This is a repository copy of *Contact and new varieties*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/75328/

Version: Submitted Version

**Book Section:**
Contact and new varieties

Paul Kerswill
Lancaster University

1 Introduction: “new” varieties

The term “new variety” implies the convergence, by a population of speakers, on a set of linguistic norms which are collectively different from previous norms. There are two epistemological issues here. The first is defining what is meant by a “population”. Trivially, this can refer to some set of people who, in our case, stop speaking Variety A and start speaking Variety B. More usefully, the term “population” can be applied in the sociolinguistic sense of “speech community” – individuals having some affinity with others through sharing linguistic norms, both in terms of linguistic structure and in terms of patterns of variation and subjective evaluation. This essentially Labovian view (Labov 1989: 2; Patrick 2002: 584–8) places the focus on collective behaviour, and therefore allows for a time-depth spanning generations – obviously essential if we assume that language change entails young speakers innovating, or at least adopting new features, differentiating them from their elders.

The second epistemological issue is how to set criteria for a “new” variety. In a “normal” speech community, subject to no more than medium rates of in- and out-migration, language change is gradual and the concept of a “new” variety is irrelevant. The formation of a new variety (which may be a language or a dialect) involves more than just changes in norms. We need to envisage a prior period of relative absence of norms followed by focusing (Le Page and Tabouret-Keller 1985) – the reduction in the number of variant forms and the increase in sociolinguistically predictable variation, that is, the (re-)emergence of norms. Importantly, new varieties lack the inherent continuity (looking backwards through time) of slowly changing speech-community norms (Kerswill 2002, pp. 695–8). To use a medical metaphor, a new variety only emerges when a speech community has experienced trauma, through the overwhelming influx of newcomers, through the shift of its members to another language, or through the transplantation of individuals from different speech communities to a new location where they, as (voluntary or involuntary) settlers, have to form a new community. This latter scenario gives rise to pidginization or creolization in cases where no language is shared by a large enough minority and where an economically or politically powerful language is a remote target. It also gives rise to koinéization in cases where mutually intelligible varieties are spoken by the settlers: According to Trudgill (1986, p. 107), koinéization starts out with the prior mixing of features from the different varieties, giving rise to a high degree of variability. This is followed by the reduction in the number forms available through koinéization. Koinéization is the leveling of variant forms of the same linguistic items (especially phonemes and morphemes), and simplification – the reduction of phonological and morphophonemic complexity. This usually, but not necessarily,
results in new, distinct, partly hybrid, partly innovative versions of the parent set of varieties, which are by definition more focussed than the original mix (Le Page and Tabouret-Keller, 1985). When focusing has occurred, the process is what Trudgill (2004: 89) calls new-dialect formation. In this chapter, much of what we will cover is concerned with this process: the formation of what Siegel (1985, p. 364) calls immigrant koines — a term synonymous with Trudgill’s “new dialects.” We will, however, extend Trudgill’s term to include another of the above “traumatic” scenarios: rapid changes in norms resulting from large-scale immigration, where a native-speaking population is numerically overwhelmed by a critical mass of incomers, to the extent that a new generation matures with a reduced exposure to the indigenous variety.

New-dialect formation (henceforth NDF) is a sub-type of change-by-contact, and so we need to delimit it from other sub-types. I have already distinguished it from pigdinization and creolization. To the extent that language change is externally motivated (and I would argue that nearly all of it is, at least in phonology — in inflectional morphology, word-formation and syntax analogy is additionally operative), contact has almost everything to do with it: Innovations, once actuated, are spread through contact, typically but not necessarily face to face. A useful position is to restrict NDF to the results of human migration. Thus, besides internally motivated changes such as chain shifts, I also exclude changes that are the result of geographical diffusion, though relocation diffusion (the transfer of linguistic features through speaker migration — Britain 2002: 622) is of course central, because in some cases new varieties (immigrant koines) crystallize as a result of it. If it can be shown that a change takes place more or less simultaneously across a particular region then this, too, is not our primary concern. Often, such pan-regional changes are part and parcel of regional dialect leveling (Kerswill 2003), or its near-synonym which emphasizes the adoption rather than leveling out of features, supralocalization (Milroy 2002): the amount of dialect/accent differentiation across a geographical region is reduced over time, giving rise to greater homogeneity. (These types of changes are dealt with in Chapter X [by Britain].) Hickey (2003, pp. 235–6; 2009) describes a further type, leading to homogeneity across regions: supraregionalization, referring to the adoption, feature by feature, of non-regional forms, without dialect contact and leveling taking place. Hickey describes this as taking place in Ireland, in whose capital, Dublin, a new “fashionable” variety is emerging as a reaction to localized Dublin English incorporating both existing and innovative non-regional forms (Hickey 2000; 2005). Supraregionalization seems to be a process related to what Sobrero (1996, pp. 106–108), in the Italian context, refers to as “koineization” — the adoption of regional varieties intermediate between (standard) Italian and the dialect of large urban centers.¹ As we shall see, all of these leveling/convergence processes interact with NDF. In fact, whether or not any of them are involved in NDF has proved controversial, as we shall see.

1.1 New-dialect formation: A provisional model

¹ The term “koine” was first used to denote the form of Greek used as a lingua franca during the Hellenistic and Roman periods (see Siegel, 1985, who lists a number of languages which have been referred to as koines). In this chapter, we restrict ourselves to “immigrant koines”, as well as related forms of language arising following what I here call “trauma.”
New varieties emerge, initially, through countless individual acts of linguistic accommodation (adjustments) performed by speakers of every age interacting with others, the adjustments being variously conscious (strategic, in a manner modeled by Communication Accommodation Theory – Giles 1973) and unconscious (through subconscious alignment – Pickering and Garrod 2004, cited in Tuten 2008). Accommodation may be responsive to context, or it may be long-term (Trudgill 1986), leading to what Labov (2001, pp. 415–7) calls “vernacular reorganization.” Accommodation takes place in a wider, but community-specific social context. Crucial community variables include:

- the proportions of children to adults in the initial stage (Kerswill and Williams, 2000, p. 90)
- degree of contact within and between age cohorts, especially in families
- relations between salient social groups (to the extent that these exist in a new community)
- the degree to which social group boundaries are, or become, sociolinguistically marked
- wider linguistic ideologies (Milroy, 1999)
- personal and social identity formation.

Against this social backdrop, much of the outline of the new variety is determined by the relative frequency of linguistic forms heard among the population in the early stages, some of those forms being the outcome of accommodation, others belonging to a speaker’s quasi-permanent vernacular. This determinism is modified by the social and demographic factors just mentioned, some early in the koineization process, some later. In particular, if in the early stages of contact social divisions are relatively absent (as we shall see in the discussion in this chapter), then accommodation, leading to convergence, will prevail, especially among the crucial first child generation (cf. Kerswill and Williams, 2000 and below). When social antagonisms emerge, whether they are new or reinforcements of divisions inherited from the “mother” country, then dissociation – the opposite of accommodation – may additionally occur (cf. Hickey, 2005).

2 *Tabulae rasae*: South African Bhojpuri and New Zealand English

In an imaginary experimental scenario, one might wish to observe the formation of a new dialect *ab initio* by depositing a population of people, carefully chosen to be speakers of different dialects and screened to avoid reflecting existing sociolinguistic divisions, on a desert island, and then return after a generation or two to see the outcome. This scenario is, of course, impossible on ethical and practical grounds. There are, however, a number of locations around the world where new dialects, isolated from their ancestral homelands, have been investigated in what Trudgill (2004) calls *tabula rasa* conditions, where speakers of the language concerned had not previously lived. These include Rosenberg (2005) on German language islands in Eastern Europe and the USA and Bhojpuri (Hindi) dialects in Fiji (Siegel, 1987, 1997) and South Africa (Mesthrie, 1991) (and elsewhere – see Trudgill 1986, pp. 99–102). Historical work in this tradition is
represented by Tuten (2003), Lodge (2004) and Dollinger (2008). Bhojpuri speakers from India arrived in these places as indentured laborers from around the middle of the nineteenth century to around 1910 (Mesthrie, 1991, p. 72), and many remained, forming distinct communities. South African Bhojpuri is a highly mixed variety, as Mesthrie points out: “SB [South African Bhojpuri] does not accord with any single language or dialect of North India, displaying – rather – a blend of features from several sources” (Mesthrie, 1991, p 104). Mixing such as this is a further characteristic of koines, which we will return to later. The Bhojpuri studies in particular show that, in relation to many of the input dialects, there is simplification in morphology and morphophonemics, with a reduction in the number of morphological categories which are marked, as well as simpler paradigms. Thus, South African Bhojpuri shows no trace of the present negative copula (Mesthrie 1991, p. 98), and lacks “respectful” forms in pronouns and verbs (op. cit., pp. 97, 100). However, these two features show a pattern which can be directly linked to the demographic and social make-up of the Bhojpuri communities. The present negative copula is a minority feature that would have been brought by Indian speakers of Eastern Bhojpuri, its absence being a consequence of the fact that these speakers were in a minority (op. cit., p. 104). Here, demography clearly played a role in a way which, as we shall see later, is of crucial importance. Secondly, “[i]t seems that this feature [respect marking] did not survive in the koine formation process in Natal, no doubt because of the levelling of social distinctions among [South African Bhojpuri] speakers” (op. cit., p. 100). Again, this is a critical observation: we will argue that the social upheaval caused by migration, and the need to establish a living in a new, often hostile environment, causes old social distinctions to be lost, and with them the linguistic marking of those distinctions.

Finally, we look briefly at the demography of the early settlers in South Africa, to see if we can find an explanation for this simplification. Mesthrie (1991, p. 6) cites statistics indicating an overwhelmingly adult, predominantly male stream of migrants. Trudgill (forthcoming) suggests that “in fact simplification is most likely to occur in situations involving language learning by adults, who are typically poor second-language learners as compared to small children, particularly so far as informal acquisition in short-term contact situations is concerned.” (See also Trudgill, this volume.) The theme of child learners vs. adult learners is one we will return to.

The Bhojpuri studies do not give us clear information on the original settlers or their immediate descendants (the first generation of children born in the new location). A study which comes close to this is the investigation of the early stages of New Zealand English from the 1840s on by Elizabeth Gordon, Peter Trudgill, and their associates (Trudgill, Gordon, Lewis, and Maclagan, 2000; Gordon, Campbell, Hay, Maclagan, Sudbury, and Trudgill, 2004; Trudgill, 2004). The uniqueness of the Gordon/Trudgill study is that it was based on oral history recordings of elderly New Zealanders made in 1946–8. Crucially, those recorded included people born in the period when New Zealand English was, according to Trudgill (2004, p. xi), being formed: 1850–1890. The recordings thus afford a window on the speech of the earliest children growing up in colonial New Zealand, albeit captured in their late adulthood. Gordon et al. (2004) presents a comprehensive report and discussion of the findings. Trudgill (2004), however,

---

2 The recordings were made by the Mobile Disc Recording Unit of the National Broadcasting Corporation of New Zealand (see Gordon et al. 2004, pp. 3–5 and Trudgill 2004, p. x).
presents the argument that new-dialect formation is almost entirely *deterministic* in the sense that the shape of a new dialect can be predicted with some precision from a knowledge of the input varieties – the dialects of the original immigrants. This theory relies on careful consideration of three *stages of new-dialect formation*, and argues that the types of accommodatory behaviour and selection of linguistic variants at each stage proceeds in a predictable way which is largely unconnected with social factors – other than demography. Before critiquing his position, I will outline Trudgill’s stages. These correspond roughly to the first three generations of speakers (Trudgill et al., 2000):

<table>
<thead>
<tr>
<th>Stage</th>
<th>Speakers involved</th>
<th>linguistic characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>adult migrants (first generation)</td>
<td>rudimentary levelling</td>
</tr>
<tr>
<td>II</td>
<td>first native-born speakers (second generation)</td>
<td>extreme variability and further levelling</td>
</tr>
<tr>
<td>III</td>
<td>subsequent generations</td>
<td>focusing, levelling and reallocation</td>
</tr>
</tbody>
</table>

At Stage I, adult migrants will level away features which are in a small minority in the mix of dialects they encounter, subject to the individual’s ability to do so (adults are generally less successful than children in modifying their language, especially phonology) (Trudgill 2004, pp. 89–93). In the New Zealand case, these would have included traditional dialect features. There would have been a great deal of inter-individual variability at this stage, and people would themselves have been inconsistent in their usage (intra-individual variability).

At Stage II, the demographic distribution of features begins to determine the shape of the focused variety which is still to appear. The Mobile Unit oral history recordings represent this stage. According to Trudgill, the absence of a stable adult norm, or even a peer-group dialect, means that children pick features to some extent “at will”, “from a kind of supermarket” (Trudgill 2004, p. 103, 108). The reason for this is that, in this tabula rasa context, they are not influenced by prestige or identity-marking functions (pp. 151–7). Trudgill states that these young speakers did not “indulge in long-term accommodation to one another” (p. 108), but rather that they selected features based on their frequency, always allowing for a “threshold rider” which filters out relatively infrequent features (p. 110). As at Stage I, speakers show considerable inter- and intra-individual variability, much more so than in a community where the transmission of norms has been “normal” in the sense of cross-generational transmission, defined by Thomason and Kaufman (1988, pp. 9–10) as taking place when “a language is passed on from parent generation to child generation and/or via peer group from immediately older to immediately younger.” Examples of this variability in two small communities are as follows (from Kerswill and Trudgill, 2005, pp. 209–10):

(1) Inter-individual variability in Arrowtown:

\[
\text{GOAT: } [o^\star] \quad [o^\nu] \quad [\o\nu] \quad [\o^\nu] \quad [\o^\nu] \quad [\o^\nu]
\]

(2) Intra-individual variability in the speech of Mr. Riddle, Palmerston:
/æt/ and /əʊ/ as in FACE and GOAT alternate between Scottish-sounding monophthongal pronunciations with [e] and [o] and very un-Scottish pronunciations with the wide diphthongs [æt] and [əʊ].

Stage III represents the focusing of the new variety, with alternate realizations leveled out, leaving only one, or two in the case of reallocation where variants are “reallocated” to a linguistic or sociolinguistic function (Britain and Trudgill, 1999, p. 245). In New Zealand, this resulted in a very homogeneous variety, apparently by 1900 (with the exception of the preservation of rhoticity in the far south), though with considerable social variation in terms of accent. Trudgill (2004, pp. 115–128) argues strongly for a purely quantitative explanation of the features adopted by the new variety, citing several vowels and consonants including the retention of /l/, the maintenance of the /n/-/w/ distinction (as in which/witch), the merger of unstressed /ə/ and /ɪ/ on /ə/ (as in rabbit), and broad diphthongs in words of the GOAT, FACE, MOUTH, and PRICE sets. In all but one of these, Trudgill adduces a simple majority principle: In the Mobile Unit recordings, these features are more frequent than their absence. In the case of the merger of /ə/ and /ɪ/, /ə/ was not in a majority, with only 32% of tokens having this vowel. Here, in a rather post hoc manner (and he admits his argument is not strong on this point – Trudgill, 2004, p. 120) he appeals to a markedness principle: /ə/ is less “marked” in this position than /ɪ/, and this helped guarantee its survival.

An important component of Stage III is the notion of “drift”, to explain apparently parallel developments in early New Zealand English and English in England. Drift, according to Trudgill, refers either to the continuation of a change in the “home” country or a tendency or propensity for a change (Trudgill, 2004, pp. 132–3). Hickey (2003, pp. 239–235) agrees with the quantitative facts, but argues that the idea of “inheriting” a change is an unwarranted reification of language. He prefers to see changes in terms of individual speakers, especially children, detecting what is innovative and what conservative, or reinterpreting small phonetic changes as a shift of phonemic status. The latter would be the case in the parallel shift in both England and New Zealand from lax /ɪ/ to tense /iː/ in words like coffee (HAPPY-tensing). Hickey reconceptualizes Trudgill’s explanation, but he does not detract from Trudgill’s essential points about the dynamic relationship between the emergent dialect and its British antecedents.

There remains the problem of how to explain the almost complete lack of regional variation in New Zealand English. Trudgill provides a partial answer. In accounting for the similarities between Australian and New Zealand English, he quotes a mixing bowl metaphor from Bernard 1981: “[T]he ingredients of the mixing bowl were very much the same, and at different times and different places the same process was carried out and the same end point achieved” (Trudgill, 2004, p. 161). Trudgill points out that the dialect mixes in the different settlements in New Zealand were not in fact exactly the same, so other factors must be brought to bear. New Zealand was a mobile society, despite the considerable distances, and Trudgill agrees with Britain, who writes, “... settlement isolation, mobility, transience and individualism led to the emergence of an atomistic society freeing people both from subservience and from the need to conform that tight-
knit local communities often engender” (Britain, 2005, pp. 164–5, referring to Fairburn, 1982). Mobility in such a society, with a lack of local speech norms, would, it is argued, quickly lead to uniformity. At the same time, social prestige played no part, as Trudgill argues for both Stage II and Stage III, because children would not have been exposed to a standard ideology in New Zealand at that time – and in any case, children align themselves linguistically with local speech, especially that of their peers. Britain (2005, 165–6) points out that, in the early decades of European settlement, there was no compulsory education and literacy was low, so that overt pressures from prestige varieties could scarcely have had any effect. Both literacy and English-style social stratification came a little later, and in any case after the earliest settlers’ children had begun the process of koineization. Moreover, New Zealand was what Belich (1996, p. 330) calls “custom shedding ... Highly overt class differences ... were leading candidates for the discard pile” – echoing the Bhojpuri communities’ leveling of social distinctions (though in their case they were in a subordinate social position *ab initio*). Together, social structure, lack of overt norms, the seemingly random choice of forms at Stage II, and a choice of forms at Stage III based solely on relative frequencies of forms used by the Stage II speakers all contributed to a rapid focusing on a set of features with origins in various British regional varieties (Trudgill 2004, pp. 114–5). Even so, the homogeneity is much greater than would be expected, and we need to account for this. First, we will take a closer look at the “social factors,” rejected by Trudgill.

3 Social factors vs. determinism in *tabula rasa* new-dialect formation

We can now relate Trudgill’s findings and arguments to the model of new-dialect formation I presented at the end of section 1. That model allowed for “social factors” to affect the process at different stages. Trudgill states that such factors did not affect children’s dialect development at either Stage II or III in *tabula rasa* new-dialect formation. His argument is that the social set-up of early colonial New Zealand meant that “prestige” (including covert prestige), “stigma”, “identity”, and “ideology” counted for nothing in what motivated children to adopt the features they did. Instead, he appeals to the notion of “behavioral co-ordination” (citing Cappella, 1997; see also Pickering and Garrod, 2004, cited by Tuten 2008). This motivates people to converge behaviorally as the default, leading to the quantitative results Trudgill claims to find. At face value all this is easy to accept – but only if we accept all of Trudgill’s assumptions about the nature of early Anglophone New Zealand. The problem lies with the conceptualization of the *tabula rasa*. The Stage II scenario as outlined by Trudgill was most likely fairly rare. Clearly it existed in the very first English-speaking settlements where children were present (and not in the earlier, male-only, transient whaling communities). After this, it would have existed only in relatively small settlements populated by immigrants arriving at roughly the same time, with relatively few newcomers arriving later. I say “relatively”, because in an isolated farmstead there would be much less variability, and because a later, massive influx could swamp the existing, small founder population. Conditions for Trudgill’s Stage II would not have existed in the larger settlements, except of course on their founding, and a complex social make-up, including institutionalized ethnic divides – much of it aimed at Irish immigrants, who needed a permit of stay (Hickey, p.c.) – quite
quickly emerged in some of the new urban communities. According to Belich (1996, p. 405), in Christchurch social stratification, some of it imported, apparently deliberately, from Britain, was present very early. Christchurch was founded as a Church of England settlement, and its ethos is shown by early complaints about Australian and “half-breed” incomers (Belich, 1996, p. 339) (see also Kerswill, 2007). Meyerhoff (2006) argues strongly for a more complex and nuanced view of the early stage. She points out that several of the variables Trudgill investigates in fact show significant effects of parents’ origins (English, Irish, Scottish or Australian) for these same Stage II speakers – the model predicts the absence of such effects. It is likely that in a diffuse speech community children will be more, rather than less, influenced by parental varieties, as has been observed in two studies of new-dialect formation in Norway (Kerswill and Williams, 2000, p. 75). Meyerhoff also points out that children were born to large families with a great age range among the children, and that parents and older siblings would, like everywhere else, use a range of contextual styles which would be detected by children. In such communities, children at both Stages II and III would have co-existed, even within the same family, a fact which strongly curtails the time-span of the social-factor-free Stage II.

Where does this leave the quantitative, deterministic account? Hickey (2003) argues that, because of the high proportion of single men and women, the impact of Irish settlers would have been less than their relatively modest numbers predict; this seems to me not to be particularly relevant, simply because Trudgill’s model excludes minority influences and the Irish were in a minority anyway. Almost all of Trudgill’s features support his model, or at least don’t contradict it. The outcome, according to Trudgill (2004, p. 157), is a composite in which Stage III children selected “... upper-class H Retention, lower-class Diphthong Shift, urban non-rhoticity, and the Rural Weak Vowel merger, from all the features available to them.” This, in his view, guarantees that they were not motivated by any prestige or identity-based factors; elsewhere, he argues: “But this kind of baggage is not relevant to 7-year-old children in the colonial situation – which is precisely why the mixing of variants from different dialects of different regional origin and different degrees of social status, with a mass of different associations and connotations, always takes place” (Trudgill, 2008, p. 279). As I have indicated earlier, I agree with Trudgill in that it is difficult to argue that normative pressures from the English-based education system and British ideas of acceptable public usage could have had more than a limited effect on this particular outcome (Trudgill excludes this possibility entirely). However, as the discussion in the previous paragraph suggests, children did not grow up in a social vacuum, devoid of adult intervention and isolated from adult norms. Here, one would have to envisage an even more extreme experiment: the depositing of a population of pre-adolescents on a desert island, Lord of the Flies-like.

A tabula rasa quickly ceases to be a clean slate (if it ever was, since adults in a new community are not entirely deracinated). It is wrong to say that social factors had no effect in the early period. The question is, which social factors? Prestige, and probably identity formation, can certainly be ruled out at the very start of the koineization process, but parents’ inherited ideas about good and bad behaviour, and acceptable linguistic practices including politeness, would have been brought to bear. The majority of New Zealand children at Stage II and certainly Stage III would not have been growing up in the idealized situation Trudgill envisages; indeed, the Mobile Unit recordings are heavily
biased towards the rural and so are not representative of how the early settlers lived, even if taken as a whole the balance across the regions of the British Isles did reflect the total immigrant population of the country. Our conclusion must be, however, that Trudgill’s three-stage deterministic model is correct, but only in the rarefied, hypothetical situation such as the imaginary experiment mentioned earlier. In the real case Trudgill presents to us, the outcome is, indeed, pretty much as predicted. However, the model Trudgill gives us is quickly “messed up”, as I have argued. Trudgill’s idealized route, with its absence of identity factors (note my restriction to this type of social factor), doubtless pertained in a few places, with the linguistic outcome closely corresponding to focussed New Zealand English. But it existed alongside a vast majority of socially complex, more “normal” situations. Thus, while accepting the numerical model, it is difficult to accept its detailed implementation “on the ground.”

4 Homogenization in new varieties

All this said, there is still a serious gap in our understanding. The geographical homogeneity of New Zealand English is not merely a relative matter (compared to, say, Great Britain), but (barring Southland rhoticity) virtually total. Phonetic divergence has only recently emerged (by all accounts) in the form of distinct Maori and Pasifika Englishes (Hay, Maclagan, and Gordon, 2008, pp. 105–109). The mixing-bowl metaphor does not in itself predict total homogeneity, but differences resulting from any variations in the ingredients. Thus, in a highly mobile and atomistic society, we might expect regional varieties to have emerged around Auckland and the other main centers, in the manner of the regional dialect leveling in southern England today (Williams and Kerswill, 1999; Kerswill, 2003). However, this is not the case. Hickey proposes a mechanism of supraregionalization to account for this: “[D]ialect speakers progressively adopt more and more features of a non-regional variety which they are in contact with. There does not have to be direct speaker contact …” (2003, p. 236). In Trudgill’s Stage III, the new variety can “be seen as a product of unconscious choices made across a broad front in a new society to create a distinct linguistic identity” (Hickey 2003, p. 215). This supraregional variety, according to Hickey, would have emerged in the “melting pot” settlements, which had mixed populations of relatively high density and size but with similar mixes, and then spread to the much more dialectally distinctive rural settlements. This process can be observed in contemporary Europe, notably in Denmark, where regional dialects have all but given way to a non-conventionally prestigious but “modern” Copenhagen-based variety (Pedersen, 2005; Kristiansen and Jørgensen 2005). But for New Zealand English, in spite of high mobility, it is extremely difficult to see how sufficient familiarity with the new variety could have come about to act as a reliable model. We cannot be certain how, or exactly when the entirely non-regional variety appeared, but it is necessarily distinct from the focusing at local or regional speech community level which presumably preceded it. We should therefore introduce a Stage IV, at which new-dialect formation is already complete at the local/regional level, and at which supraregionalization is about to set in.

Demographic and other social factors – especially gender and class – mentioned by Gordon (2008) guided the spread of New Zealand English, either promoting it or
hindering it. Crucially, they did not actually determine its form. Instead, we have an image of the fully-formed variety spreading inexorably, meeting varying degrees of resistance. Interestingly, Arrowtown (a focus of the Gordon/Trudgill studies) is one of the pockets where Trudgill’s model seems to have applied; but in the context of New Zealand it was too small to have been a center of influence in itself. What Gordon describes matches rather closely what Hickey means by supraregionalization. We have no access to its mechanism in those early days; fortunately, we have the modern sociolinguistic record of countries like Denmark and (to a more limited extent) Ireland, where the same process has been documented in recent times.

5 Koineization in New Towns

The second type of new-dialect formation we will discuss is that of the new town, which we take as representative of new settlements where there are prior speakers of the language concerned, and where, therefore, some degree of face-to-face contact with existing speakers, either in the new location or the ‘home’ location, is maintained. The term “New Town” is a designation in official UK usage for a new, planned urban settlement placed on previously more or less unoccupied land, built throughout the twentieth century but particularly since the Second World War. The earliest new towns described in the sociolinguistic literature are, however, in Norway, where several were established at the heads of fjords in the period 1910–20 to harness hydroelectric power for the smelting of various ores. These included Odda and Tyssedal (east of Bergen), Sauda (south-east of Bergen), as well as Høyanger (in the Sognefjord, north of Bergen). The first two of these were investigated by Sandve (1976), and an extensive report and interpretation of his findings appear in Kerswill (2002, pp. 674–677), with a focus on dialect mixing, interdialect (intermediate, compromise) forms, simplification, and the very marked effect of differences in the geographical origins of the in-migrants to the two towns, which are located only 5 kilometers apart. Høyanger was discussed extensively by Trudgill (1986) as a prototypical case of new-dialect formation. Basing himself on a short publication by Omdal (1977), he established the stages which he later applied to New Zealand. Omdal noted a transition from the traditional rural dialect of the oldest speakers, who grew up before the establishment of the town, through the extreme linguistic heterogeneity of the children who grew up in the new town, to a more stable new dialect spoken by the third generation – who were young adults in the 1970s. So far, the scenario is similar to New Zealand, with the important exception that some of the children growing up were the offspring of the original population. These people’s speech was mixed, but was closer to the old dialect than was that of the descendents of incomers. In 2001, Randi Solheim conducted sociolinguistic interviews in Høyanger, as well as obtaining archive recordings and previous descriptions, giving her a real-time window of 45 years (Solheim 2006; 2009 forthcoming).

As expected, the new dialect includes features from both major input varieties, West Norwegian (the dialect area in which the town is located, and from where the vast majority of in-migrants arrived) and East Norwegian (including the capital, Oslo, from where many of the managers and engineers came). However, despite the relatively small number of people from the East, a large number of high-frequency words have an eastern form. Table 1 shows the distribution of forms.
Table 1: West and East Norwegian forms in Høyanger new town dialect (Solheim 2006)

<table>
<thead>
<tr>
<th>Feature</th>
<th>West Norwegian rural dialects</th>
<th>Occurs in Høyanger new town dialect?</th>
<th>East Norwegian (=bokmål standard)</th>
<th>Occurs in Høyanger new town dialect?</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diphthong/monophthong</td>
<td>dreiv, skreiv, veit</td>
<td>✓</td>
<td>drev, skrev, vet</td>
<td>✓</td>
<td>'drove', 'wrote', 'know'</td>
</tr>
<tr>
<td></td>
<td>høyre, køyre</td>
<td></td>
<td>høre, kjøre</td>
<td>✓</td>
<td>'hear', 'drive'</td>
</tr>
<tr>
<td>Simplification of /ld/ cluster</td>
<td>forhold, gjelde [ld]</td>
<td></td>
<td>forhold, gjelde [l]</td>
<td>✓</td>
<td>'circumstance', 'concern' (vb.)</td>
</tr>
<tr>
<td>Present tense of Class 1 weak verbs</td>
<td>kasta, takka, plaga [a]</td>
<td>✓</td>
<td>kaster, takker, plager [ɔr]</td>
<td></td>
<td>'throw', 'thank', 'annoy'</td>
</tr>
<tr>
<td>Present tense of strong verbs</td>
<td>skrive, finne [ɔ]</td>
<td>✓</td>
<td>skriver, finner</td>
<td></td>
<td>'write', 'find'</td>
</tr>
<tr>
<td>Morphophonemic alternation of velars and palatals in certain grammatical categories</td>
<td>/tak/ - /tæːkəl/</td>
<td>/værə/ - /værən/</td>
<td>velar in all inflected forms</td>
<td>✓</td>
<td>'roof'–'the roof', 'book'–'the book', 'wall'–'the wall', 'time' (instance)–'the time'</td>
</tr>
<tr>
<td>Lexical forms of high frequency words</td>
<td>ho</td>
<td>✓</td>
<td>hun</td>
<td></td>
<td>'she'</td>
</tr>
<tr>
<td></td>
<td>me and vi</td>
<td>vi</td>
<td></td>
<td>✓</td>
<td>'we'</td>
</tr>
<tr>
<td></td>
<td>de /dɛː/, deke</td>
<td>dere</td>
<td></td>
<td>✓</td>
<td>'you' (pl.)</td>
</tr>
<tr>
<td></td>
<td>dei</td>
<td>de /di/</td>
<td></td>
<td>✓</td>
<td>'they'</td>
</tr>
<tr>
<td></td>
<td>noken</td>
<td>noen</td>
<td></td>
<td>✓</td>
<td>'some', 'any'</td>
</tr>
<tr>
<td></td>
<td>sjå</td>
<td>se</td>
<td></td>
<td>✓</td>
<td>'see'</td>
</tr>
<tr>
<td></td>
<td>seie</td>
<td>si</td>
<td></td>
<td>✓</td>
<td>'say'</td>
</tr>
</tbody>
</table>

As in most koines, there are interdialectal forms (Trudgill, 1986, pp. 62–5). In Høyanger, these are a compromise between the two main dialect sources, often involving the blending of a western stem with an eastern inflection – or vice versa. Table 2 shows the most striking of these.

Table 2: Interdialectal forms in Høyanger

<table>
<thead>
<tr>
<th>Feature</th>
<th>West Norwegian</th>
<th>East Norwegian</th>
<th>Høyanger</th>
<th>Comment</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eastern noun stem, western inflexional suffix (several examples)</td>
<td>/gutana/</td>
<td>/gutana/</td>
<td>/gutana/</td>
<td>Blend of E. Norwegian stem /gut/ with W. Norwegian suffix /a/</td>
<td>'boy' – 'the boys'</td>
</tr>
<tr>
<td>Umlaut (vowel change) in present tense of some strong verbs</td>
<td>komme – kjeme</td>
<td>komme – kommer</td>
<td>komme – komme</td>
<td>Høyanger simplifies by removing vowel change, but keeps western inflection (ta is a different class)</td>
<td>'come' 'sleep' 'take' (inf. and pres. tense)</td>
</tr>
<tr>
<td></td>
<td>sove – søve</td>
<td>sover – sower</td>
<td>sove – sőve</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ta – teke</td>
<td>ta – tar</td>
<td>ta – tar</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The eastern forms which also represent simplifications (particularly loss of velar–palatal alternation and umlaut in verbs) are now widespread in western Norway, as part of regional dialect leveling (Kerswill, 2002, p. 684) – a point I will return to in a discussion of the new dialect of Milton Keynes. However, most of the eastern forms, including the pronominal forms dere, de (/ðiː/) and noen, are rare elsewhere in the region, thus demonstrating that the dialect is a koine.

I mentioned earlier that the Høyanger study gives us a direct window on social influences on the outcome of new-dialect formation. We have already seen that the number of high-frequency eastern forms is disproportionate to the number of in-migrants from that region: Solheim attributes this to the high social prestige of those people. However, a number of items initially took an eastern form, only to be replaced by the original western realisation. Table 2 shows two such instances, where there has been a shift from the original dialect forms, through East Norwegian forms, back to the original variants. Solheim uses a terminology similar to that of Trudgill, labeling the stages ‘generations’ in order to link individual life experiences more directly with developments in the town itself. ‘Generation I’, however, refers not to the in-migrants, but to the existing dialect-speaking population.

Table 3: Evolution of two salient Høyanger variables

<table>
<thead>
<tr>
<th>Generation I (rural dialect)</th>
<th>Generation II</th>
<th>Generation III</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>ikke /ɪkə/</td>
<td>ikkje /ɪçə/</td>
<td>ikke and ikkje</td>
<td>ikkje</td>
</tr>
<tr>
<td>jeg /jæɪ/</td>
<td>eg /ɛɣ/</td>
<td>jeg and eg</td>
<td>eg</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>negator</td>
</tr>
</tbody>
</table>

Note: Usage in Generations II and III was variable, as expected in the early stage of a mixed dialect. By Generation III, the eastern forms had all but disappeared.

Solheim interprets this shift to the notion that the East Norwegian forms, being associated both with the former managerial class and Standard Norwegian in its prestigious bokmål instantiation, is too strong a marker to be acceptable in a West Norwegian dialect. This point of view is indeed expressed by some of the informants themselves. It is worth noting that these two items are regularly cited by Norwegians as regional and social dialect markers, doubtless partly because they are also associated with the two versions of Standard Norwegian – nynorsk (mainly western and rural) and bokmål (eastern, northern and urban). This is an extremely clear indication that language ideology may be a direct motivation for a wholesale change in linguistic usage.

There is evidence, too, of local identity formation guiding changes in usage. Two diphthongs, /æɪ/ and /øy/, originally had fairly localized realizations as [aɪ] and [ɔʏ], respectively, setting them apart from much of the rest of the country which has realizations in the region of [ɛɪ] and [ʊʏ]. Table 4 shows changes in the realizations of these diphthongs, again in terms of generations. This time, we add a Generation IV, representing the phase after the completion of new-dialect formation when the dialect is
no longer affected in its development by the community-internal process of koineization, but is oriented to external factors in the same way as the dialect area at large.

Table 4: /æɪ/ and /øy/ in Høyanger (based on Solheim 2006, pp. 168, 173, 239–241)

<table>
<thead>
<tr>
<th>Example</th>
<th>Supralocal varieties</th>
<th>Generation I (original rural dialect)</th>
<th>Generation II (first child generation – cf. Trudgill’s Stage II)</th>
<th>Generation III (fully established industrial new town – cf. Trudgill’s Stage III)</th>
<th>Generation IV (present-day young people in post-industrial town)</th>
</tr>
</thead>
<tbody>
<tr>
<td>geit ‘goat’</td>
<td>[ɛɪ]</td>
<td>[aɪ]</td>
<td>[ɛɪ] [æɪ] [aɪ]</td>
<td>[aɪ] &gt;[ɛɪ]</td>
<td>[aɪ], [ɛɪ]</td>
</tr>
<tr>
<td>höy ‘hay’</td>
<td>[øy]</td>
<td>[ɔy]</td>
<td>[øy] [œy] [ɔy]</td>
<td>[øy] &gt;[øy]</td>
<td>[øy], [øy]</td>
</tr>
</tbody>
</table>

| Note: Under Generation III, “>” indicates that these speakers are marginally increasing their use of the supralocal forms in their lifetimes (Solheim, 2006, pp. 239–241), while the youngest (Generation IV) speakers to some extent maintain the local variants, particularly [ɔy]. |

The supralocal variants of both diphthongs were in a large majority in Generation II, only to be largely replaced by the local form in Generation III. Meanwhile, Solheim has real-time evidence that Generation III speakers actually increased their use of the supralocal variant during their lifetimes, while today’s youngest speakers, Generation IV, seem to be increasing their use of the local variant of /øy/. They do this most markedly in the place-name Høyanger, insisting explicitly that [hɔyˈæŋə] is the correct pronunciation (Solheim, 2008, p. 6).

The advantage of the Høyanger studies is that it is possible to show how local social factors affected outcomes, thus confirming some of the criticisms of Trudgill’s deterministic model that it does not pay attention to the mechanism of transmission at Stages II and III. Solheim notes that

a central observation emerging from my work on more recent data from Generation II, not least from my encounters with the speakers themselves, is that individuals’ personalities and life worlds are, to a great extent, decisive influences on their language use. It is likely that background factors such as these are particularly significant for this generation, since, in their formative years, there was no stable, local linguistic norm on which they could rely [my translation]. (Solheim 2006: 237)

There is every indication that the outcomes in places like Høyanger, Odda and Tyssedal are largely predictable using a quantitative model. Solheim shows that there is a range of individual social factors, including social class, personality and even identity formation, which affect outcomes for individuals, perhaps especially in Generation II. However, collectively, individual outcomes can only affect the direction of focusing if they are
affected by the same large-scale social factor. Solheim mentions ideology in the shape of attitudes, as well as local identity – both operating at Generation/Stage III; social class actually prevented focusing until, under conditions of post-World War II social democracy, it became irrelevant. It can be argued that all this simply isn’t applicable to tabula rasa situations like New Zealand. However, as we have seen it can equally be argued that ideology (but not identity formation) played a part there as well from the very start, even within families and small communities – though naturally we have no evidence either way except for the speech of Stage II speakers in old age. Language ideologies rapidly gained a foothold in the main centers, and would have coincided even with the earliest Stage II speakers’ childhood and adolescence. That said, in the New Zealand case, I agree with Trudgill in rejecting identity, even local, as a motivation for the focusing on the mixed norm which took place, if for no other reason because there is no evidence of local dialects ever having formed before supraregionalization set in. Perhaps we haven’t been able to observe that stage – though Gordon (2008) brings us close.

Milton Keynes is the newest, and largest, of Britain’s New Towns, designated in 1967. It is situated some 80 kilometers north-west of London, in an already extensively leveled dialect area. It is situated on the boundary between what Trudgill (1990, p. 63) identifies as the South Midlands and Home Counties (i.e. south-eastern) Modern Dialect areas. We have already seen how the Høyanger dialect partly adumbrates leveling changes happening in its own still strongly dialectal region: What is the situation with a much more highly connected new town in an already leveled region?

Population statistics reveal a rapid increase, particularly from the mid-70s to the late 80s, as follows:

<table>
<thead>
<tr>
<th>Year</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>1967</td>
<td>60,000</td>
</tr>
<tr>
<td>1971</td>
<td>66,800</td>
</tr>
<tr>
<td>1977</td>
<td>96,300</td>
</tr>
<tr>
<td>1987</td>
<td>161,500</td>
</tr>
<tr>
<td>1997</td>
<td>196,920</td>
</tr>
<tr>
<td>2007</td>
<td>227,796</td>
</tr>
</tbody>
</table>

Source: Milton Keynes Intelligence Observatory (online resource)

Sociolinguistic recordings were done in 1991 and 1992, some 24 years after designation and 14 years after large-scale in-migration had got under way. Children and young people at that time were representatives of Stage II, their parents Stage I, an assumption corroborated by the fact that only 10 of the 48 adults recorded were born in Milton Keynes. The primary interest was to observe new-dialect formation as it was actually happening, rather than two or more generations later. The project could not, of course, observe the fully focused outcome, since Stage III had not been reached, though as it turned out focusing was already well under way at Stage II. The sample was composed of 8 girls and 8 boys in each of three age groups of 4, 8 and 12 years old, in addition to the principal caregiver (46 mothers, one father, one aunt). The families were selected to be classifiable as “working class”.

The project was informed by a number of “principles of koineization”:

Outcomes in post-contact varieties:

1. Majority forms found in the mix, rather than minority forms, win out.

---

3 New recordings are planned for 2009/2010 as part of a doctoral project.
2. Marked regional forms are disfavored.
3. Phonologically and lexically simple features are more often adopted than complex ones.

_The migrants and the first generation of native-born children:_
4. Adults, adolescents, and children influence the outcome of dialect contact differently.
5. The adoption of features by a speaker depends on his or her network characteristics.

_The time scale of koineization:_
6. There is no normal historical continuity with the locality, either socially or linguistically. Most first and second generation speakers are oriented toward language varieties that originate elsewhere.
7. From initial diffusion, focusing takes place over one or two generations.
8. Because of sociolinguistic maturation, the structure of the new speech community is first discernible in the speech of native-born adolescents, not young children.

Not surprisingly on the basis of our previous discussion, Principles 1 and 2 can be shown to apply in Milton Keynes. Little can be said about Principle 3, because most of the input dialects were already leveled south-eastern ones. These principles are discussed in Kerswill and Williams (2000, pp. 85–89), and will not be treated further here. Principle 4 could also not be fully addressed, though it was clear that the high proportion of children to adults relative to the general population would favour early focusing (26.1% under-sixteens compared to 20.1% for England and Wales). However, Trudgill (2002; this volume; forthcoming) has reasoned that in cases where dialect and language contact is relatively sustained and involves children, complex features may be learned more easily than when contact mainly involves adults.

Principles 5 and 8 can both be demonstrated by an examination of index scores on a phonetic variable, the fronting of the vowel of _goat_. Table 5 shows the values for this variable:

| (ou) - 0: [ɔ], [o] | score: 0 | (Northern and Scottish realization) |
| (ou) - 1: [ɔu], [ɔu] | score: 1 | (older Buckinghamshire and London) |
| (ou) - 2: [ɔY] | score: 2 | (fronting) |
| (ou) - 3: [ɔt] | score: 3 | (fronting and unrounding) |

Kerswill and Williams (2000, pp. 93–4) show, first, that it is the 12 year olds who have the highest scores, the girls exceeding the boys. Second, the older children have greater fronting than their caregivers, the conclusion being that this age group is propelling the change. Third, there appeared to be a group of low scorers and another of high scorers in this age group. The low scorers (mean score 1.3) are two boys and two girls, and appear to be socially quite isolated individuals. The high scorers are four girls (mean 2.1), and form a group of friends who are sociable and well integrated at school. Given that the fronting of this vowel is an ongoing change, the obvious conclusion is that it is female-
What this means for new-dialect formation is that “integrated” children with broad social contacts are in a position to focus the new dialect forms; in the sample, all the other children – be they 4 year olds or the less integrated older children – are linguistically more heterogeneous, showing features of their parents’ accents more strongly than the integrated children, and are clearly not in the lead in the focusing process. Interestingly, this shows that focusing can start at Stage II. It also confirms Principle 8 in showing that older children to be in the vanguard (see Kerswill and Williams 2005, pp. 1026–1032 for a fuller discussion).

We turn now to Principle 6. A marker of a new dialect is the absence of any stable, locally-based dialect to serve as a model for acquisition. This implies that there is a break in continuity between generations – not of language acquisition, which would lead to pidginization or creolization, but of the transmission of local dialect features. This is as true in new towns as in tabula rasa situations. The ”new dialect” status of Milton Keynes English can be confirmed by examining the diphthong /aʊ/ of *mouth*. There are a number of variants of this vowel, which appears to be converging on a Received Pronunciation-like [aʊ], moving away from local pronunciations such as [ɛɪ] and [eʊ].

The striking point is that there is an abrupt and complete disjunction between the variants, first between pre-Stage I (the original inhabitants of the pre-New Town district) and Stage I (the adult in-migrants), and then again between Stage I and Stage II: Note the fact that none of the Stage I or Stage II speakers use the two favored, conservative variants of the pre-Stage I people. A parallel study in Reading, a well-established town in the same region, showed a similar trend, but with the conservative tokens still in use, albeit at a low level, by the youngest speakers. This demonstrates continuity, absent from Milton Keynes.

Principle 7 can be illustrated, again using the *goat*-fronting data. Figure 1 shows the index score for all the 48 children, ranked from high to low. Against each child the caregiver’s score has been plotted. Two patterns stand out. The first is that there is no obvious link between the adults and children, with seven adults having scores close to zero (representing Scottish and northern English pronunciations). Secondly, the average score for the children is higher than for the caregivers, suggesting change (as we have seen). Further to this is the fact that three children have a score close to zero. This comes as no surprise when we realize that these are all four year olds, matching their caregiver. We assume from the data for the older children that these children will align themselves with the other children as they grow older; for the one child re-recorded 18 months later, this is indeed the case, on this feature as on others. The result is greater homogeneity among older children, a shift which is necessarily much greater in a new town than elsewhere, because almost of the adults are from elsewhere. This also, of course, supports Principle 8.
6 New varieties and migration

It is not usual to consider changes caused by immigration to an existing speech community a case of ‘new-dialect formation;’ however, in the last 20–30 years substantial and rapid realignments in the phonologies, and to a much lesser extent grammars, of urban dialects have been observed in north-west European metropolises. Recent work in Copenhagen (e.g. Quist, 2008), Stockholm (Bijvoet and Fraurud, 2006), Oslo (e.g. Aarsæther et al. 2008), as well as cities in the Netherlands and Germany, suggest such a development in highly multiethnic districts. We will take the example of London, a city with a long history of immigration (and in-migration from other parts of the British Isles). Various incoming groups are said to have influenced London English, but it is only with the very large-scale immigration post-World War II and especially since the late 50s that we see verifiable change. We will take the example of diphthongs in London English, which are usually said to be subject to Diphthong Shift (Wells, 1982, pp. 306-310). This means that the vowels of FACE, MOUTH, PRICE, and GOAT have long trajectories, a development which has been taking place over a century or more and is a further development of the Great Vowel Shift. Figure 2 shows vowel plots for an elderly working-class Londoner (Kerswill et al. 2008, p. 457). If one reads this as if it is a traditional vowel diagram, these characteristics can be clearly seen: The lines represent the trajectories of the diphthongs.

---

4 This work was carried out as part of the ESRC-funded project Linguistic innovators: The English of adolescents in London, 2004–2007, ref. RES-000-23-0680; investigators Paul Kerswill and Jenny Cheshire, research associates Susan Fox and Eivind Torgersen.
Fig. 2: Vowel plots for elderly male speaker from Hackney, born 1938.
Figure 4 shows the same vowel for a young inner-city Londoner, who is representative at least of speakers of Afro-Caribbean and West African ethnicities. The diphthongs, especially FACE, PRICE and GOAT, have much shorter trajectories, and FACE and GOAT are now high peripheral vowels. GOAT in fact shows a development which is the opposite of that found in Milton Keynes (and elsewhere in the south-east of England). The study showed that, in fact, all ethnicities, including white Anglos, could variably use this system, and that the strongest predictor was residence in the inner city (rather than the suburbs) and a friendship network which was highly multiethnic.

Contact, in this case, is in the first instance with the languages of the immigrants. However, the changes in London English as a whole are not driven directly by second-language speakers, but by the L1 English of their immediate offspring, which may well contain copies of L2 features such as near-monophthongal FACE and GOAT. It is, in other words, another case of dialect contact.

The question arises as to why this has happened now, and not at an earlier period. The answer seems to be that the rate of immigration in recent years has led to inner-city communities where children grow up learning English, but with limited access to L1 models. Instead, their models are L2-speaking adults and older children who themselves did not have an L1 model. Children from English-speaking families (of whatever ethnicity) also grow up in this environment, and acquire many of the features. This outline suggests a deterministic model. However, there is plenty of evidence that these varieties are strong social and identity markers (e.g. Rampton forthcoming).
7 Conclusion

In this chapter, I have attempted to show that new-dialect, or koine formation is a process which can be divided into several stages, at each of which speakers engage in certain behaviors and where social factors variably play a part. The overall model is broadly, but not completely deterministic. Even if outcomes are highly predictable, many different factors come into play before the outcome is achieved. Some of these, in the end, have no bearing on the outcome, while others clearly do.

References


Britain, David and Peter Trudgill 1999. ‘Migration, new-dialect formation and sociolinguistic refunctionalisation: reallocation as an outcome of dialect contact’, Transactions of the Philological Society 97, 245-256.


