Research methods in Sociolinguistics

Juan Manuel Hernández-Campoy

The development of Sociolinguistics has been qualitatively and quantitatively outstanding within Linguistic Science since its beginning in the 1950s, with a steady growth in both theoretical and methodological developments as well as in its interdisciplinary directions within the spectrum of language and society. Field methods in sociolinguistic studies have been motivated by the various research objectives pursued: sociological, sociolinguistic, or linguistic goals. The aim of this paper is twofold: (i) to provide a review of the theoretical movements within Sociolinguistics, and, on the basis of this review, (ii) to explore their consequences and implications on the research methods used in the field. This will be achieved by conducting both a retrospective synthesis of past developments and achievements, and an exploration of the current situation and of potential future developments.

1. The origins of sociolinguistics

Since Currie (1952) used the term ‘sociolinguistics’ for the first time, the field has experimented an outstanding qualitative and quantitative growth within Linguistic Science. Currently, it is widely-agreed that Sociolinguistics is an area of Linguistics concerned with the scientific study of the relationships between language and society, which entails practising a different way of doing linguistics that is very much influenced by work in the social sciences. It is empirical research — i.e. based on observation —, specifically focusing on how human beings actually use language in social interaction in real, everyday life situations and studies languages exclusively in their naturalistic social context (see Labov 1997; Trudgill 1983a).

Sociolinguistics is a multidisciplinary branch of linguistic knowledge that developed partly out of anthropology, partly out of ethnography, partly out of sociology, and partly out of dialectology as natural epistemological heritage, which conditioned both theoretically and methodologically this field and conferred it with an interdisciplinary stance. The work of Dell Hymes and his anthropological and folklore-based formation, John Gumperz and his interactional
ethnographic studies, Joshua Fishman and his sociology of language, and William Labov and his linguistic variation, constitute the disciplinary matrixes of this field. According to Figueroa (1994), the underpinnings of these three original strands in Sociolinguistics are philosophically-based: the work of Hymes is framed within relativism, Labov’s variationist approach built on realism, and Gumperz’ is portrayed in the context of intentionality and the interpretation of social meaning (see also Hymes 1972, 1974; Janicki 1990; Williams 1992).

The use of fieldwork methods in sociolinguistic research is a practice inherited from late 19th and early 20th century anthropological and ethnographically-oriented linguists — who studied Indian languages in USA (Franz Boas, Edward Sapir, or Leonard Bloomfield; see Murray 1998), native languages in Africa, Asia, or Australia (Alan Gardiner, Bronislaw Malinowski, or John Firth; see Robins 1971) — and European dialectologists, such as Georg Wenker, Jules Gilliéron, Karl Jaberg, or Jakob Jud (see, for instance, Chambers & Trudgill 1980; Davis 1983; Francis 1983). Their aim was to obtain linguistic data collected in the field, i.e.in natural environments of spoken language, just as people usually and casually meet and interact, rather than in an office (Canger 2001: 779). This new orientation reflects the Labovian’s distinction between traditional armchair linguistics — the theoretical and introspective study of language conducted in one’s own office — and real-world linguistics — going out into the real world to collect data on language as used by ordinary people in their social context in everyday life. Labov claimed that analysts must not rely on their own intuitions since experience has shown that introspective judgements — even those made by informants themselves — may easily violate the reliability and validity of a linguistic generalisation when describing the state of a language (see Labov 1975a, 1975b, 1996). Theorisation without data, especially for taxonomic-type classification building, is conceived as sterile, since theory is solely understood to emerge from empirically collected data.

The methodological rigour of sociolinguistic research to guarantee empirical validity is sustained on the observer’s paradox (see Labov 1972: 209, see also Davies 2001) and the principles of representativeness and generalisability (see Babbie 2001a; Bailey and Tillery 1999, 2004; Feagin 2002; D. Sankoff 2006; Trudgill & Hernández-Campoy 2007; Wolfram 2004) that apply to all social sciences research. A challenge in sociolinguistic research is the need to overcome the observer’s paradox effect during the data collection process: the elicitation of authentic casual speech production with no contamination from the fieldworker’s presence in order to avoid any alteration of the experiment results (Bell 2007; Cicourel 2006) and without violating ethical conditions (see Johnstone 2000a, 2001). Further, as Bucholtz (2003) points out, the ethnographic method of participant-observation has become crucial in the fieldworker’s endeavour to overcome the interference of observation on the activity being observed (see below).
No less important was the growing interest in Sociology, together with the influence of Marxist linguistics and the concern of educationalists and sociologists about poverty and social disadvantage as political issues in western industrialised countries. They tried to answer socially relevant questions, such as the relation between language and social class in Great Britain (the failure of non-standard-speaking children in schools), language and race in the United States (the failure in schools of black children for being users of Afro-American Vernacular English rather than of the standard), language and immigration in Western Europe, sexism in language, cross-cultural miscommunication (i.e., ethnic misunderstanding and racial disharmony), and the issue of language planning in multilingual post-colonial developing countries (see Bolton 1992; Dittmar 1976).

In addition to the epistemological heritage from social sciences (Ethnography, Anthropology, and Sociology) and Dialectology, Sociolinguistics was also a reaction against previous Chomskyan and Saussurean paradigms and traditional dialectological practices, which resulted in a new theoretical model, Variationism, as a prominent sociolinguistic sub-field influenced by the quantitative revolution. The origins of Variationist Sociolinguistics are to be understood as a challenge to the concept of the *idiolect*, the structuralist notions of *langue*/*parole* (language/speech) and *diachrony*/*synchrony* postulated by Ferdinand de Saussure, the later generativist concepts of *competence*/*performance* proposed by Noam Chomsky, the Bloomfieldian notion of *free variation*, and the unrealistic theories and unreliable methodologies of the dialectological tradition. Linguists focused on micro-linguistics, the systematic homogeneity of *langue* and the speaker’s *competence*, deliberately ignoring the macrolinguistic level with the orderly heterogeneous *parole* and the speaker’s *performance*, and appealing to the Bloomfieldian notion of *free variation* as an explanation of any kind of linguistic variability (see Figure 1).

For Chomsky, the focus of study was the abstract system (competence), since linguistic performance was regarded as too disorderly and chaotic to be of any value in offering an understanding of language as a system (Baxter 2010: 118). In contrast, Labov regretted the exclusion of the study of actual sociolinguistic behaviour. He mentioned four distinct difficulties in the investigation of everyday speech that made previous Saussurean and Chomskyan paradigms concentrate on *langue* and *competence*, and which condition the research methods to apply: (i) the ungrammaticality of everyday speech, (ii) variation in speech and in the speech community, (iii) difficulties of hearing and recording real speech, and (iv) the rarity of syntactic forms (Labov 1972: 183–259). In this way, it is not difficult to realise why in the past linguistic studies used to be of the armchair type.

The fact that most speech communities are to some extent socially and linguistically heterogeneous is a complexity which makes research much more difficult to any linguist wishing to describe a particular variety (Trudgill 1983a: 37).
Therefore, for many years during the end of the 19th and early 20th centuries linguists simply ignored this complexity. Dialectologists had focused on either the idiolect or the speech of rural informants, particularly that of elderly people with little education and little travelling experience, in small isolated villages, because they were very concerned with looking for the ‘real’ or ‘pure’ dialects. But, obviously, the idiolect — the speech of one person at one time in one style — was no more regular than the speech of the community as a whole, and ‘real’ or ‘pure’ homogeneous dialects turned out to be a fantasy. A monolithic linguistic system was unable to explain the fact that social structure could maintain any causal relation with the variability present in language. The obvious reaction against this theoretical model of language resulted in a shift from the fictional notion of systematically homogeneous to the orderly heterogeneous ‘speech community’.

After the Second World War, traditional dialectologists realised that confining dialect studies to rural areas meant ignoring the speech of the vast majority of the population, i.e., the speech found in large urban areas, which could not be investigated by applying the methods of traditional rural dialectology. Hence, starting the 1960s, some linguists became interested in macrolinguistics, undertaking the study of the complex reality of language in use in society considering the microlinguistic levels of analysis as linguistic variables. In this way, with the work of William Labov in the United States (see Murray 1998) and Peter Trudgill in England (see Hernández-Campoy 1993), Urban Dialectology appeared, conveying both a linguistic and a social function.
Finally, the study of the equational relationship between language and society through the correlation of extralinguistic factors (socio-demographic and/or context variables) with intralinguistic elements allowed Sociolinguistics to decipher the algorithm encrypting linguistic variation and social meaning and, consequently, to account for variability in language. The origins of Variationist Sociolinguistics did not ignore the neopositivist quantitative revolution. With their rigorous adoption of scientific methods, assuming determinism and the mechanistic nature of human behaviour, linguists’ explicit positivist desire was to develop a quantified social dialectology where extralinguistic (mostly social) factors are capable, by themselves, of explaining entirely the establishment of laws, relationships and processes (see Bayley 2002; Paolillo 2001; Rietveld & van Hout 2006; D. Sankoff 2001). Labov’s (1969) concept of variable rules and its mathematical implementation Varbrul (Cedergren & Sankoff 1974), for example, were developed to describe the predictable probability of patterns of variation (or choice) between alternative forms in language use, and the relationship between dependent (linguistic) and independent (extralinguistic) variables. These aims of explanation and prediction make sociolinguists focus not on what phenomena happen, where and how — which would solely be descriptive in intent — but on the reasons why they occur as they do. Likewise, in order to express accurately and plainly the results of their analysis, variationists demand from themselves the use of the language of mathematics and logic, validity and verifiability being the fundamental criteria and coincidence being conceived in terms of probability (see Hernández-Campoy & Almeida 2005: 10–11). Significance is here understood as the causal relationship between linguistic and extralinguistic variables in compliance of the principles of representativeness and generalisability (reliability and intersubjectivity) to pursue empirical validity. In this way, one can predictably guess the speech characterisation of speakers depending on their social background (class, age, gender, mobility, ethnicity, etc.), allowing sociolinguists to be a kind of omniscient observers in search for empirically valid sociolinguistic universals under the protection of the observer’s paradox effect (i.e. Bayley & Lucas 2007; Hernández-Campoy & Almeida 2005; Paolillo 2001; Tagliamonte 2006).

2. Development and current directions in sociolinguistics

Sociolinguistic research, like any other field of enquiry contributing to empirically-based knowledge, is designed to address explicit questions (or test hypotheses) trying to account for a given linguistic phenomenon and its rationale, which inevitably conditions the observational methodology used in terms of data collection and data analysis procedures (Gordon 2005). Field methods, as Hazen
(2001: 776) points out, require two main stages: **planning**, which is theory in practice, and **fieldwork**, which is doing. In any given study, methodology is dependent on the objectives (research aims) pursued and the theoretical framework adopted (Hernández-Campoy & Almeida 2005).

Trudgill (1978, 1983b) stressed the ambiguous nature of the term 'sociolinguistics', whose multiplicity of interpretations affected the scope of study and homogeneity of the **language and society** paradigm: “[t]he difficulty with sociolinguistics, then, is that it is a term which means many different things to many different people” (Trudgill 1978: 1). In fact, Labov was initially reluctant to the use of this label, which he considered as a “somewhat misleading use of an oddly redundant term” (Labov 1972: 183), “since it implies that there can be a successful linguistic theory or practice which is not social” (Labov 1972: xix). The reasons given by Trudgill for arguing in this way are based on how objectives of researchers in this field vary, even using the same data and methodology. Accordingly, we can differentiate between those studies that are clearly sociolinguistic in nature — since they use sociological data for linguistic purposes or for both sociolinguistic and linguistic purposes — from those that are not clearly sociolinguistic — as they use linguistic data for sociological purposes only. One thing is to investigate the relationships between **language and society** with the aim of understanding better the structure and nature of language and how languages function in social interaction, and a different thing altogether is to investigate those relationships with the aim of understanding society.

Following Trudgill (1978), on account of their aims, three different directions can be distinguished in studies of **language and society**: (i) sociological objectives, (ii) sociological and linguistic objectives, and (iii) linguistic objectives (see Figure 2).

The first group consists of studies within the **language and society** paradigm which are purely social in orientation, and thus with no linguistic objectives. That is, according to Trudgill (1978: 2), the aim of Ethnomethodology, which uses linguistic data, is to understand how language is used in social interaction (particularly talk but not speech) for sociological purposes; that is, not to tell us about language but about society (see the pioneering work of Cicourel 1974; Garfinkel 1967; Psathas 1979; or Turner 1974).

The second group of studies are, in varying degrees, both sociological and linguistic in orientation. It is in this category where the main problem with sociolinguistics, as a term, lies; i.e. the different places where scholars draw the dividing line between **Language and Society** and **Sociolinguistics**: the **Sociology of Language**, the **Social Psychology of Language**, **Anthropological Linguistics**, **Discourse Analysis**, the **Ethnography of Communication** and **Language and Gender**. Qualitative method-
ologies characterize this group of studies, as developed and applied by Dell Hymes, Gumperz, Goffman, or Sacks, amongst others (see Berenz 2001; Schuman 2001).

The third group consists of studies in the field of language and society guided by purely linguistic aims: Traditional Dialectology, Variationist Sociolinguistics, Geolinguistics, Creole Sociolinguistics, and Historical Sociolinguistics. All work in this category is made up of empirical studies of language as spoken in its social context, whose ultimate goal is to improve linguistic theory and to develop our understanding of the nature of language.

The search for the vernacular (naturally-produced) speech in sociolinguistic fieldwork and field methods have mostly drawn on survey and ethnographic research (see Table 1). Large-scale survey sampling is quantitative and was first used in Dialectology at the end of the 19th century, when Wenker conducted the first dialect survey in Germany (see Chambers & Trudgill 1980). He distributed a list of forty sentences written in standard German to nearly 50,000 schoolmasters in north Germany asking them to return the list transcribed into their local dialect. The result of the survey was the first linguistic atlas published, *Sprachatlas des Deutschen Reichs*, deposited in Marburg and Berlin in 1881. In France, in 1896, for the *Atlas Linguistique de la France*, the use of trained fieldworkers to gather data through pre-determined interviews replaced the postal questionnaire. Gilliéron designed a questionnaire that isolated about 1,500 specific items in order to elicit responses and chose a fieldworker (Edmont) to record them at each interview using a phonetic notation consistently. Similar national surveys were carried out in Italy and southern Switzerland (*Sprach- und Sachatlas des Italiens und der Südschweiz*), Catalonia...
<table>
<thead>
<tr>
<th>Discipline</th>
<th>Objectives</th>
<th>Methodology</th>
<th>Method/Technique/Tool/Data</th>
<th>Domains</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethnomethodology</td>
<td>Sociological Qualitative</td>
<td>Ethnographic fieldwork</td>
<td>Practical reasoning, common-sense knowledge</td>
<td></td>
</tr>
<tr>
<td>Sociology of Language</td>
<td>Sociological and Linguistic</td>
<td>Survey fieldwork</td>
<td>Language planning, bilingualism, diglossia, lingua franca, verbal repertoire, code-switching, multilingualism, language loyalty, linguistic ecology, etc.</td>
<td></td>
</tr>
<tr>
<td>Social Psychology of Language</td>
<td>Sociological and Linguistic</td>
<td>Survey fieldwork</td>
<td>Folk Linguistics (Perceptual Dialectology), attitudes and ideologies, Language loyalty, linguistic, accommodation, the aesthetics and prestige of dialects, gender differences, dialect distance, nativeness, etc.</td>
<td></td>
</tr>
<tr>
<td>Anthropological Linguistics</td>
<td>Sociological and Linguistic</td>
<td>Ethnographic fieldwork</td>
<td>Kinship systems, linguistic taboo, linguistic relativity, globalisation, nationalism, language commodification, ideology, language and political economy, etc.</td>
<td></td>
</tr>
<tr>
<td>Discourse Analysis</td>
<td>Sociological and Linguistic</td>
<td>Ethnographic fieldwork</td>
<td>Turn-takings, interruptions, silence, communicative competence, discourse markers, etc</td>
<td></td>
</tr>
<tr>
<td>Ethnography of Communication</td>
<td>Sociological and Linguistic</td>
<td>Ethnographic fieldwork</td>
<td>Conversation, communicative competence, writing systems, insults, irony, cross-cultural communication, etc.</td>
<td></td>
</tr>
<tr>
<td>Language and Gender</td>
<td>Sociological and Qualitative</td>
<td>Survey fieldwork</td>
<td>Androcentrism, sexism, language and equality, etc.</td>
<td></td>
</tr>
<tr>
<td>Dialectology</td>
<td>Linguistic</td>
<td>Survey fieldwork</td>
<td>NORMs, idiolects, dialect boundaries, dialect areas, pure dialects, isoglosses, bundles of isoglosses, transition areas, focal areas, relic areas, wedges, dialect continua, Neogrammarian Hypothesis, Family-tree Theory, Wave Theory, dialectometry, etc.</td>
<td></td>
</tr>
</tbody>
</table>
Research methods in Sociolinguistics

Table 1. (continued)

<table>
<thead>
<tr>
<th>Discipline</th>
<th>Objectives</th>
<th>Methodology</th>
<th>Method/Technique/Tool/Data</th>
<th>Domains</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variationist</td>
<td>Linguistic</td>
<td>Quantitative</td>
<td>Survey fieldwork Ethnographic fieldwork Sampling Natural Speech Recorded interview Questionnaires</td>
<td>Langue/parole, aggrammaticality of speech, observer’s paradox, correlational linguistics, linguistic variable, socio-demographic variable, context variable, statistical validity, representativeness, socially conditioned variables, longitudinal research, cross-sectional research, etc.</td>
</tr>
<tr>
<td>Sociolinguistics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Geolinguistics</td>
<td>Linguistic</td>
<td>Quantitative</td>
<td>Survey fieldwork Gravity models</td>
<td>Linguistic innovations, patterns of diffusion, gravity models, neighbourhood effect, etc.</td>
</tr>
<tr>
<td>Historical</td>
<td>Linguistic</td>
<td>Quantitative</td>
<td>Linguistic corpora</td>
<td>Uniformitarian Principle, Historical Paradox, bad data problem, etc.</td>
</tr>
<tr>
<td>Sociolinguistics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Creole</td>
<td>Linguistic</td>
<td>Quantitative</td>
<td>Survey fieldwork Ethnographic fieldwork Sampling Natural Speech Linguistic corpora</td>
<td>Pidgins, creoles, pidginisation, creolisation, contact varieties, etc.</td>
</tr>
<tr>
<td>Sociolinguistics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Atlas Lingüístic de Catalunya), the United States and Canada (The Linguistic Atlas of the United States and Canada), England (Survey of English Dialects) and Spain (Atlas Lingüístico de la Península Ibérica). More recently, telephone surveys have been used in studies such as Labov, Ash & Boberg (2006).

Ethnographically informed fieldwork, as Tagliamonte (2006: 20) points out, “requires that the analyst engage in research in situ” and it is qualitative: “the intensive involvement of the researcher in a given social setting in order to describe and identify, through the use of a variety of complementary research techniques, the cultural patterns and regularities that structure and perpetuate a society” (Poplack 1979: 60). The benefits of ethnographic research, therefore, according to Wolfram (2011: 307), “include the quality and amount of data and the familiarity with community practices that allow researchers to uncover the essential social and cultural factors that may affect language variation and change from the standpoint of the community itself” (see also Johnstone 2000a; Milroy & Gordon 2003). The most usual technique in ethnographic research to explore the dynamics of the speech community is participant observation and long-term involvement in a community (cf. Eckert 1989; Labov 1973; Labov et al. 1968; Milroy 1980; Wolfram & Schilling-Estes 1995).
3. Applied Sociolinguistics

Sociolinguistics has been one the most applied branches of Linguistics since its initial conception. As stated above, the development of Sociology with the emphasis on poverty and social disadvantage as political issues in western industrialized countries was an essential motivation among the origins of Sociolinguistics (see Shuy 1984). Sociolinguistics was thus foundationally sensitive to the real problems of society and their solution with the application of linguistic expertise:

In cases where real human problems are involved, theory cannot be abstract or divorced from application. If the weakness of recent applied linguistics has been its separation from linguistic theory, so has the weakness of recent theoretical linguistics been its separation from real human problems. Adequate engagement in real human problems requires the selection and development of both theory and application. Application without theory is mere methodology (an error much of applied linguistics has fallen into). Theory without application is mere speculation (an error which much of theoretical linguistics has fallen into). Theory or application absent from a real human problem is mere academic display. Life is in the doing. Faith without works is dead. Works without causes or reasons for working are frivolous.

Roger Shuy (1984: 102)

In fact, once the field was able to develop empirically-driven theory, sociolinguists such as Trudgill (1984a, 1984b) found perfectly legitimate to distinguish between Theoretical Sociolinguistics and Applied Sociolinguistics. The former is concerned with pursuing greater understanding of the nature of language — and even of society — whereas the latter would be concerned with the application of sociolinguistic theoretical findings to the solution of real-world problems.

The claims for the practical value, the social applications, of sociolinguistic research are motivated by two Labovian principles, namely The Principle of Error Correction, which postulates that “[a] scientist who becomes aware of a widespread idea or social practice with important consequences that is invalidated by his own data is obligated to bring this error to the attention of the widest possible audience” (Labov 1982b: 172); and The Principle of the Debt Incurred, according to which “[a]n investigator who has linguistic data from members of a speech community has an obligation to make knowledge of that data available to the community, when it has need of it” (Labov 1982b: 173).

Some scholars (see Cameron, Frazer, Harvey, Rampton, & Richardson 1992, 1997) argue that researchers should proactively seek ways of benefiting the community through one’s linguistic knowledge — rather than wait until a community asks for help —, which, according to Schilling-Estes (2007: 188), is also captured in Wolfram’s Principle of Linguistic Gratuity: “[i]nvestigators who have obtained
linguistic data from members of a speech community should actively pursue positive ways in which they can return linguistic favors to the community” (Wolfram 1993: 227).

In this way, in institutions and organisations outside academia variationist sociolinguists have drawn on their expertise in dialects to establish the geographical area or socio-demographic status a particular speaker comes from to investigate white-collar crime in the form of Forensic Sociolinguistics. Discourse analysts analyse interview data or other forms of evidence (texts) to help evaluate the validity of motives attributed to a speaker: syntactic analysis for untangling lease agreements that lead to litigation; stylistic analysis of written texts to determine clues to authorship, in copyright infringement, and in libel cases. In Medicine, Sociolinguistics has been helpful in therapeutic discourse and doctor-patient communication and healthcare. In Business, for intercultural communication in the world of commerce, the language of advertising and mass media communication; as well as in Education, Government, or Social Justice (cf. Hernández-Campoy 1996; Shuy 1984). This applied-oriented motivation has resulted in several collected works (cf. Ammon et al 2006; Boxer 2002; Trudgill 1984a; Wodak, Johnstone & Kerswill 2011), and in studies dealing with language and law (Conley & O’Barr 1998; Danet 1980; Finegan 1997; Gibbons 1994; Givon 2003; Levi 1982; O’Barr 1982; Shuy 1993), medicine (Curran, McGarry & Petty 1982), education (McKay & Hornberger 1996; Preston 1989; Stubbs 1980; Verhoeven 1997; Wolfson & Judd 1983), or business (Coulter & Coulter 2010; Delin 2005; Vestergaard & Schröder 1985).

4. Trends and future prospects in Sociolinguistics

The development of Sociolinguistics has been striking since its beginning in the 1950s, with a steady growth in both theoretical and methodological grounds as well as in its disciplinary directions within the study of language and society. Field methods in sociolinguistic studies have been conditioned by the research objectives pursued, namely: sociological, sociolinguistic, or linguistic goals. Data-gathering through observation, in the form of extensive interviewing, participant observation, questionnaire surveys, systematic note taking and/or record keeping has always been the primary fieldwork methods for the study, analysis and description of individual socio-cultural and linguistic systems in ethnographic research (Apte 2001). Although the methodology employed in studies whose objectives are both sociological and linguistic has usually been a qualitative one (Apte 2001: 775), quantitative approaches and quantitative analyses have also been employed more recently.
The methodological and analytical contribution of Linguistic Ethnography to the study of language and social life, as well as of language as social practice, is overtly perceptible, providing Linguistics with an eclectic interpretative stance of theoretical and methodological diversity reflected in the Ethnography of Communication and Anthropological Linguistics (see Creese 2010). Its work has managed to be empirical, offering an objective analysis of subjective meaning, without been positivistic; in fact, unlike Variationist Sociolinguistics, it does not adopt any deterministic perspective on data. As Baxter (2010: 119) states, different discourse-analytic approaches have produced varying sets of accounts of the same data, from most social to most linguistic: micro-analytical and macro-analytical approaches to generate interpretations.

The findings in Language and Gender research have been demonstrated to be socially applicable to eliminate sexual inequality and linguistic as well as social discrimination with its focus on the analysis of sex-differentiated varieties of language (sex differences in the use of particular linguistic features), the way that language reflects and helps maintain social attitudes towards women and men, sex differences in the use of conversational strategies, and sexism in language (see Cheshire & Trudgill 1998; or Holmes & Meyerhoff 2009). In studies whose objectives are entirely linguistic, Variationist Sociolinguistics has been the most widely known and prolific sociolinguistic area since the beginning, in the 1960s, of the systematic study of language in its social context quantitatively (e.g. Labov 1966, 1994, 2001, 2010; Trudgill 1974a; Chambers & Trudgill 1980; Eckert 1998; Chambers, Trudgill & Schilling-Estes 2002; Milroy & Gordon 2003; Tagliamonte 2006, 2012; Bayley & Lucas 2007; or Schilling 2013). Conceiving variation in language as socially conditioned and making use of the methods and findings from the social sciences (see Bainbridge 2001; Babbie 2001b), the aim of the sociolinguist is: i) to get a representative full picture, cross-sectionally or/and longitudinally (see G. Sankoff 2006), of the local speech variety in the population of a urban community by selecting informants randomly; and ii) correlate data obtained, pre-determined features (linguistic variables), with socio-demographic and biological parameters in the search of sociolinguistic variables and patterns of sociolinguistic behaviour. Speakers are considered as co-members of a ‘speech community’ — a community of speakers who share the same verbal repertoire (language use, style-shifting, etc.), and who also share the same norms for sociolinguistic behaviour (communicative competence, etc). The deterministic nature of speech behaviour assumes some mechanistic patterns based on the language of mathematics and logic: (i) variables and variable rules carry complex indexical meanings in the macro-sociological matrix after the empirical correlation of linguistic (dependent) and extralinguistic (independent) elements within some kind of speech community, or community of practice; (ii) dialect differentiation
is determined by the relative frequency with which particular variants are used in relation to their potential occurrence under the influence of prestige (overt/covert); and (iii) standardness is a function of those extralinguistic factors, with the vernacular having some special relevance; in this way, the use of non-standard variants correlates inversely with speakers’ socioeconomic status, gender, age, ethnicity, social networks, mobility, or level of speech formality. Empirical studies led by the pioneering work of Labov (1963, 1966, 1972) showed certain regular patterns of linguistic behaviour whose generalisations made in aggregate statistical terms raised them to the status of sociolinguistic universals, at least in the Western world. The consistency in these patterns provided irrefutable evidence against the traditional Bloomfieldian notion of free variation: free variation does not exist because linguistic variation is not free at all, but constrained by social and/or situational factors. Although research on language change nowadays is practically inconceivable without concurrent consideration of sociolinguistic variation, theoretical and procedural limitations, controversies and contradictions also exist even within the apparently empirically rigorous scientific field of Variationist Sociolinguistics, as evidenced, for example, by factors such as: (a) the unevenness with which the generalisability principle is applied or adhered to (Bailey & Tillery 2004); (b) questions surrounding age, time, and language change, including the limitations of apparent time studies of change and the difficulties of conducting studies in real time (Eckert 1997; Bayley 2002; G. Sankoff 2006); (c) ethical issues in sociolinguistic fieldwork (Labov 1982b, 1984; Murray & Murray 1996; Johnstone 2000a, 2001; Feagin 2002; or Milroy & Gordon 2003); (d) stylistic variation (Bell 1984, 1991; Coupland 2007; Eckert and Rickford 2001; Hernández-Campoy & Cutillas-Espinosa 2012; Johnstone 1996; or Schilling-Estes 2002); (e) differences in data gathering procedures and subsequent results, including differences in sampling procedures, interviewer characteristics, and elicitation strategies (such as formal elicitation frames, loosely structured interviews) (Bailey & Tillery 1999, 2004; Bailey, Wikle & Tillery 1997); and (f) issues involved in data processing and analysis, including how spoken data are transformed when transcribed into written form (Labov 1982a; Schneider 2002). Indeed, according to Bailey & Tillery (2004: 13), we cannot even know how generalisable (representative and reliable) our studies have been or may be without a body of research that examines the effects of methods on results.

When compared to the most rigorous empirical tradition characterising synchronic Variationist Sociolinguistics, Historical Sociolinguistics has sometimes been blamed for the lack of representativeness and its statistical validity has occasionally been questioned, which made Labov (1994: 11) assert that this discipline constitutes “the art of making the best use of bad data”. These methodological problems are due to the fact that the sociolinguistic study of historical language
forms must inevitably rely on linguistic records from previous periods — most of which will be incomplete or somewhat non-representative —, as well as on knowledge and understanding of past socio-cultural situations that can only be reconstructed — rather than directly observed or experienced by the researcher (Hernández-Campoy & Schilling 2012). Admittedly, it is true that this is an exercise of socio-historical reconstruction of the remote past time of a given language where the non-existence of evidence does not allow for conclusions about the non-existence of individual facts. Yet, as Schneider (2002) states, despite its limitations, Historical Sociolinguistics is not a second-best solution by inevitable necessity, but rather the best solution in those areas of study in which oral records are not available, especially when studying long-term developments of language variation and change. In addition, the historical paradox and the so-called uniformitarian principle have also been part of the controversial methodological issue: how different was the past from the present? To what extent can independent variables be reconstructed for the history of languages without running the risk of anachronism? Are sociolinguistic generalisations meaningful in terms of sociolinguistic universals? (see Hernández-Campoy & Conde-Silvestre 2012).

In overcoming these difficulties Historical Sociolinguistics has luckily been assisted by the parallel, astonishing development of other ancillary disciplines: Corpus Linguistics, on the one hand, and Social History, on the other. The links established with these two fields have provided the discipline with both ‘empirical ease’ and ‘historical confidence’ (see Hernández-Campoy & Conde-Silvestre 2012; Nevalainen & Raumolin-Brunberg 2003). It is well-known that developments in computing technology over the last twenty years have radically transformed linguistic research and that the compilation of large electronic corpora (both computer-driven and research-driven) has been instrumental in overcoming some of the problems inherent to working with ‘bad data’ from the past (see Schneider 2002; Cantos 2012). By allowing researchers to deal simultaneously with almost all the texts that have survived from a given period, Corpus Linguistics partly solves the fragmentary nature of historical material, ensures that variability in past stages can reliably be reconstructed and facilitates the selection of the variables whose analysis is worth undertaking. Similarly, the interest of social historians in the structure of groups and communities from the past, together with the reconstruction of demographic and socio-economic structures, has greatly benefited the task of historical sociolinguists, by allowing them to reconstruct — on the reliable evidence afforded by contemporary documents (thus non-anachronistically) — the socio-historical circumstances which could have affected linguistic processes in the past (see Conde-Silvestre 2007; Nevalainen & Raumolin-Brunberg 1996, 2003).

In Geolinguistics, the patterns of urbanisation along with industrialisation favoured the development of quantitative accounts for mobility and the

In data collecting, samples used from radio programme recordings have been shown to be useful sources for the study of stylistic variation in, for example, New Zealand (Bell 1982, 1984, 1991), Wales (Coupland 2001a) or Spain (Cutillas-Espinosa & Hernández-Campoy 2006, 2007; Hernández-Campoy & Cutillas-Espinosa 2010); and for the real-time measurement and analysis of linguistic variation and change (Hernández-Campoy & Jiménez-Cano 2003; van de Velde, van Hout & Gerritsen 1997; van de Velde, Gerritsen & van Hout 1996), although, as noted by Velde et al (1997), certain technical and methodological requirements are to be met.

The indexical nature of the social meaning of inter- (social) and intra- (stylistic) speaker variation in the sociolinguistic behaviour of speakers, according to Eckert (2012), has been chronologically approached from three analytic perspectives, waves or generations: First Order Variation Theory and Third Order Variation Theory. In this epistemological evolution since the beginning of sociolinguistics, there has clearly been a shift from deterministic and system-oriented approaches (language as a collective system: *langue*) to more social constructionist and speaker-oriented ones (language as individual performance: *parole*) (Hernández-Campoy & Cutillas-Espinosa 2012:7). During the 1960s, the mechanistically-based paradigm of the First Generation assumed that speech and the stylistic repertoire are ‘determined’ by the major macro-sociological categories of socioeconomic class, gender, age and ethnicity, providing us with general patterns in their aggregate data. In the 1980s, the ethnographic-based paradigm of the Second Generation set out that speech and the stylistic repertoire are ‘determined’ by social configurations — rather than categories — of multiplex relationships within the social networks of speakers and their mobility, providing us with a more concrete local (or locally-defined) perspective on the dynamics of variation and sociolinguistic behaviour at large. If the speech community was a significant element in the First Wave sociolinguistic scenario, in the Second, the community of practice is crucial. More recently, at the beginning of the 21st century, a Third Generation of sociolinguists have stressed the individuality of speakers by making use of a constructionist approach based on speaker’s agency (individual action), stance and performativity to more accurately account for the nature of the indexical relations between linguistic and extralinguistic variables. This new approach is cognizant of the fact that language acts *are* acts of identity. Therefore,
the uses of variation are now understood not simply as reflecting, but also as constructing social meaning, and the focus has shifted from speaker categories and configurations to the construction of *persona*. Not only does variation reflect the multifaceted nature of human relationships for the transmission of social meaning, but it is also a resource for identity construction and representation, and even social positioning; and accents, dialects and their styling are markers of this social meaning (Auer 2007; Podesva 2006). Like any other social stereotypes, these different ways of speaking constitute prototype categories within a wider frame that comprises not only ideological components, but also markers from a wide variety of dimensions, such as speech, physical appearance, dressing, dance, music, etc. (Kristiansen 2008: 72–73). Linguistic variation is then the instrument, or resource, for linguistic performance, rhetorical stance and identity projection, where individuals (rather than groups) and the individual voice are actively responsible for the transmission of sociolinguistic meaning in terms of a speaker’s personal and interpersonal social identity and authenticity (Giddens 1991; Johnstone 2000b).

Since the beginning of variationist studies, the ‘authentic’ speaker was a methodological condition of sociolinguistic research design in Labov’s (1972, 2001) model of determinist-based linguistic variation when observing the unselfconscious, everyday speech produced by spontaneous speakers of pure vernacular: “language produced in authentic contexts by authentic speakers” (Bucholtz 2003: 398). This was the influence from 19th century dialectological and anthropological assumptions based on romantic Philology and Folklore. For their investigations on the pure, genuine, real dialects, dialectologists had to use old male speakers living in small isolated villages as consultants in their fieldwork (NORMS, see Chambers & Trudgill 1980). Authenticity was a synonym for ‘prototypical’ (average speaker) and related to positivist ‘universals’. More recently, the ‘authentic’ speaker has become a phenomenological and theoretically paradigmatic model in social constructionist-based linguistic variation referring to a differentiating verbal positioning in society imbued with social meaning within an implicit theory of identity (Bucholtz 2003; Coupland 2003, 2007; 2010; Eckert 2003). If Labovian sociolinguistics focused on the average linguistic behaviour of the group (the statistical mean), the interest of constructionist sociolinguistics is the singularity or peculiarity of a particular speaker (the statistical deviation from the mean), with its own sociolinguistic indexicality as an *authenticity indexing*. The authentic speaker appears now as an unexpected (non-idiosyncratic) identity assumed in verbal practice creatively. Individuals’ multiplicity of social networks in post-modern society are now allowing them to develop a polyhedral image shaping and multi-faceted creative behaviour, exhibiting and aligning with different social identities for different purposes at different times and places and in different contexts of social relations and interaction (see Hernández-Campoy 2015).
Yet the unpredictable nature of informants under the socio-constructionist view raises methodological problem of observability, interpretation and reliability. Given the difficulty of observation and the unpredictable nature of speaker’s motivation in initiative speech production, interpretation may not be unique, which means that there is a validity problem, i.e., how to refute (or support) the interpretations of other analysts. As a result, reductionist generalisation from the motivations of one speaker, with a complex range of roles and conditions, to the behaviour of the whole group or to a larger community is not empirically reliable. These inconveniences are favouring the combination of qualitative and quantitative approaches in traditionally quantitative-based sub-fields such as Variationist Sociolinguistics and more recently in Historical Sociolinguistics (Hernández-Campoy & Conde-Silvestre 2015). Methodologically, it has been claimed (Cutillas-Espinosa & Hernández-Campoy 2007; Hernández-Campoy & Cutillas-Espinosa 2013), that research in this field may be enriched by interviewing the speaker in order to get a confirmation of researchers’ interpretations, as practised by the Ethnography of Speaking, the Social Psychology of Language or Folk Linguistics. Without this kind of data, the sociolinguist’s diagnostic remarks run the risk of being excessively judgmental or even subjective.

Conclusion

Sociolinguistics is therefore in a continuous process of theoretical reformulation and methodological redefinition in consonance with the epistemological evolution and the development of new fieldwork methods, data collection techniques and — in the case of quantitative approaches — statistical analyses. As Robins (1964: 319) visionarily predicted, “[t]he languages of mankind in all their fascinating detail and with all their immense power among the human faculties still present a potentially limitless field for disciplined investigation and systematic study”. New lines of inquiry are being opened up and new methods are being devised since linguistic theory must always keep pace with methodological and technological progress.

Entering the local community has always meant a challenge for the sociolinguistic fieldworker since different social-cultural situations, cities, locations in a country (urban vs rural), have conditioned the design and planning of the research (see or Feagin 2002; Schilling-Estes 2007, 2013; Wolfram 2011). As Tagliamonte (2006: 34) states, “[f]ieldwork practices from the 1960s and 1970s may not be as effective in the 2000s”. The shift from the average speaker — where the language is a collective system and authenticity is understood as ‘natural’ — to the singular speaker — where the language is individual performance and authenticity is
conceived as ‘difference’ — is undoubtedly contradicting the principles of representativeness and generalisability in the most orthodox quantitative sociolinguistic research. But, in turn, this tendency is making both quantitative and qualitative approaches, as well as sociological and linguistic objectives, converge complementarily in those areas where traditionally different perspectives of the same phenomenon are becoming feasible, significant, and revealing for language in society.

Together with this inherently dynamic activity, the integrative stance of the field, fostering its vast multidisciplinary and genuine interdisciplinary nature, has been crucial for its scholarly interest and scientific success in hermeneutic accounts of language in society, as its diverse array of high-quality international journals (see Coupland 2001b) and long-standing tradition of well-attended debate forums show. In fact, Sociolinguistics emerged with the single as well as crucial assumption of the eminently social nature of language, and, since then, it has been consolidating its foundations thanks to the interdisciplinarity practised so far. This integration based on paradigmatic complementarity has been providing us with a greater refinement and precision in analysis. It has also meant a step forward in the improvement of sociolinguistic theory, and, ultimately, of our understanding of the nature and functioning of language as a human faculty; and, crucially, its application to solve real human problems of society.

References


DOI: 10.1017/S0047404450000075


DOI: 10.1017/CBO9780511619496.010


DOI: 10.1017/CBO9780511980541


DOI: 10.1017/CBO9780511801624


**Author’s address**

Prof. Juan Manuel Hernández-Campoy
Departamento de Filología Inglesa
Facultad de Letras
Universidad de Murcia
E-30071 Murcia
Spain

jmcampoy@um.es