Speech perception across languages and writing systems – lessons for Japanese as Foreign Language from a commercial research project

Viktoria Eschbach-Szabo (Tübingen University)

Abstract

In the first part, this paper discusses our experiences developing speech-controlled car-navigation tools for Japanese, in particular problems occurring thereby. It proved to be rather difficult to deal with the linguistic complexity of written and spoken Japanese words, especially with regard to devocalization and the lack of standardized transcriptions. In the second part, lessons for Japanese as Foreign Language (JFL) evolving from our experiences are discussed.

1. Introduction

The semiotic view of language as a representational device is interconnected with other representational devices such as writing systems. This interconnection has received little attention in Japanese as Foreign Language (JFL), probably because psycholinguistic issues related to teaching *kanji* [Chinese characters] are not particularly prominent as a research subject. Rather, language educators are usually content to perceive the lexicon as being somehow connected to the writing system. Learning new words is, however, not that simple. This becomes clear when one studies new vocabulary in the context of their interconnection with different writing systems. In this paper, I will identify two concrete problems:
(1) problems pertaining to devoicing.
(2) problems related to transcribing Japanese with Roman letters.
The problems of devoicing and transcribing Japanese into the Roman alphabet were encountered in the course of a commercial research project for a Japanese car navigation system. While these problems have not emerged in the process of language education, several lessons for JFL can nevertheless be derived.

At our institute at Tübingen University, we have been developing a series of Japanese car navigation tools together with Temic, a company based in the city of Ulm. We were primarily engaged with matters of

speech recognition. In this process, we encountered several problems of applied phonology and lexicology. Since the research was based on demand, our approach was pragmatic and straightforward. Nevertheless, we faced problems on an almost daily basis. Most problems originated in the fact that existing writing systems have become so naturalized that they tend to take on a life of their own. In the case of linguistic interaction between humans and machines, as in the case of car navigation systems, such naturalization becomes a problem. Investigating these problems thus serves as a chance to deconstruct the ways in which we are influenced in our perception of language by the writing conventions we use. This is where our experiences in this commercial research project become relevant for JFL as well.

2. Description of the commercial research project

As mentioned above, the project was not planned as a scientific experiment, but grew in small steps. These steps were primarily instigated by requests from the company for which we conducted the research. Overall, the project involved the following steps.

(1) Checking texts containing a total of 900 Japanese words. Native speakers of Japanese read out sentences and word lists relevant to car navigation. These texts had already been produced in Japan. The texts were presented visually in transcriptions and over headphones to members of our team. They had the task of listening and reading the stimuli in *kanji-kana* mixed writing (*kanji kana majiribun*) and correcting the transcriptions if necessary. The transcriptions were in two writing systems: mixed *kanji-kana* writing and the Roman alphabet (*rōmaji*). Two to fifteen repeated runs were necessary to edit the texts.

(2) After the first round of corrections by one person was completed, two more members of the team verified the corrections which had been made.

(3) Next, the texts were analysed in order to obtain a phonetic transcription and in order to define the core lexical items to be used for the car navigation system.

(4) A semi-automatically derived vocabulary list was edited in more detail. For instance, the transcriptions of all long vowels in word endings *ou* were edited as *o+u*.

(5) These lexical entries were transcribed from the original *kanji-kana* mixed style into *kana* and several more writing systems and orthographies: Tu-Kana, Kunrei romanization, Hepburn romanization, In-

ternational Phonetic Alphabet (IPA) and SAMPA (Speech Assessment Method Phonetic Alphabet), a transcription system jointly developed by Temic and Tübingen University. All these transcriptions were necessary in order to allow the car navigation system to process the linguistic data: Tu-Kana was necessary for word-segmented texts, Kunrei for adapting the entries into the existing Japanese database, Hepburn for possible German and English databases, IPA for the creation of the actual spoken items and, finally, SAMPA for electronic transmission of the phonetic data.[1]

The following example illustrates the work steps described so far.

Tab. 1: Transcription process in the project

American English translation:	I wish I had a portable telephone.
Mixed *kanji-kana*	携帯電話があったらなあ。 keitaidenwa ga attara naa portable telephone I wish I had
Hiragana	けいたいでんわがあったらなあ。
(Semi-automatic identification of words)	
(Transformation to Tu-Kana)	
Hiragana	けいたい でんわが あったら なあ。 Keitai denwa ga attara naa
Tu-Kana	けいたい でんわが あったら なー。
(Transcription by Perl Script)	
SAMPA	kei:tai denwa ga a?tara na:
Hepburn	keitai denwa ga attara na
Kunrei	keitai denwa ga attara na

(6) Speech recognition of the navigation system for Japanese was tested.
(7) Product-specific spoken commands for human–machine–human communication via the auditory system were translated (from German and English to Japanese).
(8) Product-specific written commands for human–machine–human communication via the displays were translated (from German and English to Japanese).
(9) The vocabulary list was modified by adding new product-specific information.

[1] Tu-kana is a method of word segmentation and, at the same time, a master version for romanization which distinguishes between *onbin* [euphony] and long vowels.

(10) Correction of spoken and written commands with the machine interface, which allows the correction of errors that occur when translating without a fixed working system.

By now, readers might wonder why our department and 20 students from our institute were asked to develop tools for car navigation in Japanese, in spite of the fact that there are certainly plenty of well-trained Japanese computer linguists. As a matter of fact, the company in question had initially worked with computer linguists, but had encountered a plethora of problems in doing so (see below). Thus, Temic deliberately chose to consult specialists in Japanese with a philological background.

Since several variations in software systems for Japanese exist, it was necessary to engage students trained both in Japanese language and in natural language engineering in order to manage all programme versions and language encodings comprehensively. In addition to existing programmes, specific software for semi-automatic transcription and automatic error checking in Japanese and all its transcribed versions needed to be developed by our team.[2]

Our team understood Temic's problem immediately. The results of the preceding speech recognition project had not been solid enough to process the data. We decided to embark upon the project as follows: first we had to ascertain that a relatively small number of researchers, not all of whom were trained in linguistics, could contribute to the project. Since we knew that speech perception is always influenced by one's first language, in our case by native speakers of Japanese or German, we would compare their respective difficulties in identifying word and morpheme boundaries. Two problems in particular turned out to be laborious:

(1) In the case of the Japanese native speakers, the phenomenon of devoicing (*museika*) was problematic.

(2) For Germans, the biggest problem turned out to be the lack of standards for transcribing Japanese with the Roman alphabet. In particular, segmentation rules and Japanese alphabetic conventions proved to be laborious.

3. THE CASE OF DEVOICING

In step 6, the controlling of machine speech recognition, we encountered significant differences between native and non-native speakers of Japanese. In order to determine these differences more concretely, we started

[2] The software was mostly developed in Perl and Jperl (see Rich and Riechert 1998).

to compare the results systematically. The most prominent problem we encountered on the part of the native Japanese speakers was devoicing, an issue which has drawn much attention in Japanese linguistics for a long time and which can be traced back to the publications of the influential phonetician Sakuma Kanae (1929). While devoicing of vowels is clearly manifested in auditory acoustics, it is frequently not perceived psycho-acoustically. In other words, our Japanese native speakers often perceived vowels where there were none. Needless to say, this dramatically influenced the corrections of the vocabulary lists. Our Japanese consultant tended to transcribe /*ichi*/, even in cases when only the devocalized variant /*ich*/ had been articulated. This specific perception pattern originates in the fact that there are no consonant clusters in Japanese, which is manifested in the syllabic writing conventions of Japanese, in this case *i-chi*. Hence, the psycho-acoustic interpretation of /*ich*/ as /*ichi*/ was influenced by the native speaker's knowledge of existing writing conventions.

When presenting the lexical items in question over headphones, non-native speakers of Japanese initially marked the devocalized pronunciation more often than the native speaker. Non-native speakers of Japanese and German speakers without any knowledge of Japanese, too, differentiated between variants such as /*ich*/ and /*ichi*/ more clearly than native speakers. The deviance between native and non-native Japanese members of our team only vanished after we trained the native speaker about devoicing phenomena.

In the course of our working steps, the Japanese native speaker first read out a sentence and then compared it with the original written text. Since it was not possible to reduce the speed of the recording, this procedure had to be repeated several times. Then, initially, the native speaker checked his transcriptions together with a German speaker of Japanese. Needless to say, this task was extremely time-consuming. This is why we chose to train the native speaker of Japanese.

We have a clear understanding of how devoicing works from existing linguistic research.

Step 1: vocality [-consonantal] is reduced or lost.
Step 2: place of the articulation [+coronal] and continuity [+continuant] are retained.
Step 3: voicing may or may not be lost.

Phonological devoicing is thought to consist of spread of the feature (glottal spread) from the preceding obstruent to the following vowel. This results in the subsequent realignment of the spreading feature to the midpoint between the obstruent and the vowel. In this model, phonolog-

ical devoicing occurs between plosives, but phonetic loss of voicing can also occur between voiceless fricatives.

In summary, we encountered the following problem. Not differentiating between variants such as */ichi/* and */ich/* (one) or */roku/* and */rok/* (six) as a lexical entity in dictionaries gives rise to the view that such variance is irrelevant. When communicating with machines, however, this no longer holds true. Speakers frequently use devoiced variants and do not conform to written language norms. This was precisely the problem that the computer linguistics who had worked on the data before had failed to take into account.

4. The case of Romanization

The problems we encountered in the course of our project were not limited to native speakers of Japanese alone. German speakers of Japanese faced various problems concerning the Roman transcription of Japanese. The lexical entries processed for the car navigation system included the usual three types of Japanese lexical categories: indigenous Japanese words, Sino-Japanese words and western loanwords (Takeuchi 1999: 41):

Yamato Japanese	e. g. *hana* (flower)
Sino-Japanese	e. g. *happyō* (announcement)
Western loanwords	e. g. *hanbāgā* (hamburger)

Despite the existence of transcriptional techniques, it is noteworthy that the conventions for transcribing Japanese using the Roman alphabet are ambiguous. One problem is the influence from one's own language and its writing conventions. Consider a concrete example: on what basis should one decide whether "screen switch off" should be transcribed as one word (*sukuriinsuicchiofu*) or as three words (*sukuriin suicchi ofu*) following the English word boundaries? Note that the syntactic order (N-V) is Japanese and not English (V-N). There are, in addition, different transcription conventions in different European countries which reflect the writing conventions of their national languages. We thus have various European conventions for transcribing Japanese: the Portuguese system (1591), the systems used by scholars of western sciences (*rangaku*) in premodern Japan (1640–1854), the French Landresse system (nineteenth century), the Siebold system (1827), the Hepburn system (1867), the revised Hepburn system (1905), the Japanese official (*kunrei*) system (1885) and the revised Kunrei system (1937 and 1946 respectively), and the Meyer system (1971). Consider again a concrete example: the botanical name

Camellia sasanqua is an adaptation of Latin in Sino-Japanese. In the Roman alphabet, this term can be transcribed as *sazanka* or alternatively also *sasanka* (in both the Kunrei and the Hepburn system). The popular word for Camellia in Japanese is *tsubaki* (Hepburn), *tcubaki* (Portuguese), *toebaki* (Rangaku), *tsoubaki* (French), *tsubaki* (Siebold), *tubaki* (Kunrei 1946) and *zubaki* (Meyer).

As the car navigation terminology usually follows the Kunrei or Hepburn conventions, variations such as those above would rarely be possible. Nonetheless, when checking Japanese word lists by having them read by Japanese or non-Japanese, there remains some uncertainty, because certain possibilities for representing these Japanese words have been discarded. In order to account for the differences that exist between mixed *kanji-kana* writing and romanized Japanese, the method of matching Roman letters to *kana* and *kanji* must be flexible. In order to ensure such flexibility, we need to develop proper integration of linguistic knowledge and information retrieval technologies.[3] In order to avoid any ambiguities or mistakes, the most flexible solution was to use a romanized writing system *together* with the standard *kanji-kana* mixed writing.

5. Conclusions for Japanese language planning and JFL

As a result of our project we can draw attention to the fact that the creation of thorough lexicological and instructive systems for overcoming the differences in the Japanese and western perceptions of Japanese words is still urgently needed. Human beings have a highly developed and finely articulated model of the world, closely related to the vocabulary and grammar of their mother tongue. Communication depends on the congruence of such world models and of the respective language and its writing system, which have been internalized by its (native) speakers. Discovering these intermediate structures and developing them in a more organic way should also be a requirement for foreign language educators. When accounting for the complexities we faced in our project, three points of view emerged, which need to be kept in mind simultaneously:

[3] Problems of transcription are, of course, not limited to applications of car-navigation systems. The Risk Digest-Forum on Risks to the Public in Computer and Related Systems gives a detailed outline of problems involved with transcriptions (van Meter 2003).

(1) syntagmatic and paradigmatic characteristics of a particular sign system,
(2) knowledge-oriented issues when relating different structures to different cognitive strategies and
(3) using the systems as a part of human behaviour.

Sign systems must thus be treated as complex systems for particular groups of people at a particular point in time. This issue is not limited to problems of developing linguistic tools for a car-navigation system. It is relevant for all communicative interaction.

In the course of a commercial research project primarily conducted on pragmatic grounds, our team, which included students of Japanese, discovered and experienced for themselves issues of fundamental significance. The experiences of this project highlight the importance of including concrete linguistic assignments for students in order to allow them to discover and manage language problems which are encountered in the treatment of foreign languages. This, I would argue, is essentially a topic for foreign language educators.

References

Rich, Morin and Andreas Marcel Riechert (1998): Perl, MacPerl and MacJPerl. www.macperl.com/depts/articles/P_MP_MJP.html (found November 2005).

Sakuma, Kanae (1929): *Nihon onseigaku* [Japan Phonetics]. Tokyo: Kyōbunsha.

Takeuchi, Lone (1999): *The Structure and History of Japanese from Yamatokotoba to Nihongo*. London: Longman.

Van Meter, Rodney D. (2003): Japanese transcription. http://catless.ncl.ac.uk/Risks/17.03.html#subj7.1 (found 28 January 2007).