Topics for Today's Lecture
Quantitative Study of Sociolinguistic Variation
22 January 2011
Prof. Robert S. BAUER 包睿舜
Department of Linguistics
University of Hong Kong

1.
Topics for Today's Lecture
Linguistic theory and the everyday real world: Chomsky's ideal speaker-listener
Linguistic variation
Linguistic variable
Linguistic variant
Sociolinguistic variation
Quantitative sociolinguistics
Five social factors influence and correlate with speaker's linguistic behavior

2.
Topics for Today's Lecture
Presenting sociolinguistic data
Organizing and carrying out a quantitative sociolinguistics research project
Dependent and independent sociolinguistic variables
Six stages in conducting quantitative sociolinguistics research
Sample speaker population: How many subjects do you need?
Design of sociolinguistic interview for collecting speech data

3.
Topics for Today's Lecture
Linguistic insecurity
Hypercorrection
Observer's paradox
Rapid, anonymous interview
Labov's five methodological axioms of sociolinguistic research
Styles of speaking
Style-shifting

4.
Topics for Today's Lecture
Sociolinguistic study of phonetic variation and change
Conditioned phonetic variation
Unconditioned phonetic variation
Free phonetic variation

5.
Linguistic Theory According to Noam Chomsky
"Linguistic theory is concerned primarily with an ideal speaker-listener, in a completely homogeneous speech-community, who knows its language perfectly . . ."
Noam Chomsky, Aspects of the Theory of Syntax (1968:3) (emphasis added)
<table>
<thead>
<tr>
<th>Box 1</th>
<th>Speakers’ Language Behavior in the Real World</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>The ideal speaker-listener and the completely homogeneous speech-community</strong> are convenient abstractions that suit the purposes of a theoretical linguist. However, they do NOT exist in the real world. If we were to accept this limitation imposed by the Chomskyan view of language and linguistic behavior, then we could neither recognize nor study the very real, everyday phenomenon of linguistic variation, and there would be no field called sociolinguistics.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Box 2</th>
<th>Linguistic Variation and Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Every language is in a constant state of change; as we will see, variation plays an important role in the process of change. Sociolinguistic research can help us understand variation and change: “Linguistic research which is based on empirical work in secular linguistics and which is concerned to apply the data obtained in such studies to the solution of problems of linguistic theory, such as how and why language changes, and what is the cognitive status of linguistic variability.” (Trudgill 2003:83)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Box 3</th>
<th>Sociolinguistics and Linguistic Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>As already stated, sociolinguistics is basically about studying the speaker’s two or more ways of saying the same thing, that is, linguistic variation. On one occasion the speaker pronounces a word one way or uses a particular word or syntactic pattern, while on some other occasion s/he uses a different pronunciation, word, or syntactic pattern, yet the meaning associated with each of these variant linguistic items is essentially the same.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Box 4</th>
<th>Where does variation come from?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Answer:</strong> Differences (= variations) in speakers’ linguistic knowledge influence their linguistic behavior. <strong>Generalization from Lecture #1:</strong> No two people who think of themselves as speaking the same language possess exactly the same knowledge of that language, because they do not have the same experiences of their language. This statement applies to parents and their children and even to siblings growing up in the same family.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Box 5</th>
<th>Sociolinguistic Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>This is the social distribution of linguistic variation within the speech community. The study of sociolinguistic variation undertakes the following tasks: (1) identifying which individual speakers and social groups exhibit linguistic variation; (2) describing the relationship between social and linguistic variation; and (3) explaining the significance of this relationship.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Box 6</th>
<th>Theory and Methodology of Sociolinguistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>In early 1960’s William Labov began his sociolinguistic investigations of English in the USA and introduced and developed the theory, methods, and terminology for conducting sociolinguistic research. Labov’s contributions have had a major impact on the study of linguistic variation and change and developing the field of sociolinguistics. He has been justifiably called the father of sociolinguistics.</td>
<td></td>
</tr>
</tbody>
</table>
What is Quantitative Sociolinguistics?
For some linguists the quantitative study of linguistic variation is sociolinguistics. Quantitative sociolinguistics is a kind of statistical and correlational linguistics that measures linguistic variation; that is, it measures the impact of certain social variables that are associated with the speaker (sex, age, educational level, social class) on the speaker's use of particular linguistic variables.

Linguistic Variable
“...a linguistic item which has identifiable variants.” (Wardhaugh 1992:139) “A linguistic unit, sometimes known as a sociolinguistic variable, initially developed by Labov in connection with his work in secular linguistics and variation theory, in order to be able to handle linguistic variation. Variables may be lexical and grammatical, but are most often phonological.” (Trudgill 2003:82) (Secular – “of or relating to the worldly or temporal; not overtly or specifically religious”).

Linguistic Variants of Variables
The two or more forms of a linguistic variable for which variation has been observed.
- Post-vocalic “r” is a linguistic variable with two variants: [ɹ] and θ.
- “-ng” (suffix -ing) is a variable with two variants: [ŋ] and [n].
- Linguistic variables and their variants are typically written together in a formula: (r):[ɹ, θ] as in four [foː, foː], fourth [foːθ, foθ] (ng):[ŋ, n] as in singing [ˈsɪŋɪŋ, ˈsɪŋn]

Phonological Variable
“A phonological variable may be more or less than a phoneme, but will be associated with a particular lexical set or group of words in which phonetic variation has been observed to occur, where that variation can be related to social variables or to other linguistic variables.” (Trudgill 2003:82)

Two Kinds of Phonological Variables
1. Discrete: consonants
   (r): [ɹ, θ]
   (ng): [ŋ, n]
2. Non-discrete: vowels
   (oh): [ə], [ɒ], [ɔː], [ɔː], [ʌ]

Labov’s Three Kinds of Phonetic Variants
1. Indicator: variant carries little or no social significance
2. Marker: variant conveys social information about speaker; people are aware of markers which are related to social categories and speech styles;
   e.g. New York English: fourth => [faʊθ] ~ [faθ]
3. Stereotype: variant is popular and conscious characterization of particular social groups;
   e.g. some New Yorkers: bird => “boid”;
   33rd Street => “thoiy thoid street”
Secular Linguistics
“A view of sociolinguistics as a methodology – a way of doing linguistics – associated particularly with the American linguist William Labov, and sometimes also known as quantitative sociolinguistics or, less properly, **correlational sociolinguistics.**”
(Trudgill 2003:117)

Secular Linguistics (contd.)
“Secular linguistics has as its objective a series of goals which are no different from those of any other sort of linguistics, but it works on the assumption that linguistic hypotheses and theories should be based on observations and analyses of vernacular varieties as these are used by ordinary speakers (i.e. not by linguists) in everyday social contexts.”
(Trudgill 2003:117-118)

Secular Linguistics (contd.)
“The research of linguists working in their offices on their intuitions concerning their own dialect of their own language needs to be supplemented and checked by work on (usually tape-recorded) samples of) real language in real contexts.”
Trudgill 2003:117-118

Secular Linguistics (contd.)
“One of the particular concerns of secular linguistics is the attempt to achieve an understanding of linguistic change, and much work in this field is devoted to studying linguistic changes in progress.”
(Trudgill 2003:117-118)

Five Social Factors Correlate with Speaker’s Linguistic Behavior
In developing his sociolinguistic methodology, Labov discovered there are five social facors with which the speaker’s linguistic behavior is closely correlated.

1. **Contextual Style of Speech**
The speaker has available to him or her a choice or range of speech styles. These speech styles are usually determined by the social situation. A formal social situation (e.g. job interview) typically requires a formal style of speaking; an informal, casual social situation calls for an informal, colloquial style of speech.
## 2. Sex of Speaker

Much sociolinguistic research has established that the linguistic behavior of men and women differs. In Western societies women have been found to be more sensitive to prestige linguistic forms than men and to use more of them in their speech. These findings establish that sex of the speaker is an important social variable to be included in any study that investigates the correlation between social factors and linguistic behavior.

## 3. Age of Speaker

Linguistic usage is like clothing fashions in that it can change over time. Pronunciations, words, and the meanings of words can all change. Within the speech community speakers of different age groups belong to different generations, and so their experiences of learning and using their language will vary according to their ages. For this reason, age of the speaker is another important social variable.

## 4. Ethnic Group of Speaker

In Chapter 3, Language and Ethnic Group, of his book *Sociolinguistics* (2000), Trudgill has observed that the way a person speaks depends on the dominant group of people with whom that person associates. This usually means the ethnic group to which a speaker belongs can influence the speaker’s speech patterns, so that s/he typically sounds like other members of his/her ethnic group.

### 4. Ethnic Group of Speaker (contd.)

However, ethnicity is not necessarily the overriding factor that determines a speaker’s linguistic behavior. A person can belong to one ethnic group, but because they associate with another different group, their speech patterns sound as if they belong to that particular ethnic group, and so differ from their own group.

## 5. Socio-economic Status of Speaker

A number of studies have found that linguistic behavior is strongly correlated with the socio-economic class to which the speaker belongs. People with higher socio-economic status typically have more money which gives them access to better education. Speakers who have higher socio-economic status and have been better educated tend to use more forms from the standard language. In contrast, speakers with lower educational levels tend to use non-standard forms.

### Presenting Sociolinguistic Data

The quantitative data produced through sociolinguistic studies are typically presented in bar graphs, line graphs, and tables in which the use of the linguistic variants of a linguistic variable by certain social groups or in certain social contexts is displayed as percentages. The purpose of such graphic displays is to show the general trend in the use of the linguistic variable across the entire speech community.
Fig. 1. New York (r). Percentage of (r):[l] in first (I) and second (II) utterances of fourth (white) and floor (hatched) produced by assistants in three department stores (from Hudson 2000:157)

Table 1. Percentage of r-use by shop assistants in three New York City department stores.

Table 2. (ng):[ŋ] variable in Norwich. Social class: MMC Middle Class; LMC Lower Middle Class; UWC Upper Working Class; MWC Middle Working Class; LWC Lower Working Class. Style: WLS Wordlist; RPS Reading Passage; FS Formal; CS Casual (from Trudgill, The Social Differentiation of English in Norwich, 1974:94).

Fig. 2. Norwich (ng). Proportion of (ng):[ŋ] in speech of five socio-economic classes in four styles: Word-list (white), Reading-passage (hatched), Formal (dotted), Casual (solid). (From Hudson 2000:154)

Fig. 3. Percentage of [a] absence in words like farm and car in Detroit black speech. (Based on Wolfram 1969:110)

Fig. 4. Percentage [z] absence in third-person singular present tense agreement in Detroit black speech. (from Wardhaugh 1992:174)

(s):[z, ə]. For example: She comes => She come He’s good => He good He goes => He go

(Based on Wolfram 1969:136)
Fig. 5. Stylistic and social stratification of (th):[θ/ð], [t/d] as in thing, three, etc. in NYC. (From Labov 1966:260)

New York City Socioeconomic Classes:
- 0-1 Lower Class
- 2-4 Working Class
- 5-8 Lower Middle Class
- 9 Upper Middle Class
- 10 Upper Class

New York City English

Fig. 6. Pronunciation of variable (r) in New York City by social class and speech style. Note cross-over pattern of Lower Middle Class (From Labov 1966:240)

Different variables can have similar social distributions

Table 3. % of non-RP forms for three consonants in Norwich English (Trudgill 2000:37)

<table>
<thead>
<tr>
<th></th>
<th>MMC</th>
<th>LMC</th>
<th>UWC</th>
<th>MWC</th>
<th>LWC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. (ng):[ŋ]</td>
<td>31%</td>
<td>42%</td>
<td>87%</td>
<td>95%</td>
<td>100%</td>
</tr>
<tr>
<td>2. (t):[t]</td>
<td>41%</td>
<td>62%</td>
<td>89%</td>
<td>92%</td>
<td>94%</td>
</tr>
<tr>
<td>3. (h):[θ]</td>
<td>6%</td>
<td>14%</td>
<td>40%</td>
<td>59%</td>
<td>61%</td>
</tr>
</tbody>
</table>

walking [ˈwɔkn], butter [ˈbʌtə], hammer [ˈhæmə]

Linguistic Insecurity

“A set of language attitudes in which speakers have negative feelings about their native variety, or certain aspects of it, and feel insecure about its value or ‘correctness’.”
(Trudgill 2003:81)

Linguistic Insecurity (contd.)

“This insecurity may lead them to attempt to accommodate to or acquire higher status [that is, more prestigious] speech forms, and may lead to hypercorrection on the part of individuals and Labov-hypercorrection on the part of social groups.”
(Trudgill 2003:81)

Linguistic Insecurity (contd.)

“Labov has suggested that it is normally the second-highest status-group in a society [e.g. lower-middle class] that is most prone to linguistic insecurity.”
(Trudgill 2003:81)
Hyper-correction

“A form of hyperadaptation [i.e., speakers of one variety attempt to adopt features of another] in which speakers of a lower prestige variety, in attempting to adopt features of a higher prestige variety, incorrectly analyse differences between the two varieties and overgeneralise on the basis of observed correspondences.” (Trudgill 2003:59-60)

Example of Hyper-correction

“An example from English English is the faulty ‘correction’ of the north of England pronunciation of words such as look from /lʊk/ to supposedly RP /lʌk/ by analogy with correctly observed northern versus RP correspondences such as duck /dʌk/ versus /dʌk/.” (Trudgill 2003:59-60)

Labov Hyper-correction

“A secular linguistic term . . . the second highest status group in a speech community uses higher-status variants in formal styles more frequently than the highest status group. This linguistic behavior can be interpreted as being the result of linguistic insecurity. Labov-hypercorrection should be distinguished from hypercorrection, which is a feature of the speech of individuals.” (Trudgill 2003:71)

Labov Hyper-correction (contd.)

“Labov-hypercorrection is a term which is due to the British linguist J.C. Wells, who suggested that it was necessary to distinguish terminologically between individual hypercorrection and group hypercorrection . . .” (Trudgill 2003:71)

Organizing and Conducting a Sociolinguistics Research Project

- Organizing a quantitative sociolinguistics research project begins with the sociolinguist first observing a speaker’s linguistic variation and then identifying the linguistic variable for investigation.
- Conducting the research itself involves a series of stages in which different but related tasks are performed by the sociolinguist.

Dependent and Independent Sociolinguistic Variables

The dependent variable is the linguistic variable, such as (ng), (r), (z), (th), etc.

The independent variables are the social variables, such as sex, age, educational level, socio-economic status, speech style, etc.
Stages in a Quantitative Sociolinguistic Research Project

Organizing and conducting a quantitative sociolinguistic research project typically proceeds through six stages:
1. Sociolinguist enters speech community
2. Identifies linguistic variables and select speakers
3. Collects data in tape-recorded interviews
4. Transcribes tape-recorded interviews
5. Processes numerical data
6. Interprets findings and explains significance

1. Enter the Speech Community

The sociolinguistics researcher literally enters the speech community so that s/he can observe linguistic variation among its individual speakers and groups of speakers. The researcher by necessity must gain the trust of different groups of speakers so s/he can socially interact with them, talk to them, and learn about how they use their language.

2. Identify Linguistic Variables and Select Speakers as Participants

Through his/her observation of the speakers' linguistic behavior the researcher identifies the variants of the linguistic variables for sociolinguistic investigation.
In addition, the sociolinguist will select a representative sample of speakers from the speech community who will sit down with him/her and participate in sociolinguistic interviews.

Population Sample

In order for a sociolinguistic study to be able to make accurate, verifiable claims about linguistic variation among different social groups in a speech community, it is necessary for the sample of speakers to represent the social diversity of that speech community.

Representative Sample Population

It is very important that the population sample, the people who will be interviewed for the study, accurately represent or reflect the diversity of the speech community.
The population sample must include a wide variety of speakers who are young, middle-aged, old, educated, uneducated, rich, poor, etc.

Number of Subjects?

How many speakers does the sociolinguist need as subjects for his or her study?
The actual number in answer to this question depends on how many independent social variables are investigated in the study.
According to Labov (1972:38), *four to five subjects per cell* are enough to be representative of the population from which the sample of subjects has been drawn.
What do we mean by *subjects per cell*?
Sample-population Matrix with Three Independent Social Variables

<table>
<thead>
<tr>
<th>EDUCATION:</th>
<th>SEX: Males</th>
<th>females</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.14-16 yrs</td>
<td>A 18-29 yrs 4</td>
<td>A 18-29 yrs 4</td>
</tr>
<tr>
<td></td>
<td>B 30-49 yrs 4</td>
<td>B 30-49 yrs 4</td>
</tr>
<tr>
<td></td>
<td>C 50-69 yrs 4</td>
<td>C 50-69 yrs 4</td>
</tr>
<tr>
<td>2.11-13 yrs</td>
<td>A 18-29 yrs 4</td>
<td>A 18-29 yrs 4</td>
</tr>
<tr>
<td></td>
<td>B 30-49 yrs 4</td>
<td>B 30-49 yrs 4</td>
</tr>
<tr>
<td></td>
<td>C 50-69 yrs 4</td>
<td>C 50-69 yrs 4</td>
</tr>
<tr>
<td>3.0-10 yrs</td>
<td>A 18-29 yrs 4</td>
<td>A 18-29 yrs 4</td>
</tr>
<tr>
<td></td>
<td>B 30-49 yrs 4</td>
<td>B 30-49 yrs 4</td>
</tr>
<tr>
<td></td>
<td>C 50-69 yrs 4</td>
<td>C 50-69 yrs 4</td>
</tr>
</tbody>
</table>

3 Social Variables and 4 Speakers Subjects per Cell

- 2 Sexes X 3 Age Groups X 3 Educational Groups = 18 cells
- 18 cells X 4 subjects/cell = 72 subjects
- A sociolinguistic study with three social variables as indicated above would need a minimum of 72 subjects in order to be sufficiently representative.

Socially-Stratified Random Sample

The set of speakers to be studied should be a socially-stratified, random sample; that is, all social groups should be represented in the sample, and there should be no bias in selecting any one particular speaker. Depending on the size and diversity of the speech community, producing the randomized sample can be a complicated process for a sociolinguist who may need help from other specialists, such as a sociologist and statistician.

3. Collect Data through Recorded Interviews

The researcher must collect texts of natural speech from the sample population through recorded interviews for sociolinguistic analysis. The sociolinguist typically uses a tape-recorder (digital recorder) to produce a concrete record of the sociolinguistic interviews that s/he has conducted with his/her sample of speakers.

Necessity of Recording Device

The recording device is essential because the sociolinguist regards the recorded speech as a sample of the speaker’s representative speech which can be listened to, transcribed, and then carefully analyzed. The recording is also convenient because the sociolinguist can listen to it at a later date as many times as s/he needs when transcribing the speaker’s utterances for statistical analysis.

Sociolinguistic Interview

The sociolinguistic interview is structured so that it will provide the researcher with a range of different speech styles from the speakers, that is, formal, informal, and casual kinds of speech that are representative of the styles ordinarily used by speakers.
Sociolinguistic Interview
The sociolinguistic interview includes the kinds of questions that will interest the speaker and induce him or her to forget the formality of the interview and concentrate on telling stories in which he or she becomes emotionally involved. Labov (1972:49-50) has discovered one of the best questions which has this effect on the speaker is “Danger of Death” question. In responding to this question and recalling his or her past experience, the speaker forgets the formality of the sociolinguistic interview.

Speaker’s Vernacular Variety
The less attention the speaker pays to his or her speech, the more likely the vernacular variety will be used; this is the speaker’s most natural and consistent type of speech.

4. Transcribe Recorded Interviews
As researcher listens recorded interviews, s/he transcribes phonetically speakers’ texts of speech with symbols of International Phonetic Alphabet. Then s/he identifies the occurrence of the variants of the linguistic variables in the transcribed texts and notes the frequency of their distribution and the linguistic environments in which the linguistic variables occur.

5. Process Numerical Data
The researcher must next process the numerical data produced during the previous step; this typically takes the form of percentage figures. That is, if the variable occurred 30 times in the speaker’s speech, then the researcher calculates how many instances of it occurred in one variant form instead of the other.

Numerical Calculation of Frequency of Linguistic Variable
For example, if (ng) (English suffix -ing) variable occurred a total of 20 times in a section of the recorded interview, and 13 occurrences were as [ŋ] and 7 as [n], then the percentage frequency of the variable as [n] is 35% (that is, 7/20 = 35%).

6. Interpret Findings and Explain Their Significance
The researcher’s numerical data usually indicate that certain social groups (men, women, working class speakers, younger speakers, etc.), tend to use the linguistic variable in certain proportions in certain speech styles. The task of the sociolinguist is to discover/reveal the correlations between social groups and their use of linguistic variables and then identify and explain the general trends in the speech community’s linguistic behavior.
Labov's Five Methodological Axioms of Sociolinguistic Research

In the course of carrying out numerous sociolinguistic investigations over the years, Labov developed and refined his research methodology. From his research experiences he has distilled five basic methodological principles or axioms on which high quality sociolinguistic research should be based.

Labov's Five Axioms and Importance of Vernacular Speech

These five methodological axioms (from Labov 1972:9-10) help us recognize the central importance to sociolinguistics research of the vernacular, that form of speech which the sociolinguist is striving to find and record, because it “requires the least amount of monitoring by the speaker, . . . and [is] used for the most intimate and casual kind of communication” (Labov 1972:9).

1. Principle of Style-shifting

All speakers possess and use more than one style of speaking and shift from one style to another according to the sociolinguistic situation.

“Every speaker . . . shows a shift of some linguistic variables as the social context and topic change”.

2. Effect of Attention Paid to Speech

“The least attention is given in casual and emotionally involved speech; the most attention [is paid when the speaker is] reading texts and answering formal questions about the forms of words.”

3. Pre-eminence of Vernacular

 “[The vernacular,] the style of speech in which the minimum attention is given to speech, gives the most systematic data for the analysis of linguistic structure.”

For this reason the vernacular is the most important object for the sociolinguist’s study.

4. Formality of Inquiry

“Any systematic observation of a speaker defines a formal context in which more than the minimum attention is paid to speech.”

Obviously, for the speaker to pay attention to his/her speech is unnatural and so undesirable. By being a partner in a conversation with a speaker in order to observe how the s/he speaks, the sociolinguist creates what has become known as the **Observer’s Paradox**: How can the speaker be observed without the presence/participation of the sociolinguist?
5. Primacy of Data

“...accurate answers to linguistic questions on phonology and grammar can only be obtained by direct interviews with the best type of recording equipment.”

That is, the researcher collects objective linguistic data from the speakers of a particular speech community through the process of conducting and recording direct interviews with the speaker informants/subjects.

This material provides the basis for the sociolinguist’s research work.

Style

“In sociolinguistics, a variety of language which is associated with social context and which differs from other styles in terms of their formality.

Styles can thus be ranged along a continuum from very formal to highly informal or colloquial.”

(Trudgill 2003:129-130)

Style (contd.)

“In English, stylistic differentiation is most often signalled by lexical differences. Thus, in British English, to slumber, to sleep, and to kip all mean the same thing, but are stylistically different.”

(Trudgill 2003:129-130)

Styles of Speaking

“Styles are in principle distinct from dialects and from registers: non-standard dialect speakers can and do employ formal styles, and standard speakers can and do use informal styles.

Highly informal vocabulary is often referred to as slang.”

(Trudgill 2003:129-130)

Observer’s Paradox

“A term invented by the American linguist William Labov to describe the major methodological problem of secular linguistics. Secular linguistic research is based on analyses of linguistic data as obtained from people using a vernacular variety in a natural way in everyday speech situations in the speech community.”

(Trudgill 2003:97-98)

Observer’s Paradox (contd.)

“However, observing and recording such speech is difficult because as soon as people realise that their language is the focus of attention, they will tend to speak in a less natural and vernacular manner.”

(Trudgill 2003:98)
Observer's Paradox (contd.)

“The observer's paradox is thus that 'what linguists want to do is to observe the way in which people speak when they are not being observed'. A number of different methodologies have been developed to overcome this paradox." (Trudgill 2003:98)

One of the major tasks of the sociolinguist is to find ways to overcome/neutralize the observer's paradox, e.g. participant observation and rapid and anonymous interview.

Rapid and Anonymous Interview

“One of the fieldwork techniques of secular linguistics designed to overcome some of the constraints of the observer's paradox. In this technique, the fieldworker conducts brief interviews in a public place with a large number of people in such a way as to obtain appropriate linguistic information from them without their realising that their language is being investigated and without their being unduly inconvenienced." (Trudgill 2003:108-109)

Rapid and Anonymous Interview (contd.)

“The most famous series of such interviews was the one conducted by William Labov in which he investigated the speech of shop assistants in department stores in New York by asking questions designed to produce the response 'on the fourth floor', thus obtaining from the informants potential instances of nonprevocalic /r/ [also called post-vocalic [-]]]." (Trudgill 2003:108-109)

Style-shifting

“[This refers to] [c]hanging from one style to another – or, better, moving along the continuum of styles – as the formality of a situation changes, or in order to change the formality of a situation . . .” (Trudgill 2003:130)

Style Stratification

“A term from secular linguistics which refers to the correlation of linguistic variables with social context and formality.

A variable which is subject to style stratification in a speech community will show different use of different variants in different social situations.”

(Trudgill 2003:130)

Style Stratification (contd.)

“Thus, in many forms of British English, the ['?] [glottal stop] variant of the variable (t) – the pronunciation of /t/ in words such as bet and better as a glottal stop – occurs more frequently in informal than in formal styles.

Variables which are subject to style stratification are known as markers.”

(Trudgill 2003:130)
Conditioned Variation
If the variants are predictable, then it is because they occur in certain linguistic environments or under certain linguistic conditions but not others. The most easily observed conditioned variation is phonetic; that is, the variant occurs in certain phonetic environments but not in others.

Phonological Variable in HK Cantonese
Characters 五, 午, 昆, 误 are all pronounced in standard Cantonese with velar nasal syllabic phoneme /ŋ/ (romanized as “ng”).
(ng) is also a sociolinguistic variable: (ng):[ŋ, m].
Formula indicates it has two variants, velar nasal syllabic [ŋ] and bilabial nasal syllabic [m].
This is to say that some Hong Kong Cantonese speakers pronounce 五, 午, 昆, 误, etc. as [ŋ], other speakers pronounce these characters as [m], and some speakers show variation by using both pronunciations, i.e. [ŋ] ~ [m].

Conditioned Variation in Cantonese
Where did this [m] variant come from?
What caused [m] to arise as an alternate or variant form of [ŋ] in speakers’ speech?
It has been proposed that phonetic variant [m] arose through conditioned variation:
In certain phonetic environments with nearby labial sounds, such as [m, p, f, w], there was a strong tendency for some (not all) speakers to change [ŋ] to [m] through phonetic process of (anticipatory and perseveratory) labial assimilation to save articulatory effort.

Labial Assimilation in Cantonese
Examples of environments producing labial assimilation:
五百蚊 [ŋ pack men] ‘$500’ => [m pack men]
十五蚊 [sɪŋ m men] ‘$15’ => [sɪŋ m men]
Our next lecture on Cantonese sociolinguistics examines in more detail the connection between conditioned variation and phonetic environments.

Conditioned Variation and Word Frequency
• Development of [m] variant seems to have occurred first for morpheme 五 [ŋɑ́] ‘five’ which is a high frequency word.
• Later on some other items belonging to the same lexical class (i.e. morphemes also pronounced as [ŋ]) then changed to [m].
• Our next lecture on Cantonese sociolinguistics discusses the Theory of Lexical Diffusion which explains how phonetic variation affects different words at different stages of linguistic change.

Unconditioned Variation
If a variant occurs in all phonetic environments, then it is unconditioned, i.e., no phonetic condition is required to produce it.
So, for example, some Cantonese speakers always say [m] in all phonetic environments, regardless of the occurrence of nearby nasal segments, and they never say [ŋ].
This means that no phonetic condition is needed to trigger the occurrence of [m] for these speakers.
Free Variation
Occurrence of variant forms of a linguistic variable may not require any phonetic condition to trigger their original development. In this case, no general statement can be made predicting the occurrence of either variant and they are said to occur in free variation. Some Cantonese speakers use consonants [n] and [l] in syllable-initial position in free variation, e.g. some speakers pronounce ʃiː as either [niː˨˧] or [liː˨˧], and we do not know which one will occur (although there is a tendency for [l]-forms to occur more often).

Terminology for Quantitative Sociolinguistics
- Free variation
- Hypercorrection
- Independent variable
- Labov-hypercorrection
- Linguistic insecurity
- Linguistic variable
- Linguistic variant
- Linguistic variation

Terminology for Quantitative Sociolinguistics
- Linguistic variety
- Observer’s paradox
- Quantitative sociolinguistics
- Rapid, anonymous interview
- Reading-passage style
- Secular linguistics
- Social stratification
- Social variable

Terminology for Quantitative Sociolinguistics
- Styles of speaking
- Style-stratification
- Synchronic sound change
- Unconditioned variation
- Vernacular
- Word-list reading style
- Working-class speech
### References


### References
