 15783 15820% [加茂川ホテル] k a m o g a w a h o t e r u {|-} # ..|固有名詞|||
 #M 16479 15918% [放送] h o o s o o {|-} # ..|サ変名詞|||
 #M 10984 15933% [外] g a i {|-} # ..|接尾辞|||
 #M 12417 15940% [社] s h j a {|-} # ..|接尾辞|||
 #M 12361 16054% [映] u t s u # ..|本動詞|五段ラ|語幹||

#M 13375 16221% [運航] u n g k o o {|-} # ..|サ変名詞|||
 #M 13889 16347% [熱] a t s u # ..|形容詞|形容詞|語幹||
 #M 10796 16396% [兼] k e n g {|-} # ..|普通名詞|||
 #M 11635 16408% [色] s h j o k u {|-} # ..|接尾辭|||
 #M 16031 16549% [履] h a # ..|本動詞|五段カ|語幹||
 #M 17904 16574% [対] t a i {|-} # ..|普通名詞|||
 #M 10694 16616% [自国] z h i k o k u {|-} # ..|普通名詞|||
 #M 12417 16638% [車] s h j a {|-} # ..|接尾辭|||
 #M 11154 16727% [店] t e n g {|-} # ..|接尾辭|||
 #M 13859 16815% [講演] k o o e n g {|-} # ..|サ変名詞|||
 #M 17895 16907% [演台] e n g d a i {|-} # ..|普通名詞|||
 #M 16313 16953% [清浄] s e e z h j o o {|-} # ..|形容名詞|||
 #M 19575 17043% [幹事] k a n g z h i {|-} # ..|普通名詞|||
 #M 17482 17130% [社内] s h j a n a i {|-} # ..|普通名詞|||
 #M 11243 17135% [箸] h a s h i {|-} # ..|普通名詞|||
 #M 15896 17450% [向] m u # ..|本動詞|五段カ|語幹||
 #M 16412 17502% [効き] k i k i {|-} # ..|普通名詞|||
 #M 10377 17516% [補償] h o s h j o o {|-} # ..|サ変名詞|||
 #M 16033 17561% [都市] t o s h i {|-} # ..|普通名詞|||
 #M 11038 17597% [案内書] a n g n a i s h j o {|-} # ..|普通名詞|||
 #M 16203 17687% [一帯] i q t a i {|-} # ..|普通名詞|||
 #M 12766 17806% [亡くな] n a k u n a # ..|本動詞|五段ラ|語幹||
 #M 16736 17962% [盛り] m o r i {|-} # ..|普通名詞|||
 #M 16767 17966% [意] i {|-} # ..|普通名詞|||
 #M 13876 17993% [銚子] c h j o o s h i {|-} # ..|普通名詞|||
 #M 14705 18028% [館] k a n g {|-} # ..|普通名詞|||
 #M 13897 18041% [錠] z h j o o {|-} # ..|接尾辭|||
 #M 18010 18245% [最] s a i {|-} # ..|接頭辭|||
 #M 10792 18618% [就] t s u # ..|本動詞|五段カ|語幹||
 #M 14440 19542% [氏] s h i {|-} # ..|接尾辭|||
 #M 13897 19586% [疊] z h j o o {|-} # ..|接尾辭|||
 #M 13118 19629% [生] s e e {|-} # ..|接尾辭|||
 #M 15487 19971% [治ま] o s a m a # ..|本動詞|五段ラ|語幹||
 #M 17055 20184% [代] d a i {|-} # ..|普通名詞|||
 #M 13594 20267% [ママ] m a m a {|-} # ..|普通名詞|||

homonyms combined with doubles

#M 11476 #H 11475 14081% [替わ] k a w a # ..|本動詞|五段ラ|語幹||
 #M 11476 11475% [代わ] k a w a # ..|本動詞|五段ラ|語幹||
 #M 13865 #H 15897 15952% [革] k a w a {|-} # ..|普通名詞|||

#M 13865 15897% [皮] k a w a {|-} # ..|普通名詞|||

homonyms combined with doubles that are also split off as composite N-grams

#M 10323 #H 12753 20404% [き] k i # ..|本動詞|五段カ|語幹||

#M 10323 #V1000 12753% [き] k i # ..|本動詞|五段カ|語幹||

homonyms that are also split off as composite N-grams

#M 10581 #V1000 10343% [載] n o # ..|本動詞|五段ラ|語幹||

#M 10847 #V1000 10524% [日] k a {|-} # ..|接尾辞|||

#M 11735 #V1000 10624% [買え] k a e {|-} # ..|本動詞|一段|語幹||

#M 10792 #V1000 10671% [着] t s u # ..|本動詞|五段カ|語幹||

#M 10502 #V1000 10678% [線] s e n g {|-} # ..|接尾辞|||

#M 11523 #V1000 10720% [台] d a i {|-} # ..|接尾辞|||

#M 10360 #V1000 10726% [止ま] t o m a # ..|本動詞|五段ラ|語幹||

#M 10201 #V1000 10740% [速] h a j a # ..|形容詞|形容詞|語幹||

#M 18836 #V1000 10887% [もの] m o {n o | n g} {|-} # ..|普通名詞|||

#M 10984 #V1000 10933% [回] k a i {|-} # ..|接尾辞|||

#M 11735 #V1000 11151% [替え] k a e {|-} # ..|本動詞|一段|語幹||

#M 11012 #V1000 11202% [着け] t s u k e {|-} # ..|本動詞|一段|語幹||

#M 10816 #V1000 13546% [撮れ] t o r e {|-} # ..|本動詞|一段|語幹||

#M 10984 #V1000 13676% [街] g a i {|-} # ..|接尾辞|||

#M 10611 #V1000 14411% [咳] s e k i {|-} # ..|普通名詞|||

#M 11683 #V1000 14662% [送れ] o k u r e {|-} # ..|本動詞|一段|語幹||

#M 10346 #V1000 15066% [型] k a t a {|-} # ..|普通名詞|||

#M 10144 #V1000 15202% [恭] k j o o {|-} # ..|普通名詞|||

#M 10641 #V1000 15970% [周] s h j u u {|-} # ..|接尾辞|||

#M 11167 #V1000 16367% [月] t s u k i {|-} # ..|接尾辞|||

#M 10705 #V1000 16588% [時分] z h i b u n g {|-} # ..|普通名詞|||

#M 10222 #V1000 17617% [琴] k o t o {|-} # ..|普通名詞|||

#M 10295 #V1000 17985% [欠け] k a k e {|-} # ..|本動詞|一段|語幹||

#M 17382 #V1000 19748% [泊り] t o m a r i {|-} # ..|普通名詞|||

#M 14085 #V1000 19788% [あき] a k i {|-} # ..|普通名詞|||

CONCATENATIONS

#R 11823 11823 #/ 13626% [そうそう] s o o s o o {|-} # ..|間投詞|||

#R 14506 10498 #/ 14295% [はあえと] h a a e e t o {|-} # ..|間投詞|||

#R 10663 14947 #/ 14671% [じゃのー] z h j a n o o {|-} # ..|間投詞|||

#R 14555 10054 #/ 14866% [ではあのー] d e w a a n o o {|-} # ..|間投詞|||

33

```

#R 10017 10005 #/ 14874% [えーあ] e e a {-} # ..|間投詞|||
#R 10498 10950 #/ 14976% [えーつとや] e e q t o j a {-} # ..|間投詞|||
#R 10017 10263 10054 #/ 15163% [えーそうですねあの] e e s o o d e s u n e a n o {-} # ..|間投詞|||
#R 10263 12359 #/ 15194% [そうですねんーと] s o o d e s u n e n g t o {-} # ..|間投詞|||
#R 10017 11358 #/ 15321% [えーちよつと] e e c h j o q t o {-} # ..|間投詞|||
#R 11358 10054 #/ 15322% [ちよつとあのー] c h j o q t o a n o o {-} # ..|間投詞|||
#R 13227 11358 #/ 15324% [えつとですねーちよつと] e q t o d e s u n e e c h j o q t o {-} # ..|間投詞|||
#R 10017 10663 #/ 15343% [えーじゃあ] e e z h j a a {-} # ..|間投詞|||
#R 12359 10054 #/ 16337% [んーとあの] n g t o a n o {-} # ..|間投詞|||
#R 11823 11823 11823 #/ 16352% [そうそうそう] s o o s o o s o o {-} # ..|間投詞|||
#R 10017 12633 #/ 17015% [えあーと] e a a t o {-} # ..|間投詞|||
#R 10934 10005 #/ 18055% [うーんあ] u u n g a {-} # ..|間投詞|||
#R 12870 10054 #/ 18441% [であのー] d e a n o o {-} # ..|間投詞|||
#R 10017 19059 #/ 19608% [えーそれじゃあ] e e s o r e z h j a a {-} # ..|間投詞|||
#R 10017 10005 10498 #/ 19611% [えっあっえーと] e q a q e e t o {-} # ..|間投詞|||
#R 10263 10498 #/ 19681% [そうですねえーと] s o o d e s u n e e e t o {-} # ..|間投詞|||
#R 10498 10263 10498 #/ 19688% [えつとーそうですねえつとー] e q t o o s o o d e s u n e e q t o o {-} # ..|間投詞|||
#R 10498 19059 12359 #/ 19690% [えーとそれじゃあーうんと] e e t o s o r e z h j a a a u n g t o {-} # ..|間投詞|||
#R 10005 12870 #/ 19694% [あっで] a q d e {-} # ..|間投詞|||
#R 11819 10644 #/ 19699% [おーまー] o o m a a {-} # ..|間投詞|||
#R 10054 10663 #/ 19710% [あのーじゃあ] a n o o z h j a a {-} # ..|間投詞|||
#R 12518 10644 #/ 19711% [そのーまー] s o n o o m a a {-} # ..|間投詞|||
#R 10017 10644 10017 #/ 19719% [えーまあえー] e e m a a e e {-} # ..|間投詞|||
#R 12518 10054 #/ 19740% [そのーあのー] s o n o o a n o o {-} # ..|間投詞|||
#R 10017 10017 #/ 19757% [えっえー] e q e e {-} # ..|間投詞|||
#R 10054 10017 11819 #/ 19765% [あのえーお] a n o e e o {-} # ..|間投詞|||
#R 10498 10017 #/ 19773% [えとーえー] e t o o e e {-} # ..|間投詞|||
#R 10005 12518 #/ 19780% [あそのー] a s o n o o {-} # ..|間投詞|||
#R 10663 10054 10498 #/ 19812% [じゃああのーえーと] z h j a a a n o o e e t o {-} # ..|間投詞|||
#R 10017 19059 10054 #/ 19813% [えーそれじゃあの] e e s o r e z h j a a n o {-} # ..|間投詞|||
#R 10263 10017 10644 #/ 19854% [そうですねえーまあ] s o o d e s u n e e e m a a {-} # ..|間投詞|||
#R 10498 10112 10054 #/ 19900% [えーともしーあのー] e e t o m o s h i i a n o o {-} # ..|間投詞|||
#R 12633 10663 #/ 19902% [あつとーじゃあ] a q t o o z h j a a {-} # ..|間投詞|||
#R 12870 10498 #/ 19932% [でえーと] d e e t o {-} # ..|間投詞|||
#R 10498 11358 10054 #/ 19939% [えーとちよつとあのー] e e t o c h j o q t o a n o o {-} # ..|間投詞|||
#R 10934 10644 #/ 19954% [んーまあ] n g m a a {-} # ..|間投詞|||
#R 10644 10054 #/ 19979% [まああの] m a a a n o {-} # ..|間投詞|||
#R 13806 10054 #/ 19980% [あのですねあの] a n o d e s u n e a n o {-} # ..|間投詞|||
#R 10017 10644 #/ 19994% [えーま] e e m a {-} # ..|間投詞|||
#R 10005 10054 #/ 20155% [あーあの] a a a n o {-} # ..|間投詞|||

```

```

### concatenations combined with doubles

```

#H 14878 #R 10663 10054 #/ 19808% [じゃああの] zh j a a a n o {|-} # ..|間投詞|||
#R 10663 10054 #/ 14878% [じゃーあの] zh j a a a n o {|-} # ..|間投詞|||

ERRORS

errors that require an entry to be replaced

#X 10112% [もし] m o s h i {|-} # ..|副詞|||
#Y 10112% [もし] m o s h i {i} {|-} # ..|副詞|||

#X 10190% [てい] t e i # ..|助動詞|一段|語幹||
#Y 10190% [てい] t e {i}{|-} # ..|助動詞|一段|語幹||

#X 10345% [日本人] n i {h o|q p} n g z h i n g {|-} # ..|普通名詞|||
#Y 10345% [日本人] n i {h|q p} o n g z h i n g {|-} # ..|普通名詞|||

#X 10416% [て行] t e {|-} i # ..|助動詞|五段カ|語幹||
#Y 10416% [て行] t e {|-} {i} # ..|助動詞|五段カ|語幹||

#X 10503% [円] e n g {|-} # ..|接尾辞|||
#Y 10503% [円] {j} e n g {|-} # ..|接尾辞|||

#X 10568% [十七日] zh j u u n a n a n i c h i {|-} # ..|日時|||
#Y 10568% [十七日] zh j u u n a n a {n i|ng} c h i {|-} # ..|日時|||

#X 11002% [評判] h j o o b a n g {|-} # ..|普通名詞|||
#Y 11002% [評判] h j o o b a n g {|-} # ..|サ変名詞|||

#X 11044% [つて] {q|-} t e {|-} # ..|連体助詞|||
#Y 11044% [つて] {q|-} t e {|-} # ..|連体助詞|||

#X 11451% [二十九日] n i zh j u u k u n i c h i {|-} # ..|日時|||
#Y 11451% [二十九日] n i zh j u u k u {n i|ng} c h i {|-} # ..|日時|||

#X 11485% [二十八日] n i zh j u u h a c h i n i c h i {|-} # ..|日時|||
#Y 11485% [二十八日] n i zh j u u h a c h i {n i|ng} c h i {|-} # ..|日時|||

#X 11504% [二十六日] n i zh j u u r o k u n i c h i {|-} # ..|日時|||
#Y 11504% [二十六日] n i zh j u u r o k u {n i|ng} c h i {|-} # ..|日時|||

#X 11669% [二十二日] n i zh j u u n i n i ch i {|-} # ..|日時|||
 #Y 11669% [二十二日] n i zh j u u n i (n i|ng) ch i {|-} # ..|日時|||
 #X 11700% [三十日] s a ng zh j u u n i ch i {|-} # ..|日時|||
 #Y 11700% [三十日] s a ng zh j u u (n i|ng) ch i {|-} # ..|日時|||
 #X 11701% [三十一日] s a ng zh j u u i ch i n i ch i {|-} # ..|日時|||
 #Y 11701% [三十一日] s a ng zh j u u i ch i (n i|ng) ch i {|-} # ..|日時|||
 #X 11860% [十五日] zh j u u g o n i ch i {|-} # ..|日時|||
 #Y 11860% [十五日] zh j u u g o (n i|ng) ch i {|-} # ..|日時|||
 #X 12571% [病氣] b j o o k i {|-} # ..|サ変名詞|||
 #Y 12571% [病氣] b j o o k i {|-} # ..|普通名詞|||
 #X 12915% [なれ] n a r e {|-} # ..|本動詞|一段|語幹||
 #Y 12915% [にもなれ] n i m o n a r e {|-} # ..|補助動詞|一段|語幹||
 #X 14367% [問診] m o (ng|o) sh i ng {|-} # ..|サ変名詞|||
 #Y 14367% [問診] m o (ng|o) sh i ng {|-} # ..|普通名詞|||
 #X 14391% [処方] sh j o h o o {|-} # ..|サ変名詞|||
 #Y 14391% [処方] sh j o h o o {|-} # ..|普通名詞|||
 #X 14692% [通話] ts u u w a {|-} # ..|普通名詞|||
 #Y 14692% [通話] ts u u w a {|-} # ..|サ変名詞|||
 #X 15006% [可否] k a h i {|-} # ..|サ変名詞|||
 #Y 15006% [可否] k a h i {|-} # ..|普通名詞|||
 #X 15255% [二十一時] n i zh j u u i ch i zh i {|-} # ..|日時|||
 #Y 15255% [二十一時] n i zh j u {u} i ch i zh i {|-} # ..|日時|||
 #X 15483% [二十七日] n i zh j u u {|-} n a n a n i ch i {|-} # ..|日時|||
 #Y 15483% [二十七日] n i zh j u u {|-} n a n a (n i|ng) ch i {|-} # ..|日時|||
 #X 15687% [日光浴] n i q k o o j o k u {|-} # ..|サ変名詞|||
 #Y 15687% [日光浴] n i q k o o j o k u {|-} # ..|普通名詞|||
 #X 15688% [運だめし] u ng d a m e sh i {|-} # ..|サ変名詞|||
 #Y 15688% [運だめし] u ng d a m e sh i {|-} # ..|普通名詞|||

#X 15878% [意見] i k e n g {|-} # ..|サ変名詞|||
 #Y 15878% [意見] i k e n g {|-} # ..|普通名詞|||

 #X 16228% [内装] n a i s o o {|-} # ..|サ変名詞|||
 #Y 16228% [内装] n a i s o o {|-} # ..|普通名詞|||

 #X 16259% [着払い] ch j a k u b a r a i {|-} # ..|サ変名詞|||
 #Y 16259% [着払い] ch j a k u b a r a i {|-} # ..|普通名詞|||

 #X 16627% [値打ち] n e u c h i {|-} # ..|サ変名詞|||
 #Y 16627% [値打ち] n e u c h i {|-} # ..|普通名詞|||

 #X 16734% [参考] s a n g k o o {|-} # ..|サ変名詞|||
 #Y 16734% [参考] s a n g k o o {|-} # ..|普通名詞|||

 #X 16886% [京都運輸] k j o o t o o n g j u {|-} # ..|固有名詞|||
 #Y 16886% [京都運輸] k j o o t o {|-} u n g j u {|-} # ..|固有名詞|||

 #X 17306% [その上] s o n o o e {|-} # ..|接続詞|||
 #Y 17306% [その上] s o n o u e {|-} # ..|接続詞|||

 #X 17748% [所用] sh j o j o o {|-} # ..|サ変名詞|||
 #Y 17748% [所用] sh j o j o o {|-} # ..|普通名詞|||

 #X 17799% [堤防] t e e b o o {|-} # ..|サ変名詞|||
 #Y 17799% [堤防] t e e b o o {|-} # ..|普通名詞|||

 #X 17989% [衝撃] sh j o o g e k i {|-} # ..|サ変名詞|||
 #Y 17989% [衝撃] sh j o o g e k i {|-} # ..|普通名詞|||

 #X 18205% [おっきな] o o k i n a {|-} # ..|連体詞|||
 #Y 18205% [大きな] o o k i n a {|-} # ..|連体詞|||

 #X 18306% [ん] {ng|n i} {|-} # ..|格助詞|||
 #Y 18306% [に] {ng|n i} {|-} # ..|格助詞|||

 #X 18493% [てゆう] {q} t {o|e} {j u|i} u {|-} # ..|連体助詞|||
 #Y 18493% [という] {q} t {o|e} {j u|i} u {|-} # ..|連体助詞|||

 #X 18650% [向かえ] m u k a i {|-} # ..|普通名詞|||
 #Y 18650% [向かい] m u k a i {|-} # ..|普通名詞|||

#X 18806% [とお] t (e {-}) o # ..|助動詞|五段ラ|語幹||
 #Y 18806% [てお] t (e {-}) o # ..|助動詞|五段ラ|語幹||

#X 18836% [もん] m o {n o|ng} {-} # ..|普通名詞| ||||
 #Y 18836% [もの] m o {n o|ng} {-} # ..|普通名詞| ||||

#X 18839% [執] t o # ..|本動詞|五段ラ|語幹||
 #Y 18839% [取] t o # ..|本動詞|五段ラ|語幹||

#X 18988% [い] r u {-} # ..|語尾|特殊ラ|終止||
 #Y 18988% [る] r u {-} # ..|語尾|特殊ラ|終止||

#X 19125% [そのまんま] s o n o m a m a {-} # ..|副詞| ||||
 #Y 19125% [そのまんま] s o n o m a {ng} m a {-} # ..|副詞| ||||

#X 19447% [そん] s o n o {-} # ..|連体詞| ||||
 #Y 19447% [その] s o n o {-} # ..|連体詞| ||||

#X 19479% [しょうがありません] s h i j o o g a a r i m a s e n g {-} # ..|本動詞|特殊|終止||
 #Y 19479% [しょうがありません] s h {i} j o o g a a r i m a s e n g {-} # ..|本動詞|特殊|終止||

#X 19772% [制服] s e e h u k u {-} # ..|サ変名詞| ||||
 #Y 19772% [制服] s e e h u k u {-} # ..|普通名詞| ||||

#X 19894% [インターナショナル] i n g t a a n a s h j o n a r u {-} # ..|サ変名詞| ||||
 #Y 19894% [インターナショナル] i n g t a a n a s h j o n a r u {-} # ..|形容名詞| ||||

#X 19956% [おっき] o o k i # ..|形容詞|形容詞|語幹||
 #Y 19956% [大き] o o k i # ..|形容詞|形容詞|語幹||

#X 20220% [でっか] d e q k a # ..|形容詞|形容詞|語幹||
 #Y 20220% [でっか] d e {q} k a # ..|形容詞|形容詞|語幹||

errors in concatenated morphemes

#X #R 10005 19059 10017 #/ 19803% [あっそれではえー] a q s o r e d e h a e e {-} # ..|間投詞| ||||
 #Y #R 10005 19059 10017 #/ 19803% [あっそれではえー] a q s o r e d e w a e e {-} # ..|間投詞| ||||

errors that can be solved by deleting the entry

#X2 10190 10022% [て] t e {-} # ..|助動詞|一段|語幹||
 #X2 16871 12909% [あら] a r a {-} # ..|間投詞| ||||

#X2 10836 14735% [もらえ] m o r a e {-} # ..|補助動詞|一段|語幹||
#X2 18414 14839% [そこら] s o k o r a {-} # ..|副詞|||
#X2 11309 16800% [もしもし] m o s h i m o s h i {-} # ..|問投詞|||
#X2 10498 17205% [えーと] e e t o {-} # ..|感動詞|||
#X2 10416 18513% [て] t e # ..|助動詞|五段カ|語幹||
#X2 10112 19603% [もしー] m o s h i i {-} # ..|問投詞|||
#X2 16295 19665% [やや] j a j a {-} # ..|問投詞|||

#X2 #S 12806 16376% [く] k u {-} # ..|問投詞|||

REPARANDA

(not listed)

END OF FILE