1 Language Variation

James A. Walker

1.1 Introduction: Defining Language Variation

Language variation can be defined as "different ways of saying the same thing," where "different ways" refers to differences in the **form** of language (sounds, words, sentences, ways of speaking) and "the same thing" refers to the intended **meaning** conveyed by those forms: talking about things or events in the world, marking distinctions required by the language's grammar, conveying the speaker's intention, or indicating something about the speaker's social position or relationship to the listener. Let us start by distinguishing variation *between* languages (**interlinguistic** variation) from variation *within* languages (**intralinguistic** variation).

1.1.1 Interlinguistic Variation

Different languages say the same thing in different ways. The words *jī* (Chinese), *frango* (Portuguese), and *kuku* (Swahili) are different sequences of sounds that all refer to the same animal as the English word *chicken*. Since the relationship between form and meaning differs arbitrarily from language to language, and there is no way to figure out the meaning of a word based *only* on its sounds, learning a language involves making the connection between forms and the meanings conveyed by those forms.

Interlinguistic variation goes beyond the words that each language uses to refer to things. The human vocal apparatus is capable of producing many different sounds, but each language uses only a subset of these sounds for the purposes of speech and uses them in ways that differ from the way that other languages use them (**phonology**). The English sounds represented by the spelling *th* (as in *thin* or *the*) are not used in other languages, while sounds such as German *ch* (as in *Bach*) or the "click" sounds in languages such as Xhosa are not used in English. These differences make it difficult for some learners of English to pronounce *th* or for speakers of English to pronounce German *ch* or Xhosa clicks.

Languages differ from each other in the way that they put words together (**morphology**). English uses a number of word-formation strategies to indicate

grammatical distinctions and to create new words. Plural can be indicated by adding a suffix to the noun (*cat* ∼ *cat-s*) or by changing its stem vowel (*goose* ∼ *geese*). New words can be formed by adding a prefix to another word (*read* ∼ *re-read*) or by compounding two words into one (*black* + *board* > *blackboard*). Word-formation strategies vary quite a bit across languages: languages like Turkish and Finnish use many prefixes and suffixes to indicate grammatical distinctions, while languages like Thai and Chinese do not use word-formation strategies to indicate grammatical distinctions but use compounding extensively to form new words.

Languages differ in how they group words to form clauses and sentences (**syntax**). In English, subjects tend to come before verbs, which come before their objects (subject–verb–object). In Japanese, the verb comes after everything else (subject–object–verb), while in Gaelic the verb comes first (verb–subject–object). Morphology and syntax are interrelated: some languages indicate the roles of nouns in the sentence through their position, while others mark these roles on the nouns, which allows more freedom in the syntax.

Finally, languages differ from each other in the strategies used to indicate the speaker's belief in, source of, or attitude toward what they are saying or to indicate elements of their relationship to the hearer (**discourse** and **pragmatics**). Japanese indicates levels of politeness through the choice of pronouns and marking on the verb, and some languages indicate how the speaker came to hear about what they are stating. In English these functions are usually not conveyed grammatically but through discourse strategies, such as indirectness ("*Could you open the door?*") and across verbs in multiple clauses ("*I heard that …*").

1.1.2   Intralinguistic Variation

Variation within languages is less apparent, largely because we learn to filter out a lot of that variation when we learn to speak a language. Speakers sometimes refer to differences between the way people speak their language as a matter of "dialect," but from a linguistic perspective the distinction between dialect and language is not straightforward. Some languages are similar enough to each other to be mutually understandable, but cultural, social, or political attitudes can influence the extent to which their speakers think of them as different. Norwegian, Swedish, and Danish are so similar to each other that their speakers can understand each other, but they are considered to be languages rather than dialects because they are spoken in different countries. In contrast, the many types of "Chinese" spoken in China are very different from each other and not mutually understandable, but they are considered dialects of the same language because of their speakers' shared history, culture, and political unity. Dialects differ from each other on a number of levels, just as languages do.

Focusing on a single dialect (or even a single speaker), we still find variation in form. The English plural suffix *-s* has different realizations: a voiceless sound [s] (*cats*), a voiced sound [z] (*dogs*), or a full syllable [ɪz] (*bushes*). The object of a verb with a particle sometimes occurs after the particle (*he picked up **the children***) and sometimes before it (*he picked **them** up*). However, we do not normally think of these differences in form as variation because they can be predicted on the basis of the linguistic context: the plural alternation is triggered by the type of consonant that comes before the suffix, and the object's position depends on whether it is a pronoun or a noun phrase. The goal of linguistic analysis is to correlate differences in form with differences in meaning or with elements of the linguistic context.

Nevertheless, in linguistic analysis there always remains a certain amount of variation that we are unable to predict, traditionally referred to as "free variation." The use of the word *free* implies that the realization of form is completely random, due to forces outside of language and therefore not of interest to linguistics.

In the 1960s, William Labov and his associates and students began to develop an approach to linguistic analysis that saw variation as an inherent feature of language rather than something to be eliminated or ignored (Labov 1963, 1966). The goal of the **variationist** approach is to conduct linguistic analysis through correlating quantitative patterns of variation with elements of the linguistic and social context.

Central to the variationist approach is the **linguistic variable**, which embodies our definition of linguistic variation as "different ways of saying the same thing": the "different ways" are the **variant** forms and "the same thing" is their common meaning or function. English speakers exhibit variation in the pronunciation of the consonant in the suffix -ing, alternating between a velar form [ɪŋ] and an apical form [ɪn] (*singing* ∼ *singin'*). These forms are in "free variation," as the same speaker can produce either variant under the same circumstances. Determining the circumstances under which a speaker is free to vary between forms is a crucial component of the variationist analysis, known as defining the **variable context** (or the envelope of variation). In the case of (ING), we must restrict the variable context to word-final unstressed -*ing*, as the variation is not observed outside of this context (that is, the word *ring* is never pronounced *rin'*). Defining the variable context determines how we go about looking for examples of the variable (occurrences, or tokens) to count.

1.2   Types of Linguistic Variables

Since variation occurs at all levels of the linguistic system, variables can exist at each of these levels. At the lexical level, different words can refer to the same thing (synonyms). The English words *running shoes*, *runners*,

and *sneakers* may all be used to refer to the same type of athletic footwear. Phonetic variables reflect different pronunciations of the same underlying sound. Consonants can be deleted or inserted or can change in their place or manner of articulation or their voicing, or they may vary in secondary articulation. A couple of well-studied consonantal phonetic variables in English are (t/d)-deletion, which involves the variable pronunciation of /t/ and /d/ at the end of words (*west* ∼ *wes'*, *sand* ∼ *san'*), and (ING), which involves the variable realization of word-final *-ing* with a velar [ɪŋ] or coronal [ɪn] nasal. Phonetic variables define their variable context in terms of structural positions – (t/d)-deletion occurs word-finally in consonant clusters, and (ING) occurs only in unstressed word-final syllables. Variation in vowels involves alternations in height, frontness/backness, and rounding; realization as a monophthong or a diphthong; and devoicing. In Canadian English, the vowels /ɪ/, /ɛ/ and /æ/are variably lowered and retracted. The variable context for vowels may be defined on the basis of their underlying sounds, as above, or may be defined through classes of words that act similarly and are identified by keywords (the KIT, DRESS, and TRAP vowel classes). Consonant variables are normally considered to be categorical – each variant can be classified into a different category (deleted or not deleted, velar or coronal). Since vowel variants form a continuum, they may be classified impressionistically (lowered or not lowered) or the properties of their soundwave may be measured for identifying formant frequencies. Suprasegmental variables involve considerations of pitch (tone, intonation) or rhythm (prosody, stress). In some varieties of English a declaration may be uttered with a rising intonation, making it sound more like a question. As with vowels, classification of variants may be made impressionistically or acoustically.

Morphological variables concern the variable occurrence of forms that function to indicate grammatical differences. Normally these functions reflect inflectional properties such as number, gender, or case for nouns or tense, mood, aspect, or person–number agreement for verbs. For morphological variables the definition of the variable context is usually made on the basis of grammatical function – since the function itself may alternate between overt and unmarked forms, the grammatical function sometimes must be inferred from the wider discourse context. If a speaker alternates between saying "three cats" and "three cat" or "Yesterday she walked a mile" and "Yesterday she walk a mile," we can infer that there is variation in the formal marking of plural or tense, respectively. Syntactic variables concern alternation in the position of constituents in the sentence or the presence or absence of grammatical words. The given example of verbs with particles demonstrates predictable variation with pronoun objects, but if the object is a noun phrase, the particle placement is variable: *he picked up **the children** ∼ he picked **the children** up*. English shows alternation in the occurrence of the complementizer *that*: *she knew that he was lying ∼ she knew*

he was lying. Note that there is often no neat division between morphological and syntactic variation, and in fact some variables may cut across this division. In French, reference to future time alternates between inflection on the verb (*j'acheterai*, "I will buy") and a multiword construction (*je **vais** acheter*, "I'm going to buy").

Extending the study of linguistic variation into the realm of morphology and syntax brings us into considerations of meaning, which creates problems for defining the variable context. If linguistic variation is "different ways of saying the same thing," can we interpret differences in morphology and syntax as referring to the same thing or are all such differences indicative of changes in meaning?

This question occupied the variationist approach in its early years (Lavandera 1977; D. Sankoff 1988), but recent research has relaxed the requirement of strict equivalence of meaning as a criterion for defining variables. Instead, variables are defined on the basis of their shared function. The different forms of the French future are often said to indicate subtle distinctions of meaning (such as "near future" vs. "far future"), but in practice these distinctions are not always evident. If we define the different forms as fulfilling the same (grammatical) function – referring to states or events after the time of speaking – we sidestep the need to determine whether in fact there are subtle differences of meaning.

The question of equivalence of meaning becomes even more controversial when we move into variation at the level of discourse or pragmatics, both because differences in form are here themselves normally taken to indicate differences in meaning and because the meaning of forms is not always clear or may be multilayered. People often comment on the prevalence of the word *like* as a discourse marker in English, but it is not clear exactly what its meaning is (or meanings are), where it can (and cannot) occur, and what (if anything) it varies with (D'Arcy 2017). Nevertheless, there are variables for which a discourse function can be isolated. One use of *like* that has accelerated in recent years in English is in reporting speech (*she was like, "How are you?"*), where it alternates with verbs of saying as well as the verb *go* (*she went, "How are you?"*). All of these forms have in common the discourse function of introducing speech made outside of the current speech event, which can then serve as the variable context (Buchstaller 2014).

1.3   Dimensions of Variation

The variationist approach is concerned not only with the simple fact of linguistic variation or with how frequently each of the variants occurs but also, and more importantly, how variants are distributed across contexts. Some contexts are language internal – for example, we might ask how frequently the deletion of a consonant occurs when the following sound is a consonant or when it is a vowel.

Examining the distribution of variants according to elements of the linguistic context (which simply represents a quantitative extension of linguistic analysis) can provide us with further evidence of how the language works.

Contexts that are external to language are just as important to the variationist approach. Some of these distributions may reflect underlying linguistic differences – groups of speakers may show differences because they speak different dialects or languages – but more often they demonstrate that the choice of variant has social meaning. Speakers of a language use linguistic variation to construct and express how they view themselves and their social world.

1.3.1   Region (*Dialect*)

Long before the beginning of the variationist approach, variants were noted to be distributed differently according to geographical region. Regional differences arise because people tend to talk like the people that they talk to: people who live near each other are more likely to interact, meaning that their speech will resemble each other's more closely.

The systematic study of regional variation in language (**dialectology**) began in the middle of the nineteenth century, when researchers in France and Germany who were concerned about the loss of traditional dialects under pressure from language standardization began to carry out **dialect surveys**. Traveling to different parts of the country, they would ask inhabitants for the local words and pronunciations of common items, or they would send these questions by post to be answered by local literate officials such as schoolteachers. Recent dialect surveys have relied on advances in technology to ask questions over the telephone or by internet or social media, but the tool of the **dialect questionnaire** remains the same (Chambers and Trudgill 1989).

The goal of a dialect survey is to map the distribution of variants according to their geographical location. Dialect maps allow for graphic interpretation of regional distribution, revealing geographic concentrations of speakers who agree in their choice of variant. Where groups of speakers differ, a line (**isogloss**) can be drawn between geographical areas. The isoglosses for a number of variables often coincide, revealing dialect areas that are characterized by differences across a set of features (Kretzschmar 2017). For example, the traditional distinction between Low German and High German is made on the basis of a number of lexical and phonological variables whose isoglosses coincide.

Traditional dialect surveys are limited by providing a single response to each question. Proportions represent the number of speakers or locations within a region who agree on the choice of variant. As we have seen, individual speakers may vary in their pronunciation, in which case a single response would not capture that variation. A larger number of responses from each speaker or locale would more accurately reflect linguistic behavior. Recent

work in the study of regional variation has incorporated larger datasets that use extended recordings of speech rather than responses to questionnaires, as well as employing more sophisticated statistical and mapping techniques (**dialectometry**) (Szmrecsanyi 2013). The goal remains the same: to understand the geographic distribution of linguistic variation.

1.3.2   Style

As noted, individual speakers do not behave the same under all circumstances. Instead, they alter their way of speaking according to the social situation in which they find themselves (**style**). On a broad level, we can distinguish formal and informal contexts. In certain situations or topics, speech is expected to be more formal or informal. Discussions in a church or about religious topics call for more formal variants than do discussions in a pub or about what you did on the weekend. In addition to the setting or topic, the type of relationship that exists between speakers and their audience calls for different levels of formality. Speaking to a family member involves less formal variants than speaking to an outsider – although even within the family, speaking to a relative of an older generation may call for more formal speech than speaking to a relative of the same generation. Levels of formality may be asymmetric, with lower-status speakers required to use more formal variants with higher-status interlocutors. It is normal for employees to address their employer more formally than vice versa.

In some languages, these distinctions of style related to situation, topic, and status are conveyed grammatically. In Japanese, the choice of pronoun and verb marker reflects these considerations. In other languages, the choice of word depends on situation and status. In Javanese, which word you use for "eat" is determined by who you are talking to and what situation you are in. Even in English, the choice of word that we use or how indirect we are in making a request or command depends on the situation, the topic, or the audience. This type of variation, in which there is a clear-cut and qualitative relationship between the variants and elements of the social situation, we can refer to as differences of **register**. In other words, given our knowledge of the social context, we can make a pretty firm prediction about which variant will occur.

However, other types of variation may show differences of a quantitative nature. I note that my use of *-in'* is higher when I am talking to my friends than when I am making a presentation at work. Over the years, studies of variables that are sensitive to style show that speakers exhibit different rates of variants depending on the context (**style shifting**).

The observed effects of style shifting have been attributed to a number of different explanations. In his work in New York City in the 1960s, Labov (1966) argued that speakers shift according to how much attention they pay to *how*

they are speaking (rather than what they are saying). Under this view, asking the speaker to engage in tasks that focus more and more attention on language (such as reciting a list of words or pairs of words that differ minimally) rather than content should elicit increasingly formal speech. Conversely, more informal or casual speech is elicited by focusing the speaker's attention on content. In relating narratives of personal experience (such as "danger of death" narratives, in which speakers are asked to relate a time when they thought that they were about to die), speakers tend to get so caught up in reliving the experience that they focus less on their speech. Most research since then has either followed or challenged Labov's techniques or explanation.

The effects noted in Labov's study were acknowledged by Allan Bell (1984), but he identified a common theme to the different contexts in which speaker behavior changed: the characteristics of the person or people being spoken to. His audience design model argues that variants that are perceived as being formal or informal derived that interpretation from their association with the social characteristics of the people who use them. In other words, speakers change the way they speak to adapt to the perceived characteristics of their audience. Bell provided an overview of different studies showing how differences in the composition of the audience could be related to greater or lesser degrees of style shifting.

These views of style characterize speakers as reacting or responding to elements of the social context (situation, topic, audience). In contrast, more recent work led by Penelope Eckert (2000) and others has offered an alternative view in which speakers take an active role in defining the sociolinguistic situation. In contrast to responsive or reactive theories of style shifting, agentive approaches view speakers as acting together to define the sociolinguistic situation through their use of linguistic variants. Under this view, linguistic variants have potential rather than fixed meaning, and the interpretation of their meaning depends on the context in which they are used in conversational interaction. Rather than passively relating social meaning to variants, speakers and their interlocutors co-construct the meaning of variants.

1.3.3   Social Group (*Sociolect*)

Since speakers make use of the fact of linguistic variation to construct, express, and interpret social meaning through language, an important task of the variationist approach is to determine which divisions are socially meaningful, whether and how those divisions are expressed linguistically. While the most straightforward approach would be to ask people directly, language speakers are not always consciously aware of their own behavior. Paradoxically, people can exhibit quite complex sociolinguistic behavior without being able to discuss

this behavior! For this reason, the variationist approach relies primarily on observations of linguistic behavior in natural contexts.

Within the variationist approach, there is often a tension between "macro-level" and "micro-level" (or local) definitions of groups. Macro-level definitions involve grouping speakers in ways that are measurable and objective and can be replicated and compared across studies (**etic**). Categories such as occupation, amount of income, level of education, biological sex, and racial-physical characteristics are relatively easy to determine and measure and can be compared across studies. However, these categories, while objectively useful, may not be meaningful to the community or the individuals being studied and may not be conveyed linguistically. An alternative approach, more commonly adopted in recent years, is to seek explanations for sociolinguistic divisions in micro-level or local categories (**emic**). Local categories, being more subjective, are less easy to code and may require supplementing observations of linguistic behavior with long-term ethnographic analysis or social-psychological questionnaires or experiments, but they often provide more socially motivated explanations for the observed patterns of variation than do the macro-level categories (Eckert 2000; Hoffman & Walker 2010).

1.3.3.1   Social Class – Social Network – Community of Practice   In New York in the 1960s, Labov noticed a correlation between style shifting and the speaker's socioeconomic status: speakers with lower status tend to use more informal variants and speakers of higher status use more formal variants (Labov 1966). This finding reflects a common observation across different communities that a person's social status is correlated with his or her linguistic behavior.

Distinctions in socioeconomic status were operationalized by Labov as **social class**, defined on the basis of three (etic) measurements: occupation, income, and education. Together these measurements placed people on a scale which could then be divided into sectors such as "middle class" and "upper working class." The correlation between socioeconomic status and linguistic behavior has been established across a number of studies conducted in English-speaking communities, but the relative effect of each of the components of the socioeconomic scale differs across communities. Labov traced the differences in social-class behavior to two competing pressures: pressure from above (we try to sound like those of higher social classes when considerations of power or prestige are invoked) and pressure from below (we try to sound like those of lower social classes when we want to display solidarity).

Social class has not figured as prominently in subsequent variationist research, which has taken a couple of different directions in accounting for sociolinguistic divisions on the basis of power and status. In a study of Montreal

French, David Sankoff Suzanne and Laberge (1978) adapted Pierre Bourdieu's idea of language as a form of **social capital** to reimagine social class as access to the "linguistic market." In their analysis of different variables in Montreal French, they found correlations between the formal variants associated with standard French and the degree of access that speakers had to this form of capital.

Another direction of research in sociolinguistic stratification has looked at the relationships that exist among speakers. As I noted, geographical variation forms dialect regions because people tend to talk like the people they talk to. In a study of English in Belfast, Northern Ireland, James and Lesley Milroy made use of the notion of **social networks** (Milroy 1987). Under this approach, individuals form social ties with others in order to carry out the different functions of life. Social networks can be characterized by a couple of dimensions: density (the extent to which people you know know each other) and multiplexity (the number of types of relationships). Urban working-class speakers tend to have dense and multiplex networks, which leads to reinforcement and preservation of informal and nonstandard variants, whereas middle-class speakers tend to have looser and more simplex networks, leading to pressure to use more standard and formal variants. The Milroys found a high degree of correlation between the type of social network and the use of local Belfast variants. Because social classes are characterized by different kinds of social network, this approach can be seen as complementary to, rather than in competition with, the socioeconomic approach. Later work has looked in more detail at the types of social networks speakers engage in. Penelope Eckert and Sally McConnell-Ginet (1992) introduced the notion of the "**community of practice**" (CoP) from education theory. The social networks studied by the Milroys were formed on the basis of voluntary (friendship) and involuntary (family, neighborhood, work) bases. In contrast, the CoP is formed by groups of speakers who come together purposefully to achieve certain ends. Their work and subsequent studies have shown that the volitional component of networks can have a strong correlation with patterns of linguistic behavior.

1.3.3.2 Sex/Gender   As sex is a basic biological distinction that figures prominently in most social systems, it should not be surprising that it also serves to distinguish speaker behavior in many communities. We should be careful to note that although there are physiological differences between men and women that produce differences in the speech signal – men tend to have longer vocal cords than women, which lowers overall pitch– not all sex-based differences in linguistic behavior can be traced to physiology. In many languages the sex of the speaker is conveyed through their choice of pronouns or the use of particular discourse markers. English does not convey sex distinctions grammatically, but certain discourse strategies and lexical choices are popularly believed to be more common among men or women.

Speaker sex has been a social factor since the earliest days of variationist studies. Most studies have agreed in finding that women tend to use more formal variants than men do. Early studies suggested that, because women traditionally lack power and capital, they compensate for this lack through their manipulation of the "symbolic capital" of formal or standard language. These studies were later criticized for framing women's behavior as deviant from men's, and therefore in need of explanation. More recent studies view any sex-based differences as arising from different processes of socialization during childhood, which (stereotypically) orients girls and boys to different behavioral norms (Eckert 1989; Labov 1990). In their study in Belfast, the Milroys (Milroy 1987) found that men and women in urban working-class communities tend to have different types of social networks: men's networks are denser and more multiplex, while women's are looser and more simplex. These differences can be seen to result from patterns of employment, in which men worked locally, whereas women had to travel outside of the neighborhood to find work, coming into contact with a greater variety of people.

1.3.3.3 Ethnicity Ethnicity is a fundamental social division that also entails differences in linguistic behavior, but defining ethnicity as a general concept is made difficult because it arises from different sources. Ethnicity is sometimes defined on the basis of physical characteristics, such as skin and hair color and eye and nose shape (race). In the United States, the major ethnic division is between black people (most of whom trace their origins to Africa) and white people (most of whom trace their origins to Europe). Ethnicity may be defined on the basis of adherence to a particular religious tradition. In many societies around the world, people are divided according to whether they are Muslim, Christian, Jewish, Buddhist, etc. Ethnicity may be defined on the basis of place of origin, even if that origin lies in the distant past. In an immigrant society such as Australia, people are divided ethnically according to where their ancestors originally migrated from. Ethnicity may be defined on the basis of a heritage or community language, even if that language is no longer spoken. In the United States there is a sizable Hispanic population whose ethnicity is based on Spanish as their heritage or community language, although many people categorized as Hispanic do not speak Spanish.

Ethnicity serves as the basis for a number of qualitative and quantitative differences in linguistic behavior (ethnolects). A common assumption is that ethnolects stem from a process of language acquisition from the heritage or community language to the majority language, but tracing ethnolectal features to the original language is not straightforward after the first generation. Much of the variationist work on ethnicity in the United States has been concerned with the black–white division. In the way they speak English, many African Americans differ on all linguistic levels from Americans of other origins

(Labov 1972). Some researchers have suggested that these differences arose from an earlier African American creole (a mixture of English and African languages with some second-language features) (Rickford 1998), while others have suggested that the long period of social segregation led to linguistic divergence (Poplack 2000). For other groups, variationist research has shown that ethnic boundaries, whatever their source, can serve to maintain different linguistic systems over time and impede participation in ongoing change. The variants that make up ethnolects may also serve as a way of constructing and expressing a particular ethnic identity, particularly when the original heritage or community language is no longer spoken and cannot serve this function (Hoffman & Walker 2010).

1.3.3.4 Age and Time   As with sex, age is a biological reality that divides people into groups, whether expressed as generations (cohorts of people born at the same time) or life stages (periods of life characterized by different social roles and responsibilities). Age groups may correlate with different patterns of linguistic behavior, but these patterns may result from different sources: changes going on in the community or social expectations of membership in each age group.

Quantitative differences in linguistic behavior may be inferred to reflect ongoing change in the language (**apparent time**) if we assume that individual speakers do not change the way they speak once they acquire their language. People born in 1966 should speak the same whether they are studied in 1986 or 2016, and their speech should reflect the state of their language as spoken in the community in the late 1960s. On the other hand, age-based differences may reflect patterns of linguistic behavior associated with each life stage, repeated by each generation as they go through each stage (**age grading**). A person born in 1966 studied in 1986 and a person born in 1996 studied in 2016 should behave the same.

The only way to definitively decide whether age-based differences reflect ongoing change or age grading is to conduct a longitudinal study of the same community over an extended period of **real time**, either by resampling the same community similarly each time (**trend study**) or by rerecording the same individuals over time (**panel study**). Until recently, real-time studies were difficult to conduct. A real-time study of spoken data is obviously limited by the lack of recorded speech before the twentieth century. Written data can serve as a proxy for spoken data, but we are limited to using whatever documents have happened to survive, and since the rate of literacy in previous centuries was very low, the documents are unlikely to represent the range of variation that existed in the community.

Nevertheless, there is now enough spoken-language data recorded for the purposes of variationist research from the 1960s to the present that real-time

studies have become more feasible. Real-time studies suggest that the assumptions of apparent time are largely valid: most speakers do not substantially change their rate of use of incoming variants across the course of their life. The few cases where change has been observed across the life span are correlated with changes in the individual's life situation, and speakers tend to change in the same direction as the community (G. Sankoff & Blondeau 2007). These studies suggest that we can continue to conduct apparent-time studies, so long as we understand the social characteristics of ongoing change.

When mapped across time or across age groups, changes in progress show an S-shaped curve, starting out slowly, accelerating in the middle, and slowing at the end. Changes in progress are also characterized by particular types of distribution across social categories. Although I noted that women tend to be more conservative in greater use of standard forms, they are at the same time more innovative in ongoing changes. Women, especially young women, are likely to adopt or lead in changes ahead of men. Ethnic groups exhibit different rates of participation in ongoing changes. African Americans have been shown not to participate in changes going on in the wider English-speaking community. Changes will show different patterns of distribution according to the origins of the change. As seen in the case of Belfast, changes involving the introduction of standard variants into the community are usually initiated through the type of social network that characterizes middle-class speakers. Such **changes from above** begin with higher social classes and filter down the social-class hierarchy. In contrast, **changes from below** are initiated by lower social classes and percolate to higher levels of the socioeconomic scale.

1.4   Conclusion

Variation is a fact of language, from differences between languages and dialects to differences between speakers and even within an individual speaker. The variationist approach to the study of language allows us to use the fact of linguistic variation to understand the workings of linguistic systems and the way that speakers make use of variation to construct and express aspects of their social life.

References

Bell, A. 1984. Language style as audience design. *Language in Society* 13, 145–204.

Buchstaller, I. 2014. *Quotatives: New Trends and Sociolinguistic Implications*. Oxford: Wiley-Blackwell.

Chambers, J. K., and Trudgill, P. 1998. *Dialectology*. 2nd edn. Cambridge: Cambridge University Press.

D'Arcy, A. 2017. *Discourse-Pragmatic Variation in Context: Eight Hundred Years of Like*. Amsterdam: John Benjamins.

Eckert, P. 1989. The whole woman: Sex and gender differences in variation. *Language Variation and Change* 1: 245–267.

Eckert, P. 2000. *Linguistic Variation as Social Practice*. Oxford: Blackwell.

Eckert, P., and McConnell-Ginet, S. 1992. Think practically and look locally: Language and gender as community-based practice. *Annual Review of Anthropology* 21, 461–490.

Hoffman, M. F., and Walker, J. A. 2010. Ethnolects and the city: Ethnic orientation and linguistic variation in Toronto English. *Language Variation and Change*, 22, 37–67.

Kretzschmar, W. A. 2017. Linguistic atlases. Pages 57–72 of: C. Boberg, J. Nerbonne, and D. Watt (eds.), *The Handbook of Dialectology*. Oxford: Wiley-Blackwell.

Labov, W. 1963. The social motivation of a sound change. *Word*, 19, 273–309.

Labov, W. 1966. *The Social Stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.

Labov, W. 1972. *Language in the Inner City: Studies in the Black English Vernacular*. Philadelphia: University of Pennsylvania Press.

Labov, W. 1990. The intersection of sex and social class in the course of linguistic change. *Language Variation and Change* 2: 205–254.

Lavandera, B. 1977. Where does the sociolinguistic variable stop? *Working Papers in Sociolinguistics*, 40, 6–19.

Milroy, L. 1987. *Language and Social Networks*. 2nd edn. Oxford: Blackwell.

Poplack, S. (ed.). 2000. *The English History of African American English*. Oxford: Blackwell.

Rickford, J. R. 1998. The Creole origins of African-American vernacular English: Evidence from copula absence. Pages 154–200 of: S. Mufwene, J. R. Rickford, G. Bailey, and J. Baugh (eds.), *African-American English: Structure, History and Use*. London: Routledge.

Sankoff, D. 1988. Sociolinguistics and syntactic variation. Pages 140–161 of: F. J. Newmeyer (ed.), *Linguistics: The Cambridge Survey. Vol. IV: Language: The Socio-Cultural Context*. Cambridge: Cambridge University Press.

Sankoff, D., and Laberge, S. 1978. The linguistic market and the statistical explanation of variability. Pages 239–250 of: D. Sankodff (ed.), *Linguistic Variation: Models and Methods*. New York: Academic Press.

Sankoff, G., and Blondeau, H. 2007. Language change across the lifespan: /r/ in Montreal French. *Language*, 83, 560–588.

Szmrecsanyi, B. 2013. *Grammatical Variation in British English Dialects: A Study in Corpus-Based Dialectometry*. Cambridge: Cambridge University Press.