

Sociolinguistic typology

Social determinants of linguistic complexity

PETER TRUDGILL

OXFORD
UNIVERSITY PRESS

OXFORD

UNIVERSITY PRESS

Great Clarendon Street, Oxford ox2 6up

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Cape Town Dar es Salaam Hong Kong Karachi

Kuala Lumpur Madrid Melbourne Mexico City Nairobi

New Delhi Shanghai Taipei Toronto

With offices in

Argentina Austria Brazil Chile Czech Republic France Greece

Guatemala Hungary Italy Japan Poland Portugal Singapore

South Korea Switzerland Thailand Turkey Ukraine Vietnam

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York

© Peter Trudgill 2011

The moral rights of the author have been asserted
Database right Oxford University Press (maker)

First published 2011

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the appropriate
reprographics rights organization. Enquiries concerning reproduction
outside the scope of the above should be sent to the Rights Department,
Oxford University Press, at the address above

You must not circulate this book in any other binding or cover
and you must impose the same condition on any acquirer

British Library Cataloguing in Publication Data
Data available

Library of Congress Cataloging in Publication Data
Data available

Typeset by SPI Publisher Services, Pondicherry, India
Printed in Great Britain
on acid-free paper by
MPG Books Group, Bodmin and King's Lynn

ISBN 978-0-19-960434-0 (Hbk.)

ISBN 978-0-19-960435-7 (Pbk.)

1 3 5 7 9 10 8 6 4 2

Mechanisms of complexification

In the previous chapter I presented evidence to suggest that there might be a link between type of community and the spontaneous development of linguistic complexification. It is, however, one thing to produce data to indicate that this might be the case, and another matter altogether to explain *why*, if there is such a link, this should be so.

Let us now therefore examine, in turn and briefly, the four different types of complexification, as defined and illustrated in Chapter 3, and see how this kind of development might work in a little more detail. What follows will not be an in-depth account of the linguistic mechanisms and processes which lead to complexity-development. I will not ask *exactly which* but rather *what sorts of* mechanisms might in principle be involved in particular types of complexity-development. I will do this in order to be able to ask what sorts of different sociolinguistic conditions might be of relevance in permitting or encouraging the operation of such mechanisms. While we will be looking at only one or two examples of each of the four different types of complexification, the mechanisms that appear to be involved in their development will hopefully give us some indication of what social conditions might favour the operation of such mechanisms generally.

Increase in opacity

We begin with increase in opacity. This type of increase can of course take a number of forms, but let us consider initially just one kind. We saw earlier that low morphological transparency is typical of fusional languages such as Faroese, as opposed to agglutinating and isolating languages (Dixon 2010a: 226ff.).

This point can be illustrated using well-known phenomena from archetypical exemplars of the three morphological types—isolating Vietnamese, agglutinating Turkish, and fusional Latin:

	'from a man'	'from men'
Vietnamese:	<i>tu nguoi</i> from man	<i>tu nguoi</i> from man
Turkish:	<i>adam-dan</i> man-abl.	<i>adam-lar-dan</i> man-pl.-abl.
Latin:	<i>homine</i> man + abl. + sg.	<i>hominibus</i> man + abl. + pl.

A comparison of paradigms shows that the forms of *adam* 'man' in Turkish are totally transparent:

	sg.	pl.
nom.	<i>adam</i>	<i>adam-lar</i>
acc.	<i>adam-i</i>	<i>adam-lar-i</i>
gen.	<i>adam-in</i>	<i>adam-lar-in</i>
dat.	<i>adam-a</i>	<i>adam-la-ra</i>
loc.	<i>adam-da</i>	<i>adam-lar-da</i>
abl.	<i>adam-dan</i>	<i>adam-lar-dan</i>

In Latin, on the other hand, segmentability is low, as was pointed out in Chapter 2:

	sg.	pl.
nom.	<i>homo</i>	<i>homines</i>
acc.	<i>hominem</i>	<i>homines</i>
gen.	<i>hominis</i>	<i>hominum</i>
dat.	<i>homini</i>	<i>hominibus</i>
abl.	<i>homine</i>	<i>hominibus</i>

The genitive plural of *homo* is *hominum*, but it is not possible to separate out distinct genitive and plural morphemes.

It is in fusional languages such as Latin that Kusters's (2003) *transparency principle*—the demand that the relation between form and meaning should be as transparent as possible—is most heavily violated. Any linguistic change, therefore, which leads a language in the direction of the development of a more fusional character will have as a consequence increased

opacity, and therefore complexification. And indeed one does not necessarily have to accept in full the validity of the concept of the morphological cycle (Hodge 1970), in which languages are said to move from an isolating type to agglutinating to fusional and back to isolating again, to agree that agglutinating languages can indeed become less agglutinating and more fusional over time.

As an example of this kind of change, we can take the following forms from Finnish, a well-known agglutinating language, in comparison to the same paradigms from the closely related but less agglutinating Finnic language Estonian:

	Finnish		Estonian	
	'leg'	'flag'	'leg'	'flag'
nom.sg.	<i>jalka</i>	<i>lippu</i>	<i>jalg</i>	<i>lipp</i>
gen.sg.	<i>jala-n</i>	<i>lipu-n</i>	<i>jala</i>	<i>lipu</i>
part.sg.	<i>jalka-a</i>	<i>lippu-a</i>	<i>jalga</i>	<i>lippu</i>
part.pl.	<i>jalko-j-a</i>	<i>lippu-j-a</i>	<i>jalgu</i>	<i>lippe</i>

The Finnish word-forms are readily segmentable because of the agglutinating structure. It can easily be seen that *-j-* is the plural morpheme, *-n-* the genitive, and *-a* the partitive. The stem remains constant throughout the paradigm, moreover, except for the small amount of allomorphy provided by the alternation in the genitive singular between *k* and \emptyset in *jalka*, and *pp* and *p* in *lippu*. Crucially, these alternations are entirely predictable. The Estonian paradigms, on the other hand, are much less transparent. The forms are not readily segmentable—the separate partitive and plural morphemes in Finnish have been fused in Estonian, for instance; and the alternations—for instance, between partitive plural *-u* and *-e*—are unpredictable.

It is known that the Estonian forms are more innovative than the Finnish, and that they developed out of originally more agglutinating forms similar to those of Finnish. We can see then that Estonian has experienced changes which have not occurred in Finnish, and that these changes have had the complexifying effect of obscuring an originally more transparent morphological structure. When we come to consider what processes might have been implicated in this development, it is clear that the changes involved were phonological in origin: all

original word-final segments, for instance, have been lost in Estonian.

If it is the case that fusional inflections develop out of the (originally phonological) fusing of earlier, more analytic agglutinating structures, as in Estonian—"fusional structures depend diachronically upon agglutinating ones" (Dahl 2004: 184); and if it is the case, as has been argued, that agglutinating morphology may in turn in some cases develop out of clitics, which have in turn developed out of independent lexemes through grammaticalization processes, then Comrie (1980) has given us some indication of how this might happen. He illustrates, from the Finnic languages, a typical path of development: noun > postposition > clitic > suffix > case ending. In the Balto-Finnic of the first millennium BC (Campbell 1997b), the form *kanssa* was an independent noun meaning 'people'. In modern Finnish it has become a comitative postposition:

hyvā-n poja-n kanssa 'good-GEN boy-GEN with'

Note that its independent lexical status is demonstrated in the fact that there is no vowel harmony in constructions with *kanssa*. In the related language Vepsian, on the other hand, a further stage of development has occurred, and the corresponding form *ka* has become a bound comitative suffix:

lahse-ka 'with the child' *lehmä-kä* 'with the cow'

The bound status of the form is indicated by the vowel harmony alternation between *ka* and *kä*.

In the Finnic language Karelian, on the other hand, the comitative form *kela* has become more than just a bound suffix: here there is not only vowel harmony but also agreement within the noun phrase (see below on the development of agreement):

kolme-n-kela lapše-n-kela 'with three children'
tütö-n-kelä 'with the girl'

The final stage in the development of fusional forms, where segmentability is lost, can be seen in the example already given above where Estonian *lippe* has lost the segmentability of Finnish *lippu-j-a*.

These few examples just hint at the sorts of developments that are required for the development of fusion and inflection; but what interests the sociolinguistic typologist is what kind of social matrix is most likely to accompany and produce this kind of opacity-producing development. In what kind of social conditions are we most likely to see the development of highly fusional, inflecting languages like, say Proto-Indo European, or Latin, or Polish, perhaps out of an earlier more analytic, agglutinating ancestor of Proto-Indo-European?

In fact, it is easy to see, when it comes to considering what kind of sociolinguistic conditions are necessary here, that relatively long periods of time will be required for these grammaticalization processes to run their course (see Chapter 6 on Dahl's (2004) introduction of the notion of the "mature phenomenon"). As Lichtenberk (1991) points out, grammaticalization is *gradual*. There can be no guarantee that other forms of opacity-development, such as increases in allomorphy, can be regarded in exactly the same way. But from the sociolinguistic-typological perspective, we can suggest that processes which lead to an increase in fusional structures are such that they must be allowed to run for a long time, uninterrupted by periods of significant adult language contact, in order to result in the type of outcome we are discussing.

Of the two sociolinguistic parameters we have so far mentioned in this book, *contact vs. isolation* and *stability vs. instability*, both would seem to be crucial. High levels of contact would lead to more, not less, transparent forms; and only stable situations will permit the uninterrupted development of mature, inflectional morphology in a language over the long periods of time required. At least this particular form of opacity-development, then, is more likely to develop in relatively stable, low-contact communities.

Irregularization

Let us now turn to irregularization and ask how exactly this, too, comes into being linguistically.

The answer to this question, as far as certain forms of irregularization are concerned, is already well known from the historical linguistics literature. One of the examples we

used in Chapter 3 to illustrate increase in irregularity and opacity came from Faroese. For ease of exposition I repeat the data here. Older Faroese had irregular nominal paradigms such as *dagur* 'day':

	sg.	pl.
nom.	[dagur]	[dagar]
acc.	[dag]	[dagar]
gen.	[dagás]	[daga]
dat.	[degi]	[døgum]

The complexity of the modern nominal system is now even greater:

	sg.	pl.
nom.	[dɛavur]	[dɛ:ar]
acc.	[dɛa]	[dɛ:ar]
gen.	[dagʂ]	[dɛ:a]
dat.	[de:ji]	[dø:vun]

We saw that Braunmüller pointed out that only the consonantal onset remains constant throughout the paradigm.

If we now wish to investigate the kind of mechanisms involved in producing this type of complexity-development, it is apparent that the increase in morphological opacity and irregularity in Faroese is due, once again, to phonology—to sound change. It is of course a historical-linguistic commonplace that regular sound changes can lead to grammatical irregularity (see Hock 2003: 450 on Sturtevant's Paradox); but from the point of view of sociolinguistic typology we can now expand this point somewhat further by considering the actual nature of the particular sound changes involved here in Faroese.

Most linguists would probably agree that the sound changes which have occurred in continental Scandinavian in recent centuries have been rather more natural and expected, and rather less complex, than many of those which have occurred in Faroese. For example, the vowel shift undergone by Swedish and Norwegian such that /a:/ > /ɔ:/ > /o:/ > /u:/ > /y:/ > /ɥ:/ strikes no linguist as being at all strange, especially after Labov's work on vocalic chain shifts (1994). And the lenition of intervocalic voiceless stops in Danish (and in some varieties

of Norwegian, Swedish, and Icelandic) is also a very common type of change:

p > b > β > w	Norw. /pi:pə/ vs. Dan. /pi:wə/
t > d > ð > ɖ	Norw. /ga:tə/ vs. Dan. /ge:ɖə/
k > g > ɣ > j	Norw. /ka:kə/ vs. Dan. /ke:jə/

This lenition is a change which has been repeated with high frequency in the world's language families, witness for example Latin *-atum*, Italian *-ato*, Spanish *-ado* > [-aðo] > [ao].

On the other hand, the Faroese sound changes implicated in the development of irregularity and opacity in the above paradigms are intuitively felt by many historical linguists to be rather unusual. And I would suggest that such intuitions are important and valid—when Bailey (1982) writes of languages with “incredible” consonant clusters, this is a legitimate observation (see further Chapter 5). The Faroese changes involved in this increase in morphological opacity include:

(1) the well-known “Verschärfung” or ‘sharpening’, whereby rather startling changes occurred which led to correspondences between modern Danish and modern Faroese such as:

Danish	Faroese	
/ku:/	/kigv/	‘cow’
/ny:/	/nudʒ/	‘new’

Sandøy (2005: 1867) explains that this development occurred during the 18th century. In diphthongs where [i] or [u] was the second element in contexts before an immediately following vowel, these second elements became respectively a palatal affricate and a “velar cluster”. Thus *oyin* ‘the island’ became *oyggin* [ɔjʒ:ɪn], and *bua* [bæu:a] ‘live’ became *bugva* [bɪgva]. Then, by analogy, “many words transferred the sharpened form to the final position in the other forms of the word, cf. *oyggi* ‘island’. We can understand up to a point how this happened, but there is perhaps still something slightly mysterious about vowels becoming consonants in the case of [i] > [jʒ] and [u] > [gv].

Note also that this change is not totally dissimilar to the change often known as Holtzmann's Law, whereby a stop was

introduced into geminate *j* and *w* clusters in Gothic and in Old Norse, giving ON *tryggva*, cf. OHG *triuwa*, 'troth'; ON *tveggja*, cf. OHG *zweio* 'of two' (Scardigli 2002: 557). The sociolinguistic typologist cannot help but note, of course, that this change took place at a time when the ancient Germanic languages would also have been spoken only by relatively small groups of speakers.

(2) Diphthongizations also took place in Faroese in the period 1350–1550 (Haugen 1976: 256) which gave rise to forms such as /luyk/ from earlier /li:k/ 'like'. Note that some of these diphthongizations led to mergers:

i: > ui u: > yu
y: > ui ø: > øe
æ: > æa o: > ou
e: > æa â: > âa

(3) *ð* was lost, and *g* was lost intervocally and finally, by about 1600. This gave rise to many new instances of vowel-vowel hiatus. Subsequently a new "intrusive glide" developed in many such cases of hiatus, with the identity of the glide depending on the quality of the two vowels in contact:

maður > [mæa:vʊɹ] 'man'
blair > [blæa:jɹɹ] 'blue' (pl. masc.)
skógur > [skou:wʊɹ] 'forest'.

Again, the development of the intrusive consonants /j/, /w/, and especially /v/, is not as expected as the earlier loss of /ð/ and /g/.

It is possible to argue, then, that not only is the increased morphological complexity of Faroese due to sound change, as is so often the case, but perhaps more especially that in this case, at least, it is also due to the fact that at least some of these sound changes can be classified as "unusual" (see more on "unusual" sound changes in Chapter 5).

Furthermore, it is very encouraging for a sociolinguistic-typological study of the social matrices in which irregularization and opacity-development occur to note that Henning Andersen (1988) has also proposed a serious sociolinguistic correlate of the development of marked as opposed to unmarked sound changes (see further Chapter 5). Andersen points

to unusual sound changes in dialects which "are located in peripheral dialect areas, away from major avenues of interdialectal communication", and his hypothesis is that "there is a connection between the limited socio-spatial function of a dialect, its relative closeness, and its ability to sustain *exorbitant phonetic developments*" (1988: 70; [my italics]).

It may indeed be the case, then, that isolated communities are more likely to produce changes that could be labelled, in Henning Andersen's (1988: 70) careful words, "slightly unusual". Andersen argues that socially and geographically peripheral speech communities are more prone to develop phonologies with elaborate phonetic norms and the proliferation of low-level pronunciation rules: "dialects that serve predominantly local functions are more prone to elaborate phonetic detail rules than dialects with a wider sphere of use".

Notice that while we are clearly dealing here with the relevance of our first social parameter *contact vs. isolation*, we now also seem to be focusing as well on an additional parameter, so far not overtly discussed in this book. In an important paper, Braunmüller (1984) has argued that morphological opacity is a typical characteristic of *small languages*. Hymes, too, says that "the surface structures of languages spoken in small cheek-by-jowl communities so often are markedly complex, and the surface structures of languages spoken over wider areas less so" (1974: 50). Thurston, too, has argued (e.g. 1989—and see also Chapter 2) that there is a lower degree of variation in small traditional communities characterized by *esoterogeny*—that is, they are not acquired by non-native speakers—and that this promotes the development of morphological irregularity. And indeed there is further evidence in the literature—see for example Haudricourt (1961)—that community size may be linguistically important. There is a statistical, if inconclusive, discussion of the relevance of this parameter in Wichmann et al. (2008). More helpfully, in a detailed statistical paper which takes Trudgill (2004c) as one of its starting points, Sinnemäki (2009: 139) concludes that "by and large, the present paper indicates that language complexity is not necessarily independent of sociolinguistic properties such as speech community size".

Certainly, a number of the communities we pointed to in previous chapters as being isolated and stable, and therefore demonstrating slow rates of change and maintenance of complexity, also have languages that have always been spoken by relatively small groups. For example, there are today approximately only 45,000 Faroese and 225,000 Icelandic speakers, as opposed to 5 million Norwegians, 6 million Danes, and 8 million Swedes. Sandøy (2001: 127), moreover, writes that the population of the Faroes from the Middle Ages until almost 1800 was only about 5,000, divided into 40 or 50 settlements, giving about 100–125 people per settlement; and in mediaeval times, the population of each Icelandic settlement averaged only 7–10 people.

I therefore now add to this study of the relationship between social and linguistic typology a third, apparently relevant social parameter, namely *small vs. large community size* (in terms of numbers of speakers).

Nettle suggests a statistical explanation for this parameter. He points out that, for example, object-initial constituent order is exceedingly rare in the world's languages. Indeed, it was unknown to academic linguistics generally until the 1970s, when Desmond Derbyshire, with the crucial encouragement of Geoff Pullum, published a short but exceedingly important paper on OVS order in the Amazonian language Hixkaryana in the widely read *Linguistic Inquiry* (Derbyshire 1977). Nettle points out that all the languages which have this feature are spoken by small or very small numbers of speakers: the median number is 750, compared to 5,000 for the world's languages as a whole. He suggests that non-optimal orders are "more likely to be found in small communities than in large ones, since these would be more vulnerable to drift away from optimal states" (1999: 139). In population genetics, Nettle says, the effects of random change are known to be greater when the population is small, for statistical reasons. "This is because the probability of a slightly deleterious variant becoming fixed in a population is inversely related to the population size. The smaller the community, the greater the stochastic chance of changes in gene frequency" (Nettle 1999: 139). Nettle then hypothesizes that the same might be true of linguistic communities and linguistic features.

If we wonder what might be "non-optimal" about OVS and OSV order, it is relevant that Givón (1984: §7.3) argues that SOV is in some sense the basic order—and indeed the earliest pattern to be found in human language—and is favoured by factors to do with the role of the position of agent/topic and goal/object in the origins of human communication; he also argues that diachronic development to SVO, VSO, or VOS, where this has occurred, has been favoured for reasons of a discourse-pragmatic nature (1984: §7.11). OVS and OSV, however, are not favoured in either way.¹ (We can perhaps also extend this statistical account of Nettle's to other "non-optimal", "less usual", or "exorbitant" linguistic features.)

The three different social parameters that have now been established so far—contact, stability, and size—can very readily be linked. Although our discussion of simplification in Chapter 3 focused on contact vs. isolation, my proposal now is that in fact simplification is most likely to be found in communities which are characterized not only by high levels of contact but by a particular *complex* of all three of these social parameters.² Linguistic simplification, I suggest, is most likely to be found in communities which demonstrate *high contact* (of the post-critical threshold type), *social instability*, and *large size*. Correspondingly, therefore, spontaneous complexification is most likely to emerge in communities characterized by *low contact*, *social stability*, and *small size*.

But why should this be so? The answer, I propose, lies in the seminal work of James and Lesley Milroy (and work following on from theirs—see for example Ross 1997). A key social distinction adumbrated in Milroy and Milroy (1985) and J. Milroy (1992b) is between communities with dense, multiplex social networks—in layman's terms, communities where it is common for everybody to know everybody else, and where your neighbour and your second cousin and your workmate may be one and the same person—and communities with loose networks, where the reverse is the case. Zabrocki (1963), too,

¹ Dixon's (2010b: 75) caveat should be noted here: many languages do not have a fixed constituent order, and even for languages which do the matter is "of only marginal interest for basic linguistic theory".

² It is important to stress this point because certain commentators on my earlier work have focused on one or other of these factors to the exclusion of the other two—see Chapter 6.

distinguishes between tight and loose communicative communities. How tight a communicative community is, he says, depends on the speed and frequency of the communication between its members.

From the work of the Milroys, notably in Belfast, it can be concluded that dense networks lead to strong social ties, which then lead to closer maintenance of community norms—in language as in other forms of behaviour. On the other hand, loose networks lead to weaker social ties and so to a relative lack of maintenance of community norms. The Milroys argue, surely correctly, that “linguistic change is slow to the extent that the relevant populations are well established and bound by strong ties, whereas it is rapid to the extent that weak ties exist in populations” (1985: 375; see also L. Milroy 2000).

Dense social networks are most likely to be found in small, stable communities with few external contacts and a high degree of social cohesion. As Croft (2000: 192) points out, “a close-knit network is of necessity small in size: it is difficult for everyone to know and talk to everyone else in a large society”. Loose social networks are more liable to develop in larger, unstable communities which have relatively many external contacts and a relative lack of social cohesion. Linguistic change is liable, other things being equal, to be faster in larger than in smaller communities;³ and instability is also associated with faster rates of change, as we saw in Chapter 1.

We can now see that the role of instability is due to the weakening of social network ties in situations of social breakdown and chaos. Raumolin-Brunberg, for example, in her discussion of the rapid changes in English we discussed in Chapter 2, specifically claims that the period during and after the English Civil War “must have witnessed a considerable increase in weak ties” and actually gives chapter and verse from the content of her data: “many of the letters in the CEEC [*Corpus of Early English Correspondence*] in fact testify to the splitting up of families and broken neighbourhood ties”. In her work on the loss of multiple negation in English, Nevalainen (1998: 281) also ascribes a role to social networks: “a Milroyan type of weak-ties

³ Taking size as their only criterion, however, and without looking at networks, Wichmann et al. (2008) find no evidence for or against this hypothesis.

network structure could well have been the means of spreading the loss of multiple negation”. This is based on research also involving data from the *Corpus of Early English Correspondence* using texts from the period 1417 to 1681.

This insight into the social-network causes of differential *speeds* of change, as discussed in Chapter 1, can now also be extended to differences between *types* of change. Given that there is a strong tendency for relative density of social networks to correlate with community size, we can note the comment made by Grace (1990: 126), who writes:

A language exists in the people who speak it, but people do not live very long, and the language goes on much longer. This continuity is achieved by the recruitment of new speakers, but it is not a perfect continuity. Children (or adults) learning a language learn it from people who already speak it, but these teachers exercise considerably less than total control over the learning process.

We must accept that no ‘teachers’ exercise total control, but this perspective does also suggest that, because of differences in social network structure, there is a possibility that in smaller communities the ‘teachers’ have more control than in larger societies. Because of this, small tightly-knit communities are better able to encourage the preservation of norms and the continued adherence to norms from one generation to another, *however complex they may be*; and the absence of external contacts and social instability will also strengthen a community’s ability to maintain its own linguistic complexity.

If this is so, then it is not unreasonable to suppose that these same societal factors may also assist in the *production* of complexification. In small, isolated, stable communities, linguistic change will be slower. But when it *does* occur, there is a greater chance that it will be of the complexification type—the other side of the coin from high contact and loose networks leading to rapid change and simplification. Small isolated communities are more able, because of their network structures, to push through, enforce, and sustain linguistic changes which would have a much smaller chance of success in larger, more fluid communities—namely changes of a relatively marked, complex type. Indeed, it may even be that, as Nettle

(1999: 138) says, "if a group consists of just a few hundred people, the idiosyncrasies of one very influential individual can spread through it very easily".

So not only is there less simplification in low-contact situations, there is indeed also more complexification, as argued in the previous chapter. That is, it may well be that *innovations* of a complexification type occur with roughly equal frequency in all types of community, but that it is simply the case that these innovations are likely, perhaps much more likely, to succeed and become established linguistic *changes*, i.e. innovations which are accepted and become permanent, in small isolated communities. For opacity, for example, we can say that innovations which render forms less transparent may well develop in all types of community, but they are much more likely to be successful ultimately in communities with tight social networks. And we can see why it is in larger language communities that irregular verbs are more likely to become regular, while the reverse may occur in smaller communities—irregularization is just less likely to succeed in larger societies.

This opens up the possibility of applying the above general account of the social correlates of complexification to particular cases. Regular sound change can lead to grammatical irregularity, but the more "exorbitant" the sound change, the greater the irregularity is likely to be. The sort of remarkable irregularization found in Faroese is the result of the kind of "exorbitant phonetic development" that is more likely to occur in small isolated communities because it is in such communities, as I just argued, that social network structures make for greater likelihood of complex changes being pushed through, enforced, and sustained from generation to generation.

Maybe other forms of irregularization such as Norfolk dialect *wrapped* > *wrop* will not be so susceptible to the same kind of analysis, but it does now seem that we have found an additional, fourth social parameter to add to our list of factors that may be implicated in producing different types of linguistic structure, namely *dense vs. loose social networks*.

Growth of morphological categories

If we now turn to the spontaneous growth of morphological categories, we saw in the previous chapter that there seemed to be a greater likelihood of this type of complexification, too, occurring in certain types of community. But what mechanisms can be involved in this kind of development? And what social contexts would favour the operation of such mechanisms, and how would they do so?

If we turn to the examples I used in Chapter 3, we can see that the same kind of mechanisms were at work in a number of the cases discussed there. One of these was the development, as an innovation in southwest of England dialects, of the distinctive marking of intransitive as opposed to transitive infinitives: intransitives take *-y*, transitives are unmarked: *he can't swimmy* vs. *he can't hit it*.

How this situation came about diachronically seems to be reasonably clear. Once again, phonological change was crucially implicated in what eventually became a grammatical development. The intransitive *-y* ending is considered to be a relic of the Middle English infinitive ending *-en*, later *[-ə]* or presumably, in the southwestern dialects in question, *[-ɪ]* (Ihala-lainen 1991). Originally, all infinitives would have carried this vowel. Later, as is well known, the final unstressed vowel was eventually lost in most dialects. Presumably, while this loss was occurring, there was a period of variability of the type that normally accompanies change, with alternation between older forms and newer forms with and without the final vowel, respectively. A useful hypothesis would be, however, that during this period the vowel was originally lost less often in utterance-final position than when another word, e.g. an object noun, followed.

Support for this interpretation can be gained from observations of similar developments in Scandinavian dialects. Torp (2003: 249) shows how in certain of the Swedish dialects of Bohuslän, the unstressed final vowels of infinitives preceding a direct object have been subject to reduction to *[-ə]*, while in certain other contexts, such as before an adverbial phrase, the original unreduced vowel *-a* is retained:

Ve feck järe de vi kunne
 we got to-do that we could
 'We had to do what we could'

vs.

Va feck du jära där på Strändräng?
 what got you to-do there at S.
 'What did you have to do there at S.?'

Similarly, Torp also cites Christiansen (1948: 191), who shows that in the Norwegian dialect of Vefsn a similar process has gone even further: infinitive endings alternate between -Ø and -a depending on sentential context, just as I am supposing happened in the English southwest:

Han sku kām ijæn 'He should come back'

vs.

Kanj du kāmā? 'Can you come?'

We can then suppose that it was at this kind of variability stage, in the southwestern English dialects in question, that the difference between forms with and without the vocalic ending eventually became reinterpreted in such a way that a difference without a distinction became a difference with a distinction. The difference, which was originally phonologically conditioned and variable, came to be reinterpreted as grammatico-semantic—transitive vs. intransitive—and categorical. Crucially, it was the frequent presence (vs. absence) of following object nouns which led to this transitive (vs. intransitive) reinterpretation of forms without -y.

This is a form of reanalysis (Harris and Campbell 1995: 30) or *exaptation* (Lass 1990; 1997): formerly meaningless differences end up being employed to make meaningful distinctions. Exaptation, as Lass (1997: 318) says, can lead "to the development of new grammatical categories", and he further argues (1997: 319) that exaptation is "conceptual invention" and "what was once a predictable alternation ... can be ... reanalysed as a new primary categorical marker"—which is precisely what has happened here.

A similar argument can be made for the development of the distinction between the two different sets of definite articles in

North Frisian. It will be recalled from Chapter 3 that in the Mooring dialect the two sets of forms are:

masc.	<i>di</i>	<i>e</i>
fem.	<i>jü</i>	<i>e</i>
neut.	<i>dât</i>	<i>et</i>
pl.	<i>da</i>	<i>e</i>

and that the -e/-et forms are proximal and/or refer to a unique referent, such as *e moune* 'the moon', while the *di/jü/dât/da* forms are distal and/or are context-bound and apply to definite but non-unique referents.

The forms in the left-hand column descend directly from the nominative definite article forms of Old Frisian: *thi, thiū, thet, tha* (Bremmer 2009: 54). As Hoekstra (2001: 777) indicates, the forms in the right-hand column are in origin "weak variants" of the forms in the left-hand column, i.e. they were originally phonologically reduced forms which occurred in unstressed environments. So in origin they were simply phonologically conditioned and therefore predictable variants, as they still are in many dialects. However, in the Mooring and other innovating dialects, the two types of variant have been reanalysed as having distinct semantic-grammatical functions, and refunctionalized as applying to non-unique (etc.) vs. unique (etc.) referents (Löfstedt 1968). Of course, "have been reanalysed" does not really answer the question as to why the exaptation took the form that it did—why the variants have been reanalysed in this particular way: strong forms = distal, weak forms = proximal. But we can perhaps hazard a guess that it might have had to do with the more demonstrative-type function of the stronger unreduced, perhaps more frequently stressed, variants.

Given that this type of diachronic exaptational refunctionalization of variants is common enough in linguistic change generally, the question then is: why might reanalysis, leading to morphological-category development, be more likely to occur in some social matrices than others?⁴ The answer would appear

⁴ I do not intend to suggest that reanalysis is necessarily the only mechanism leading to new morphological categories, although I currently have no suggestions as to what the other mechanisms might be.

to be that this is the type of change which naturally takes place when a language is not in contact with other languages—changes which Bailey (1982) refers to as *connatural*. We can suppose that it is more common in low-contact, smaller, stable communities because it requires some generations of uninterrupted native-speaker development for completion. As Croft (2000: 193) puts it, “linguistic isolation allows for processes of change to evolve to an elaborate degree that would otherwise be curtailed by levelling or simplification in a larger, more loose-knit society”.⁵

Exaptation of this type, however, will be less common in high-contact communities because it makes for greater L2 difficulty.

Increase in syntagmatic redundancy

Finally, let us now turn to the fourth type of complexification—the tendency for syntagmatic redundancy, or repetition of information, to become more common in what we can now describe as being smaller, more isolated, low-contact communities with dense social networks.

This, it has to be said, is less obviously explicable. Indeed, it might have seemed more logical if it were the other way round. In small tightly-knit communities where there is a higher degree of shared knowledge and information, one would have thought that less repetition would be necessary, not more. The clue, however, may lie in the word “necessary”. Linguistic change is not teleological (Lass 1990; 1997: 340ff.); changes are never “necessary”—they just happen. And the development of agreement, the most common form of syntagmatic redundancy, does seem to be an extraordinarily common process in linguistic change. According to Corbett (2006: 1), agreement is “a widespread and varied phenomenon”.

Part of this problem of agreement-development being less obviously explicable than other changes may have to do with the fact that, as Corbett says, the functions of agreement are by no means totally clear, and “agreement often appears to

⁵ For more on the social matrices that favour the growth of morphological categories, see Chapter 6.

involve a lot of effort for a questionable payoff” (2006: 274). Acuña-Fariña (2009: 390) also writes that “maybe the greatest puzzle about agreement systems lies in their apparent arbitrariness and uselessness”. And so it is hardly surprising that pidgins, creoles, and other high-contact language varieties tend to manage without it, and that adult language contact tends to reduce it.

In this chapter we are trying to produce insights into the social settings which favour complexification by considering some of the linguistic mechanisms involved in its production. We therefore need to consider how agreement develops. There is actually a considerable consensus in the literature about the nature of the development of at least certain types of agreement. Lehmann (1988: 59), for example, argues that “the most important and most regular diachronic source of agreement is pronominal anaphora (including cataphora). More precisely, agreement markers usually stem from pronouns”. Corbett (2006: 264) agrees, saying that “it has long been accepted that pronouns provide a major source of agreement morphology, progressing from full pronouns, to clitics, to inflections along a well-established grammaticalisation path”, with the result that referential pronouns become agreement markers. Number and gender agreement arise specifically from the grammaticalization of third person pronouns, which get cliticized and then become affixes. From original structures where there was, say, a cliticized subject pronoun as well as an overt subject, the grammaticalization process produces a situation where the overt subject now agrees with the verb by means of the newly grammaticalized marker.

Where the development of agreement markers out of pronouns is relatively recent, there may be clear evidence that this is what has happened. In Manambu, a Ndu language of Papua New Guinea with about 2,500 speakers investigated by Aikhenvald (2008b: 51), there is phonotactic evidence for this hypothesis. In monosyllabic monomorphemic words, “if C_1 and C_2 have the same place of articulation, the coda (C_2) cannot be more voiced than the onset (C_1). That is, words like ***pab* and ***tad* are not well-formed”. However, this constraint is not operative when the final consonant is an agreement marker, thus *ta:d* ‘he stands’, which is bimorphemic

and where the *-d* is the third person masculine singular marker, is well formed. Aikhenvald argues (p.c.) that words which contain a morphological boundary behave differently from monomorphemic words because agreement markers such as *-d* are derived from independent pronouns. It was not so long ago since *ta:d* was actually two words, she suggests, so

a synchronic violation of the constraint on distribution of voiced and voiceless consonants within a monosyllabic word of CVC structure can be used as a piece of evidence in favour of the origin of agreement markers from independent pronouns.

Indeed, the pronominal origins of agreement markers can be seen not only in cases of verb agreement but also in some instances of the remarkable phenomenon of complementizer agreement that was noted in Chapter 3. Haegeman (1992), for instance, gives Flemish examples such as:

K'peinzen dan-k ik morgen goan
I think that-1sg. I tomorrow go
'I think I'll go tomorrow'

where it is clear that the *-k* postposed to the conjunction *dan* is historically derived from the pronoun *ik* 'I'.

Not all cases of agreement-development, however, involve pronouns. Aikhenvald argues that different types of agreement develop in different ways diachronically. She concurs that verbal agreement "usually arises from an anaphorically-used independent pronoun which is subsequently cliticised to the verb; it then becomes incorporated as an agreement marker. The mechanism is basically pragmatic, and it involves reanalysis and grammaticalisation of independent members of a closed class" (2000: 392).

But the development of gender agreement within a noun phrase can be the result of two different mechanisms, as proposed by Heine and Reh (1984: 230–1). In the first, the path for creating gender agreement does indeed once again include pronouns: it "involves the pragmatic use of third person pronouns as a kind of 'afterthought' to specify additional information" (Aikhenvald 2000: 393). In the second, however, "agreement may arise if demonstratives develop into subordination markers, e.g. relative clause markers, as the

result of their grammaticalisation" (2000: 392). She cites an example from the work of Heine and Reh (1984: 231) from the Nilotic language Maasai as follows:

in-kishú na-á-apshana na-á-ibor
PL-cattle REL-PL-seven REL-PL-be.white
'seven white cows'

The key mechanism in all these cases, then, seems to be some form of grammaticalization, although there are many aspects of developments of this type which are not fully understood. The extent to which grammaticalization is the result of pragmatic and cognitive (Heine et al. 1991), discourse (e.g. Comrie 1980), morphological, syntactic (see Li 1975), or phonological (e.g. Comrie 1980) processes is an important and unresolved question, although factors of all these types would appear to be involved. And Joseph (2001, 2004) has argued that grammaticalization is not really a "process" as such at all. Rather,

most of what is observed and labelled as such under this rubric is best taken as a result of independently needed and recognised processes of language change, namely (phonetically driven) sound change, (morphologically and conceptually driven) analogy, semantic change (especially metaphor), reanalysis, and borrowing, and not as a separate process in and of itself. (Joseph 2009: 199)

Fischer (2009) agrees about grammaticalization: "as a process, it is an analyst's generalization, a convenient summary but not something that has actually 'happened'" (2009: 18). (For a critique of these views, see Wischer 2006.) Dahl (2004) argues that phonological change plays a crucial role in grammaticalization: "I see changes in structure as typically driven by phonological changes". And Harris and Campbell (1995: 67) agree that grammatical reanalysis can be triggered by phonological change: they show, for example, that in some Nuclear Micronesian languages such as Ponapean (Pohnpeian), noun incorporation has developed in certain constructions as a result of the phonological erosion of verb forms. (See more on the importance of phonology in Chapter 5).

To return to the genesis of agreement, however, it turns out that there are many other more perplexing cases where gram-

matalization does not seem to be involved at all. For example, the Bavarian examples of complementizer agreement cited in Chapter 3 do not fit into this scenario. In forms such as:

... obst du noch Minga kummst
whether you-sg. to Munich come
'... whether you are coming to Munich.'

the agreement marker *-st* on *ob* 'whether' is clearly not the result of the grammaticalization of a pronoun—it is simply a repetition of the person marker that originally appeared only on the verb *kumm-st*. So how and why does this repetition develop? Why would a verbal person-marker end up attached to something which is not a verb?

The same question can be asked of the scenario outlined by Marchese (1988). She examines the development of agreement systems via noun class markers in the Kru languages of Liberia and Ivory Coast, and suggests that the trajectory is: subject pronoun > definite marker > (noun) class marker, with the class marker "eventually collapsing into the stem" (1988: 339). Beginning with original proto-Kru topic-comment structures such as

li ε ku mɔ
spear it is over-there
'The spear, it's over there'

Marchese hypothesises that the subject pronoun then undergoes cliticization and reanalysis as a definite marker on the noun. For example, in Godié the outcome is:

li-ε ku mɔ
spear-DEF is over-there
'The spear is over there'

But how exactly does this definite marker turn into a noun class marker—where does the actual *agreement* come from? The degree to which agreement-development is such a poorly understood phenomenon is nicely illustrated by Marchese's account of how this happened: "once the reanalysis had taken place in this position, the definite suffix would eventually make its way onto other NPs in the clause" (1988: 339). Yes, but how and why do such markers "eventually make their

way" onto further elements, something which has to happen if agreement is going to develop? Similarly, why did Karelian speakers start saying *kolme-n-kela lapše-n-kela* 'with three children'?

Some mechanisms that may lie behind these developments have been proposed. For the very rare Bavarian, and similar, complementizer agreement cases, De Vogelaer and van der Auwera, (2010) argue for "analogical extension". And De Vogelaer (2010) describes 'pathways' via which such extensions occur, again appealing to analogy. But do we understand why and how this "extension" would happen? And what exactly is "analogical" about it? If we are to claim that person markers are added to complementizers by analogy with person markers on verbs, we have to be clear that this is surely not analogy as it is usually understood, since this normally involves proportional relations.

Andersen similarly suggests that agreement-development of this type stems from what he calls *expression doubling*—"expression doubling is presumably the typical source of concord and agreement" (2008: 26).⁶ But is this actually an explanation? Why are expressions "doubled" in this way? What exactly happens?

Similarly, Dahl writes of the development of gender agreement in Indo-European languages that one of the major problems is how to explain "how gender markers come to be obligatory parts of e.g. adjectives and verbs" (2004: 199). The standard theory, he says, and as we saw above, is that gender markers derive from pronouns. "Then one has to assume a stage where there is obligatory 'redundant' or doubled use of these pronouns" (2004: 199). But what he does not say is why and how this doubling happens.

The answer seems to be that experts in the field do not yet fully understand what happens. Corbett's book (2006) on *Agreement* consists of 284 pages of text, only five of which are actually devoted to origins. As Corbett states, "agreement remains deeply puzzling" (2006: 1).

Perhaps we can content ourselves for the time being with noting that human languages are "like that". And, happily, it is

⁶ Comrie (p.c. 2009) has similarly suggested that apposition might be involved.

rather clear why languages should be “like that”. Aikhenvald has hit the nail on the head when she points to “pragmatic use” and to the fact that “the mechanism is basically pragmatic”. As she explains it, a prerequisite to the emergence of some forms of grammatical agreement is likely to be a repetition technique, whose function is “to clarify what is being talked about” (p.c. 2009). Redundancy, in the form of “expression doubling”, develops because redundancy is very helpful to human speakers in actual, real-life discourse. We have already noted in Chapter 3 the optional repetition that occurs to different extents in various Swiss German dialects (Berthele 2006), as in

und iez tuet er ufe baumstamm uufchlädere (p. 183)
and now does he up-a tree-trunk up-climb
'and now he climbs up a tree trunk'

And I already suggested there that this helpful discourse strategy could be a stepping stone towards obligatory, grammatical repetition.

In the examples cited in Chapter 3 from Berthele (2006), the repetition illustrated in Swiss German cases such as:

de hirsch hed ne da übernes bord abbegrüärt ine täich ine
(p. 184)
the stag has him then over-a bank down-thrown into-a pond into
'the stag then threw him over a bank down into a pond'

is simply a discourse phenomenon, where the speaker apparently has some choice in the matter. On the other hand, the Flemish examples of person agreement on answer particle cited from De Vogelaer (2005), such as

Zullen we gaan? Jom. 'Shall we go? Yes [1pl.]'

represent a later stage of agreement-development (with the *-m* derived diachronically from the clitic form of the first person plural pronoun *me*) where the person marking has become compulsory.

Taking all this into account, the widespread occurrence of agreement in the languages of the world implies that, if linguistic change proceeds without the large-scale involvement of

adult non-native speakers in contact situations, agreement is rather likely to develop over time even in constructions where there was none before. And if repetition—redundancy, expression doubling—for the purpose of “clarifying what is being talked about” occurs frequently enough in discourse, then it may become grammatical and thus compulsory. But, it seems, it will only do that if, as it were, it is given enough time to do so. That is, agreement, too, can be considered to be a *mature phenomenon* (see Chapter 6).

Agreement-development is once again, then, the type of change described by Bailey (1982) as *connatural*—the sort of change which takes place when languages are not in contact with other languages. We can suppose that this type of development is less common in large fluid high-contact communities because, as we said above, repetition of information means that there is more for non-native speakers to learn—it makes for L2 difficulty/outsider complexity. And it is more common in smaller, stable communities because complexity-development of this kind needs many generations uninterrupted by contact to come to fruition.

In Chapter 6 there will be a further discussion of the implications for sociolinguistic typology of these complexity-producing mechanisms.