

# S-shaped Curve of Phonological Standardization – Six Surveys in the Tsuruoka and Yamazoe Areas

FUMIO INOUE

*Meikai University, /1 Akemi, Urayasu, Japan,  
innowayf@yahoo.co.jp*

---

SCN III/1 [2010], 101–112

---

Prispevek obravnava hitrost jezikovnih sprememb. Jezikovne spremembe zaznamo kot otipljiv družbeni pojav, vendar se pri tem pojavlja teoretično vprašanje razlikovanja med realnim oz. dejanskim in navideznim časom. Opazovanje jezikovnih sprememb zahteva večkratno ponovitev raziskave, prav japonska dialektologija pa se s tem še posebno natančno ukvarja – pri tem krivulja navideznega časa ustreza krivulji realnega časa, in sicer je za jezikovno spremembo potrebnih skoraj dvesto let.

This paper is concerned with the speed of linguistic change. If social variation is a reflection of linguistic change, the scene of the linguistic change can be caught in action as a palpable phenomenon. However there is one theoretical question, the distinction of “real time” and “apparent time”. In order to observe linguistic change repeated surveying is necessary. The tradition of Japanese dialectology provides us with several examples of repetitive surveys. The conclusion of the analysis is as follows. The curve of apparent-time change corresponds to that of real-time change. The total number of years necessary for a linguistic change to be completed is nearly 200 years.

**Ključne besede:** realni in navidezni čas, standardizacija, S-krivulja, hitrost difuzije

**Key words:** real time and apparent time, standardization, S-curve, speed of diffusion

---

## 0 Introduction

This paper is concerned with the speed of linguistic change. First, the theoretical background of this paper will be discussed. This study is intended as a development of the field of computational dialectology, and is also a cornerstone of sociolinguistics. The starting point of this kind of research is William Labov's treatment of social stratification as one aspect of linguistic change in progress. If social variation is a reflection of linguistic change, the scene of the linguistic change can be grasped or caught in action as a palpable phenomenon. The speed of the linguistic change can also be inspected on the spot or observed at the scene.

However there is one theoretical question which has been pointed out by Labov, that is, the distinction of "real time" and "apparent time" (Labov 1972). The investigation of age differences makes use of apparent time, but apparent time is not a simple and direct reflection of linguistic change. In order to observe linguistic change repeated surveying is necessary. The tradition of Japanese dialectology luckily provides us with several examples of repetitive surveys, the representative one being the three surveys of Tsuruoka City and the three surveys of one of its outlying areas, Yamazoe.

The study of phonological standardization was widely conducted after the Second World War in Japan. One noteworthy investigation was a field survey which was carried out in 1950 in Tsuruoka City, in the northern part of Japan (see Figure 8). The results were so clear that two consecutive surveys were conducted,<sup>1</sup> one about 20 years later and the other about 40 years later (Yoneda 1997). When the main results of these three surveys were brought together, linguistic change was found to form an S-curve, as is shown in Fig 1. The process of standardization discussed here is just the reversal of decline of dialect.<sup>2</sup>

<sup>1</sup> Approximately from 400 to 600 Samples (informants, consultants) were selected randomly for the three field surveys in 1950, 1971 and 1991. Field workers are trained linguists and dialectologists. Items investigated are shown below in approximate Hepburn Romanization System in the order of "standard Japanese / Tsuruoka dialect".

**Phonology** (31 items): Velarization 1: kayobi/kwayobi, suika/suikwa; Velarization 2: hebi/febi, hyaku/fyaku, hige/fige; Palatalization: zeimusho/jeimusho, senaka/shenaka, ase/ashe; Voicing: matsu/mazu, kaki/kagi, kutsu/kuzu, hato/hado, hata/hada, hachi/hazi, kuchi/kuzi, neko/nego; Nasalization: obi/onbi, mado/mando, suzu/sunzu; Centralization 1: kitsune/kizune, karasu/karasi, sumi/simi, chizu/chizi; Centralization 2: shima/sima, uchiwa/uziwa, chiji/chizi, karashi/karasi; i and e: entotsu/intotsu, eki/iki, ito/eto, iki/eki. **Pitch accent** (5 items): neko, hata, senaka, karasu, uchiwa.

<sup>2</sup> Answers in standard pronunciation were given 1 point, and in dialectal pronunciation, 0 point. Average ratio was calculated for the informants dividing the total by 31 (phonology) and 5 (pitch accent).

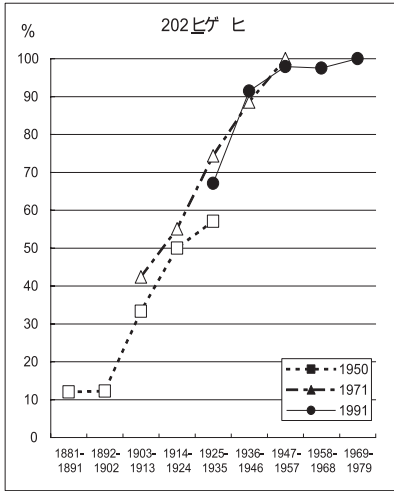


Figure 1. Ratio of standard pronunciation [hi] for “hige” (beard) in three surveys of Tsuruoka showing a typical S-curve over one century

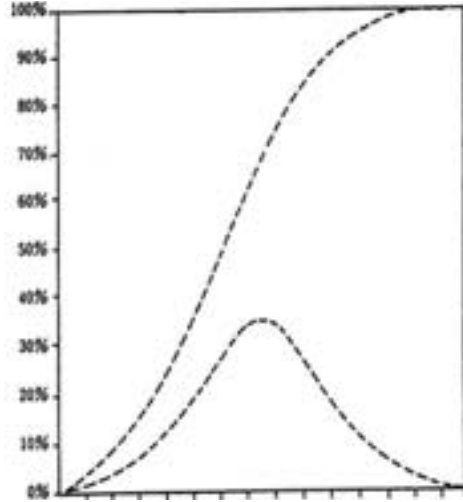


Figure 2. S-shaped curve as a model of diffusion

## 1 Background

The S-curve is often observed in the process of diffusion and spread of various phenomena. There are various theoretical bases for the hypothesis of the S-curve, and several calculating formulae. As is shown in Figure 2, it is often assumed as the accumulation curve of normal distribution in statistics.

## 2 Material: Three surveys in Tsuruoka

Figure 3 shows the aggregate results of the three Tsuruoka surveys.<sup>3</sup> Three surveys of the phonetic items compose the upper half of the S-curve, and the three surveys of accent compose the first lower half of the S-curve.

<sup>3</sup> The informants are divided into five generations (eleven years apart) by their birth year in order to more exactly calculate real time linguistic changes. They are also plotted by their birth year in Fig 3. Upper three lines represent phonological phenomena, and lower three lines, phonological phenomena,

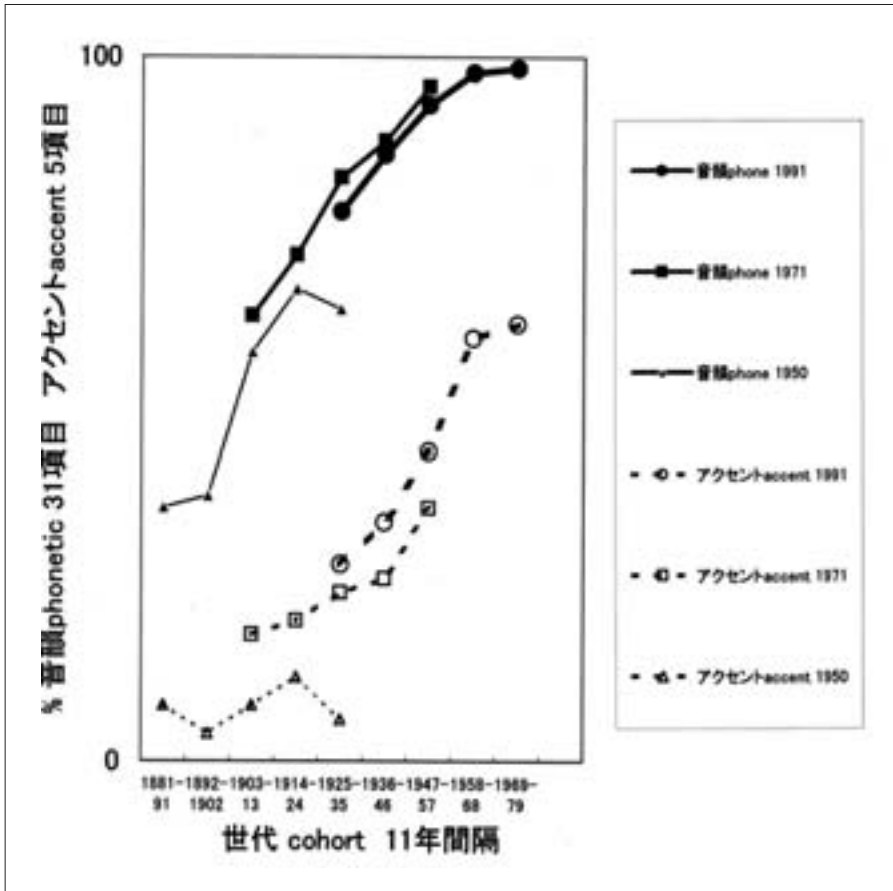


Figure 3. The phonetic items and accent of three surveys in Tsuruoka according to age group

### 3 Application of a mean value to the Tsuruoka data

#### 3.1 Social survey and age-linguistics

Inoue (1997) applied the absolute age movement method making use of the actual year of birth of informants (or cohort). The numerical values of standardization of the phonetic items and accent of the three surveys were moved on the graph by observation on a trial and error principle based on the hypothesis that accent composes the first (lower) half and the phonetic items composes the second (upper) half of the S-curve. Afterwards, a technique for more accurate graphing was felt indispensable. If it were possible to locate the curves by some appropriate numerical value, the connected curves of accent and the phonetic items could be arranged in succession to form an S-curve.

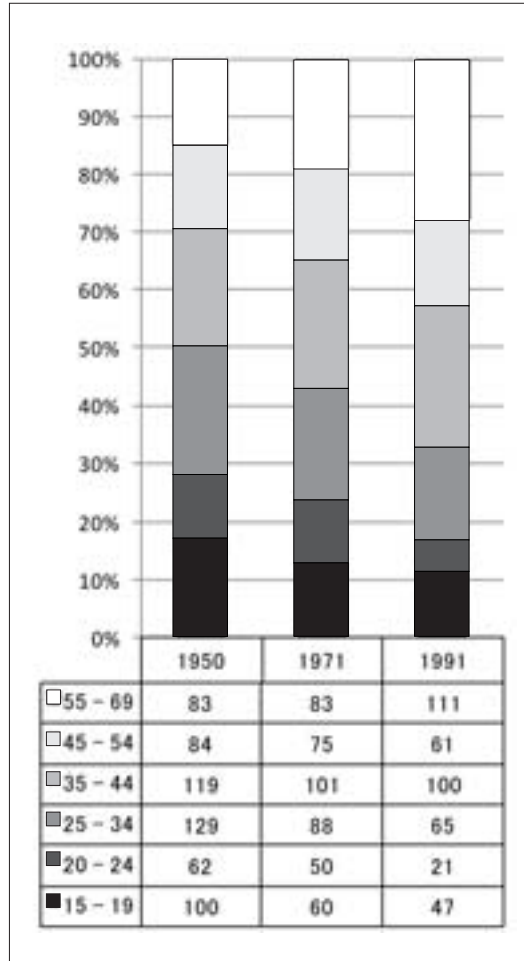


Figure 4: Ratio of informants according to age group of three surveys of Tsuruoka

First of all, the use of the average score of the whole group of informants in each of the surveys was considered. However, there were some problems. The mean values of the entire sample were different depending on the time of the survey, as is shown in Figure 4. In this graph, the ratio of informants of the three surveys in Tsuruoka city is shown according to age group. In the first survey, the numerical values of young informants greatly influenced the mean, because the number of older informants was low. In the third survey, the influence of the numerical values of the older informants was stronger because the number of younger informants was low and the number of older informants was high. Appropriate measures of dealing with age in linguistics are necessary in such a situation in order to advance academic sociolinguistic research.

### 3.2 Average of age group mean values

As the ratio of young people differs depending upon when the survey was carried out, a new concept of “the average of mean value” was examined as follows. It is basic knowledge of statistics that “the average of the mean values of age groups” does not represent the whole picture. Therefore, it is necessary to calculate the total mean value of an entire resident group. “Mode value” and “median” (not a mean value) are sometimes used according to the data, and the “moving average method” is sometimes used. However to look at the data of social surveys, a newer concept is necessary (Inoue et al. 2009).

In a sociolinguistic investigation which makes use of age differences as a cue to *apparent* time linguistic change, it is necessary to consider the value of each age group equally. In other words each age group should be treated and weighted so that they are equal in weight, as it would be inappropriate for them to be weighted according to the age group composition of the whole population of the region at the time of the survey.

This idea of weighting each age group as equal is actually the same thing as “average of the mean value”.<sup>4</sup>

### 3.3 Application of mean value method

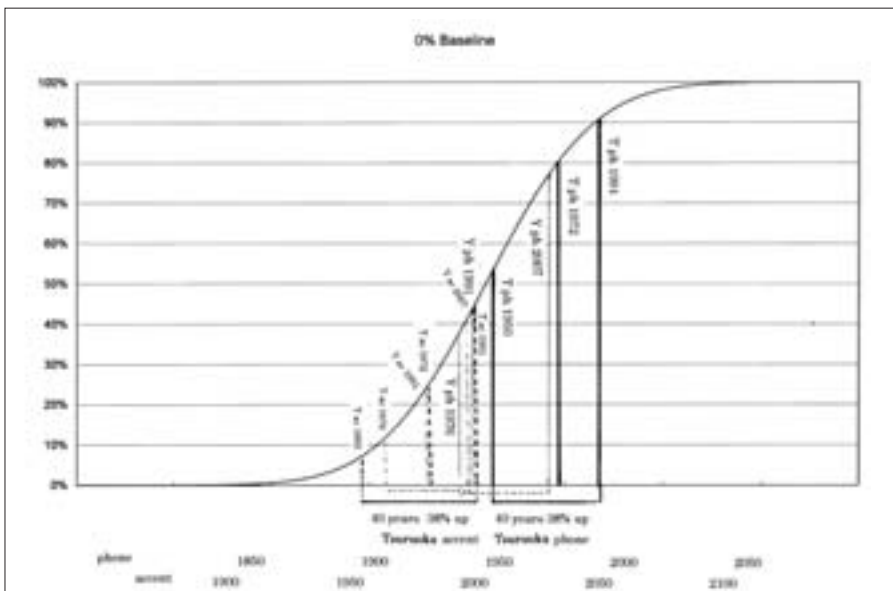


Figure 5. S-curve of three surveys of Tsuruoka and Yamazoe: mean values of the phonetic items and accent

<sup>4</sup> The common sense of statistics is thus overturned. It is, so as to say, a Copernicus conversion, or Columbus’s egg. This idea is the substance of this paper.

Ideally, the numerical value of each age group at the time of survey should form a part of the S-curve. In that case, the averages of the mean values of each age group would appropriately represent the usage pattern of the community. According to this concept the results of various surveys of actual conditions would be represented by one numerical value of an average, and this average would not be influenced by variations in the number of informants in any age group.

### 3.4 Explanation of graphing

In Figure 5, the average of the mean values of each age group was plotted on a theoretical S-curve. The position of each survey was located at the corresponding value position on the S-curve as this technique aims at correspondence with the mean values of the age group. In order to make the correspondence with the year of survey easily understood, perpendicular lines using different shapes and thickness were drawn from the S-curve to the bottom line. The phonetic items are shown by solid lines and pitch accent is shown by dotted lines. Tsuruoka is shown with heavy lines and Yamazoe is shown with thin lines. Yamazoe will not be explained here because of the length limit. In this graph, the year of the survey was the main clue. This method allows us to see how the mean values have moved on the S-curve over the 40 actual years of the three surveys (or **real time**). The result as a whole conformed to the theoretical S-curve very well.

### 3.5 Results of the three successive surveys

The time intervals on the graph of Figure 5 correspond closely with numerical values for the phonetic items and accent of the three surveys in Tsuruoka. The rise of almost 38% for both the phonetic items and accent was observed over the 40 years from the first survey to the third survey of Tsuruoka. The length in years on the horizontal axis corresponds with the 38% rise for phonetic items and accent. This result shows that the technique of arranging the results of the three surveys on the theoretical S-curve by simple mean values was appropriate. Language standardization seems to have advanced at the rate of 1% a year on average over the 40 years.

The number of years necessary for a linguistic change can be calculated from the above results. The years of the Tsuruoka surveys for the phonetic items and accent are respectively shown beside the perpendicular lines of the graph. As this S-curve is an accumulation curve of normal distribution, it shows a change from an infinite smallness to 100%. When the increase and decrease of 40 years in Tsuruoka are moved on the S-curve and extended, then, 1% becomes 99% in a time length of a little less than 200 years.

#### 4 Pasting the Tsuruoka data to the S-curve

The above-mentioned attempt can be developed further as shown in Figure 6. It is possible to test whether the age differences of the three surveys fit the theoretical S-curve. We have only to paste the actual graphs of Figure 3 (adjusted to the same ratio as Figure 5) to the theoretical S-curve of Figure 5. A graph with smoothing technique was applied so that a more appropriate result is performed by the S-curve. It was presumed from Figure 3 that the results for the phonetic items and accent would respectively conform to the S-curve. The data was divided into two; the phonetic items and accent, and they were pasted together on Figure 5. Actually values for the phonetic items and accent pasted on Figure 5 were based on mean values. As can be seen, they fitted beautifully.

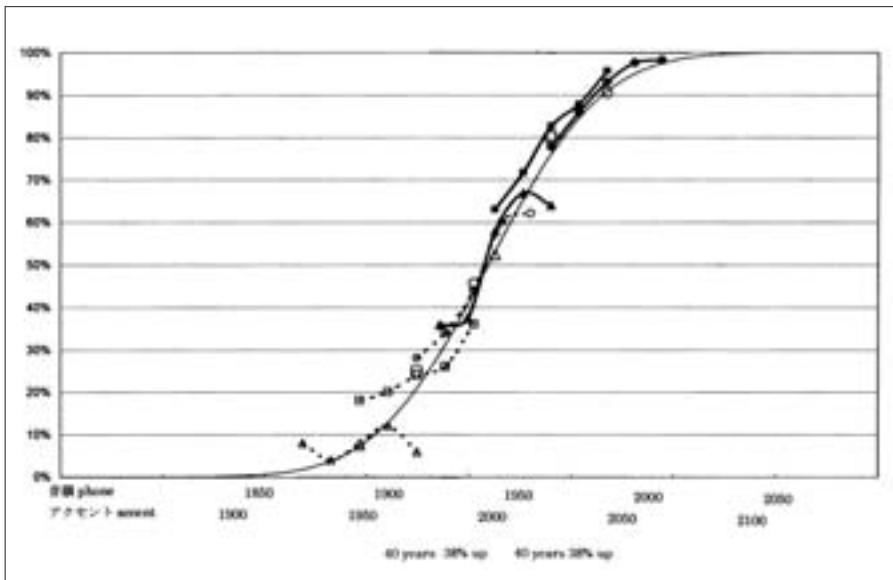


Figure 6. Mean values of the phonetic items and accent: age groups of three surveys in Tsuruoka

In Figure 6, the phonetic items are shown by the solid lines and accent by the dotted lines. The three surveys are indicated differently: the first survey by triangles, the second by rectangles and the third by circles. The result shown in Figure 6 of pasting the phonetic items and accent by mean values of the three surveys in Tsuruoka showed the beautiful pattern of an S-curve, by the rhythm of slow – quick – quick – slow. Some age curves became similar to a part of the S-curve, and they became easier to interpret. On the whole, age difference is great at the center of the graph which reveals quick change. Age difference is small in value at the far left side of the graph and the far right. Change is slow at the positions of extremely high (more than 80%) or low (less than 20%) usage rates.



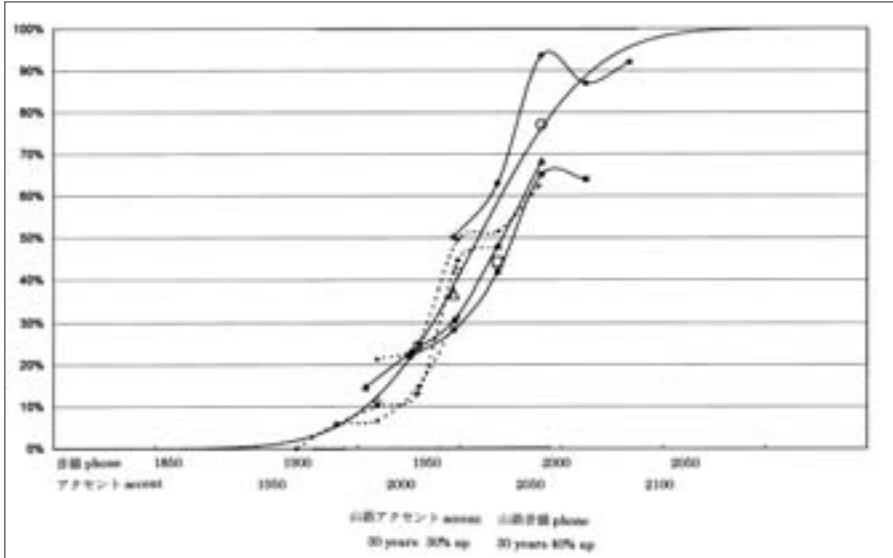


Figure 7. Mean values of the phonetic items and accent: age groups of three surveys in Yamazoe slow at the positions of extremely low (less than 20%) or high (more than 80%) usage rates.

### 5 Analysis of the Yamazoe Data

Comparing the Tsuruoka results with the results of another survey would allow us to look at geographical or *geolinguistic* dimensions. To look at geographical diffusion over forty years, data of a similar survey of the Yamazoe area on the outskirts of Tsuruoka City was used.



Figure 8. Map of Tsuruoka and Yamazoe

The corresponding graph of the Yamazoe surveys is shown in Figure 7. Though a minute analysis of the Yamazoe data was omitted here, a difference of 20 to 30 years for language standardization was observed between Tsuruoka and Yamazoe.<sup>5</sup> This means a difference of about 20 to 30 years of diffusion speed when the same generation is compared. The distance between Tsuruoka city and the Yamazoe area is about four kilometers as is shown in Figure 8. The speed of language change seen here was very different from the tendency of “a speed of one kilometer a year” of nationwide dialect diffusion (Inoue 2003), which was advocated on the basis of many linguistic changes in progress using the technique of glottograms or age-area graphs (Inoue forthcoming). This may be because of differences due to the regional character of the midsize cities and farm villages. There are considerable differences between the cities and the farm villages, not only in language but also in folkloristic customs, and in contrast, there are a lot of common features among the farm villages in that part of the plains. Differences in the occupations of residents might also have influenced language use. Social and geographical differences in honorifics have also been pointed out in the previous studies (Inoue 1989). The difference in the phonetic items is also great.

## 6 Discussion

It was thus confirmed that the idea of the S-curve could be applied to the three surveys of Tsuruoka and the three surveys of Yamazoe. The tendency that the age curves generally conform to the S-curve was made use of to draw graphs. The S-curve was applied to the 40 year time-span of the three surveys, and the number of years necessary for phonological change was speculated. A little less than 200 years was calculated as being necessary from the beginning to the end of a change towards language standardization.

This is an aggregate numerical value of the synthesis of a lot of changes, and an individual change might proceed faster. For instance, as the graph in Figure 1 shows, the dialectal pronunciation of [ɸige] (beard) of Tsuruoka was completely replaced by the standard pronunciation of [hige] within three generations (or in about one century). This tendency of language change became clearer when the age differences of the three surveys were connected to the year of birth of the informants. The time needed for change differs greatly between individual words, and if other words and phenomena were included together, the time needed for change may be longer.

---

<sup>5</sup> It is because the curves of Yamazoe survey 20 or 30 years later show similar acquisition patterns to Tsuruoka survey.

## 7 Conclusion

The analysis above has shown that when the S-curve model of linguistic change was applied to the data, the curve of *apparent time* change corresponded to that of *real time* change, and that the total number of years necessary for a linguistic change is more than 100 years and nearly 200 years. The collection of concrete survey data is important in order to see the actual situation of dialects. We should not however stop at mere collection of data. We should proceed further to reach a more profound theory of dialect or language change. When actual age differences are observed by field surveys, the tendencies can be applied using the theoretical pattern of the S-curve, like in Figure 5. Then, we can guess or presume the magnitude, time, and necessary years of the change. It may be possible to apply this conclusion to other languages. By collating this tendency with studies dealing with age differences in language usage in other parts of the world, a more general theory on the speed of change may be accomplished.<sup>6</sup>

The study of dialectology is domestic if only the differences or variations within one language are investigated. However, the methodology, technique and general theory of dialectology can be applied to any language, and are thus universal. This International Congress is a wonderful place for the exchange of ideas, for us to learn from and stimulate each other. I am sure that the development of dialectology in the 21<sup>st</sup> century will become more fruitful owing to this international congress.<sup>7</sup>

## REFERENCES

- Fumio INOUE, 1989: *Kotobazukai Shinfukei (New scenery of speech)* (Akiyama Shoten).
- , 1997: S-shaped curves of language standardization. Alan R. Thomas (ed.) *Issues and Methods in Dialectology*. pp. 79–93.
- , 2003: *Nihongo wa nensoku 1 kiro de ugoku (Japanese language moves 1 km per year)* Kodansha Gendai Shinsho).
- , forthcoming: Real and apparent time clues to the speed of dialect diffusion. *Dialectologia 5 (revista electronica) (Barcelona)* <http://www.publicacions.ub.es/revistes/dialectologia/>

---

<sup>6</sup> Some related topics could not be discussed here because of shortage of time. All the English papers by me are now open to the public through the internet. If you use yahoo or Google, please search using key words like, “Fumio Inoue, dialect (sociolinguistics, papers)”.

<sup>7</sup> The original paper was read at the 6<sup>th</sup> SIDG (International Society for Dialectology & Geolinguistics) congress, held at University of Maribor, Slovenia, on 16 Sept 2009.

Fumio INOUE, Kiyoshi EGAWA, Ryoichi SATO, 2009: "S-shaped curve of phonological standardization -- six surveys in Tsuruoka and Yamazoe areas. *Mathematical linguistics* 26–8: 269–289.

William LABOV, 1972: *Sociolinguistic Patterns* (Philadelphia).

Masato YONEDA, 1997: "Survey of standardization in Tsuruoka, Japan" *Nihongo Kagaku* 2.

#### S-KRIVULJA FONOLOŠKE STANDARDIZACIJE – ŠEST RAZISKAV NA PODROČJIH TSURUOKA IN YAMAZOE

Široko zastavljena študija fonološke standardizacije na Japonskem se izvaja že od druge svetovne vojne, zelo natančno predvsem v mestu Tsuruoka in njegovih predmestjih (Yamazoe) – za vsako lokacijo so bile izvedene in teoretično obdelane tri raziskave. Raziskava kaže, da jezikovne spremembe oblikujejo krivuljo S-oblike. Kadar uporabimo teoretično krivuljo, ki temelji na tej zamisli, lahko izračunamo število let, potrebnih za jezikovno spremembo. V skladu s predhodnimi študijami je od začetka do konca jezikovne spremembe potrebnih več kot sto let. Tokrat je bilo pridobljenih več podatkov, uporabljene pa so bile tudi novejša tehnika izračuna za preverjanje rezultatov prejšnjih študij.

Rezultati kažejo, da posamezna krivulja razlike v letih sovпада s teoretično oblikovanimi grafi S-oblike in da je potrebnih skoraj dvesto let, da se zgodijo temeljne jezikovne spremembe.

---