4 Variationist Sociolinguistics

JUNKO HIBIYA

0 Introduction

The present chapter is intended to provide a synthesis of the field called variationist sociolinguistics that began in the 1960s. Within the framework proposed by Weinreich et al. (1968), attempts have been made to study linguistic change and variation in different speech communities around the world with a perspective which views the linguistic system in a fundamentally different way from any of the preceding models (Labov 1963, 1966). In this research tradition, the most important development towards an understanding of linguistic change has been made by recognizing that historical change in progress is manifested as synchronic variation.

Variationist research conducted in the past thirty years has shown that variation found in linguistic data is highly systematic and furthermore variation which occurs in the speech of individuals exhibits a regular pattern. Viewed from a traditional standpoint, linguistic variation might be regarded as mere chaos. Labov and his associates, on the other hand, have proposed the notion of orderly heterogeneity. In this school of linguistic analysis, it is believed that the investigation of the linguistic analysis of individuals within their speech community is directly relevant to diachronic linguistics.

The next section begins with the discussion of the standard methodology. The section describes the data collection procedures and reviews the study of /-t, d/ deletion phenomenon in English as a concrete example of what has been conducted in the field. Section 2 presents the results of the quantitative analyses of voiced velar nasalization in Japanese as it is spoken in Tokyo. The section concludes with a discussion of how diachronic and synchronic variation observed in Tokyo Japanese can be represented in a theoretical framework. Section 3 summarizes the chapter and outlines directions for further research.

1   Methodological Principles

1.1   Overview

The field of sociolinguistics probably is best defined in terms of its methods and goals, rather than its subject matter. Thus, in this section, we will review the progress which has been made in the area of methodology in the course of a number of studies of sound change in progress in the past three decades (Abdel-Jawad 1981, Bailey 1973, Baugh 1979, Cedergren 1973, Cofer 1972, Guy 1981, Haeri 1991, Harris 1985, Herold 1990, Hindle 1980, Hong 1991, Laberge 1977, Labov et al. 1972, 1980, Lavandera 1975, Lennig 1978, MacCaulay 1977, L. Milroy 1980, Modaressi 1978, C. Paradis 1985, Payne 1976, Poplack 1979, Tarallo 1983, Trudgill 1974).

The success of the study of language in its social context depends heavily on selecting appropriate methods of data collection and analysis. The traditional field methods developed in dialect geography are inappropriate for studying natural speech. Chambers (1992: 674) states:

> Dialect geography studied regional speech patterns qualitatively, and concentrating on the speech of nonmobile, older, rural, predominantly male consultants – known acronymically as NORMs – elicited relatively stable, regionally distinctive, highly differentiated speech samples. Variability was peripheral, and only occasionally drew comment from the investigators.

Technological progress in sound recording in the past several decades has enabled the student of language variation to take a different methodological path. Instead of asking the informant, she or he can record spontaneous speech and capture its systematicity.

Previous research has shown that the vernacular is the best database to reveal the most systematic and regular character of language. Labov (1984: 29) states the reasons for preferring the vernacular used in casual and sponta-neous interaction to more formal styles as follows:

> The vernacular is defined as that mode of speech that is acquired in pre-adolescent years. Its highly regular character is an empirical observation. The vernacular includes inherent variation, but the rules governing that variation appear to be more regular than those operating in the more formal "superposed" styles that are acquired later in life. Each speaker has a vernacular form in at least one language; this may be the prestige dialects (as in the case of "RP"), or a non-standard variety.

In his New York City study (Labov 1966), Labov conducted a series of face-to-face, tape-recorded interviews with individuals based on a random sample. Although one can get a representative sample of a speech community most efficiently by this method, it does not facilitate the elicitation of vernacular

for two reasons: first, the setting in which an informant is being interviewed by a stranger is more likely to elicit careful style; second, we can get no view of the interaction of speakers since each individual is isolated from his or her social networks. To overcome the problem of direct interviewing, Labov first attempted to minimize the constraints of the setting which may promote speakers' consciousness of their speech style based on topic; general interview topics such as place of birth, residential history, education, and occupation are considered to elicit relatively careful and formal style, while a specific set of topics such as danger of death, childhood rhymes, and customs (and any speech outside the formal interview, with a third person, or not in direct response to questions) override the negative effects of the interview situation and are likely to lead the interaction to the other end of the spectrum, i.e. casual and informal style. This approach, however, is not always a successful means of approximating the ideal; it sometimes fails to enable researchers to elicit an unreflecting way of speaking.

In their Harlem study, Labov and his associates took a completely different approach: participant observation techniques and group sessions (Labov et al. 1968). A number of pre-existing peer groups were located and their members were recorded as a group. This way of obtaining data is superior to an individual interview, for the undesirable effect of observation and recording is reduced by the control of social interaction. The method of eliciting data through group sessions, however, lacks representativeness. No one can claim that the speech collected through this method is representative of the entire speech community. Labov et al. (1968) overcame this shortcoming by supplementing the data from the group interviews with material collected from individual interviews with each member of these groups as well as with interview data obtained from a random sample of the community.

The Project on Linguistic Change and Variation (Labov et al. 1980) was an attempt to aim at both depth and breadth by combining both methods. It incorporated a telephone survey, which is comparable to a conventional random sample survey, and a series of neighborhood studies, through which a large number of linguistic and social data on the major social networks of the neighborhoods under investigation are obtained. No one method can be perfect in all respects. Labov (1984: 50) lists seven criteria to rate the methods of collecting data:

> the possibility of obtaining a representative sample; the demographic data obtained; the comparability of the data obtained; success in minimizing the effects of observation; the quality of the sound recorded; the volume of data obtained; and the feasibility of including field experiments.

Either of the methods employed in the Project on Linguistic Change and Variation has its own strengths and weaknesses, and hence, agreement in the results provided the basis for confidence in the findings of the project and presented a well-balanced view of linguistic change in progress.[1]

1.2　/-t, d/ deletion in English

The field of variationist sociolinguistics originally started from an interest in diachronic change. Labov (1963, 1966) analyzed primarily the effect of social factors such as age, gender, socioeconomic status, and ethnic or other characteristics of the speakers on linguistic variables which represent linguistic change in progress. The variationist approach has been modified in a significant way by recognizing that a particular variation is maintained across many generations with no sign of change. The case which brought this home is the /-t, d/ deletion phenomenon in English. It is characterized by the absence of change.

The /-t, d/ deletion is a phonological process which results in the absence of a final apical stop (/-t/ or /-d/) when it is the last member of a consonant cluster. The variable omission of /-t, d/ yields pronunciations such as *didn' wan' us* (< didn't want us) and *the firs' chil'* (< the first child). It is the best-known variable process that has been studied in the variationist paradigm of socioliguistic studies.

Three decades of empirical research carried out for a wide variety of English dialects has repeatedly shown that the variable process of /-t, d/ deletion is constrained not only by the social factors but also by the linguistic internal constraints (Bayley 1991, 1994, Fasold 1972, Guy 1980, 1991a, 1991b, 1993, 1994, 1997, Guy and Boyd 1990, Hibiya 1997, Labov 1975, 1989, Labov et al. 1968, Neu 1980, Patrick 1991, 1992, Roberts 1994, Romaine 1984, Santa Ana 1991, 1992, Wolfram 1969, 1972, Wolfram and Christian 1976). The linguistic constraints include the following:

(1)　a.　whether the syllable containing /-t, d/ is stressed or not:
　　　　　e.g.　different, island, thousand, interest > hard-working
　　　b.　consonant cluster length:
　　　　　e.g.　next > list
　　　c.　the phonetic features of the segment preceding /-t, d/:
　　　　　/s/ > stops > nasals > other fricatives > liquids
　　　　　e.g.　first > except > recent > left > bolt
　　　d.　the phonetic features of the segment following /-t, d/:
　　　　　obstruents > liquids > glides > vowels
　　　　　e.g.　didn't go > didn't learn > didn't watch > didn't eat
　　　e.　the grammatical feature of /-t, d/:
　　　　　part of *-n't* morpheme > part of stem > irregular verb
　　　　　past tense or past participial form which are marked by both suffixation and ablaut > regular past tense or past participial suffix
　　　　　e.g.　isn't > mist > felt > missed
　　　f.　whether the segments preceding and following /-t, d/ agree in voicing or not:
　　　　　homovoiced > heterovoiced
　　　　　e.g.　except swimming > except running

Among them, the effects of (1d) and (1e) have been unanimously found to be the most important in every dialect. The relative ranking of the two varies from dialect to dialect: in European American English dialects (1e) is more significant (Neu 1980) while in the so-called ethnic or minority varieties, i.e. African American, Puerto Rican, Chicano, Appalachian, and Japanese-Canadian English, (1d) is stronger (Labov et al. 1968, Wolfram 1972, Santa Ana 1991, Wolfram and Christian 1976, Hibiya 1997).

Let us examine these two constraints in detail. For (1d), the deletion is promoted in the following order: obstruents > liquids > glides > vowels. In other words, when a segment which follows the final /-t, d/ is lower in sonority, the /-t, d/ is likely to be deleted. Interestingly, the effect of a following pause is dialect-specific (Guy 1980, Hibiya 1997).

With respect to (1e), the /-t/ of -*n't* produced by negative contraction is deleted most frequently. The deletion is promoted to a larger extent when the /-t, d/ is part of the stem, as in *past*, *mist* (M-words), than when it is the regular past tense or past participial suffix, as in *passed*, *missed* (P-verbs). The final /-t, d/ of verb forms like *left* and *sold* (S-verbs) fall between these two in deletion rate.

Two different hypotheses have been proposed to account for the ranking. The first one is in functional terms. According to this hypothesis, the /-t/ of *missed* is deleted less frequently than the /-t/ of *mist* since the former is an inflectional suffix that carries the grammatical feature of past tense. In other words, semantically relevant information is more likely to be retained (Kiparsky 1971, 1982). The second one is based on lexical phonology (Kiparsky 1982, 1985, Mohanan 1986). In this theory, lexical derivation is organized into two principal levels, i.e. lexical and postlexical. The lexicon consists of at least two ordered levels. At each level, the morphological and phonological processes alternate with each other. It is possible for rules to apply at more than one level. In the case of the /-t, d/ deletion phenomenon, M-words have their final consonant clusters underlyingly. S-verbs acquire their final /-t, d/ by undergoing affixation at level 1 of the lexicon while the regular inflectional affixes are attached at level 2. Adopting the theoretical framework of lexical phonology, Guy (1991a) proposes that /-t, d/ deletion is a variable phonological rule that applies both within the lexicon and in the postlexical level whenever its structural descriptions are met. M-words undergo deletion three times whereas S-verbs twice and P-verbs do so only once each. Therefore, the difference in derivational history accounts for the ranking M-words > S-verbs > P-verbs.

The deletion is subject to stylistic and social factors as well. Earlier studies have repeatedly found that in formal speech /t/ and /d/ tend to be retained. The gender, socioeconomic status, and educational background of speakers also constrain the variable process.

As has been mentioned at the beginning of this section, the /-t, d/ deletion is a typical case of a stable variable which shows no sign of change, and among the linguistic variables examined so far, stable ones are relatively common. The field of linguistics has been concerned with invariance for a long time; but the

analyses of this phenomenon have established the notion that variation is the fundamental problem of synchronic linguistics which is worthy of systematic study.

2   The Tokyo Project

The /-t, d/ deletion has been considered as a showcase variable of inherent variation which is constrained by both social and linguistic internal factors. Studies of /-t, d/ deletion in English have set standards for the detailed analysis of variation. This section presents the results of the quantitative investigation of one variable in Japanese (Hibiya 1988). It is an attempt to describe the language as it is spoken in Tokyo. The study was carried out with the purpose of building upon the progress which had been made since the 1960s in the area of data collection and analysis.

2.1   Background

The purpose of this subsection is to discuss the methodology used in this study. I will describe in detail the specific methods used for: (1) selecting informants in such a way that the sample would provide both detail and representativeness of linguistic behavior; and (2) eliciting data which would illustrate the informants' full range of speech behavior. Let us first describe the speech community under investigation.

Tokyo (previously Edo), one of the major urban centers of the world, has been inhabited since 2500 BC. It used to be a peripheral village until Tokugawa Ieyasu established a feudal government in Edo in 1603. Since then, the city has developed greatly as a political center of the country. By the end of the eighteenth century, it had grown into a city with a population of 1,200,000. During the Edo era (1603–1868), the city was inhabited by members of the samurai "military" class, merchants, and artisans. In 1868, the emperor defeated the Tokugawas and regained power. Since then, it has been the capital of the country. The city was renamed Tokyo, which means "eastern capital."

Its importance as the political, commercial, financial, and cultural center of Japan has steadily increased. Although Tokyo was badly damaged by the earthquake of 1923 and air raids during World War II, the city was rebuilt within a short period of time after the destruction.

Tokyo is divided into two subareas, *yamanote* "uptown" and *shitamachi* "downtown." The former consists of middle- and upper-class residential areas, whereas the latter is mostly a commercial or industrial area where the lower middle-class and blue-collar workers live. Historically, *yamanote* was inhabited by the samurai class, and *shitamachi* by merchants and artisans during the Edo era. This social difference resulted in the development of two different versions of the dialect.

Today standard Japanese is based on a variety of the Tokyo dialect typical of educated, middle-class natives of the *yamanote* area. It is the language of the government, education, and media. The *yamanote* dialect is considered to have more prestige than the *shitamachi* dialect (Vance 1987).

The main target speech community of Hibiya (1988) was one neighborhood in Tokyo named Nezu. It is a typical *shitamachi* neighborhood located on the border between the two subareas. At the time of data collection (1986), the population was 7,514, and the number of households was 2,913 (census data).

In Nezu, a total of 97 people were interviewed.[2] The first major source of data came from a neighborhood study. We entered the community in July 1986, to establish contacts and develop relationships with members who later allowed us to make tape-recordings. The Nezu neighborhood study began in mid-July 1986 with a series of interviews with a peer group of retired people, which lasted for three months. All of the interviews with old men and women were conducted in the center for retired people run by Bunkyo Ward. During this period, the interviews with other peer groups and individuals from different generations of the area were carried out at their homes.

In the neighborhood study, we interviewed a total of nine people, five males and four females, ranging in age from 16 to 83 years old. Because they were few in number and the procedures employed to select them were not adequate to achieve representativeness, the results from this study would not be considered as typical of the neighborhood as a whole.

In addition to the above neighborhood study, we carried out a random sample survey of the population to evaluate the data from the neighborhood study. This second sample consisted of 88 people. There were 41 males and 47 females. The sampling procedure was as follows. First, we selected one out of every 15 people from the list of residents kept at the Nezu branch of Bunkyo Ward Office, yielding a sample of 402. We then excluded 50 who were below 10 years of age, reducing the sample to 352 people. It turned out that 42 had moved out of the neighborhood, and 16 could not be interviewed because they were sick or handicapped. Of the remaining 294, we were able to carry out standard sociolinguistic interviews with 88 residents: 30 refused to be interviewed, while 176 were never located although we tried to reach each one as many as three times. Among the 88, the speech of 62 speakers was analyzed. They were native speakers of the dialect; 48 were born in Nezu, while 14 grew up in other areas of the city. They ranged in age between 14 and 81 years old at the time of the data collection (1986). Table 4.1 shows the distribution of the informants by age and sex.

A supplementary group of ten native speakers from other areas of Tokyo (five males and five females) was contacted in the course of data collection. For this series of interviews, we worked with families; the two home sites made it possible to obtain data across three generations. The informants ranged in age from 15 to 81.

Each individual interview in the neighborhood study consisted of several modules of questions referred to as Q-GEN-II, which were developed at the

Table 4.1   Distribution of informants by age and sex

Sex	10–19	20–9	30–9	40–9	50–9	60–9	70+	Total
Male	7	5	2	9	6	1	1	31
Female	5	2	7	4	5	4	4	31
Total	12	7	9	13	11	5	5	62

University of Pennsylvania. It is a network of modules covering a wide range of topics which is designed to allow the interviewer to get at speakers' areas of major interest. Each module contains a dozen questions which are meant to be open-ended. Questions were adapted to the speech communities in Tokyo.

The individual interview of the Nezu random sample survey included readings of texts and words and a short interview. The recordings were made using a Sony TCM 5000 EV tape recorder and a Sony ECM 150 lavaliere microphone.

2.2   *Voiced velar nasalization as a linguistic variable*

The study of variation and change over the past 30 years has focused on the social context of language use, the linguistic constraints of variation and change, or both. In this section, we will discuss the interaction between linguistic and social processes, focusing on one well-known linguistic variable in Tokyo Japanese: voiced velar nasalization.

2.2.1   *Facts*

In the conservative dialect of Tokyo Japanese, voiced velar plosives are nasalized word-medially. Some examples are listed in (2). (The symbol [¥] represents a voiced velar nasal.)

(2)   *Word-initial [g]*:   *Word-medial [¥]*:
       gin   "silver"   hi¥asi   "east"
       geki   "play"   ka¥e   "shadow"
       gan   "cancer"   ka¥i   "key"
       goma   "sesame"   iti¥o   "strawberry"
       gun   "county"   hu¥u   "blowfish"

As a first approximation, it can be stated that for speakers of this dialect, underlying /g/ is realized as [g] in word-initial position, but as [¥] in other

positions; in other words, [g] and [¥] are in complementary distribution. However, there are several systematic exceptions where word-internal [¥] must or may be blocked (Sakuma 1929, S. E. Martin 1987, Arisaka 1959, Vance 1987). These exceptions will be discussed in detail later.

The pronunciation with consistent use of non-initial [¥] has been considered to be standard and has official status. Before World War II, it was an emphasized point of elementary school education that /g/ should be realized as [¥] in non-initial position. Today, many trained television and radio broadcasters use non-initial [¥].

2.2.2   Variation in apparent and real time: social aspects

It has widely been observed that the non-initial nasal [¥] is undergoing a change and being pronounced as a plosive [g] (Hibiya 1988, 1995a). For speakers of this newer dialect of Tokyo Japanese, voiced velar plosives are never nasalized regardless of their position within a word.

Figure 4.1 shows the temporal distribution of word-internal [g]. For each speaker of the Nezu random sample survey, the number of [g] tokens were counted and their percentage was plotted against age individually. A clear pattern of age stratification emerges. Among the speakers in their teens and twenties, the [g] variant is predominant while the percentage of [g] drops off as age increases. The correlation is highly significant ($R^2 = 0.651$, $p < .001$).
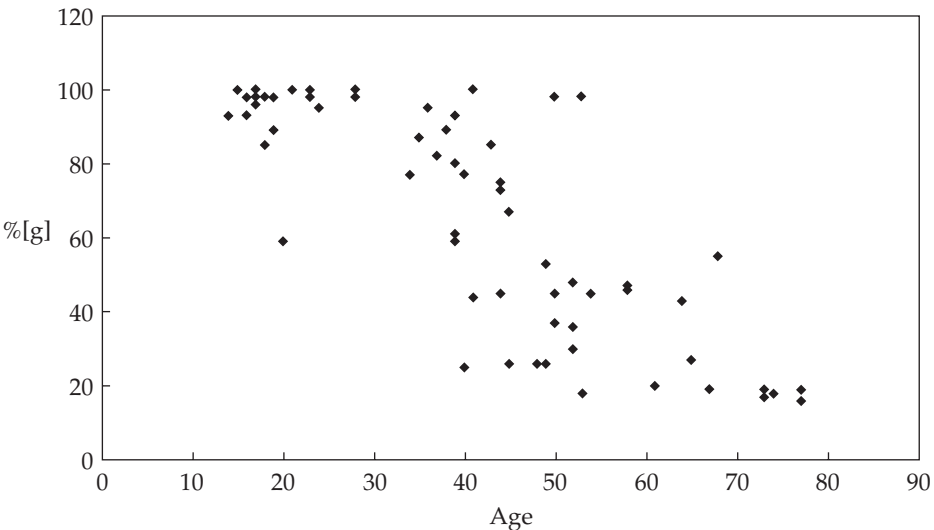

Figure 4.1   Age distribution of [g]
Source: Hibiya 1995a: 144

When a clear correlation between a given linguistic variable and age is found, it is assumed to reflect a linguistic change, in this case a shift from word-internal [¥] toward [g]. This way, the trajectory of linguistic change within a community can be followed in apparent time. Correlation with age, however, may be a case of age-grading, a pattern which repeats in each generation instead of a sign of change (Hockett 1950). To determine whether a given relationship between age and the linguistic variable represents change in progress or age-grading, real-time observations have to be consulted. Differences among sets of data collected at discrete points in time can confirm that a true change has occurred (Labov 1994).

For the linguistic variable under consideration, there are a number of concrete sources of real-time data (Hibiya 1995b). In early nineteenth-century literature (Ukiyoburo written by Shikitei Samba), it is specifically mentioned that speakers in Edo had the nasal word-internally while rural speakers had the plosive. The descriptions of the language by foreigners who came to Japan in the late nineteenth century (Hepburn 1872, Chamberlain 1888) also indicate that the native speakers then had the initial [g] and internal [¥].

A study carried out in 1941 by Kindaichi bears directly on this phenomenon (H. Kindaichi 1967). By means of a word list, he examined how 70 native middle-school students who were 14–15 years of age pronounced underlying /g/. The results showed that about one-third of the informants had consistent use of word-internal [¥], and about another one-third had consistent use of word-internal [g]. The rest used both variably. Although he did not include any speakers from other age groups in his study, he also mentioned that native speakers over 30 at that time all had consistent use of word-internal [¥]. His results and observation clearly show that the shift from an older variant to a newer one was in progress more than five decades ago.

In addition to these written records, we were able to locate two different sets of earlier tape-recordings. In the 1950s, NHK (Japan Broadcasting Association) carried out a project in which it located a few aged informants from each major dialect area of the country and recorded their natural conversation. The tape prepared in 1952 for the Tokyo dialect included the speech of two native speakers: a man born in 1868 and a woman born in 1891. The examination of the recording shows that both of them had word-internal [¥]. Interestingly, they overtly talked about voiced velar nasalization during their taped conversation; both agreed that the word-internal [g] which was occasionally heard in their daily life then was not typical of genuine Tokyo dialect. This kind of comment indicates that the people were aware of the change.

The second set is a collection of recordings of a series from a radio program. Seventeen speakers who were born between 1875 and 1888 were interviewed at home for the program in the 1940s and 1950s. They acquired their vernacular in the late nineteenth century when the above-mentioned foreigners were describing the language. Everyone had the pattern which was consistent with the rules and their exceptions for voiced velar nasal of the conservative dialect.

The analyses of the real-time data have all indicated that native speakers had word-internal [¥] until at least the late nineteenth century. The change must have started in the early twentieth century and is headed toward completion in the youngest generation of today.

As compared to the age stratification which has been discussed, the gender difference is not so striking. Both male and female speakers participated in the variation in a similar way.

Among the social factors, the degree of one's contact with *yamanote* has also turned out to be important in constraining the variability and change. As described earlier, the city consists of two subareas: *yamanote* and *shitamachi*. According to the study conducted by Kindaichi, it was the natives of the former that were leading the change 50 years ago. An examination of the speech of the Nezu random sample survey shows that not only those who were born and brought up in the *yamanote* area but also those who had daily contact with *yamanote* in their adolescence definitely favored word-internal [g]. As can be seen in figure 4.1, there are three informants who use [g] almost categorically among the speakers between 40 and 59 years of age. They use [g] much more frequently than the rest of the same age group. What these three have in common is that they all went to high school in the *yamanote* area. In figure 4.2, all speakers that have had an extensive contact with *yamanote* are coded as [+*yamanote*]. These speakers favor [g], and this pattern holds for all age groups but the youngest, in which this factor does not have an effect.
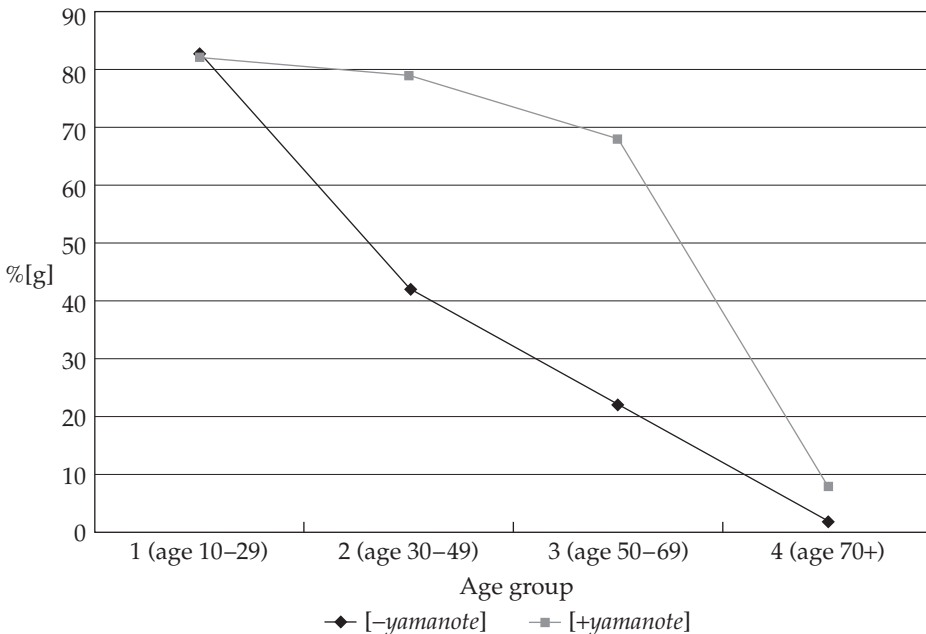

Figure 4.2 Age distribution of [g] by contact with *yamanote*

The results of these analyses, along with the earlier findings and observations, show that the change of word-internal [¥] to [g] in Tokyo Japanese originated in *yamanote* in the early twentieth century and has spread to the entire Tokyo speech community. An overall increase in the use of word-internal [g] in the vernacular is observed over several decades. The change is near completion, and the word-internal [¥] has been replaced by [g] at a very rapid rate within three generations.

2.2.3   *Linguistic internal factors*

As is clear from figure 4.1, except for some speakers in the youngest age group, there are only few who have categorical word-internal [g]. In this subsection, the effect of various linguistic internal factors will be analyzed and inter- and intra-speaker variation will be examined. The data used for the analyses of linguistic internal factors were drawn from the neighborhood studies in Nezu and other areas: a large body of vernacular speech from each speaker was obtained.

As mentioned earlier, there are several systematic exceptions to voiced velar nasalization in the conservative dialect. There are items for which the nasalization is blocked, and word-internal [g] occurs (Vance 1987).

The /g/ in relatively recent loanwords such as (3a) is pronounced as [g], while older ones including (3b) are more likely to have [¥].

(3)   a.   yooguruto   "yogurt"
      b.   i¥irisu   "England"

In the corpus, there were 56 loanword tokens; the percentage of [g] for this category is 77 percent. This result clearly shows that loanwords favor the [g] realization.

The reduplicated /g/ at the beginning of the second part in mimetic forms is always realized as [g].

(4)   gara-gara   "rattle-rattle"

In the corpus, there were 21 mimetic tokens; the percentage of [g] for this category is 100 percent.

Earlier research has indicated that morphological structuring plays a crucial role in determining whether nasalization is obligatory or not. In compound words consisting of two elements, there is a tendency for medial [g] to appear in initial position of the second element. This is almost categorical when the second element is a loanword,

(5)   doku#gasu   "poison gas"

or consists of two Sino-Japanese morphemes. In what follows, the tokens which belong to this category are classified as "#compound."

(6)   sekai#ginkoo   "World Bank"

It should be noted that these second elements (e.g. gasu, ginkoo) also occur with [g] in isolation because they will be in the word-initial position.

When the second element is a single Sino-Japanese bound root, however, the root-initial /g/ is realized as [¥]. The tokens of this type are referred to as "+compound."

(7)   ho+¥o   "protection"

Based on the description of several pronunciation dictionaries, Vance (1987) observed that [¥] is favored when the word-internal voiced velar is induced by sequential voicing. These tokens are abbreviated as "voicing."

(8)   hana+¥asa   "flower umbrella"
        *hana+kasa

Finally, it should be noted that there is a very strong tendency for the nominative case particle -*ga*, which marks the subject, to be pronounced [¥] (Hibiya 1988). This type is classified as "nom." The tokens belonging to this category make up 44 percent of the neighborhood study database. Only 14 percent of them were pronounced [g].

In the following, we will select six individuals from the neighborhood study and examine their data in detail; a substantial number of tokens were obtained from each of them. Since they belong to different age groups, some are typical speakers of the conservative dialect whereas some exhibit a wide range of variation in their voiced velar nasalization. From the analyses below, the loanword and mimetic tokens are excluded because the effect of these factors is obvious. Consider table 4.2.

For each speaker, the percentage of [g] tokens was calculated. As is expected from their age, the first two individuals, speakers 1 and 2, are consistent [¥] speakers. [g] appeared in word-initial position or in compounds of which the second element is either a loanword or consists of two Sino-Japanese morphemes, while [¥] appeared in all other cases.

Let us now examine the results of speakers 3 and 4. They pronounced "#compound" /g/s as [g] almost categorically. Those which belong to the other four categories, i.e. "+compound," 'voicing," "no boundary," and "nom," however, were not always realized as [¥].

Speakers 5 and 6 use [g] much more often. This tendency is especially clear for speaker 6, who pronounced /g/s as [g] almost categorically except for "nom" and "no boundary" tokens.

The foregoing analyses show that the levels of the [g] usage increase steadily as age decreases. Linguistic internal factors strongly influence the pronunciation of /g/. The plosive pronunciation appears least often with "nom" tokens.

Table 4.2 Effects of linguistic factors

	#compound	+compound	Voicing	No boundary	Nom	Total
Speaker 1 (male, 80s):						
[g]/N	26/32	4/88	0/17	1/122	1/308	32/567
[g]%	81	5	0	1	0	6
Speaker 2 (male, 70s):						
[g]/N	8/10	0/106	0/1	3/124	4/303	15/544
[g]%	80	0	0	2	1	3
Speaker 3 (female, 40s):						
[g]/N	10/10	16/60	0/10	23/109	14/191	53/380
[g]%	100	27	0	23	7	14
Speaker 4 (female, 30s):						
[g]/N	10/11	15/44	5/15	18/46	33/137	81/253
[g]%	91	34	33	39	24	32
Speaker 5 (male, 20s):						
[g]/N	3/3	20/32	4/5	48/70	49/75	124/185
[g]%	100	63	80	69	65	67
Speaker 6 (female, 10s):						
[g]/N	3/3	38/39	3/3	38/43	41/52	123/140
[g]%	100	97	100	88	79	79

2.3   An optimality-theoretic approach to variation and change

As is obvious from the discussion of the /-t, d/ deletion phenomenon in English in section 1.2, the variationist approach has benefited greatly from formal linguistic theory. Over the past several decades, the field has formulated the findings of variation studies in terms of the Sound Pattern of English format and the lexical phonology.

The present section is another attempt to account for the quantitative data by a particular formal model. The theory and analysis of variationist sociolinguistics in the 1990s have been enriched in a significant way by drawing on the current emergence of Optimality Theory (McCarthy and Prince 1993a, 1993b, Prince and Smolensky 1993, Reynolds 1994, Reynolds and Nagy 1994, Nagy 1996, Sells et al. 1996, Nagy and Reynolds 1996, 1997, Zubritskaya 1994, 1997). Optimality Theory is a constraint-based approach and a number of constraints are relevant for the realization of the voiced velar. As will be shown, these constraints are ranked and violable, as is standard in Optimality Theory.

Adopting the framework of McCarthy and Prince (1995), Itô and Mester (1997a: 4) have proposed the following three constraints to account for the voiced velar nasalization among words consisting of a single element in the conservative dialect. The abbreviations which will be used for these constraints in this section are given in parentheses.

¥ is prohibited in word-initial position (*[¥). The form [geki] satisfies this constraint because the initial sound is [g], while the form [¥eki] does not because it starts with [¥].

Voiced dorsal obstruents are prohibited (*g). [¥eki] and [ka¥e] satisfy this constraint because they do not contain [g]. Given the choice between [geki] and [¥eki], or [kage] and [ka¥e], this constraint chooses the latter forms.

The underlying form and surface form are identically specified for [nasal] (IDENTLS). When the input forms are /geki/ and /kage/, [geki] and [kage] satisfy this constraint while [¥eki] and [ka¥e] do not.

Another constraint is necessary to account for both the optionality of voiced velar nasalization in certain compounds, i.e. words consisting of two free elements (*hana-gara* and *hana-¥ara* "flower pattern"), and the obligatory nasalization of bound stems (*doku¥a* "poison fang"). The bound form of a stem must be segmentally identical with its corresponding free form (IDENTSS).

In Itô and Mester (1997a), these four constraints are ranked in the following hierarchy. The place for IDENTSS is exactly the same as *g (free ranking).

(9)   1   *[¥
        2   *g
            IDENTSS
        3   IDENTLS

/hana-gara/ Surf [gara]	*[¥	IDENTSS	*g	IDENTLS
→ [hanagara]			*	
[hana¥ara]		*!		*

/hana-gara/ Surf [gara]	*[¥	*g	IDENTSS	IDENTLS
[hanagara]		*!		
→ [hana¥ara]			*	*

Given the choice between [hanagara] and [hana¥ara], the former is the winner with the 1 *[¥, 2 IDENTSS, 3 *g, and 4 IDENTLS ranking, since it satisfies the second constraint. On the other hand, if the ranking is 1 *[¥, 2 *g, 3 IDENTSS, and 4 IDENTLS, the latter is selected because avoiding the voiced velar [g] is more important. In the case of *dokuga*, [doku¥a] is chosen with either ranking. This candidate satisfies *g. IDENTSS is irrelevant here because, being a bound form, [-¥a] does not have a corresponding free form.

Finally, Itô and Mester (1997a) have proposed one more constraint. In a compound word, the second element must begin with a [+voice] segment (SEQ VOI). For instance, [iro+kami] "colored paper" does not satisfy this constraint, whereas both [iro+gami] and [iro+¥ami] do. No other constraint dominates SEQ VOI. Since it is a constraint regarding the noninitial element, *[¥ becomes irrelevant whenever SEQ VOI plays a role in determining the winning candidate. This is illustrated by the vertical line between *[¥ and SEQ VOI. The final constraint ranking is given in (10):

(10)   1   *[¥ | SEQ VOI
       2   *g
           IDENTSS
       3   IDENTLS

Itô and Mester (1997a) have specifically included no social and geographic variation in their analyses by focusing on the conservative dialect. How can the inherent inter- and intra-speaker variability found in the speech community be accounted for within the framework of Optimality Theory?

Reynolds (1994) summarizes three ways for Optimality Theory in its current state to handle variation. (1) Multiple candidates may be equally optimal because more than one form satisfies exactly the same constraints; (2) several different rankings may exist within a given language; or (3) two or more constraints may be unranked with respect to each other.

In order to account for the patterns of diachronic change, Reynolds (1994) has added the notion of Floating Constraint (FC) to Optimality Theory. FCs are those which can appear anywhere within its domain on the constraint hierarchy. The ranking of normal constraints is fixed while that of FCs may vary in a principled manner. The addition of FCs produces several different rankings under which different forms may be optimal.

Let us adopt this approach to account for the variation and change discussed in the above sections. The constraints are ranked as (11) for the single stem word:

(11) 1 *[¥
      2 { *g
         { IDENTLS }

*[¥ is anchored and always ranked over the other two constraints, while *g and IDENTLS are FCs which can appear in any position with each other. As has been seen, *g is ranked above IDENTLS in the conservative dialect. The relative ranking of these two constraints is reversed in the newer dialect and IDENTLS is ranked above *g.[3] Consider /kage/ as an example. Neither [kage] nor [ka¥e] violates *[¥. [kage], not [ka¥e], is selected because it does not violate IDENTLS.

Let us include IDENTSS and SEQ VOI in the constraint hierarchy.

(12) 1 *[¥ | SEQ VOI
      2 [......   IDENTSS   ......]
         [{ *g   }   ]
         [{ IDENTLS }   ]

IDENTSS is also an FC which can appear in any position within its domain. As has been stated, *g and IDENTLS are FCs with respect to each other. This hierarchy, consisting of both ranked and floating constraints, thus produces different hierarchical orders. The newer dialect has the order shown in (13), which makes the right prediction for the words with three different types of morphological structuring. Tableaux for /hana+gara/, /doku-ga/, and /iro+kami/ are given in (14), (15), and (16) respectively.

(13) 1 *[¥ | SEQ VOI
      2 IDENTSS
      3 IDENTLS
      4 *g

(14)

/hana+gara/ Surf [gara]	*[¥	SEQ VOI	IDENTSS	IDENTLS	*g
→ [hanagara]				*
[hana¥ara]		*!	*	

(15)

/doku-ga/ Surf –	*[¥	SEQ VOI	IDENTSS	IDENTLS	*g
→ [doku ga]				*
[doku ¥a]		*!		

(16)

/iro+kami/ Surf [kami]	*[¥	SEQ VOI	IDENTSS	IDENTLS	*g
[irogami]		*	*	*!
→ [iro¥ami]		*	*	
[irokami]	*!			

First let us examine /hana+gara/. *[¥ and Seq Voi are irrelevant. The form [hana-gara] is the winning candidate because [hana-¥ara] violates IdentSS. In the case of /doku-ga/, for which *[¥ and Seq Voi are again irrelevant, neither [doku ga] nor [doku ¥a] violates IdentSS. IdentLS is crucial in selecting [doku ga]. For /iro+kami/, [iro-kami] is excluded because it violates Seq Voi. IdentSS is violated both by [iro-gami] and [iro-¥ami], because neither is identical with /iro+kami/. IdentLS cannot distinguish between the two either. Finally, [iro-¥ami] is selected because it is the only candidate that fulfills *g.

The third case is particularly interesting in that it accounts for the data presented in table 4.2. Both speakers 3 and 4 disfavor the [g] pronunciation more when /g/ is induced by sequential voicing than when it occurs at nonboundary position. However, there are too few "voicing" tokens from speakers 5 and 6 to support a detailed analysis of this constraint.

The change from [g] to [¥] which has been discussed earlier in this chapter can be accounted for by positing that (1) IdentSS floats at the higher end of its domain, i.e. right below *[¥|Seq Voi; and (2) IdentLS is ranked higher than *g in the grammar of the newer dialect.

3   Synthesis: Variationist Approach to Linguistic Heterogeneity

In the development of theoretical linguistics, it has generally been assumed that language is homogeneous and that linguistic variation falls outside the scope of the field. Variationist research carried out in the past three decades, however, has shown that variation found within spontaneous speech exhibits a highly regular pattern whether it represents a change in progress or not. The systematic nature of inherent variation has repeatedly been shown in numerous studies. The voiced velar nasalization alternation can be viewed as another supporting evidence for the notion of "orderly heterogeneity."

In review, the analyses of the natural conversation data have shown that in Tokyo Japanese the word-internal /g/ is undergoing a change from [¥] to [g]. This diachronic change is manifested as synchronic variation, which is constrained by both social and linguistic factors. Among the former, the effect of two factors, i.e. age and one's contact with *yamanote*, have been crucial. As to the latter, the variation is constrained by the environments in which /g/ is found.

The goal of variationist sociolinguistics is to give an answer to the question about where and how variation is located in speakers' grammars. In the present investigation, we have shown that one way of solving this problem is to adopt Optimality Theory enhanced with FCs, in analyzing the data obtained by means of the methodology which has been developed within the tradition of variationist sociolinguistics. This approach to heterogeneity, which attempts to account for phonological variation in terms of constraint ranking, has proved

fruitful. It provides new ways for relating quantitative patterns to the formal principles and will link empirical studies of variation and formal linguistic theory.

In the present chapter, I have discussed several assumptions underlying the framework first proposed by Weinreich et al. (1968) and recognized the value of variationist sociolinguistics. The data on which the findings in this investigation are based have mostly been collected in one neighborhood in Tokyo. It is desirable that similar projects which rely on comparable methodological principles be carried out at different sites.

To conclude, a few suggestions can be made for future research. In the history of Japanese phonology, pitch accent is by far the most widely studied topic. Like most phonological aspects, it changes over time. The wealth of material collected within the field of traditional dialectology serves as a starting point in searching for variables to be investigated in detail.

Traditionally it has been said that adjectives are divided into two classes in Tokyo Japanese: accented and unaccented. The former has a lexical accent on the penultimate mora in the present tense form; the latter does not.

(17) *Accented*   *Unaccented*
      ao'i   "blue"   akai   "red"
      uresi'i   "happy"   akarui   "bright"
      omosiro'i   "interesting" muzukasii   "difficult"

All the conjugational forms of the former have an accent on the penultimate vowel of the stem while only some of the latter have an accent on the final vowel of the stem.

(18)   *Accented*   *Unaccented*
      stem   aok-   akak-
      preverbal   a'oku   akaku
      past   a'okatta   aka'katta

Preliminary analyses (Hibiya 1990, 1991, 1993, Nakao et al. 1997) show the apparent and real-time evidence for the loss of this distinction.

(19)   *Accented*   *Unaccented*
      present tense   ao'i   aka'i
      preverbal   ao'ku   aka'ku
      past   ao'katta   aka'katta

Vowel devoicing is another topic for future investigation. In some regional dialects, high vowels are devoiced when they occur between two voiceless consonants, or at the end of the word and preceded by a voiceless consonant. The process interacts in an interesting way with accent (Nakao et al. 1997). It applies only when the syllables containing high vowels are unaccented.

In order to avoid an accent on a devoiced syllable, accent shift occurs as in (20) and (21).

(20)   compound noun tookyooti'hoo   "Tokyo region"
 $\rightarrow$ tookyootiho'o (accent shift)

(21)   present tense   ayasi'i (accented adjective) "dubious"
 preverbal   ayasi'ku $\rightarrow$ aya'siku (accent shift)

Recent observations, however, have found that the accent shift has been disappearing in Tokyo Japanese (tookyooti'hoo, ayasi'ku). More analyses are called for to account for these changes.

   Recent studies (Labov 1989, Guy and Boyd 1990, Kerswill and Williams 1992, Roberts 1994, 1996, 1997a, 1997b, Kerswill 1996, Roberts and Labov 1995) have illuminated the acquisition of variation by young children. Among them, Labov (1989), Guy and Boyd (1990), and Roberts (1994, 1997b) have shown that first-language speakers as young as 3 or 4 years of age acquire variable constraints on /-t, d/ deletion, discussed in section 1.2. This is an area which should have attracted more attention in the history of variationist sociolinguistics. More studies on this topic will give us some clue to the problem of how constraints on variables are transmitted from generation to generation.

NOTES

1   Until recently most research of the field has relied on random sampling in a speech community under investigation. Two different approaches, namely social network analysis (L. Milroy 1980, 1987, J. Milroy 1992) and principal components analysis (Horvath 1985), have been proposed as alternatives to the conventional methodology.

2   The fieldwork was conducted by the author and Kenjiro Matsuda in 1986. His cooperation is much appreciated.

3   This ranking is exactly the same as lexically marked ranking, which Itô and Mester (1997a: 23) have proposed to account for nonnasalized g in loanwords.