Published in: Y. Asahi, M. Usami & F. Inoue (Ed.), *Handbook of Japanese Sociolinguistics* (pp. 19-40).
Berlin, Boston: De Gruyter Mouton, 2022. https://doi.org/10.1515/9781501501470-002

Florian Coulmas

1 Research methodology

1 Introduction

The methods used to investigate language and explicate patterns of structure and usage are sometimes conceptualized as neutral tools anyone can handle to study any language. This is not so. As a matter of fact, and not surprisingly, research methodologies are as historically contingent as are theories, reflecting in one way or another the traditions, theories, and sociocultural settings that gave rise to their gestation. Although causal relationships between social circumstances and theory formation are not always direct or obvious, it cannot be denied that in sociolinguistics as in all social sciences, theories and methods are subject to influences of the societies in and for which they were developed. Since Sociolinguistics as it is understood by mainstream handbooks, journals and textbooks emerged as a new field of scholarship in and for the investigation of language in Western industrial societies, it is characterized by intense Western dominance (cf. Smakman and Heinrich (eds.) 2015). This means that most sociolinguists are Western-educated, that European languages have been more often subject to sociolinguistic investigations than non-Western languages, and that Western scholars are overrepresented on the editorial boards of journals in the field of Sociolinguistics.

In the 1960s, when the work of Basil Bernstein, in Britain, and William Labov, in the United States, provided a major impulse to studying language in industrial society in a new way that differed markedly from both the description of language as a generalized atemporal structure in the Saussurean sense of synchrony and the description of regional distinctions in the vein of dialectology, it was common to make a distinction between sociolinguistics and ethnolinguistics. The former dealt with industrialized societies of the West, the latter with post-colonial societies where scholars and missionaries had been or were active. By and large this division of labor corresponded to that between sociology and ethnography or cultural anthropology. Sociology was self-centered and based in Western societies, while ethnography studied from a Western point of view how "premodern" societies, alternatively called "primitive", "oral", "autochthonous", or "indigenous groups", differed from the West (Said 1978).

A principal point of difference between traditional ways of studying language and the new paradigm of sociolinguistics had to do with social change. Industrialization induced population movements and concentration. Urbanization created major differences between city versus country life that undermined the notion of relative linguistic homogeneity in small spatial extension. Sociolinguistics emerged as the "linguistics of the city". Since the industrial revolution, urbanization progressed most forcefully

https://doi.org/10.1515/9781501501470-002

and rapidly in Western countries. The study of language in an urban setting therefore first developed in the West and thus quite naturally initially became a Western science (Coulmas 2013: 21). Research about the nexus of language and society that developed in other parts of the world without much participation by Western scientists was largely ignored.

The fact that this was also so in regards to research about the Japanese language in its social setting can be attributed to:

(1) the general theoretical Eurocentrism prevalent in the social sciences,
(2) the language barrier, and
(3) the scarcity of publications about Japanese outside Japanese philology and in languages other than Japanese.

Reflecting actual power relations and Western dominance, the language barrier was one-way. The reception of Western scholarship in Japan and Western academics lecturing in their own language or in English in Japan (and anywhere else outside the Western world) were taken for granted, while active command of Japanese outside Japan was confined to a small circle of specialists.

In the case of Japan, the history of sociolinguistics thus described, however briefly and simplistically, is remarkable for two reasons.

Firstly, even though Japan was a latecomer to industrialization, urbanization had begun to transform life in Japan earlier than in most countries and by the end of the nineteenth century had reached a level commensurate with European countries. Already in the eighteenth century, Tokyo was one of the biggest cities of the world.

Secondly, when in the 1960s scientists in advanced Western countries began to explore correlations between social and linguistic characteristics, this way of thinking about language was not new in Japan. As early as in the 1930s, sociologist Juri Tanabe had coined the term *gengo shakaigaku* ‘sociology of language’ for a new field of scholarship (Tanabe 1936), while during the same period anthropological linguist Kyōsuke Kindaichi established “Language Life” (*gengo seikatsu*) as an empirical research paradigm that guided Japanese social linguistics for decades (see Chapter 4 (Inoue, this volume)). However, the first wave of sociolinguistic research in the English-speaking world focused on how class distinctions found expression in different codes, social class being a concept that was predicated on city life in Western industrial societies and, therefore, could not be assumed to be readily applicable to Japanese society. Although Karl Marx’s and Max Weber’s conceptions of social class were known and debated in Japan, social class analysis did not inform language life research.

2 Empirical research about language life

The language life paradigm exemplifies the point made above about the historical contingency of research methodologies.[1] After being initially theory-driven based on Motoki Tokieda's utterance process model, language life research turned empirical after World War II. *Kokuritsu Kokugo Kenkyūjo* (National Language Research Institute, NLRI hereafter)[2] founded in 1948 became the hub of socially relevant linguistic research. As the name indicates, this is a national institute charged with particular tasks deemed relevant for functions of the state, notably education. In conjunction with the many social changes of the post-war period, four issues became topical: literacy assessment, language standardization, honorific speech, and a general survey of regional forms of Japanese.

3 Literacy

The first representative literacy survey in Japan ever was carried out in 1948 under the direction of Kikuo Nomoto who would later become director of NLRI. It was a massive empirical inquiry collecting and processing data by some 17,000 informants in 270 locations (Yomikaki Nōryoku Chōsa Iinkai (ed.) 1951). The primary intent of the survey was to establish literacy levels in order to enable informed decisions about adjusting language planning and educational policies. Japan's modernization was accompanied by deliberate attempts to reduce the reliance on Chinese characters and narrow the gap between spoken and written language (*genbun itchi*). The defeat in the Second World War revitalized the writing reform discussion, and to this end, data about proficiency in reading and writing were needed.

In addition to literacy levels, the survey collected social and demographic data, such as gender, age, education, and occupation of the respondents. The results of the survey suggested a higher overall literacy level than had been expected. According to the findings, only 1.7 per cent of the respondents were illiterate. The statistical analyses of the data furthermore revealed increasing literacy levels with declining age: The younger the informants the higher the proficiency. This led to the conclusion that literacy was not a pressing problem. These findings effectively terminated the writing reform debate. In particular, Romanization, the strategy favored by the US occupation, was no longer seriously considered. In the absence of any reliability tests, the findings of the survey were taken at face value and gave rise to the cherished belief

---

1 In this section I follow, by and large, Heinrich's (2001: 171–75) account of the development of language life research.

2 Renamed National Institute for Japanese Language and Linguistics (NINJAL) 2009.

that illiteracy did not exist in Japan. Consequently, no further literacy assessment studies were undertaken for the next half century.

The participants' literacy skills were tested by means of a huge instrument consisting of 90 questions, but the statistical techniques for processing the data were rather crude. Methodologically the survey was straightforward, making use of questions in the manner of fill-in school exams. These questions fell into eight categories, as follows:

(1) writing kana,
(2) writing numerals,
(3) reading kana,
(4) reading Chinese characters,
(5) writing Chinese characters,
(6) context-specific comprehension of Chinese character words,
(7) comprehension of synonyms,
(8) sentence comprehension.

The completed questionnaires were graded like school exams. Since only respondents who failed on every category were counted as illiterate, the survey was later criticized for obscuring rather than elucidating the true situation of literacy in Japan (Sumi 2005).

4 Speech community and the diffusion of the standard language

One of the researchers involved in the fieldwork for the literacy survey was Takesi Sibata. He transferred research methods he had acquired in that context to another large survey conducted in 1950 about speech communities, *gengo chiiki shakai*. The Japanese term literally translates as "local speech communities", indicating what the survey was about: the local variation of linguistic behavior that distinguishes one community from another.

Japanese dialects had been studied at least since the 1920s when Kunio Yanagita, pioneer of Japanese folklore studies, developed his theory of concentric circles (*hōgen shūken-ron*) which posited Kyoto as the origin from which linguistic innovations spread to other parts of the country (Yanagita 1963). Yanagita was an empiricist at heart. He relentlessly roamed the country from north to south collecting linguistic data and observing folk events and practices in remote places. His concentric model of language diffusion did not eventually stand the test of verification (Inoue 2011), but it inspired empirical research about dialects. In post-war Japan, this kind of research flourished.

At the time, Sibata was the leading researcher about geographic language variation. However, his many studies were not just about the geographic distribution of varieties but tackled theoretical and methodological issues that arose in conjunction with

demographic changes, such as the diminishing population density in rural areas, on one hand, and increasing urban sprawl, on the other. Increasing life expectancy and intensifying social mobility further changed the social fabric of Japan. One of the overarching questions the NLRI addressed in a succession of surveys concerned the diffusion of the standard language[3] (*hyōjungo*) defined as the educated dialect of Tokyo. For the purposes of a uniform system of education throughout the country and since widespread exposure to mass media began to have an effect on language behavior throughout Japan, early surveys concentrated on peripheral rather than metropolitan regions, such as the island of Hachijōjima, Shirakawa City in Fukushima Prefecture, and Tsuruoka City in Yamagata Prefecture. The questions of how fast and to what extent the standard variety of Japanese would penetrate regions outside the capital raised methodological issues as it involved language change rather than just a static description of the distribution of regional dialects. Of particular importance was the question of how linguistic data were to be correlated with the age of the respondents.

The question Sibata and other researchers at NLRI debated was whether and how the age structure of linguistic communities and aggregate findings about the ability to use the standard language and actual usage could be taken as a measure of the degree of language standardization. They therefore designed instruments for assessing proficiency in the standard language both in speech and with respect to reading and writing. Findings of the surveys in Shirakawa City and Tsuruoka City revealed a consistent tendency of declining proficiency values with increasing age, which Sibata interpreted as indicating progressing standardization (Sibata 1999 [1952]). In the event, the observed differences were understood not as age cohort-specific speech behavior but as language change over time, anticipating as it were the apparent time concept that was later developed in mainstream variationist sociolinguistics (Bailey 2002).

The apparent time method, which assumes that linguistic differences that vary in a regular way with age differences in a sample population was a makeshift solution to a fundamental methodological problem: How to study language change if actual observation is confined to the moment? Incremental language change is continuous and so slow that it defies direct observation. Prima facie language change can moreover mean two different things,

(1) that individual speakers change their habits (of pronunciation, word choice, or phrasing) and
(2) that the aggregate variety that defines a community changes.

---

3 The terms *kokugo* 'national language', *kyōtsūgo* 'common language', and *hyōjungo* 'standard language' are sometimes used interchangeably. However, there is also a theoretical discussion about how they differ in meaning. While a standard depends on explicit definition by an agency authorized to set rules, "common language" just refers to a variety that does not exhibit any regional accent. Cf. Tokugawa (1988).

How (1) and (2) interact, how sociodemographic characteristics of individual speakers and socially defined subgroups impact the process of change are genuinely sociolinguistic questions. The Shirakawa and Tsuruoka surveys addressed questions like these. However, that age differences in regards to a given linguistic feature observed in data recorded at the same time can be justifiably interpreted as representing language change over time is a hypothesis that can be verified or falsified only in retrospect, after a sufficiently long period has elapsed. The practical problems that stand in the way of longitudinal studies spanning several decades are such that studies of this sort are exceedingly rare.

It is here that the NLRI surveys made significant contributions to sociolinguistic scholarship. The point of departure was to conceptualize language change as frequency of occurrence in a given population, rather than the presence or absence of certain features. Using this concept in order to measure stability and change, researchers at NLRI replicated the original 1950 Tsuruoka study after two decades and then again in 1991 and 2011 (Kokuritsu Kokugo Kenkyūjo 2007; Tōkei Sūri Kenkyūjo and Kokuritsu Kokugo Kenkyūjo 2014). In the second and even in the third wave of the survey it was possible to re-interview many of the original respondents making it possible to measure aggregate and individual change. There were 497 respondents in the first wave, 107 of which could be contacted and re-interviewed in the second wave. In addition to these 107, new respondents were added in order to have a comparable sample of 401, as indicated in Table 1.1. The same procedure was followed in waves three and four. Each wave included a sizeable number of the previous wave's respondents. Four respondents of the original random sample survived in the fourth wave of 2011.

Table 1.1: The sample of interviewees of the Tsuruoka language survey, as it changed over time. (Kokuritsu Kokugo Kenkyūjo 2013: 3).

	1950	1971	1991	2011
Original random sample	496	107	53	4
Total sample	496	401	405	466

The wealth of data collected in the four waves of the survey makes the Tsuruoka study unique. Its main features can be summarized as follows. It treats language as a statistical vector rather than essentializing it. Combining insight into individual and aggregate change, it contributes to our understanding of what a speech community is and how it moves through time. And by totaling rates of change of individual features it provides a measure of progressing language standardization.

Methodologically the Tsuruoka study built on the literacy survey, but it was much more sophisticated and became more so from wave to wave, not least thanks to the

cooperation with the Institute of Statistical Mathematics.[4] The first wave used a random sample of subjects chosen from the city registry. They were interviewed by the researchers one by one on the basis of a questionnaire that solicited information about gender, age, domicile, and education. Focusing on the pronunciation of segmental sounds as the most reliable indicator of language standardization, the 1950 questionnaire included 31 words the respondents were asked to pronounce. No attempts were made to solicit unmonitored speech, the assumption being that it is difficult to change one's pronunciation at will. Responses were judged by the investigators as to whether or not they conformed to the standard. For instance, if the word *eki* 'station' was pronounced with initial [e], the standard pronunciation, it was rated one, and if the onset sound was [i] or [ɨ], zero. A respondent's total score on all items was the measure of that speaker's degree of standardization, and in this way the added scores of age cohorts and other subgroups as well as of the whole sample gauged the degree of standardization.

This method combined keen observation and the selection of phonetic features that were indicative of regional speech and likely to give way to the pressure of standardization with statistical techniques to measure the prevalence of these features in a sample. Since the data allow for the analysis of language change of both individual speakers and of a speech community, they have been processed and made available to the scientific community in highly elaborate form (Abe 2010) and used for many secondary studies about cohort-specific language behavior, ageing and language change, and language standardization as a process over time, among others (Yokoyama 2011). Thanks to the four respondents who took part in all four waves of the survey, the data also stimulated research into individual linguistic life courses (cf., e.g. Yokoyama and Sanada 2010).

From the point of view of research methodology, the sixty years of the Tsuruoka survey must be recounted as a story of steadily progressing sophistication from the beginnings of rather simple, hands-on data collection that relied on the investigators' judgement to the construction of highly elaborate statistical models of individual and collective language change.

5 Honorific speech

Another issue that figured prominently in early Japanese sociolinguistics was politeness or "honorific speech". There were three reasons for this: One, the elaborate system of linguistic means the Japanese language affords for expressing respect, deference, social distance, and formality; two, the assumption that politeness expressions somehow reflect social relations; and three, the desire to monitor the effects on actual usage of the 1952 recommendations issued by the National Language Council on honorific language. These recommendations were intended to support the democratiza-

4 http://www.ism.ac.jp/index_e.html (accessed 13 December 2021).

tion of Japan, stating in article 2: “In the past, honorific language was mainly based on superior inferior-relationships. From now on, it shall be based on mutual respect for each individual’s personality” (Bunkachō 1952). If previous usage was indexical of a strongly hierarchical society, changes in honorific usage might reveal insights about social change. What is more, they might shed some light into the complex interaction between social change and language change more generally. Sibata (1999 [1973]: 159) explicitly stated that the sociolinguistic investigation of honorific language is a suitable tool for grasping the social structure.

In passing, it may be noted that research about honorific language in Japan resulted in a huge literature that preceded the “politeness turn” and the investigation of gendered speech in Western sociolinguistics by several decades.

Linguistic politeness research has to begin by identifying the object of inquiry. In the early 1950s, when systematic surveys were first conducted, they focused on terms of address and self-reference, as use of these expressions is most obviously conditioned on social relationships and speaker status. Many studies of this period are part of the wider agenda of studying language standardization and were accordingly carried out in remote places where speech patterns could be expected to be at variance with standard usage.

For example, a 1952 inquiry described the use of kinship terms in a small village in Machino Town, Ishikawa Prefecture (Sibata 1999 [1973]). In the event, rather than recording unsolicited conversations through participant observation, interviews were the preferred method of collecting data. Although Sibata discusses the methodological differences between structural and social linguistics, pointing out that there is a danger that “interrogation” for a formal description “skews the data” (Sibata 1999 [1973]: 159), he justified such an approach on the grounds that members of small rural communities know about each other’s habits. Moreover, usage of kinship terms was dependent on a household’s position in the community and was the same for members of the household and other members of the community. Thus, if the members of a given household would call the head of the household *danna-sama*, other members of the community would also call him that, thereby recognizing that individual’s and his household’s ranking in the community. In order to examine the address kinship terms used in the community and the social relations underlying them, so the argument went, it was sufficient to question an informant or a small number of informants. The informants’ responses were taken at face value. No checks or reliability tests were included in the research design.

In the event, the study of the Kami-Tokikuni community interviewed 70 people who had 4830 (= 70 × (70 − 1) + ×)) potential addressees. They were asked how they addressed each member of the community all of whom were identified by their first name. 130 different terms of address were found, and after first and second person pronouns were discounted, 110 terms remained. The complex system of terms of address that resulted was descriptively interesting, but how to interpret these findings in regards to social structure remained unclear. As in the case of the Tsuruoka study, the Kami-Tokikuni survey was repeated after almost two decades, in 1970, revealing a

number of changes in honorific usage. However, the conclusions that could be drawn from the data collected by interviewing informants with standardized questionnaires were limited. After the Second World War a land reform was carried out, which in effect eliminated the old system of community ranking. The reform was a rare case of social engineering, and it was a matter of considerable interest whether it caused the observed changes in address behavior. Yet the survey did not yield any hard evidence to this effect, because it was not designed to do so. Tellingly, Sibata considers the possibility that the honorific system in the 1952 survey "might have been based on a preconception. Informants and investigators might have forced [the] *structuralization*" (Sibata 1999 [1973]: 162), seeing more structure than there was.

In the present context, the Kami-Tokikuni survey can be cited as an example of a highly interesting research project about honorific language carried out with too little attention being paid to its methodological design. The unwelcome outcome of this disregard is that the value of the collected data is compromised. It is a descriptive case study that does not allow for generalizations.

Other studies of honorific language made use of more elaborate methods of collecting data that are amenable to statistical analysis. A 1954 survey of the occurrence of the honorific prefix *o-* in Tokyo speech illustrates. The objective of the survey was to establish usage patterns of honorific speech among the residents of Tokyo. To this end, a representative sample of 617 informants was formed by systematic random sampling from official resident registers. Demographic factors such as age, education, occupation, and characteristics of the area (residential, industrial, and agricultural) were taken into consideration. Eventually 476 respondents were interviewed, all female, for practical and theoretical reasons. At the time, women tended to be homemakers and were thus easier to interview than men were. They were moreover thought to be more reliable judges of honorific speech.

Ten investigators interviewed the 476 informants, presenting them with 49 words of which they were asked to say whether they used them with the prefix *o-*. If an informant judged all 49 words positive with regard to attaching the prefix, she would obtain 49 points, and if she used none with the prefix, she would obtain 0 points. In this way, it was possible to discover any sociodemographic influence on the use of *o-*. Social conditions governing the use of *o-* were found. Age differences were negligible, while education had an effect, higher levels of education being positively correlated with higher frequency of use. Social class was also said to be a significant factor: "the use of *o-* increased with higher social standing" (Kunihiro, Inoue and Long (eds.) 1999: 116). However, no robust conclusions about class-specific use of *o-* can be drawn from this survey, because the informants' social class "was judged subjectively by the interviewers [on the basis of] the structure of the house, condition of the interior, informant's clothing" (Kunihiro, Inoue, and Long (eds.) 1999: 116). The absence of a well-defined model of social class in this survey is typical of sociolinguistic research in Japan in the 1950s. This is more of a theoretical than a methodological deficit, demonstrating once again the close interaction between theory and methodology.

A quarter century later, the same deficit still makes it impossible to interpret survey results about honorific language use in terms of social class. Sibata (1977: 113) voiced the opinion that "in modern Japan . . . differences in social classes seem to have already disappeared". His understanding of social class was closely connected with feudal society, and he thought that remnants thereof remained only in Okinawa. There he investigated linguistic differences between "gentry" (*shizoku*) and "commoners" (*shima no hito*, 'island people'). That was the extent of social stratification he considered relevant as a predictor of differences in language use.

During the same period, Ogino developed methods for the quantification of politeness in a number of publications (e.g. 1980). While previous studies of politeness were mostly limited to describing the structural means used for the expression of respect, deference, etc. without revealing any information about actual behavior, Ogino proposed a method "to study objectively, from the viewpoint of sociolinguistics, how the honorific varieties appear in large quantities of actual, self-reported data" (Ogino 1986: 38). For this purpose, he applied mathematical quantification techniques making use of cross-tabulations.

The data Ogino (1986) used were self-reported statements of honorific usage rather than actually observed usage. 488 people of a random sample were asked how they would affirmatively answer the question, "Do you know so-and-so's telephone number?" The respondents were asked to imagine eight different addressees as follows: (1) a close friend of about the same age, (2) a close acquaintance a little older, (3) a close acquaintance a little younger, (4) a not-too-close acquaintance of about the same age, (5) a not-too-close acquaintance a little older, (6) a not-too-close acquaintance a little younger, (7) the family member to whom the respondent talks most casually, (8) the person to whom the respondent talks most politely. 115 different ways of saying "yes, I know" (the telephone number) were obtained from face-to-face interviews. The relative politeness of each expression was calculated on the basis of the frequencies with which it was used toward each type of addressee. Ogino then correlated the sociodemographic data of the respondents with expressions variously rated as more or less polite and thus arrived at a quantitatively founded representation of politeness speech by Tokyo residents.

Two problems with this method stick out: the use of self-reported data and the paucity of the social categories used for describing the social structure. Subjects were assumed to be fully conscious of, and to reliably report on their own speech behavior. Social distinctions between them were limited to sex, age, familiarity, and inside vs. outside the family. In light of the combination of these two limitations, Ogino's method for the quantification of politeness is an interesting mathematical exercise, which, however, produces little insight about how politeness behavior is socially indicative in Japan.

Numerous surveys about how the standard language spread from Tokyo to the provinces were carried out in many parts of the country, but conceptually these were surveys about the (linguistic) unification of the country rather than differentiation. Dialect differences were well known and taken for granted; however, that in the process

of language standardization new linguistic distinctions could potentially emerge that were indicative of social stratification was rarely taken into consideration.[5] Instead of looking at social hierarchy as a predictor of language variation, more sociolinguistic research was done about other distinctions, notably gender and group language, such as, argot, occupational language, housewife language, student slang, etc.

It should be pointed out that this lack of interest in class structure is not limited to sociolinguistics. In Japan, social class has long been associated with the old order of feudal estates the Japanese nation state since the Meiji Restauration worked hard to overcome. The age of reconstruction after the defeat in World War II then gave rise to a self-image of Japan as a universal middle class society, which took root even among many sociologists. Only few of them have consistently and systematically studied social class and social mobility, like for instance Hiroshi Ishida, and even he uses a wider concept of social class which transcends the traditional concept that is limited to economic aspects (e.g. Ishida and Slater (eds.) 2010). That no methods for correlating language and social class have been developed in Japan thus reflects the fact that for a long time social class was not very salient in Japanese society and Japanese sociology.

About other social factors such as education, gender, and the urban-rural division Japanese sociolinguists have produced a rich and varied literature. Its most characteristic feature is strong emphasis on empirical research, often involving large-scale costly surveys that only institutions employing permanent staff, such as NLRI/NINJAL can carry out. Methodology has always played an important role in preparing surveys and other research projects. The summary Table 1.2 Tokugawa composed still provides a useful overview.

6 Dialect geography

The diffusion of the standard language in Japan was a socio-political project closely related to the above-mentioned abolition of the feudal order; the relative immobility of the peasant population of the old feudal domains (*han*) was thought to contribute to the habituation of local dialects. Unifying the country linguistically was seen as essential for its modernization, and the dispersion of the standard language as an indicator of overcoming inequality. Such was the primary motivation for another huge project undertaken by NLRI, directed for nine years by Takesi Sibata, the National Linguistic Atlas project (Kokuritsu Kokugo Kenkyūjo 1955–1975). Data were collected by means of a questionnaire comprising 230 items.

5 As late as 1975, Honna disputed the applicability of social stratification models to Japan, arguing that “it seems almost a taboo nowadays to discuss social stratification or social class, even worse to mention its scale” (Honna 1975: 194).

Table 1.2: Research methods used in linguistic research projects (adapted from Tokugawa (1988: 1003)).

Subfields of dialect research / Methods of approaching object of research			Records and descriptions								Theory of dialect division	Geolinguistics	Comparative dialectology	Social dialectology
			Descriptions					Translations, recordings						
			General overview	Lexicon, dialect groups	Grammar	Phonology	Word usage	Translations	Recordings of natural conversations	Impromptu skits				
Introspection			○	○	○	○	○	○	×	Δ	×	×	Δ	Δ
Natural observation	natural conversation		◊	◊	◊	◊	◊	×	○	×	×	×	×	◊
	documents	recorded data	Δ	◊	Δ	Δ	◊	×	×	×	Δ	×	×	◊
		research literature	×	Δ	×	×	×	×	×	×	○	Δ	Δ	×
Stimulus observation (survey)	surreptitious investigation		×	×	×	×	◊	×	Δ	◊	×	◊	×	Δ
	interview	individual	○	○	○	○	○	Δ	○	○	Δ	○	○	○
		group	×	×	×	×	◊	×	×	×	◊	○	×	Δ
	questionnaire	postal	×	◊	◊	×	◊	◊	×	×	◊	○	×	◊
		telephone	×	◊	◊	◊	◊	×	◊	◊	×	Δ	×	◊

○ normally used; Δ sometimes used; ◊ occasional ancillary; × almost never used.

As these figures indicate, the hallmark of this project and Japanese dialect studies in general is the collection of data from a large number of informants. Tahara (1991) reviewed the advantages and disadvantages of various fieldwork methods. As he points out, while the semi-structured face-to-face interview is preferable for the purpose of soliciting natural, unselfconscious speech, this method is very time consuming, since the interviewee may not produce the desired lexical item or pronunciation for a long time. In order to collect a sufficiently large corpus for the data to be statistically processable, fill-in questionnaires are often preferred, especially for dialect maps, which inevitably require a large number of informants. The Linguistic Atlas of Japan, for instance, used 2400 localities for data collection. Depending on the nature of the investigation, the pros and cons of direct and indirect methods of gathering information about the state of the language must be weighed carefully. From the very beginning of dialect studies until the present, questions of methodology were an important part of the development of the discipline. Matters such as questionnaire design, informant selection, interview techniques, data recording and management, and the identification of pertinent research objects – lexical items, accent patterns, honorific morphology, syntactic structures, etc. – received a great deal of attention (Tokugawa 1988: 1003).

Linguistic geography had been pioneered in the late 1920 by folklorist Kunio Yanagita's study of how snails are called throughout the country (Figure 1.1) which kindled an interest in the multiformity of dialects (Yanagita 1930, Sato 1979: 14). The distribution patterns he discovered made him speculate that radiating from the center to the periphery, the oldest forms were to be found in outlying regions. With this idea of concentric spreading Yanagita raised for the first time the intricate question of how space and time interact in determining the distribution of linguistic features. It was around this question that Japanese dialectology developed, quite independently from similar research in the West. Writing in the mid-nineteen sixties, Grootaers observed:

> Very few important works in this field are written in English, the dialectal studies in both Britain and the United States being definitely not ahead of Japanese dialectology. The languages necessary for dialectologists are German, French, Italian, and Dutch in that order. Few scholars of the "national linguistics" field now alive can consult works in these languages with any proficiency. (Grootaers 1967: 589)

Willem A. Grootaers played an active role in the development of Japanese dialectology, which was from its inception part of the language life approach and as such motivated by practical rather than theoretical concerns. After a decade of linguistic fieldwork in China, Grootaers arrived in Japan in 1950 and from then on collaborated in projects, including the Linguistic Atlas of Japan, as well as in the development of geolinguistic research tools. Linguistic maps depict the spatial distribution of linguistic forms. At the time it was standard practice to focus on elderly men leading a land-bound life as informants because their speech was considered least affected by outside influence and in this sense most authentic.

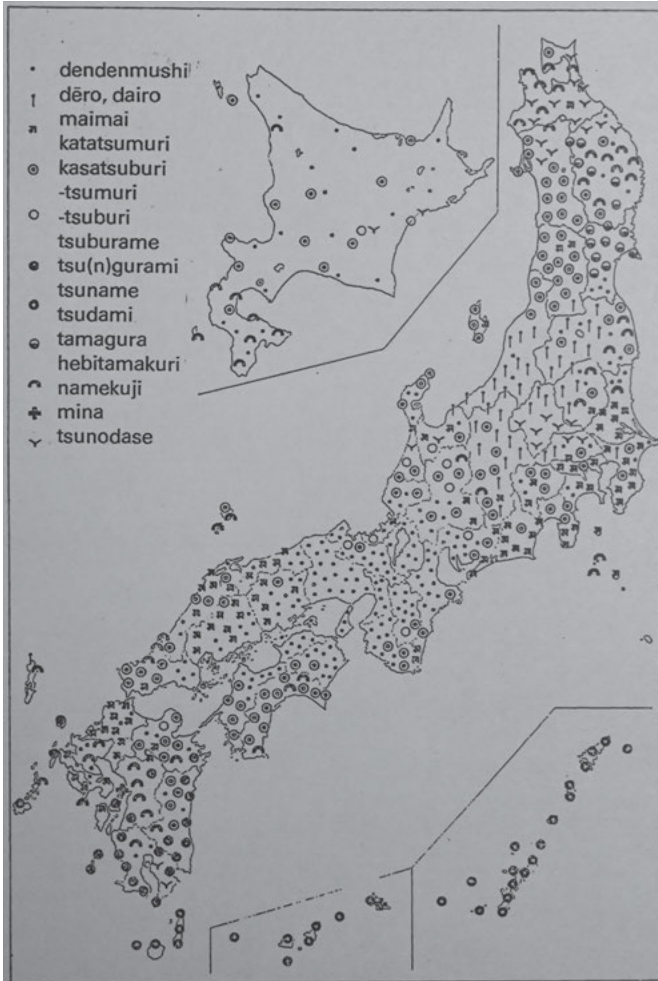

Figure 1.1: Kunio Yanagita's words for 'snail' (Sato 1979: 15).

This selective way of data collection is a prime example of the theoretical importance of methodology, illustrating as it does how the research instrument creates the object of inquiry. As an aggregate phenomenon in constant flux and subject to the influence of geographic, social and economic forces, language is not an object waiting to be dissected that reveals different aspects of its immutable existence just depending on the strength of the microscope lens through which we look at it. Rather, by cataloguing utterances or judgements about utterances by some speakers instead of others, we create the object we purportedly depict. Accordingly, most linguistic maps produced in Japan during the first half of the twentieth century represent a rather limited view on small fragments of the actual linguistic situation (Tokugawa (ed.) 1979). While this does not invalidate these maps and is inevitable for every

description of social phenomena, the question was eventually raised how representative elderly men were, and how the rest of the speech community should be represented on linguistic atlases. It gave rise to a methodological discussion about visual representations of language that can combine the special and temporal dimensions of its existence.

The principal result of these deliberations that took place within the big NLRI geolinguistic survey mentioned above was the "glottogram" as described by Sanada (2002), a mathematical method of correlating survey locality and informant age (Figure 1.2).

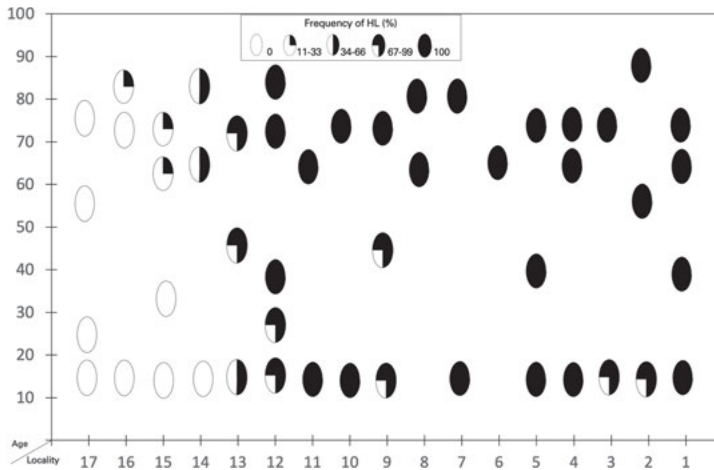

Figure 1.2: Glottogram, from "old" accent (HL) to "new" accent (LH). (Sanada 1971).

By crossing localities with speakers' ages, the glottogram catches the dynamics of language change. The survey from which the glottogram in Figure 1.2 is derived was conducted in 17 locations (x-axis) among speakers of five age groups (y-axis). The investigated item is the occurrence of a rising pitch accent in certain words that was found spreading geographically. In the event, there is a clear division between the extreme points on the location axis, all of the informants in location 1 using the HL accent, while in location 17 all informants use the LH accent. In the middle ground of the graph at locations 12 through 16 age differences in the use of either accent appear which are interpreted as representing ongoing language change. The apparent time method mentioned above also uses speakers' ages of data collected at one point in time as evidence of ongoing language change, but it ignores variable frequencies of use in space. The glottogram has its own limitations, as it can only depict the linear diffusion of linguistic change, rather than its spreading on a plane. However, by correlating variation across age groups and localities glottograms show that there is a better way of depicting language change than comparing linguistic maps based on data solicited from the

old generation at successive points in time, but also that language change does not proceed simultaneously throughout a language area.

As of the 1980s, Fumio Inoue made extensive use of the glottogram method. After a half century of studying dialects and language change, he published a book with the stunning title, Japanese moves at one kilometer per year (Inoue 2003). In it he takes stock of the research about language standardization based on lexical items of the Linguistic Atlas of Japan and other surveys. In his words: “As [this research method] simultaneously depicts regional and age differences, the process of language change is made beautifully visible” (Inoue 2003: 34, 36). Inoue does not ignore the above-mentioned limitation of the glottogram to linear diffusion of linguistic innovation. Yet, for him this method is part of a set of instruments that can elucidate the complex process of language change.

For one thing, Inoue argues that by taking the age of informants into account, comparisons of the available data from the Linguistic Atlas of Japan, whose oldest informants were born around 1895, with more recent data of informants born around 1985, make it possible to devise a linear model of the process of language standardization spanning more than 200 years (Inoue et al. 2009). Extrapolating from the actually existing data, he predicts that speakers born around 2035 will bring the diffusion of the common language to completion. Evidently, this is a hypothetical conjecture beyond verification or falsification at present. However, rather than the veracity of the prediction, the important point is that quantitative methods of linguistic analysis have been elaborated to the extent that testable predictions can be formulated. Glottochronology, a procedure entirely reliant on written documents of the past, has thus been supplemented by a method that is much closer to the spoken language (although it still depends on transcription) and extends to the future. This method has helped to bring the aggregation of individual acts that result in the social phenomenon recognized as language change closer to objective investigation. This process is further stimulated by more research based on spontaneous rather than elicited speech, as for instance, Asahi’s study of new speech forms in a Kobe suburb (Asahi 2008).

Harking back to the beginnings of systematic research of dialectal variety in Japan, Inoue is able to demonstrate based on quantitative data that Kunio Yanagita’s model of concentric diffusion of linguistic features falls short of explaining the multifaceted interaction of change that has driven language adaptations of various sorts in modern Japan. “Modern” must be emphasized in this context, because language standardization was an important part of Japan’s modernization efforts that interfered with unintended, quasi-natural language change. Education, deliberate dialect eradication campaigns, increased mobility, and expanded exposure to the mass media have rendered the concentric diffusion model obsolete. For in the course of investigating the advance of common language in peripheral regions it became clear that not only was the flow of linguistic features from the center, Tokyo, to other parts of the country

uneven, there was also a counter current of new features flowing towards Tokyo from where they then spread throughout the country.

This exchange between center and periphery gave rise to what has been called “new dialects” (*shin-hōgen*) and with them an entirely new subfield of social dialectology (Tokugawa and Sanada (eds.) 1991). The forms first identified as exemplifying this new phenomenon, such as *–ra* omission (*-ra-nuki kotoba*) (Inoue 1998: 2–15) among many others, provoked much public criticism and were branded as language confusion (*kotoba no midare*). They were produced predominantly by young speakers who used non-standard forms in informal domains. Careful examination of the evidence reveals that the “confusion” is part of the process of change, which the glottogram method allows to depict as incremental increase of frequency of occurrence progressing through several stages from misuse through uncertainty and common use to accepted correct use. Glottograms can moreover reveal the direction of innovations that move to and from Tokyo. Inoue has tried to visualize the mutual influence of regional dialects and common language in a model he calls the “umbrella model of language change” (Figure 1.3).

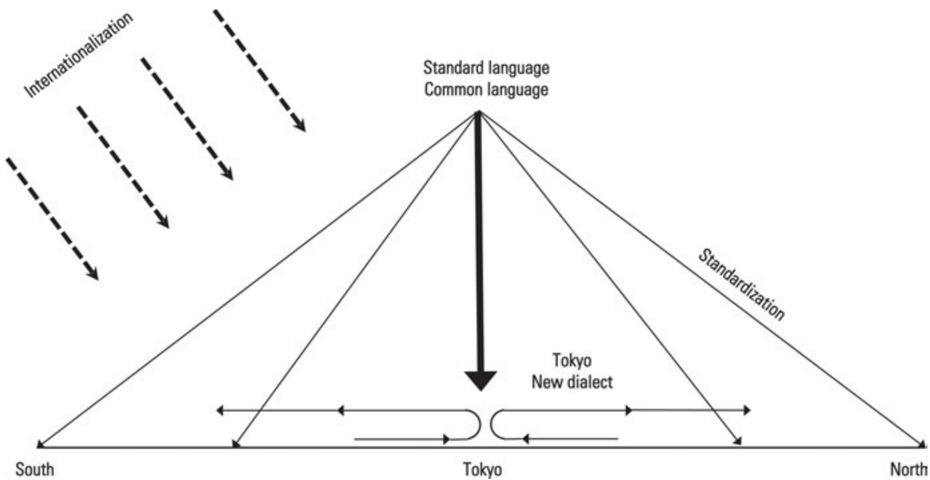

Figure 1.3: Inoue’s Umbrella Model of language change (Inoue 1998: 202).

The picture is complicated by the mutual influence of spontaneous and planned language change that has characterized the dynamics of Japanese throughout the twentieth century. The process of standardization promoted by the Ministry of Education and other agencies – change from above – continues and is at the same time counter caricatured by the emergence of new dialects “in the street” – change from below. A third force that makes itself felt Inoue calls “internationalization”, that is, increasing exposure to language contact, translation and the absorption of loan words. The language continues to change, and uncovering the factors of change and weighing

their relative impact poses many challenges. Clearly, since Kunio Yanagita's field notes about snails, the research tools have been sharpened and the rigor of empirical investigation greatly increased, steadily expanding the realm of that which can be subjected to systematic study.

Combining large data sets that have considerable time depth with linguistic maps and glottograms, Inoue (2003) attempts to make the interaction of spatial and temporal language variation calculable. His bold claim cited above: Linguistic innovations in twentieth and twenty-first century Japan progress at the speed of about one kilometer per year. Once again, the exactitude of this assertion is of minor interest only. The true significance of his claim lies in the implicit, more general claim that the speed of language change can be measured and that, accordingly, based on the available empirical evidence, falsifiable hypotheses can be formulated.

Since the linguistic innovations summarily labelled "new dialects" travel from the periphery to the center and interact with forms that progress in the opposite direction, the very notion of dialect is to be reconceptualized. Just like glaciers that may appear stationary, but in fact always flow, dialects do not freeze. Computer-assisted multivariate analysis of the frequencies of occurrence of multiple items in multiple locations across several age groups of speakers makes dialects appear as moving entities, rather than geographically fixed remnants of antiquated speech forms (Inoue 2003: 212). Thanks to new research tools it has become possible to investigate these multidimensional entities as constitutive elements of the continuous process of language change. In this sense, progress in research methodology has changed the logic of scientific discovery.

7 Conclusions

To a greater extent than the natural sciences, the social sciences create the object of their investigation. The development of the methods employed in sociolinguistic scholarship described in this chapter confirms this general insight. How the object of investigation is conceptualized and delimited depends on the available research instruments; and more than that. While an analytic research tool is designed for studying certain phenomena in accordance with certain theoretical provisions, it is not necessarily limited to a subservient role, but may work back to the theory that informs it.

The methods developed for investigating the progress of language standardization since the first linguistic atlas of Japan in the early 1900's and gradually elaborated in the course of the twentieth century have not just strengthened the empirical foundations of linguistics but improved our understanding of language as a social entity that exists in space and time. As the study of language variation progressed from interviewing old men living in the place where they were born to probing the conditions under which linguistic innovations traverse "dialect boundaries" for-

merly seen as stable, our notions of how dialects relate to each other and to the supra-regional common language have been transformed. This interplay of theory and method is further modulated by ideological changes. The traditional methods of dialectology that produced linguistic maps portray language as an idealized stable entity, a notion that was in accord with the promotion of the national language early in the last century. By contrast, research tools such as the glottogram that at once capture the spatial and temporal dynamics of language evolution are more attuned to the relaxing of standards and language contact induced hybridization characteristic of the present time.

References

Abe, Takahito. 2010. *Tsuruoka Chōsa Kaitō Dēta Setsumeisho Toraiaruban* [A Guide to Tsuruoka Survey Data: A Trial Version] *Ver.01*. Tokyo: Kokuritsu Kokugo Kenkyūjo.

Asahi, Yoshiyuki. 2008. *Nyūtaun Kotoba No Keisei Katei Ni Kansuru Shakai Gengogakuteki Kenkyū* [A Sociolinguistic Study on the Formation of “New Town” Language]. Tokyo: Hituzi Syobo.

Bailey, Guy. 2002. Real and apparent time. In J. K. Chambers, Peter Trudgill, and Natalie Schilling-Estes (eds.), *The Handbook of Language Variation and Change*, 312–332. Oxford: Blackwell.

Bunkachō. 1952. *Korekara no keigo* [Honorific language from now on]. http://kokugo.bunka.go.jp/kokugo_nihongo/joho/kakuki/01/tosin06/index.html

Coulmas, Florian. 2013. *Sociolinguistics. The Study of Speakers’ Choices*. Cambridge: Cambridge University Press.

Grootaers, Willem A. 1967. Dialectology. In Thomas A. Sebeok (ed.), *Current Trends in Linguistics*, vol. 2, 585–607. The Hague: Mouton.

Heinrich, Patrick. 2001. *Die Rezeption westlicher Linguistik im modernen Japan bis Ende der Shōwa-Zeit* [The Reception of Western Linguistics in Modern Japan until the End of the Showa Period (1926–1989)]. Munich: Iudicium.

Honna, Nobuyuki. 1975. A note on social structure and linguistic behavior: A case study of a Japanese community. In Fred C. Peng (ed.), *Language in Japanese Society*, 193–214. Tokyo: University of Tokyo Press.

Inoue, Fumio. 1998. *Nihongo uotchingu* [Japanese Language Watching]. Tokyo: Iwanami.

Inoue, Fumio. 2003. *Nihongo wa nensoku ichikiro de ugoku* [Japanese Moves at the Speed of One Kilometer per Year]. Tokyo: Kodansha Gendai Shinsho.

Inoue, Fumio. 2011. Standardization and de-standardization. In Patrick Heinrich and Christian Galan (eds.), *Language Life in Japan. Transformations and Prospects*, 109–123. London: Routledge.

Inoue, Fumio. This volume. Style, prestige, and salience in language change in progress. In Yoshiyuki Asahi, Mayumi Usami and Fumio Inoue (eds.), *Handbook of Japanese Sociolinguistics*. Berlin and Boston: De Gruyter Mouton.

Inoue, Fumio, Kiyoshi Egawa, Ryoichi Sato and Masato Yoneda. 2009. On’in kyōtsūgoka no S-ji kābu [S-shaped curve of phonetic language standardization]. *Keiryō Kokugogaku* [Mathematical Linguistics] 26. 269–289.

Ishida, Hiroshi and David Slater (eds). 2010. *Social Class in Contemporary Japan: Structures, Sortings and Strategies*. London: Routledge.

Kokuritsu Kokugo Kenkyūjo. 1955–1975. *Nihon Gengo Chizu* ((kokuritsu Kokugo Kenkyūjo Hōkoku 30) [Linguistic Atlas of Japan (National Language Research Institute Report 30)]. Tokyo: Kokuritsu Kokugo Kenkyūjo.

Kokuritsu Kokugo Kenkyūjo. 2007. *Chiiki Shakai No Gengo Seikatsu: Tsuruoka Ni Okeru 20 Nen Kankaku Sankai No Keizoku Chōsa* [Language Life of the Speech Community: The Tsuruoka Survey, Three Times with Twenty Year-intervals]. Tokyo: Kokuritsu Kokugo Kenkyūjo.

Kunihiro, Tetsuya, Fumio Inoue and Daniel Long (eds.). 1999. *Takesi Sibata: Sociolinguistics in Japanese Contexts*. Berlin: Mouton de Gruyter.

Ogino, Tsunao. 1986. Quantification of politeness based on the usage patterns of honorific expressions. *International Journal of the Sociology of Language* 58. 37–58.

Said, Edward. 1978. *Orientalism*. New York: Pantheon.

Sanada, Shinji. 1971. Toyamaken Togadani ni okeru akusento no dōtai [The dynamics of accent patterns in Togadani, Toyama Prefecture]. *Bungei Kenkyū* [Studies in Japanese Literature and Language] 68. 39–45.

Sanada, Shinji. 2002. Shakai hōgenchirigaku no jissen, grottoguramu shōshi [Doing social dialectology. A concise history of the glottogram]. In Yoshio Mase, Ryoichi Sato, Takuichiro Onishi, and Takashi Kobayashi. (eds.), *Hōgen Chirigaku no Kadai* [Topics in dialect geography], 345–356. Tokyo: Meiji Shoin.

Sato, Ryoichi. 1979. Hōgen no bunpu [The distribution of dialects]. In Munemasa Tokugawa (ed.), *Nihon no Hōgen Chizu* [Dialect maps of Japan], 3–52. Tokyo: Chuko Shinsho.

Smakman, Dick and Patrick Heinrich (eds.). 2015. *Globalising Sociolinguistics. Challenging and Expanding Theory*. London: Routledge.

Sibata, Takesi. 1999 [1952]. The age structure of the speech community. In Tetsuya Kunihiro, Fumio Inoue and Daniel Long (eds.), *Takesi Sibata: Sociolinguistics in Japanese Contexts*, 207–215. Berlin: Mouton de Gruyter.

Sibata, Takesi. 1999 [1973]. Honorifics in a community. In Tetsuya Kunihiro, Fumio Inoue and Daniel Long (eds.), *Takesi Sibata: Sociolinguistics in Japanese Contexts*, 157–180. Berlin: Mouton de Gruyter.

Sibata, Takesi. 1977. Toshika to gengo no kaikyūsa [Urbanization and class differences in language]. In F. C. Peng (ed.), *Kankyō to Kotoba* [Environments and language], 113–126. Hiroshima: Bunkahyōronshuppan.

Sumi, Tomoyasu. 2005. Nihonjin no yomikaki nōryoku chōsa 1948 no sai-kenshō [Reinspection of the 1948 survey of Japanese reading and writing proficiency]. *Tenri Daigaku Gakuhō* [Bulletin of Tenri University] 56(2). 105–124.

Tahara, Hiroshi. 1991. Dēta no shūshū to shori [Data collection and management]. In Munemasa Tokugawa and Shinji Sanada (eds.), *Shin-hōgengaku o Manabu Hito no Tame ni* [For those who want to learn about of new dialects], 264–282. Kyoto: Sekaishisōsha.

Tanabe, Juri. 1936. *Gengo-shakaigaku Josetsu* [Introduction to the Sociology of Language]. Tokyo: Jichōsha.

Tōkei Sūri Kenkyūjo and Kokuritsu Kokugo Kenkyūjo [The Institute of Statistical Mathematics and National Institute for Japanese Language and Linguistics]. 2014. *Dai 4 kai Tsuruoka ni okeru gengochōsa randamu sanpuringu chōsa no gaiyō: Shiryōhen* [Synopsis of random sampling survey of 4th linguistic survey in Tsuruoka: material]. Tokyo: NINJAL. https://www2.ninjal.ac.jp/longitudinal/files/tsuruoka4_report01.pdf (accessed 27 September 2020)

Tokugawa, Munemasa (ed.). 1979. *Nihon no Hōgen Chizu* [Dialect Maps of Japan]. Tokyo: Chuko Shinsho.

Tokugawa, Munemasa. 1988. Hōgen to kyōtsūgo [Dialect and common speech]. In Haruhiko Kindaiichi, Oki Hayashi, Takesi Sibata (eds.), *Nihongo Hyakka Daijiten* [An encyclopedia of the Japanese language], Part 17, 913–1006. Tokyo: Taishūkanshoten.

Tokugawa, Munemasa and Shinji Sanada (eds.). 1991. *Shin-hōgengaku o Manabu Hito no Tame ni* [For Those Who Want to Learn about of New Dialects]. Kyoto: Sekaishisosha.
Yanagita, Kunio. 1930. *Kagyūkō* [Considering Snails]. Tokyo: Iwanami Shoten.
Yanagita, Kunio. 1963. *Kokugoshi* [History of the National Language]. Tokyo: Chikuma Shobo.
Yokoyama, Shoichi. 2011. Gengohenka Wa Keinen Chōsa Dēta Kara Yosoku Kanō ka?[Is Prediction of Language Change Based on Longitudinal Survey Data Possible?] *Kokuritsu Kokugo Kenkyūjo Ronshū* [NINJAL Project Review] 6. 27–37.
Yokoyama, Shoichi and Haruko Sanada. 2010. Gengo no shōgai shūtoku moderu ni yoru kyōtsūgoka yosoku [Predicting language standardization on the basis of a model for life-long language learning]. *Nihongo no Kenkyū* [Studies in the Japanese Language] 6. 31–45.
Yomikaki Nōryoku Chōsa Iinkai (ed.). 1951. *Nihonjin no yomikaki nōryoku* [Japanese People and Their Literacy]. Tokyo: Tokyo Daigaku Shuppanbu.

DuEPublico

Duisburg-Essen Publications online

UNIVERSITÄT DUISBURG ESSEN

Offen im Denken

ub | universitäts bibliothek

This text is made available via DuEPublico, the institutional repository of the University of Duisburg-Essen. This version may eventually differ from another version distributed by a commercial publisher.

DOI: 10.1515/9781501501470-002
URN: urn:nbn:de:hbz:465-20231107-145021-6

In: Y. Asahi, M. Usami & F. Inoue (Ed.), *Handbook of Japanese Sociolinguistics* (pp. 19-40). Berlin, Boston: De Gruyter Mouton. https://doi.org/10.1515/9781501501470-002

The publisher version is available at www.degruyter.com.

© 2022 Walter de Gruyter Inc., Boston/Berlin. All rights reserved.