Linguistic coherence; systems, repertoires and speech communities.

Introduction to *Lingua* Special Issue on Coherence, covariation and bricolage

1. Terminology

In every linguistic system there are some elements, structures and processes which vary, i.e. cases of *intra-systemic variation*. Each system also contains, clearly, other components that do not vary; as Berruto observes "in language not everything is variable – on the contrary, the system has a large core which is not subject to variation" (Berruto 1987:264, our translation).

Similarly, language varieties – dialects, ethnolects, etc. – always share some features with related varieties, while other features are unique to a variety, yielding *inter-systemic* variation. Both the features which are and those which are not unique for a given variety can be either categorical or (intra-systemically) variable. For the variable ones, the quantitative patterns (say, the probability of occurrence of the dialect variant in the relevant linguistic and extralinguistic conditions) may or may not be unique to that dialect.

When specific linguistic systems incorporate multiple variable phenomena – *linguistic variables* – questions arise as to the ways in which they interact or cohere: Do variants of different linguistic variables tend to co-occur, or do they rather tend to exclude each other, or do they display independent and unrelated patterns of occurrence? To the extent that linguistic variables systematically *co-vary*, they can be characterized as displaying *coherence*. In connection with language variation, coherence concerns to what extent multiple co-existing linguistic variables have similar distributions, both internally and in the speech community at large. At the level of the speech community the question is whether (socially) similar speakers have similar patterns of usage of the variables – are the sociolinguistic ‘isoglosses’ more or less coincident, as it were. Coherence thus concerns both the language and the speech community. Le Page & Tabouret-Keller (1985) distinguish communities with focused and communities with diffuse linguistic norms; the latter characterize unstable and highly dynamic sociolinguistic situations, such as in cases of verbal repertoires around a creole language. This distinction seems partly parallel to more versus less coherent linguistic systems.

The concept of coherent linguistic systems is problematized by the body of work on language and identity that argues that speakers actively and idiosyncratically select from a palette of variants available in their communities of practice to construct identities, stances, and styles – a process that Eckert (2008) characterizes as ‘*bricolage*’. The concept of more or less coherent linguistic systems has also been questioned by scholars such as Geeraerts (2010), who doubts that, in a situation with multiple variable phenomena, each and every one with its specific own distribution, there can still be something like a language system. Of course, given that this is in a sense true of every

living language, one could argue that *a fortiori* there has to be a unifying, albeit maybe slightly more abstract overall system.

What does coherent variation, or the lack thereof, imply about the organization of the linguistic system and, more generally, of linguistic competence including wider conceptions of competence (cf. Hymes' 1966 '*communicative competence*')? What does the system, and competence, and the speech community consist of, if coherence in this sense is absent – if each variable has its own distribution and each speaker makes idiosyncratic selections among available variants?

These issues have implications for the position of the study of language variation and change with respect to both other areas of linguistics (including theories of grammar). and to the broader social sciences, including sociology, ethnography and anthropology. Even more generally, these issues have implications for the place of the study of language and linguistic dynamics in the area of tension between nomothetic and idiographic approaches to research and explanation (cf. section 2 below).

On a more specific level, the answers to these questions and their ramifications will be interpreted primarily –though not exclusively– in connection with an emerging debate which spans several, relatively distinct areas of modern sociolinguistics and dialectology. Are the varieties of language that both linguists and laymen give names to (languages, dialects, speech styles, standard and nonstandard varieties, ethnolects, etc.) coherent objects, or diffuse abstractions? They are typically characterized in terms of clusters of linguistic elements: entire grammars and lexicons in the case of languages and dialects, or sets of linguistic variables in the case of sociolects, ethnolects and speech styles. Do such sets in fact cluster and co-occur, and if they do not, in what sense do the specific varieties exist?

2. Conceptual and theoretical backdrop

The members of every community of speakers share a great number of linguistic characteristics – the elements of their phonology, syntax, and lexicon as well as the organization of these elements. Concurrently and collectively, these characteristics are seen as defining the identity of the language or dialect of the community. But at the same time, it is clear that individuals show considerable diversity: every speaker is unique in their linguistic experience, and in at least some details of their usage, and every individual exercises a great deal of choice in the linguistic performance of speech style, identity, and stance. This contrast – between the mutual and the idiosyncratic aspects of language – presents challenges for the speakers, and for the linguists who study them. For the speakers, the issues are how to communicate effectively, how to understand and be understood, how to adapt to different circumstances or audiences, in a universe in which other people's usage, practice, and intentions may differ from one's own. The issues that the linguists face reflect and elaborate on the speakers' challenges. A central problem is the factual one: how much and in what respects the usages of different speakers in a community are similar or different?

Clearly many components of a language, dialect, or variety are essentially invariant across the community of speakers. But studies of the variable elements,

especially those that are understood in a community to differentiate styles and groups of speakers, or to index social meanings, have led scholars to distinct, and sometimes conflicting, understandings of how sociolinguistic diversity works. One view emphasizes 'orderly heterogeneity', which is neatly indicated in Labov's pioneering demonstration of the systematic usage of sociolinguistic variables in New York City: a number of different varying elements – coda /r/ deletion, plosive variants of interdental fricatives, raised nuclei of the vowels in BOUGHT and BAD – all showed simultaneous social and stylistic stratification, suggesting that they all indexed levels of formality and status. Labov argued that it was the appropriate use of these variables and the knowledge of these social meanings that defined membership in the New York City English speech community.

This vision of orderly diversity implies that speech communities are sociolinguistically coherent, in the following sense: the orderly variables that define the community should collectively behave in parallel: variants (or rates of use of variants) that index a given style, status, or a social characteristic should co-occur. Coherent middle class speakers would use all the variants associated with their status, and speakers who are coherently signaling a 'casual' style would use all the 'casual' variants. The collocations of variants of different linguistic variables are social conventions, but they may be at least partly internally (structurally) motivated.

The multiple social dimensions of diversity are, of course, a further complication of the question of community coherence. Speech communities -- and the language varieties that they encompass, define, or are defined by -- are internally diverse in a number of ways. Speakers are differentiated by gender and age, by education, status or class, by ethnicity, by geography (neighborhood, place, rural vs. urban, etc.), by 'localness' (long established vs. new arrivals), and so on. Linguistic variables may have indexical associations, of various orders, with any of these social characteristics. Individuals may command an assortment or range of varieties or registers that they use in different contexts or for different social purposes. Variants or varieties may be the object of different kinds of evaluation: overt and covert prestige, stigmatization, national or international 'standards'. The question of coherence is thus repeated along these various dimensions of diversity. The more coherent a set of coexisting linguistic variables, the bigger the chances that a change in the variant of one of the variables will trigger a switch to another variety – like a falling domino can make a row of neighboring dominoes fall. See also Auer (1999).

Another perspective on linguistic community and diversity emphasizes individual agency and purpose. In this point of view, every speaker in every social interaction is inevitably revealing, or deliberately expressing, something about who they are and what they are trying to do – constructing an identity and conveying a purpose, stance, or intent. The tools at their disposal for doing so include linguistic variables that have socially meaningful interpretations or indexicalities, i.e. socially symbolic meanings. Given that indexicalities, identities, and purposes are all profoundly diverse, the particular mix of variants used by a particular speaker on a particular occasion will be drawn from a selection that is essentially infinite. Hence there is no reason to expect, *a priori*, that such mixes will be coherent in the sense defined above. Each variant a speaker uses in an interaction could potentially display a distinct indexical intent, and any given

combination is not coherent or incoherent but rather expressive and nuanced. This view is closely related to the concept of style proposed by the California Style Collective (1993:14), who defined style as “a clustering of linguistic resources and an association of that clustering with an identifiable aspect of social practice”. Eckert (2004) characterizes the speaker’s choices of variants as a practice of *bricolage* – drawing from “a range of existing resources to construct new meanings or new twists on old meanings”. In this view collocations of variants of different linguistic variables are ‘constructions’ which can be more or less temporary and situation-specific.

The distinction between coherence and covariation on the one hand and bricolage on the other is paralleled by several related dichotomies, the first one being a leading comprehensive insight from the philosophy of science, viz. Dilthey’s (1883) conception of the broad methodological distinction between *Naturwissenschaften* and *Geisteswissenschaften*, i.e. science (*sensu stricto*) and humanities. In Dilthey’s view, which has become an *idée reçue*, science in the narrow sense aims at *erklären*, ‘explaining’, with causality and probability as central concepts, whereas the humanities target at *verstehen*, understanding, with meaningfulness as one of the central concepts.¹ This general distinction is related to the overall watershed between nomothetic and idiographic approaches: while the former focus on recurring phenomena, inducing the regularities, laws or principles they are subject to, the latter deal with unique phenomena and their specific properties.

In late 19th century linguistics the dichotomy was manifest in the debates between the Neogrammarians (with their central claim of grammatically blind and lexically exceptionless sound laws, the effects of which can be obscured by analogy or borrowing) on the one hand, and Schuchardt and mainstream dialectology, on the other; the latter explicitly took idiosyncrasies of individual lexical items into account – cf. Jaberg’s (1908:6) dictum ‘in reality each word has its own particular history’.

Interestingly, modern linguistics unites subdisciplines, theories and methods of both orientations. In syntax, for example there is a contrast between formal approaches that aspire to universal status, such as generative theory (à la Chomsky, e.g., 2006), and approaches that begin from the assumption that syntax is rooted in meaning and social interaction, which are often culturally-specific (cf. “Language as social semiotic”, Halliday 1978). Both the sociolinguistically oriented approaches to processes of linguistic innovation and change (which, unlike traditional dialectology, have a synchronic orientation), and the formal theoretical approaches seek to apply scientific methods, the difference being that the former seeks probabilistic explanation and the latter categorical ones.

Contemporary studies of ethnolectal variation show yet another facet of this dichotomy, namely the distinction between the language-centered and the ethnographic approach. Whereas the ethnographic approach conceives language systems as infinite resources from which speakers may freely choose to shape their identity, the language centered approach tries to disentangle the ‘rules’, generalisations and restrictions on these resources.

¹ The interpretation of people’s behavior in their own terms is what the anthropologist Clifford Geertz (1973), following Gilbert Ryle, has labelled ‘thick description’.

Is sociolinguistics thus developing into a new battlefield of an old intellectual split? With regards to the dichotomy between coherence and covariation on the one hand and bricolage on the other, we would like to stress that the two approaches are not equally strong in every respect. Thus traditional quantitative sociolinguistics is not particularly illuminating about language contact, dialect borrowing and the appropriation of variants across related dialects or about the local discursive significance of a variant choice – the latter two of which play an important role in the performance of identity through bricolage. Bricolage, on the other hand, is linguistically a superficial mechanism: it can't go beyond phenomena which are subject to conscious manipulation. This would exclude many late phonological and phonetic processes which have been fruitfully investigated in quantitative variationist research. One would not typically find bricolage involving the manipulation of, say, postlexical assimilation, or of variable morphosyntactic phenomena related to the internal structure of NP.

Perhaps, therefore, the two perspectives are essentially complementary? Bod (2013) argues that drawing a distinction between the sciences, in their quest for universal laws, and the humanities, zooming in on unique events and hermeneutic approaches, is shortsighted, as it fails to appreciate the fact that searching for patterns is what unites both traditions. So in the case at hand, what type of synergy might genuine collaboration lead to?

3. The main questions

Although the visions of community underlying identity construction/bricolage and orderly heterogeneity may have different implications for concepts of coherence, it is not clear that they are incompatible. To successfully communicate the “new meanings” Eckert refers to, speakers depend on a shared understanding with their interlocutors of the established or potential indexicalities of the “existing resources” they are utilizing. As Eckert (2004) notes, “since a stylistic move is to be put out into a community for the purpose of being interpreted, speakers select resources on the basis of their potential comprehensibility in that community”. Such shared understanding is a central element of the Labovian definition of orderly heterogeneity. But exploring the compatibility or conflict between these visions will depend crucially on a deeper insight into the facts. Empirically, the issue can be illuminated by examining the extent to which clusters of variants co-occur in speech (cf. Guy 2013). If varieties are coherent, the variables associated with them should co-vary in the usage of individuals. If co-variation is absent, the concept of the language variety may need redefinition.

The questions that arise in this regard can be summarized as follows:

1. Which features correlate and which do not? To what extent, and in what ways, do the characteristic variables associated with a dialect or a speech community co-vary? Which covarying linguistic features/domains are involved in change in progress and which ones tend to be constant?

2. To what extent and in what ways do speakers select and combine variants in usage? How freely can they select and combine elements from the “range of existing resources” in their usage?

3. How similar or different are the indexicalities of particular linguistic forms? Are there clusters of variables that coherently index, or are associated with, the dimensions or subdivisions of a community discussed above? What role does the indexicality of the features play in existence or absence of clusterings?

4. By what linguistic means do speakers reflect or perform their sociolinguistic identities on all these dimensions?

5. Are there socially identifiable leaders of change who tend to use all the innovative variants together, or are different innovations subject to differentiated social interpretations and individuated patterns of usage? When change is underway, how much do members of a speech community really share, and what does this mean for the very definition of a speech community, a dialect, or a language? At which point do we want to talk about a separate variety?

(This touches on the issue of continuity versus discreteness – which can paradoxically apply to the same situation: for the linguist smooth transitions may exist between language varieties, but the speakers may perceive discrete boundaries between groups; cf. Herson Finn 1996).

6. When correlated usages are encountered, are they better understood as indicators of the social coherence of a variety, or as (perhaps inevitable) consequences of structural factors or historical factors?

7. Are some kinds of language varieties (e.g., local dialects) more coherent than others (e.g. speech styles)? If so, why? What roles do perception and categorical perception play?

One important caveat in the search for covariation is that certain *internally* generated linguistic changes may correlate because of structural relations between them. Thus multiple syntactic changes are sometimes argued to be triggered by a single parametric change; vocalic chain shifts alter several vowels at once. Numerous processes of vocalic chain shifting have been discussed in the literature (in, e.g., chain shifting in North American English and Austrian German with respect to both dialectal and social differentiation; a contemporary Dutch case involves diphthongization of tense mid vowels in spoken standard Dutch and the lowering of the prominent first element of the diphthongs /ei, œy, au/).² All of these are diachronically unidirectional and coherent, but exactly how coherent is the relevant part of the segment inventory in any single synchronic phase in the process (cf. Gordon 2002: 257-264)?

² In all three cases there are considerable cross-dialectal differences; cf. Labov 2010, Moosmüller & Scheutz 2013, and Jacobi 2009, respectively.

Is there a parallelism between Neogrammarian versus lexically diffuse sound change on the one hand and coherent versus a bricolage-type process on the other? In the same line: do phonetic and postlexical processes maintain different patterns of variation and cohesion than sound changes which have been incorporated in the morphology (so-called lexical rules such as allomorphy or supra-segmental phenomena such trisyllabic shortening in English) or in the lexicon (lexicalized, historical sound change)?

Bricolage in Eckert's sense can involve the incorporation of elements *from other varieties*³ or languages. The implication is that social and stylistic usage is not organized in discrete coherent varieties, but is fluidly recast by individuals adapting to changing circumstances and self-expressions. Compare, for example, Benor's (2010: 160) concept of ethnolinguistic repertoire "as a fluid set of linguistic resources that members of an ethnic group may use variably as they index their ethnic identities". The picture is complicated by the co-existence of different sources of variation: multiple ethnolectal features may have common origins in substrate or L2 acquisition effects. Are 'borrowed' innovations which add new variants to existing linguistic features (as explored in e.g. Hinskens 2011) more likely to be used in idiosyncratic ways?

A special test of these issues occurs in those situations that are described as involving lectal continua, where intermediate varieties emerge (e.g. between standard and nonstandard, or regional and supraregional dialects, cf. Cornips 2006). In dialect/standard continua more or less separate intermediate phases, such as koinés or regional varieties of the standard variety, evolve. They result from the transition from a diglossic situation (in which the national standard language, serving as the H code, and traditional dialects, serving as L codes, were kept neatly apart) into a situation with a more fluid repertoire, labelled 'diaglossia' by Bellmann (1998). According to Hoppenbrouwers, dialect/standard continua of the type that he has labelled 'regiolects', are "intermediary systems which no longer constitute an autonomous language system" (1983:1, our translation). The question of how stable diglossic intermediate varieties can become has hardly been addressed, just as little as the question how coherent their constituent dialect features can become. Work by Auer (1997) argues that such settings show implicational relationships between certain types of variables, motivated by structural relations among the variants, but other scholars have suggested that bricolage occurs (Eckert 2012). In strongly levelled old dialects as well as in young dialects, where most features are still in flux and structural relations between variable phenomena have not yet fully crystallized, the strict co-occurrence (cf. Möller 2013), conjunction or disjunction between variants is probably rare.

In young ethnolectal varieties, linguistically the situation is probably comparable to that of young dialects, which are still in the course of developing and which are thus very much in motion. Like young dialects, emerging ethnolects either develop in or create sociolinguistically diffuse situations. More empirical investigation is warranted: are the relevant variables correlated in such settings (cf. Hinskens 2007) or involved in statistical implication relations (Lameli 2004)? Are intermediate varieties constructed with arbitrary collections of variants (standard or nonstandard, local, regional or national), or do they show systematic patterns of co-occurrence of variables?

³ Analogous to Bloomfield's (1933) 'dialect borrowing'.

4. Some relevant older research and the studies in this Special Issue

The question of whether and how in a given speech community (or in a coherent sector of a speech community) multiple variable phenomena are interrelated has received relatively little attention until recently.. No doubt this was partly a practical matter: the affordable hardware and user-friendly software that facilitate the quantitative analyses required for studying such relationships have only become widely available in the last decade or so. But another factor was the development of theoretical approaches in sociolinguistics to linguistic community and diversity and individual identity, which have brought the questions considered here to the forefront of contemporary concerns. Even so, several pioneering studies that addressed these matters merit our attention.

One of the first studies of coherence in a linguistically heterogeneous setting was Ma & Herasimchuk (1972), which looked at the use of 14 phonological variables (6 from English, 8 from Spanish) by bilingual Puerto Rican speakers of English and Spanish in New York City. Factor analyses identified five clusters of variables constituting bilingual speech styles such as 'standard conversational Spanish' and 'accented conversational English'. Q analyses of the speakers identified four speaker groups with distinctive style profiles. Remarkably, three of the five variable clusters consist of both English and Spanish variable phenomena.

In his study of language variation in the Burträsk area of northern Sweden, Thelander (1982) focussed on twelve linguistic variables, each with two variants labeled dialect and standard. Two distinct apparent time patterns emerged: decrease versus increase of the proportion of dialect variant. Thelander refers to the two sets as 'macro-variables'. The macro-variable which increases in apparent time includes dialect features with wide geographical dispersion in northern Sweden, whereas the decreasing macro-variable involves dialect features with a relatively small geographical spread. Measuring by a 'cohesion coefficient', Thelander shows that rates of use of the two macro-variables defined three language varieties: traditional dialect, regional standard language (which included only features of the macro-variable with a relatively wide areal spread) and standard Swedish. The community thus had a heterogeneous repertoire of varieties, but each variety was relatively coherent.

Horvath & Sankoff (1987) investigated the use of 20 vocalic variants among speakers of Sydney English. A principle components analysis identified clustered loadings of the variants that generated several dimensions of linguistic similarity or dispersion among the speakers. The first two components differentiated a discrete group of immigrant speakers from the rest of the sample. Additional dimensions were correlated with social status, but in a relatively continuous distribution rather than discretely segregated lects. Hence the non-immigrant community is fairly coherent in terms of evaluation and use of these vocalic variables. Broadly speaking, the variables covary; the variants that tend to be accorded higher status in the community co-occur in the usage of higher status individuals. But the central contribution of this work is to take the linguistic facts as the point of departure; the study investigates how speakers are

distributed in a space of linguistic similarity or difference, rather than how linguistic variables are distributed in a social space defined by class, ethnicity, etc.

In a study that stimulated some of the work in this Special Issue, Guy (2013) investigated intercorrelations among four binary variables in a group of working class speakers of Popular Brazilian Portuguese. Each of the variables opposed variants with standard and non-standard social evaluations. Several strong correlations between variables were found, along with some notable cases of non-correlation. Across the entire sample covariation occurred at a level considerably better than chance, but some speakers had distinctly idiosyncratic behavior. One variable showed a strong gender association, which meant that the correlations within the male and female samples were better than in the entire sample. This study also sought to distinguish socially motivated covariation from correlations that might be linguistically driven by structural relationships among variables. Overall, variables with potential structural connections correlated better than those without, but some such cases showed no significant correlation, and there were cases of robust covariation without structural motivation. This indicates that although linguistic structure may facilitate covariation, sociolinguistic coherence emerges independently of structural factors.

Further work on Brazilian Portuguese by Oushiro & Guy (2015) examines how coherence is affected by another structural factor – perceptual salience. In a set of 6 variables (including phonetic, phonological and morphosyntactic cases), these authors distinguished contexts where the variants were more saliently distinctive (e.g., occurring in stressed syllables, irregular forms, etc.). These salient subsets of the variables regularly showed weaker and less significant intercorrelations than the unsalient cases, which suggests that the individuated choice of variants postulated to occur in *bricolage* may focus on tokens that have perceptual prominence, while the general patterns of occurrence reflecting community norms govern the forms that speakers have less conscious awareness of.

The papers in this Special Issue seek to address the kinds of questions sketched in the preceding section. All the authors present new data on multiple variables in various *language communities*: on English and Spanish in New York City (Becker; Erker & Otheguy; Newlin-Lukowicz), on Danish spoken in four locations in Denmark (Gregersen & Phrao), on Dutch in both Belgium and the Netherlands (Grondelaers & van Hout; van Meel, Hinskens & van Hout), on Greek in Cyprus (Tsiplakou, Armostis & Evripidou), Portuguese in São Paulo (Oushiro), and German (Kiezdeutsch) in Berlin (Wiese & Rehbein).

The linguistic variables that are addressed span all *language components*. Among the phonetic and phonological variables investigated here are vowels and diphthongs in English, Portuguese, Dutch, and Danish, and consonant lenitions or changes in place, manner or voicing in Spanish, Portuguese, English, Dutch, and Greek. Morphosyntactic variables include agreement phenomena in Portuguese, Dutch, German; tense marking in Greek, pronominal expression or placement in Spanish and Greek, case marking in Dutch and German.

The papers explore several of the *extra-linguistic dimensions of diversity* we have mentioned. Thus several papers address the issue of ethnic diversity and ‘ethnolects’ – whether associated with long-standing ethnic diversity (Becker), or arising from recent

immigration, from Turkey and Morocco in the Netherlands (van Meel, Hinskens & van Hout), from Poland in New York (Newlin-Lukowicz), and from various countries in Berlin (Wiese & Rehbein). Other papers look at issues of dialect diversity and contact, between varieties of Spanish in New York (Erker & Otheguy), between Cypriot Greek and Modern Standard Greek in Cyprus (Tsiplakou, Armistis & Evripidou), and across Denmark (Gregersen & Pharao). Geography and place are important in Oushiro's study and also in Gregersen & Pharao's study of four locations in Denmark. The influence of a 'standard' variety is an issue for speakers of Belgian Dutch, Cypriot Greek, and Kiezdeutsch speakers in Germany. Local identity – whether a speaker is authentically 'from around here' – is at issue in the studies of São Paulo, New York, Cyprus, the Netherlands, and Germany. Social class and educational level is an important issue in Oushiro's study of São Paulo Portuguese. Becker's speakers deal with the stigmatization of New York City English vs. the greater prestige associated with other American English dialects.

Ongoing language change is a systematic disrupter of community coherence, as some speakers lead and others lag or do not participate in the change. Issues of change arise in most of these papers. Besides the changes arising from language contact that are discussed in several papers (Wiese & Rehbein; van Meel et al., Newlin-Lukowicz), others examine language-internal change, both from above and from below. One of Oushiro's variables in São Paulo is a phonetic innovation, which is socially distributed in a completely different pattern than the more established variables. Gregersen & Pharao's variables are reflexes of a 'change from above' in which a relatively uniform national Danish has supplanted historic dialects. New York English shows a change from above towards exogenous variants that is reflected in Becker's data.

Perception plays a major role in Grondelaers & van Hout's approach to a newly emerging Belgian regional standard variety of Dutch as well as in Gregersen & Pharao's discussion of the findings (little coherence on every level) of their study.

Finally, with respect to the *methods of analysis*, Pearson correlations (between continuous variables) play a role in several contributions. Moderate to strong use of correlations is made in the papers by Ecker & Otheguy; Gregersen & Pharao; Oushiro and Tsiplakou and co-workers. Multivariate analysis does not play a major role in this volume. Only Oushiro applies that technique with the aid of the Rbrul statistical package. Van Meel et al. begin their analyses with ANOVA, then go a few steps further with the aid of cluster analysis (which is not a correlational technique), bringing their results to bear on some of the questions raised above. Newlin-Lukowicz begins with a cluster analysis of the social data, and then analyzes the distribution of the linguistic variables across each social cluster. Becker deals with two continuous variables by comparing speaker means for each, but a different approach is required for a third variable (the tensing of /æ/), for which three discrete lects emerge, each tensing in a different set of phonological contexts. Here, Becker assigns each speaker to one of the tensing lects, and then calculates the mean use of the other variables for each lect. Finally, interesting insights are achieved using percentage tabulations in Wiese & Rehbein's paper on Kiezdeutsch, and even with very little quantitative treatment, as in the contribution by Grondelaers & Van Hout.

The studies in this Special Issue therefore provide a broad overview of community coherence and diversity in seven languages on three continents, and of the multifaceted problem of what is shared and what is individual in language use. In what follows, we give a brief introduction to each paper.

In her study of the Lower East Side, Becker revisits the site of Labov's work on New York City English (NYCE), some 50 years later, in a neighborhood which is now considerably more ethnically diverse than it was in 1963. She presents data on three of the variables Labov (1966) described: coda (r) and the two vocalic variables in BAD and BOUGHT. Her speaker sample is distributed across the dimensions of age, sex, class, and ethnicity. Her results show a marked withdrawal from the NYC variants that predominated in Labov's findings, and broad correlations in the patterns of use of the three variables. Thus speakers with more open BOUGHT vowels also produce coda /r/ at a high rate and diverge maximally from the traditional NYC pattern of tensing of the BAD vowel. But Becker also shows that many speakers are somewhat inconsistent in usage of the three variables, and that individuals vary their stylistic usage in ways that suggest purposeful *bricolage*, for example, by using more traditional variants when casting themselves as local Lower East Siders facing the influx of non-local residents as the neighborhood gentrifies.

Erker & Otheguy look at Spanish speakers in New York City, who make up fully a quarter of the city population, but come from many different countries speaking some quite different dialects. The city thus constitutes a melting pot of linguistic contact – among different dialects of Spanish and between Spanish and English. The study population includes speakers of different dialect backgrounds and contrasting lengths of residence in NYC (and therefore, levels of exposure to English and to other dialects). The variables investigated included measures of –s lenition, subject pronoun use and placement, subjunctive morphology, and copular constructions. The results show consistent evidence of convergence arising from contact; the community is effectively becoming more coherent. But the variables behave in different ways: in some cases convergence with English is implicated, but others suggest leveling of the different dialects of Spanish, and variables with the lowest social salience show the greatest convergence. Notably, the linguistic constraints on the variables are also affected; with longer exposure to contact, speakers reconfigure and attenuate the contextual conditions on variation. The authors argue that convergence advances cognitive economy by extending the range of applicability of linguistic generalizations.

On the basis of recent data from a sample of 187 speakers (balanced for generation, gender, and social class) from four sites in the extremely dialect-levelled speech community of Denmark, Gregersen & Pharao investigated whether systematic co-variation existed among three structurally unrelated variable phonetic phenomena. After they established that co-variation does not depend on site and that it does not hold in general for idiolects, the authors relate coherence as co-variation of features to the perception of variation in Danish lects. They argue that lects are perceived as such precisely because some features are foregrounded while others do not seem to attract much attention and are thus neglected.

Grondelaars & van Hout look at the emergence of Tussentaal, a variety of Belgian Dutch that diverges from the traditional standard language by incorporating nonstandard variants, but has a “dynamic prestige” driven by its use in popular culture and media,

indexing a “cool and streetwise” identity. Flemish evaluation of varieties of Dutch are further complicated by the exogenous source of the traditional standard, which was based on the Dutch of the Netherlands; hence certain ‘nonstandard’ features index local identity – Flemishness – as well as dynamism. The authors contend that perception and evaluation is central to the coherence of a language variety, proposing the concept of ‘perceptual harmony’ as the limiting factor on bricolage in performance.

Van Meel, Hinskens & van Hout present a study of patterns of co-variation of a range of linguistic variables in emerging Moroccan and Turkish varieties of Dutch spoken in the Dutch cities of Amsterdam and Nijmegen. They address the question whether the features (which have different origins) can be randomly mixed or, rather, show co-occurrence restrictions on their use. Are correlated usages better understood as consequences of internal factors, or as indicators of social (specifically ethnic) coherence? Although in the outcomes the structural relations are the first one to catch the eye, more refined analyses yield more complex patterns. Whereas one cluster of features shows no social differentiation whatsoever, another cluster of features appears to be areally defined and yet two others by the speakers’ ethnic background in interaction with both their age and geographical background.

Newlin-Lukowicz looks at a community with ongoing language contact, Polish-American speakers of English living in New York City. She finds that her speakers appear to be organized around three reference points: Poland, the New York City Polish community, and American society in general. These groupings emerge from her ethnographic analysis of speakers’ descriptions of their interests and answers to her interview questions. Her linguistic analysis considers three variables which have historically characterized New York City English, but which are also potentially affected by contact with a Polish substrate: the vowels of BOUGHT and BAD, and the stopping of interdental fricatives.

Tsiplakou, Armostis & Evripidou observe that contemporary Cypriot Greek has been called a ‘mixed system’. Not only does it show heavy code-mixing between Cypriot and Standard Modern Greek, but at the same time local varieties undergo structural dialect loss. The authors test the hypothesis that the pool of variants that speakers have at their disposal is used strategically, depending on indexical purposes, and that therefore one ought not to expect consistency in, e.g., rates of occurrence of variants. The hypothesis is tested for two phonological and two morphosyntactic variables in data from sociolinguistic interviews. While there is overall coherence across variables in the sense that higher rates of the Cypriot variant of one variable correlate with similar higher rates in other variables, the behaviour of different variants also depends on situational as well as social factors (viz. the speakers’ sex and education), resulting in different ‘coherent’ constellations depending on particular constellations of extralinguistic factors. In the case of Cypriot Greek, ‘coherence’ hence turns out to be a complex notion.

Oushiro analyzes the social distribution and patterns of co-occurrence among six sociolinguistic variables – three phonological and three morphosyntactic – in São Paulo Portuguese. Several of these have local significance as markers of the city and/or its surrounding rural areas. One is a recent innovation in which young middle class women are the leaders. Others occur widely in Brazilian Portuguese with social stratification, but some (verbal non-agreement) are more stigmatized than others (coda –r deletion). The picture of cohesion that emerges is nuanced: the innovative form shows no significant

correlation with any other variable, but the socially stratified variables and those that are structurally related (the agreement phenomena) are strongly correlated in use.

Wiese and Rehbein report on a substantial corpus of varieties of German used in multilingual urban neighborhoods where various immigrant communities are present, including speakers whose heritage languages are Arabic, Turkish, and Kurdish. They contrast speech in these multi-ethnic settings with another neighborhood that is mostly mono-ethnic, consisting mainly of residents with monolingual German backgrounds. A range of syntactic variables are examined, including omission of determiners and pronouns, expressions of locality with bare NPs as opposed to those marked with prepositions, and new directive particles and light-verb constructions. They demonstrate the connectedness of these variables at the levels of linguistic structure, but also see them as socially coherent, in a way that locates Kiezdeutsch at the interface between the immigrant communities and broader German society: on the one hand, most of these traits are used more in the multi-ethnic setting, but all are present to some degree in the monolingual German neighborhood.

5. Sizing up

We will conclude this introduction to the volume with a few observations on continuity and progress. How do the papers build on and advance the results of the pioneering work of Ma & Herasimchuk, Thelander and the others mentioned above? There are important expansions of the methodological repertoire in this volume, to include both correlational and non-correlational techniques, notably distance-based methods such as cluster analyses, and linguistic groupings based on shared structural systems. In the theoretical domain, these papers respond to and investigate the ongoing debate about the scope of individual agency and bricolage, and the performance of style and identity. They also extend the typological range of sociolinguistic settings, including situations of social stratification, dialect vs. standard, dialect leveling, ethnolectal differentiation, change in progress, and systemic restructuring. Finally, these works make it abundantly clear how the search for coherence among multiple variable phenomena goes beyond what has been achieved by looking at single linguistic variables in isolation; this approach deepens our understanding of linguistic variation and diversity in many ways, especially by facilitating and empirically confirming the identification of sectors and subdivisions of the speech community, each with their own characteristic variational DNA, and by illuminating the affiliations of individuals with these community domains.

Acknowledgements

We thank the reviewers of this paper for their very valuable remarks and suggestions. Only the authors are responsible for any shortcomings. Generally, we should like to thank (1) the contributors to this special issue for their patience, trust and cooperativeness, (2) Peter Auer, Richard Cameron, Peter Gilles, Kirk Hazen, Michol Hoffman, Roland Kehrein, Paul Kerswill, William Labov, Miriam Meyerhoff, Naomi Nagy, Dimitris

Papazachariou, Panos Papas, Mats Thelander, Hannes Scheutz, Scott Schwenter, Sali Tagliamonte, Reinhild Vandekerckhove, Hans van de Velde, Juan Villena Ponsoda and James Walker for their thorough and cooperative reviews of the various contributions, and (3) *Lingua*'s Anikó Lipták and Johan Rooryck for their trust and support.

References

- Auer, Peter (1997). Co-occurrence restrictions between linguistic variables: A case for social dialectology, phonological theory, and variation studies. In: Hinskens, F., R. van Hout & W.L. Wetzels, eds. *Variation, change, and phonological theory*. Amsterdam/Philadelphia: Benjamins. 69-99.
- Auer, Peter (1999). From Code-switching via Language Mixing to Fused Lects: Toward a Dynamic Typology of Bilingual Speech. *International Journal of Bilingualism* 3, 4, 309-332
- Bellmann, Günter (1998). Between base dialect and standard language, in: P. Auer (ed.), *Dialect levelling and the standard varieties in Europe*, *Folia Linguistica* 32, 1-2, 23-34
- Benor, Sarah (2010). Ethnolinguistic repertoire: Shifting the analytic focus in language and ethnicity. In: *Journal of Sociolinguistics* 14(2), 159-183.
- Berruto, Gaetano (1987). Varietät. In: U. Ammon, N. Dittmar & K. Mattheier (eds.), *Soziolinguistik. Ein internationales Handbuch zur Wissenschaft von Sprache und Gesellschaft*. Berlin/NewYork: De Gruyter. 263-267
- Bloomfield, Leonard (1933). *Language*. New York: Henry Holt.
- Bod, Rens (2013). *A New History of the Humanities. The Search for Principles and Patterns from Antiquity to the Present*. Oxford: OUP
- California Style Collective (1993). Variation and personal/group style. Paper presented at the annual New Ways of Analyzing Variation Conference, Ottawa
- Chomsky, Noam (2006). Bilingualism and the human capacity, in: *Language and Mind*, Third Edition. Cambridge: CUP, 173-185.
- Cornips, Leonie (2006). Intermediate Syntactic Variants in a Dialect - Standard Speech Repertoire and Relative Acceptability. In: G. Fanselow, C. Féry, M. Schlesewsky & R. Vogel (eds), *Gradience in Grammar. Generative Perspectives*. Oxford: OUP, 85-105.
- Dilthey, Wilhelm (1883). *Einleitung in die Geisteswissenschaften. Versuch einer Grundlegung für das Studium der Gesellschaft und der Geschichte*. Band 1. Leipzig: Duncker & Humblot
- Eckert, Penelope (2004). The meaning of style. *Texas Linguistic Forum* 47, 41-53.
- Eckert, Penelope (2008). Variation and the indexical field. *Journal of Sociolinguistics* 12, 453-476.
- Eckert, Penelope (2012). Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology* 41:87-100.
- Geeraerts, Dirk (2010). Schmidt redux: How systematic is the linguistic system if variation is rampant?, in: K. Boye & E. Engberg-Pedersen (eds.), *Language Usage and Language Structure*. Berlin/New York: Mouton de Gruyter, 237-262

- Geertz, Clifford (1973). *The interpretation of cultures: selected essays*. New York: Basic Books
- Gordon, Matthew (2002). Investigating chain shifts and mergers. In: J. Chambers, P. Trudgill & N. Schilling-Estes (eds), *The Handbook of Language Variation and Change*. Oxford: Blackwell, 244-266
- Guy, Gregory (2013). The cognitive coherence of sociolects: How do speakers handle multiple sociolinguistic variables? *Journal of Pragmatics* 52, 63-71.
- Halliday, Michael (1978). *Language as a Social Semiotic: The Social Interpretation of Language and Meaning*. Maryland: University Park Press.
- Herson Finn, Viktoria (1996). *What is Nas? Toward a theory of ethnolect in the south Slavic dialect continuum*. PhD-thesis, Ohio State University
- Hinskens, Frans (2007). New Types of non-standard Dutch. In: Fandrych, C. & R. Salverda eds. *Standard, Variation and Language Change in Germanic Languages*. Mannheim & Tübingen: IDS & Narr, Francke, Attempto, pp. 281-300
- Hinskens, Frans (2011). Emerging Moroccan and Turkish varieties of Dutch: ethnolects or ethnic styles? In: Kern, Fr. & M. Selting (ed.). *Ethnic Styles of Speaking in European Metropolitan Areas*. Amsterdam/Philadelphia: Benjamins, 2011, 103-131
- Hoppenbrouwers, Cor (1983). Het genus in een Brabants regiolect. In: *TABU, Bulletin voor Nederlandse Taalkunde* 13, 1-25
- Horvath, Barbara & David Sankoff (1987). Delimiting the Sydney speech community. *Language in Society* 16, 179-204.
- Hymes, Dell (1966). Two types of linguistic relativity. In W. Bright (ed.), *Sociolinguistics*. The Hague: Mouton. 114-158.
- Jacobi, Irene (2009). *On Variation and Change in Diphthongs and Long Vowels of Spoken Dutch*. PhD-thesis, Universiteit van Amsterdam
- Jaberg, Karl (1908). *Sprachgeographie*. Aarau: Sauerländer
- Labov, William (1963). The social motivation of a sound change, *Word* 19, 273-309
- Labov, William (1966). *The social stratification of English in New York City*. Washington DC: Center for Applied Linguistics.
- Labov, William (2001). *Principles of linguistic change. Vol II. Social Factors*. Oxford: Blackwell
- Lameli, Alfred (2004). Hierarchies of dialect features in a diachronic view: implicational scaling of real time data. In: B.-L. Gunnarson et al. (eds), *Language Variation in Europe. Papers from the Second International Conference on Language Variation in Europe (ICLaVE 2)*, 253-266.
- Le Page, Robert & Andrée Tabouret-Keller (1985). *Acts of identity. Creole-based approaches to language and identity*. Cambridge: Cambridge University Press
- Ma, Roxana & Eleanor Herasimchuk (1972). Speech styles in Puerto Rican bilingual speakers: a factor analysis of co-variation of phonological variables. In: J.A. Fishman (ed.), *Advances in the sociology of languages*, vol. II. The Hague: Mouton. 268-295
- Möller, Robert (2013). *Erscheinungsformen rheinischer Alltagssprache Untersuchungen zu Variation und Kookkurrenzregularitäten im „mittleren Bereich“ zwischen Dialekt und Standardsprache*. ZDL Beiheft 153. Stuttgart: Steiner.

- Moosmüller, Silvia & Hannes Scheutz (2013). Chain shifts revisited: the case of monophthongisation and E-merger in the city dialects of Salzburg and Vienna, in: P. Auer, J. Caro Reina & G. Kaufmann (eds), *Language variation – European perspectives IV*. Amsterdam/Philadelphia: Benjamins. 173-186
- Oushiro, Livia & Gregory R. Guy (2015). The effect of salience on co-variation in Brazilian Portuguese. *Pennsylvania Working Papers in Linguistics* vol. 21:2, no. 18
- Thelander, Mats (1982). A qualitative approach to quantitative data of speech variation. In: S. Romaine (ed.), *Sociolinguistic Variation in Speech Communities*. London: Edward Arnold, 65-83