



# A Free Online Accent and Intonation Dictionary for Teachers and Learners of Japanese.

Hiroko Hirano<sup>‡</sup>, Ibuki Nakamura<sup>†</sup>, Nobuaki Minematsu<sup>†</sup>, Masayuki Suzuki<sup>†</sup>, Chieko Nakagawa\*  
Noriko Nakamura\*, Yukinori Tagawa\*, Keikichi Hirose<sup>†</sup>, Hiroya Hashimoto<sup>†</sup>

<sup>‡</sup> Northeast Normal University, Jilin, China      <sup>†</sup> The University of Tokyo, Tokyo, Japan  
\* Waseda University, Tokyo, Japan      \* Tokyo University of Foreign Studies, Tokyo, Japan

## Abstract

We have developed the very first free online and free framework for teaching and learning Japanese prosody with features for word accent and phrase intonation. This framework is called OJAD (Online Japanese Accent Dictionary) [1], which provides three functions. 1) Visual, auditory, systematic, and comprehensive illustration of patterns of accent change (accent sandhi) of verbs and adjectives. Here only the changes resulting from twelve kinds of fundamental conjugation are focused upon. 2) Visual illustration of the accent pattern of a given verbal expression, which is a combination of a verb and its postpositional auxiliary words. 3) Visual illustration of the pitch pattern of an any given sentence and the expected positions of accent nuclei in the sentence. The third function is implemented by using an accent change prediction module that we developed for Japanese text-to-speech (TTS) synthesizers [2]. Subjective assessment by teachers shows very high pedagogical effectiveness of the framework.

**Index Terms:** language education, Japanese prosody, accent sandhi, OJAD, TTS synthesizer, assessment experiment

## 1. The three functions of OJAD

### 1.1. Comprehensive illustration of accent changes

Japanese is a pitch accent language and along with the conjugation of verbs and adjectives, their accent patterns also change regularly and systematically. If a learner desires to speak sounding less foreign-accented, he or she will need to follow the accent rules. However, as existing word dictionaries merely list “dictionary form” before conjugation of a verb/adjective, and even accent dictionaries just describe the accent rules of conjugation with a few samples at the end, learners do not have a resource available to them at present. Therefore, we engineered a system that shows the accent changes due to conjugation of these words. Users type verbs and/or adjectives of interest to know their accent changes. Here, twelve kinds of fundamental conjugation were adopted and their accents are displayed in a table. Fig. 1 shows an example. Eleven widely-used textbooks were selected and all the verbs and adjectives found in them were manually extracted for building OJAD. The total number of words is about 3,500 and that of their conjugated forms is about 42,000.

Each of these forms was read aloud in the Tokyo dialect, which is widely accepted pronunciation in public speaking, by a voice actor and a voice actress. About 84,000 speech samples were recorded and they were segmented semi-automatically using voice activity detection techniques. In Fig. 1, by clicking a blue/pink icon, users can listen to a male/female speech sample

of each form, respectively. A series of the samples on a row or on a column can be heard by clicking an icon of that purpose. These samples can be downloaded onto users’ PCs or portable devices.



Figure 1: Illustration of the accent patterns of conjugated forms

### 1.2. Illustration of the accent of long verbal expressions

The first function only shows the accent patterns of the twelve fundamental conjugated forms of verbs and adjectives. Since Japanese is an agglutinative language, a verb can be combined with multiple postpositional and auxiliary words. For example, verb “倒れる” (fall) can be concatenated to “そう”, “に”, “なった”, “こと”, “か”, and “ある” in this order. By conjugation, “倒れる” is finally changed into “倒れそうになったことがある” and this kind of long verbal expressions can be found even in a textbook for beginners. This means that with only the first function the system cannot instruct the learner how to produce natural-sounding intonation. So, we developed a second function to integrate into this system to show the accent pattern of a given long verbal expression.

Fig. 2 shows several examples of model pitch contours for long verbal expressions illustrated with this function. It is easily expected that a verbal expression including an unknown postpositional expression can be typed as input. Our database contains the information of the 320 expressions only. If the system can find the postpositional expression in the database, it shows the accent pattern in a red rectangle. If the system cannot, it shows the accent pattern of the most similar expression found in the database instead, which is found in the database, in a pink rectangle. Availability of the system response is indicated by color.

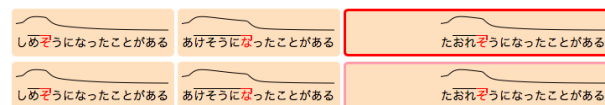


Figure 2: Illustration of the accent patterns of long expressions

### 1.3. Illustration of the pitch pattern of any input sentence

The first and second functions only focus upon verbs and adjectives. Word accent changes are not only found in these words but also in other words such as nouns. So, as a third function, we developed a prosodic reading tutor to support learners by presenting the pitch pattern of an any given sentence.

This function is realized easily by using several internal modules developed for TTS synthesizers. They are morphological analysis (linguistic analysis) [5], accent phrase boundary detection from text [2], accent nucleus location for an detected accent phrase [2], and pitch pattern visualization by the Fujisaki model [4]. Direct visualization of the output of these modules is not good pedagogically because it is sometimes too complicated for learners to learn. It is important to present pitch patterns that are as simple as possible but maintain naturalness. In tight collaboration with Japanese teachers, we designed simplification rules [3]. Generation of the original and simplified pitch patterns is schematized in Fig. 3.

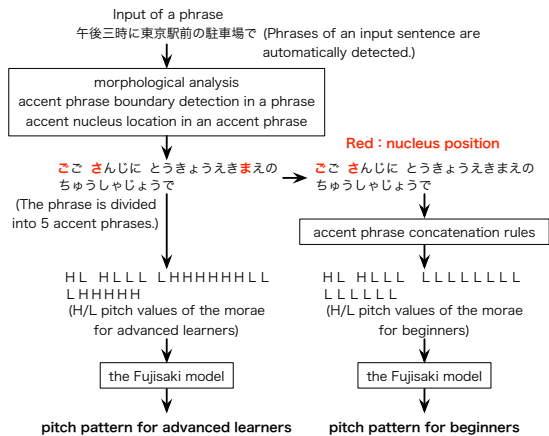


Figure 3: Generation of original and simplified pitch patterns

An input sentence is automatically divided into phrases by punctuation marks and phrase boundary marks (/), which are explicitly given by users. Three types of morphological analysis are run for each phrase: accent phrase boundary detection, and accent nucleus location. Then, the input phrase is divided into multiple accent phrases, in each of which the accent nucleus position is estimated automatically. For advanced learners, these accent phrases are directly used to visualize the pitch pattern. An example is shown at the top of Fig. 4. For beginners, simplification is needed. In some accent phrases, there is no accent nucleus. [3] claims that learners should focus on the first nucleus found in a phrase (not an accent phrase). The remaining accent nuclei found in the phrase after the first nucleus can be ignored to generate a very simple pitch pattern with sufficient naturalness. An example is shown at the bottom of Fig. 4.

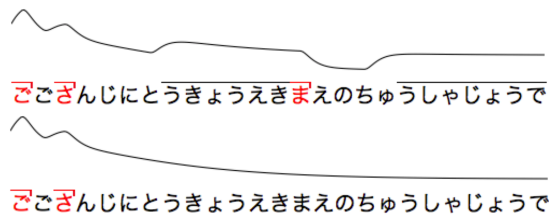


Figure 4: Original and simplified pitch patterns

## 2. Subjective assessment

We asked teachers of Japanese to join a subjective assessment test after instructing them how to use OJAD. Eighty teachers joined the test, two thirds of whom were teaching Japanese outside Japan. Although the subjective assessment was composed of a series of questionnaire items, we show in Tab. 1 the results of the two fundamental questions: a) How useful do you think the system is for learners? and b) Do you want to use the system in your class? Considering that teaching Japanese prosody is just only one aspect of Japanese language education, we regard these results as being evidence that the 80 teachers of Japanese see very high pedagogical effectiveness in the proposed framework.

Table 1: Assessment of the three proposed systems (%)

a) How useful do you think the system is for learners?			
	1st system	2nd system	3rd system
Very useful	71.0	54.8	62.7
Rather useful	29.0	45.2	28.8
Not so useful	0.0	0.0	8.5
Not useful at all	0.0	0.0	0.0

b) Do you want to use the system in your class?			
	1st system	2nd system	3rd system
Yes, definitely	38.7	29.0	42.6
Yes, if needed	59.7	64.5	50.0
No	1.6	6.5	7.4

## 3. Conclusions

In this paper, we developed the very first online and free framework for teaching and learning Japanese prosody with features for learning word accent and phrase intonation. The results show that teachers view this framework as being pedagogically effective.

This framework, called OJAD, was released to the public in August 2012 in an international conference of Japanese education [6]. After that, by using Google Analytics, all the accesses to OJAD have been recorded in a history file. The file shows that the number of accesses is around 28,000 as of May 2013 and about half of them are from outside Japan. Considering a fact that 72% of teachers of Japanese are non-native outside Japan, we can say that not a small number of learners are using OJAD as it is the only information source for learning how to control word accent when trying to speak natural sounding Japanese.

## 4. References

- [1] OJAD: <http://www.gavo.t.u-tokyo.ac.jp/ojad/>
- [2] M. Suzuki, R. Kuroiwa, K. Innami, S. Kobayashi, S. Shimizu, N. Minematsu, K. Hirose, "Accent sandhi estimation of Tokyo dialect of Japanese using conditional random fields," *Trans IEICE*, J96-D, 3, 655–654 (2013 in Japanese)
- [3] C. Nakagawa, N. Nakamura, S. Ho, *Japanese pronunciation drills for advanced oral presentation*, published by Hitsuji-Shobo (2009 in Japanese)
- [4] H. Fujisaki, K. Hirose, "Analysis of voice fundamental frequency contours for declarative sentences of Japanese," *J. Acoust. Soc. Japan* (E), 5, 4, 233–242 (1984)
- [5] MeCab: <http://mecab.sourceforge.net/>
- [6] N. Minematsu, M. Suzuki, H. Hirano, C. Nakagawa, N. Nakamura, Y. Tagawa, K. Hirose, "Development of an online Japanese accent dictionary with speech output facility," *Proc. Int. Conf. on Japanese Language Education (ICJLE)*, pp.94 (2012)