

ETHNIC, RACIAL AND TRIBAL

The language of racism?

Ramesh Krishnamurthy

Source: C. R. Caldas-Coulthard and M. Coulthard (eds.), *Texts and Practices: Readings in Critical Discourse Analysis*, London: Routledge, 1996, pp. 129–149.

Each one of us is exposed to a wide variety of language input in our daily lives, some of it of our own choosing and some not. This input helps to shape our knowledge and understanding of both the language and its users. However, the predominant attitudes and opinions expressed in the language may also shape our thinking.

This chapter is divided into three sections, reflecting three major language sources: the media, dictionaries, and a large language corpus. The media are responsible for enormous amounts of language output, which must have a substantial influence on the language community they serve. Written texts have a great impact because they can be read and re-read by the consumer, shared with friends and colleagues, photocopied or faxed, and once they are archived, acquire permanency and public accessibility. Spoken media – radio and television – are not as permanent or as readily accessible, though some resources do exist, such as sound archives, and television and film libraries. However, one cannot discount their influence: 20 million people in Britain might watch a popular television programme, whereas the best-selling national newspaper, the *Sun*, has daily sales of around 3.5 million (*Guardian*, 12 July 1993: 15). Another feature of spoken media is that, for example in news broadcasts, the same terminology and phraseology is often used several times a day, and may be further recycled in current affairs programmes, weekly reviews, and so on. Many programmes are repeated in their entirety, and their content, vocabulary and style are frequently the subject of both public discussion and private conversation. The first section of this chapter therefore looks at four newspaper articles, as a small sample of media output.

Dictionaries are widely regarded as arbiters of linguistic usage. Until the late 1980s, they were based largely on the intuitions of lexicographers,

supplemented by reference to existing dictionaries, and sometimes bolstered by citational evidence. Lexicographers' judgements dictated the inclusion policy, the amount of space allotted to each entry, the division into discrete senses, etc. Most dictionaries are therefore far from being the objective records of the language that they are popularly conceived to be, although they may vary in the degree to which their editorial motivations are transparent or concealed, conscious or subconscious.

With the advent of large-scale computerised language corpora, lexicographers now have access to numerous examples of usage for thousands of the commonest words in the language. The *Collins COBUILD English Language Dictionary* (CELD, 1987) made a start in this direction, using an 18-million-word corpus of modern English texts. However, the assumption that corpus-based dictionaries must of necessity be less subjective is open to question. From my own experience at COBUILD, I suspect that even with a large corpus no 'purely objective' account of a language is possible.

COBUILD quite openly adopted a non-neutral position on sexism and racism. Placing the 'homosexual' meaning of 'gay' first went against corpus frequency, but was felt to be appropriate, partly because corpus evidence is always slightly out of date, and partly because English language learners (at whom the dictionary was primarily aimed) needed to be warned off using the word in its older meaning. With contentious usages (such as 'hopefully' as a sentence adverb) the evidence was recorded neutrally, but a cautionary note was added to the effect that some people (especially teachers and examiners) might regard this as incorrect usage, however well-attested. However, as with the other dictionaries, many decisions on inclusion, space allocation, and sense discrimination were still a matter for personal lexicographic judgement by the editors.

The second section therefore looks at the entries in several dictionaries, both for learners and for native-speakers. The third section looks at a 121-million-word sample corpus from COBUILD's current Bank of English corpus-building initiative, and the fourth section reviews the extent to which the newspaper articles and the dictionary entries conform to or deviate from patterns attested in the data.

Newspapers

The starting-point for this chapter was my initial reaction to two brief newspaper articles reporting on similar conflicts in former Yugoslavia and Kenya. The first was in a local Birmingham newspaper:

Belgrade civil war looms as crisis worsens

BELGRADE – Yugoslavia's leaders held a crisis meeting yesterday as the multi-ethnic state continued its slide toward anarchy and possible civil war.

It was a last-ditch bid to end a row over whether Yugoslavia should remain a federation dominated by Serbia or become a collection of independent states.

Nationalists in Croatia and Slovenia are threatening to break away and paramilitary units in the regions have refused to surrender arms, sparking fears of civil war.

A deadline of January 19 has now been set for the groups to lay down arms.

Dominated

Yugoslavia's eight-man presidential council, dominated by Communists, was appealing for unity combined with economic and political reform.

But it seemed a forlorn hope among the nation's 23 million people, who belong to six main ethnic groups and three major religions and are suffering economic hardship.

They were thrown together when the map of Europe was redefined at the end of the First World War.

There are fears of a civil war or a crackdown by the presidency's military, a bastion of hardline Marxism committed to a federation based on Socialism.

(Birmingham Daily News, 11 January 1991)

Several general journalistic features and devices are apparent even in this short text. Items like 'looms', 'crisis meeting', 'slide toward anarchy', 'last-ditch bid', 'sparking fears', and so on are typical of the genre. So is the way that the text of the headline is repeated in the following paragraphs. 'Civil war looms' is echoed by 'possible civil war', and two occurrences of 'fears of civil war'. Note that the journalistically mundane 'civil war' is repeated, whereas the more emotive 'looms' is re-lexicalised as 'possible' and 'fears of'.

The political orientation of the text is indicated in the way that 'dominated by Serbia' in the second paragraph is highlighted in the subheading, and then reapplied as 'dominated by Communists'. The message is underlined by 'a bastion of hardline Marxism' in the last line.

However, let us concentrate on the references to ethnicity. Yugoslavia is described as a 'multi-ethnic state', and of its people it is said that they 'belong to six main ethnic groups'. Unfortunately, the six groups are not listed, but note that the journalist specifically adds 'and three major religions'.

The following week, I came across this article in the *Guardian*:

Africa round-up

Weekend of tribal violence leaves 13 dead in Kenya

Thirteen people were killed in a weekend of renewed tribal fighting in western Kenya, local newspapers reported yesterday.

The Daily Nation said that eight people were killed on Sunday near Bungoma, 250 miles north-west of Nairobi, bringing the number of those killed in ethnic violence since Friday to 13. Police said 153 houses were burnt down in two separate areas and 45 cattle were stolen.

At least 120 people have been killed in recent months in Kenya's worst tribal fighting since independence.

Residents told reporters three of those killed were fellow Kalenjin tribesmen of President Daniel arap Moi, who have been blamed for fuelling the clashes.

(*Guardian*, 15 January 1991)

The final phrase, 'fuelling the clashes', is typical of the journalistic genre. 'Ethnic violence' is referred to once in the second paragraph, but the 'tribal violence' of the headline is repeated twice as 'tribal fighting' in the article and the reference is further strengthened by 'Kalenjin tribesmen'.

The alternation between 'ethnic' and 'tribal' gives rise to several questions. If the two terms are genuinely synonymous, is 'tribe' ever used of the people of former Yugoslavia? If the term for 'ethnic groups' in relation to Africa is 'tribe', why is 'ethnic' used here? Is 'ethnic' the superordinate term, with 'tribal' available only for subsets of the human population such as Africans? Are there other subsets to whom the term applies? These questions will be taken up in the final section of the chapter, as corpus data may provide some answers.

A reinspection of the local Birmingham newspaper yielded two further articles that seemed relevant. The first was about the British police force:

Police chiefs told to match city's ethnic recruiting rate

Britain's police chiefs have been condemned for their poor record on ethnic minority recruitment only days after West Midlands Police were praised for their own efforts.

The Commission for Racial Equality says forces up and down the country should implement a policy of 'positive action'.

West Midlands Police's campaign includes a careers caravan which visits inner city areas, leaflet drops and nationwide appeals for black recruits.

The CRE's employment officer, Jim Gribbin, said yesterday: 'We need to encourage and help applications from the ethnic minorities because they are severely under-represented at the moment'. Around 11 per cent of Britain's ethnic minority police officers serve with West Midlands Police, making it the force's highest employer of black and Asian people.

But that figure still only amounts to 192 officers from a total establishment of nearly 7,000 – a mere 2.7 per cent in an area where ethnic minorities account for around 20 per cent of the population.

(*Birmingham Daily News*, 11 January 1991)

The headline uses 'ethnic', and 'ethnic minority' occurs four times in this short article. Note that the term 'racial' is used in the name of the British institution 'The Commission for Racial Equality', but is not used by the journalist at all.

When talking about Britain, 'ethnic' seems to mean something different again. It certainly does not refer to indigenous minorities such as the Irish, the Scots or the Welsh, nor to any group of European origin, such as the Italian, Greek or Polish communities resident here. In Britain, according to the article, 'ethnic' means 'black' (as in 'black recruits', paragraph 3), or sometimes 'black and Asian' (paragraph 5).

The use of the abstract noun 'application', in the phrase 'encourage and help applications', is evidence of a depersonalised perspective, even though the speaker is an officer in the CRE. Why not encourage and help 'applicants'? The whole problem is seen very much from an impersonal stance: notice the emphasis on statistics rather than personalities. The aim is to 'boost the numbers', not to assist individuals against discrimination. In this context, compare also the significant use of numbers in the article on Kenya.

The second article in the *Daily News* was about South Africa. Below the caption 'An engaging smile' was a photograph of a young woman (one knee inexplicably raised nearly to her chin), and underneath that a very brief story introduced by a large bullet point:

● PRETORIA – Erica Adams (above), daughter of a prominent South African mixed-race politician, is rumoured to be engaged to a son of President F. W. de Klerk.

Cape Town newspapers said Ms Adams, aged 22, and Willem de Klerk, 24, met in college in the city.

Willem is the reformist president's second eldest son.

'I can't deny there has been a long-time friendship between Willem and myself, but I can't say we are engaged,' said Ms Adams.

Inter-racial marriage and sex was formerly a crime in South Africa and was only legalised in 1985.

(*Birmingham Daily News*, 11 January 1991)

Note the strikingly different focus here: the concentration on the human aspects, relationships, meeting-places, quotations, and so on. What is the rationale behind this shift of focus? Is it a matter of what journalists and editors consider to be of interest to the British public?

Although the report suggests a disapproving distance from apartheid ('was only legalised in 1985'), it slips into familiar South African usages such as 'mixed-race' and 'inter-racial'. Would similar reports about Britain use the same terms, or would they talk about 'mixed marriages'? Again, I shall return to this question later, when looking at COBUILD data.

On the basis of these four articles alone, one might sum up the evidence as follows:

- 1 Former Yugoslavia is definitely an ‘ethnic’ zone, with ‘six main ethnic groups’, and three major religions.
- 2 Kenya is equally definitely a ‘tribal’ zone, but its inhabitants may become victims of ‘ethnic violence’.
- 3 Britain is an ‘ethnic’ zone, but not in the same way as Yugoslavia. In Britain, ‘ethnic’ means ‘black’ or ‘black and Asian’.
- 4 South Africa is unequivocally a ‘racial’ zone, despite any recent reforms.

Thus three apparently near-synonymous terms are selected according to the part of the world that is being talked about.

Dictionaries

Let us now turn to the acknowledged authorities on language, dictionaries, and see whether they distinguish between these three terms, and if so, what criteria are involved.

ethnic or **ethnical** adj. **1.** relating to or characteristic of a human group having racial, religious, linguistic, and certain other traits in common. **2.** relating to the classification of mankind into groups, esp. on the basis of racial characteristics. **3.** denoting or deriving from the cultural traditions of a group of people: *the ethnic dances of Bosnia*. **4.** characteristic of another culture, esp. a peasant culture: *the ethnic look; ethnic food*.

(*Collins English Dictionary* (CED), 1991)

Note that two of the definitions use the word ‘racial’. Unlike the news article about Yugoslavia, the entry includes religion as a component of ‘ethnicity’, and also linguistic identity. There follows the vague reference to ‘certain other traits’. The example ‘the ethnic dances of Bosnia’ implies that the differences between the ‘ethnic’ groups in Yugoslavia is ‘cultural’.

The entry later gives the noun (‘used mainly in the US’) and a brief etymology: entered English in the fourteenth century, via Late Latin, derived from Greek ‘ethnos’ (glossed as ‘race’). The entry for ‘ethnic minority’ says: ‘an immigrant or racial group regarded by those claiming to speak for the cultural majority as distinct and unassimilated’. Note the distancing devices employed: ‘regarded as’, ‘claiming to speak for’.

The prefix ‘ethno-’ is defined as ‘indicating race, people, or culture’, and is followed mainly by technical, scientific and academic terms: ‘ethnography’ is ‘the branch of anthropology that deals with the scientific description of

individual human societies', and 'ethnomethodology' is 'a method of studying linguistic communication that emphasizes common-sense views of conversation and the world'. However, 'ethnocentrism' means 'belief in the intrinsic superiority of the nation, culture, or group to which one belongs, often accompanied by feelings of dislike and contempt for other groups'.

The range of terms used in these definitions is surely confusing, even to a sophisticated native-speaker of English. 'Race, religion, language, culture, immigration, distinctness, assimilation, nation, society' are concepts as difficult to fathom as 'ethnicity'. The variation in scale from 'mankind' at one end to 'groups' at the other, and distinctions as vague as 'certain other traits', sanction the use of the word in almost any context.

Let us look at some other native-speaker dictionaries:

1 of or being human races or large groups of people classed according to common traits and customs <~ *minorities*> <~ *group*> **2** (characteristic) of a traditional, esp. peasant, culture <~ *music*> <*the ~ look in fashion*> **3** *archaic* heathen

(*Longman Dictionary of the English Language* (LDEL), 1991)

concerning nations or races: pertaining to gentiles or the heathen: pertaining to the customs, dress, food, etc. of a particular racial group or cult: belonging to a particular racial group; foreign; exotic.

(*Chambers 20th Century Dictionary* (CTCD), 1983)

fr. Gk *ethnikos* national, gentile, fr. *ethnos* nation, people **1** HEATHEN **2 a:** of or relating to large groups of people classed according to common racial, national, tribal, religious, linguistic, or cultural origin or background <~ *minorities*> <~ *enclaves*> **b:** being a member of an ethnic group <~ *Chinese in Vietnam*> **c:** of, relating to, or characteristic of ethnics <~ *neighborhoods*> <~ *theater*> <~ *foods*>

(*Webster's Ninth New Collegiate Dictionary* (W9), 1983)

1. Of or pertaining to a religious, racial, national, or cultural group.
2. Pertaining to a people not Christian or Jewish; heathen . . . <*ethnos*, nation

(*American Heritage Dictionary* (AHD), 1982)

LDEL starts with a comprehensive, 'neutral' definition, invoking 'races', 'traits' and 'customs', and supplying the collocate 'minorities'. It then combines the 'traditional' with the 'peasant' sense (separated by CED), before listing the archaic use for 'heathen'. CTCD also begins 'neutrally' ('nations or races'), but has 'gentiles/heathen' next. It uses another difficult term: 'cult', and ends with the patronising 'exotic'. There are American/British differences: W9 has the noun first and glosses the Greek etymon as 'nation,

people' rather than CED's 'race'. AHD gives the gloss 'nation' here, but 'people' at 'ethnic'. All the dictionaries suggest that the derived forms and compounds of 'ethnic' have a scientific, academic or technical status.

What about learners' dictionaries?

1 of a racial, national, or tribal group: *ethnic art/traditions – ethnic minority groups* **2** interestingly unusual because typical of such a group: *This music would sound more ethnic if you played it on steel drums*
(*Longman Dictionary of Contemporary English* (LDOCE), 1987)

1 **Ethnic** means connected with or relating to different racial groups of people, especially when referring to the native people of a particular region or to racial minorities within a particular country or city. EG . . . *the ethnic composition of the voters of New York ethnic minorities.* **2** **Ethnic** clothes, music, food, etc., are characteristic of a particular ethnic group, and very different from what is usually found in modern Western culture; used showing approval. EG *She's really into ethnic music these days.*

(*Collins COBUILD English Language Dictionary* (CCELD), 1987)

1 of a national, racial, or tribal group that has a common cultural tradition: *ethnic minorities, groups, communities, etc.* **2** (typical) of a particular cultural group: *ethnic clothes, food, music, an ethnic restaurant.*

(*Oxford Advanced Learner's Dictionary* (OALD), 1990)

Unlike the native-speaker dictionaries, there are no references to the archaic use for 'heathen', and more emphasis on the 'exotic' connotations acquired since the mid-1970s. But general vagueness and synonymic definitions are common to both types: the frequent use of 'racial' and 'tribal' obscures any distinctions. The collocation with 'minorities' (noticed earlier in the article about Britain) is also generally supported: only CTCD and AHD omit it.

Let me give a more abbreviated overview of the entries for 'racial' and 'tribal'.

1 denoting or relating to the division of the human species into races on grounds of physical characteristics **2** characteristic of any such group **3** relating to or arising from differences between the races: *racial harmony* **4** of or relating to a subspecies . . . from Italian *razza*, of uncertain origin

(CED)

1 of or connected with a person's race: *racial pride/customs* **2** existing or happening between different races of people: *racial violence/discrimination/harmony/segregation*

(LDOCE)

1 An unpleasant act that is **racial** is done to people because they belong to a particular race. EG... *the fight against racial discrimination the crudest kind of racial prejudice an alarming rise in racial harassment.* **2 Racial** is also used to describe **2.1** things that happen between people who belong to different races. EG... *our message of racial reconciliation a struggle which transcended racial barriers the racial inequality in our society.* **2.2** things that affect or relate to people who are members of a particular race. EG *An old racial memory is stirred in us.*

(CCELD)

characteristic of race³ (1a); due to or resulting from race: *a racial feature, type, difference, etc – racial conflict, harmony, hatred, pride – racial discrimination.*

(OALD)

Although CED optimistically gives ‘racial harmony’ as its sole example, the learners’ dictionaries include a large proportion of negative collocations: violence, discrimination, prejudice, harassment, conflict, hatred. Note that the etymon ‘razza’ is ‘of uncertain origin’ and relatively recent (Italian). Interestingly, ‘race’ is commonly found in less elevated combinations (race relations, race riots), rather than in scientific, academic and technical compounds like ‘ethnic’. This pattern is not unusual: ‘town, city, metropolis’: the term from Greek is used for the greatest dimension or scope, and in more formal registers.

There are differences between the learner’s dictionaries: CCELD puts the ‘unpleasant’ sense first, LDOCE and OALD give the ‘neutral’ meaning before the negative connotations. OALD relies on the entry for ‘race’ to explain ‘racial’. A more detailed investigation would entail the entries for all the related forms: ‘racism’ and ‘racist’ would be obvious candidates.

In most dictionaries, ‘tribal’ requires reference to ‘tribe’:

1 a social division of people, esp preliterate, defined in terms of common descent, territory, culture, etc **2** an ethnic or ancestral division of ancient cultures (esp Rome, Israel, Greece) **3** *informal/humorous* large number of persons, animals, etc **4** Biology . . . **5** Stockbreeding . . . ~ tribal *adj* . . . from Latin *tribus* probably related to Latin *tres* three

(CED)

of a tribe or tribes: *a tribal dance – a tribal chief – tribal warfare divisions* **1** a social group made up of people of the same race, beliefs, customs, language, etc., living in a particular area often under the leadership of a chief: *the tribes living in the Amazonian jungle – a member of the Zulu tribe* **2** a group of related plants or animals: *the cat tribe*

(LDOCE)

tribal is used to describe things relating to or belonging to tribes and to the way they are organized. EG . . . *political and tribal leaders*. . . . *Her father had recently died in a tribal war*.

A **tribe** is **1** a group of people of the same race, who share the same customs, religion, language, or land, especially when they are not considered to have reached a very advanced level of civilization. EG *Mr Otunnu is a member of the Acholi tribe*. . . . *This attitude still remains in some primitive tribes*. **2** a group of related animals, especially ones that live or hunt together. EG . . . *the tribe of cheetahs*. **3** a group of people who do the same activities or job. EG *There was a tribe of schoolchildren coming up the path*. **4** a family; an informal and humorous use. EG *Good to see you, John! How's the tribe?*

(CCELD)

of a tribe or tribes: *tribal loyalties, dances, gods, wars*.

1 racial group (esp in a primitive or nomadic culture) united by language, religion, customs, etc and living as a community under one or more chiefs: *Zulu tribes – the twelve tribes of ancient Israel*
2 group of related animals or plants **3** (*infml esp joc*) large number of people: *tribes of holiday-makers – What a tribe (ie large family) they've got!* **4** (*usu derog*) set or class of people: *I hate the whole tribe of politicians*.

(OALD)

LDOCE's use of *chiefs* in both definition and example seems excessive. Do 'races' and 'ethnic groups' not have leaders? Oxford eschews the example, but agrees that 'chiefs' are essential. The dictionaries unanimously concede the pejorative connotations of 'tribal' and its related forms, both in its core meaning ('esp preliterate', 'not considered to have reached a very advanced level of civilization', 'esp primitive or nomadic') and in its extended meanings, which relate to plants and animals, stock-breeding, etc. CCELD tries to specify a 'neutral' and 'pejorative' meaning in the same definition, linked by 'especially', and the binary definition is matched by the examples. However, an unwary learner might well miss this subtlety and tar the Acholi with the 'primitive' brush.

The main point that emerges from these dictionary entries is that 'ethnic', 'racial' and 'tribal' do not have the same connotations. 'Ethnic' and 'racial' are not used pejoratively or humorously at all, so we cannot use 'tribal' derogatively and humorously to refer to 'politicians', 'schoolchildren' and 'overlarge families' on some occasions, and still expect it to be 'neutral' when we use it to refer to 'ethnic groups'. And if we use it only of some 'ethnic groups' and not others, this strengthens the suspicion of racism.

Corpus data

COBUILD has been collecting data in the form of modern English written and spoken texts since 1980. In 1987, when the *Collins COBUILD English Language Dictionary* was published, COBUILD's core corpus data amounted to about 18 million words (for details see Sinclair, 1987).

The corpus in daily use at COBUILD now is a sample of the Bank of English, and stands at 121 million words. This is mostly post-1985, and includes British and American books, newspapers and magazines, and international and national radio broadcasts, as well as local radio phone-in programmes, informal conversations, lectures, meetings, and so on.

Note that in all the statistical evidence from corpus data, words with initial capital letters are standardly lower-cased for ease of programming.

Frequency

The first stage of corpus analysis is to assess the frequency of occurrence of the words under investigation. 'Ethnic' occurs 5,128 times in the 121-million-word corpus, 'racial' 2,924 times, and 'tribal' 1,362 times. So 'ethnic' is almost twice as common as 'racial', and four times as common as 'tribal'.

Comparing the rates of occurrence in the mainly pre-1985 18-million-word corpus and the mainly post-1985 121-million-word corpus, we find that 'ethnic' has increased in this respect since the mid-1980s. In the 1970s and early 1980s 'racial' was the predominant term. 'Tribal' has not changed much in rate of occurrence.

	18 million	121 million	
ethnic	9.33	42.73	occurrences per million words
racial	13.83	24.17	occurrences per million words
tribal	10.61	11.26	occurrences per million words

If we look at the number of forms associated with each word in the two corpora, we find that 'ethnic' had six related forms ('ethnic, ethnicity, ethnically, ethnics, multi-ethnic, inter-ethnic') in the 18-million and has eleven in the 121-million corpus (adding 'ethnical, ethnicide, ethnicisation', and so on). 'Racial' had seventeen related forms in the 18-million and has fifteen in the 121-million corpus. 'Tribal' had ten in the 18-million and has thirteen in the 121-million. So in morphological productivity, 'ethnic' has increased most, but still has fewer related forms than 'racial' or 'tribal'.

Distribution

The 121-million-word COBUILD corpus is divided into twelve sub-corpora, and the average rates of occurrence of the terms in these

sub-corpora shows up some interesting 'genre' or 'text-type' differences. All three terms are least frequently used in the general spoken sub-corpus (local radio phone-ins, informal conversation, domestic phone calls, meetings, lectures): 'ethnic' occurs six times per million words, 'racial' four times, and 'tribal' only once per million. Most of the general spoken data is British in origin, so it would be interesting to compare it with similar American data, but as yet none is available.

However, we can compare other British and American data. Broadly speaking, the radio sub-corpora use the terms most frequently: the BBC World Service uses 'ethnic' 135 times per million words, 'racial' 38 times and 'tribal' 24 times. National Public Radio from Washington USA uses 'ethnic' 53 times per million words, 'racial' 56 times and 'tribal' only 11 times. So whereas British broadcasters overwhelmingly favour 'ethnic', US radio broadcasters prefer 'racial' to 'ethnic'. 'Tribal' is far less frequent in both, but the Americans seem to use it even less than their British counterparts. These two sub-corpora are almost identical in sample dates (1990–1), but the US data is mainly for domestic consumption, so that may have some influence.

Comparing *The Economist* with the *Wall Street Journal*, the British publication (like the BBC) favours 'ethnic' by a substantial margin: 61 per million, as compared with 37 for 'racial' and 21 for 'tribal'. The *Wall Street Journal* uses 'ethnic' and 'racial' almost equally (19 and 18 per million words), but rarely uses 'tribal' (3 per million). American books and British books seem to show a reversed distribution: both use 'tribal' 8 times per million words on average, but British books use 'racial' 17 times and 'ethnic' 14 times, whereas American books use 'ethnic' 18 times and 'racial' only 12. However, the ratio of fiction to non-fiction is substantially higher in the British books sub-corpus, so it may be that 'ethnic' is the prevalent term in non-fiction.

Comparisons can also be made within the same variety of English as well as within the same 'genre' or 'text-type'. Comparing three British newspapers, *The Times* heavily favours 'ethnic' (50 per million) over 'racial' (20 per million), the *Independent* is almost even-handed in its use (33 per million for 'ethnic' and 29 for 'racial'), but *Today* uses 'racial' slightly more than 'ethnic' (13 for 'racial', 10 for 'ethnic'). 'Tribal' is rare in all three (9 per million in *The Times* and *Independent*, and 2 per million in *Today*).

Surprisingly, the 'ephemera' sub-corpus (information leaflets, tourist brochures, 'junk mail', letters, diaries, etc.) has the highest incidence of 'tribal': 62 per million. These occurrences are mainly accounted for by travel brochures (informing us about trekking in remote areas of the world with the assistance of 'tribal' guides, and tourist attractions such as 'tribal' markets and 'tribal' dancing), and newsletters from environmental organisations (warning of the dangers to 'tribal' people of the rainforest timber trade). That is to say, 'tribal' is frequent only in these two topic-specific text-types.

Collocation

Finally, let us look at the collocational profiles of the three terms in the 121-million-word sample corpus from the Bank of English. There are several different statistical methods currently being used at COBUILD to indicate collocation, but I have selected just one ('T-scores') for the purpose of this chapter, and applied it to the three terms. Table 1 shows the 'top ten' strongest collocates for each, in order of their T-scores.

Table 1 Ordered list of collocates for *ethnic*, *racial* and *tribal*.

ethnic	racial	tribal
groups	discrimination	assembly
minorities	non	grand
and	multi	leaders
minority	equality	chiefs
of	and	and
violence	of	groups
in	south	killings
cleansing	africa	navajo
group	commission	in
albanians	ethnic	a

Some clear differences are apparent: 'ethnic' collocates with 'group/s' and 'minority/ies', associated activities are 'violence' and 'cleansing', and major participants are 'Albanians'. 'Racial' has more abstract, ethical and bureaucratic associations (as suggested by the third newspaper article's reference to the Commission for Racial Equality): 'equality, non-racial, multi-racial, commission'. The principal related activity is 'discrimination' and the main geographical location is South Africa. 'Tribal' shares with 'ethnic' the collocate 'groups', but the focus is more on modes of organisation: 'assembly, leaders, chiefs'. The more abstract activity of 'violence' associated with 'ethnic' is replaced by the very specific term 'killings'. 'Navajo' are the featured peoples.

The presence of 'and' in all three lists suggests that each term is often accompanied by other adjectives indicating other taxonomic dimensions. Looking further down the list of collocates for 'ethnic', we find 'religious, political, racial'. For 'racial', 'ethnic' is in the top ten listed in the table, but there are also 'religious, sexual, cultural, political' lower down the order. For 'tribal', we also find 'religious, political, ethnic'.

Looking at the associated activities lower down the collocate lists, we see with 'ethnic': 'conflict/s, tension/s, unrest, clashes, problems' but also 'studies' (as the dictionaries suggested, 'ethnic' is the preferred term in academic contexts). With 'racial': 'tension/s, harassment, violence, prejudice, segregation, abuse, hatred, attacks, preferences, conflict' but also 'quotas,

constitution, democratic, policies, law'. With 'tribal': 'conflict, fighting, violence, warfare'.

The nationalities and geographical areas that occur in collocation with each term also vary: with 'ethnic' we find, as well as 'Albanians', 'Germans, Serbs, Hungarians, Soviet, Turks, Yugoslavia, Chinese, Kosovo, Romanians, Serbians'. With 'racial' we find, in addition to South Africa, no other nationality or region! With 'tribal', after 'Navajo', we come across 'African, Western, Geneva'. This apparently rather strange set of collocates is explained by the prominence given by the newspapers that are included in the corpus to the arrival in Geneva of tribal chiefs from the Western Sahara for peace talks.

Review

Let us now look briefly at the extent to which the usages in the newspaper items in the first section are reflected in a broader range of English texts, and the extent to which the information given in the dictionary entries in the second section can be considered to be reasonable summaries of actual usage as evidenced by the corpus. I shall also try and answer some of the questions I raised earlier in this chapter.

In the first newspaper article about Yugoslavia, featuring the word 'ethnic', the repetition of 'civil war' was noticed. Of the 5,128 concordance lines for 'ethnic' in the 121-million-word corpus, nine also mention 'civil war', referring to situations in Burma, Eritrea, Yugoslavia (three), Moldavia, the Soviet Union, and Sudan.

The collocation 'tribal fighting' highlighted in the second article is borne out by corpus evidence: 'fighting' is not a significant collocate of 'ethnic' or 'racial'. In fact there are twelve occurrences of 'ethnic fighting' in the corpus, but significantly none of them is connected with Africa. Georgia, Soviet Union, Croatia, Nagorno-Karabakh, Azerbaijan, Yugoslavia, even Pakistan, Afghanistan, and Sri Lanka, but not Africa.

Is 'tribe' ever used in the context of Yugoslavia? In nearly 9,000 lines for 'yugoslavia/-an/-ans', 'tribal' and 'tribe' do not occur even once. There is a solitary use of 'tribes' in *The Times*:

Serbian nationalist theory does not have much time for the remaining three tribes of Yugoslavia.

How common is 'ethnic' in connection with Africa? Of the 5,128 lines for 'ethnic', fifteen mention Africa, five of these are to do with South Africa, two with the West Indies, one with Britain, and five talk about ethnicity in global terms, with Africa one of several areas referred to. Two lines are for the 'exotic' sense of 'ethnic' mentioned in the dictionaries: 'ethnic fabrics' and 'ethnic-inspired clothes'.

What clues are there to the other principal subsets of humanity to whom the term 'tribe' can refer?

Sahara Krahn Venda Navajo Bedouin Chadian Mohawk Sumatra
Liberia Afghanistan Kabul Indians Pakistan Africa African Indian
American.

What does 'ethnic' mean in a British context? Let me quote a few of the forty-one lines for 'ethnic' that refer to Britain:

these findings don't vary much across different ethnic groups.
Britons of Asian or Afro-Caribbean origin . . .
(BBC)

Mr Tebbit devised his famous cricket test when he asked which
team was cheered for by West Indians and other ethnic minorities
living in Britain.
(*Today*)

But in Britain we could not find one ethnic Father Christmas in
Birmingham, Bradford or Manchester. So, is there really a black
Santa?
(*Today*)

The term 'ethnic' extends beyond 'blacks and Asians' in a British context
only when 'exotic' cultural items are under discussion: arts, music, food,
clothing:

Its aim is to offer programmes specifically to listeners from
ethnic backgrounds, immigrants to Britain, and their children –
Italians, Hispanics, Greeks, Jews, Asians, Arabs, Chinese and
Afro-Caribbeans.
(BBC)

One of the problems of cooking ethnic Italian dishes in Britain has
always been to find the ingredients.
(*Vogue Magazine*)

Are 'mixed-race' and 'inter-racial' ever used in a British context? Of
102 occurrences of 'mixed-race' in the corpus, thirty-six from the BBC refer
mainly to South Africa. Of the four from *The Times*, two relate to South
Africa, but two are British references:

Police seek two men in their twenties, one white with a ponytail and
the other of mixed race.

No MPs today would call some of their constituents ‘half-casts’ [*sic.*] They would say ‘of mixed race’, for ‘half-caste’ implies contempt.

Of the five in the *Independent*, two refer to South Africa, and the other three are from a play review (about adoption), a sociological/medical survey (‘31 per cent of the white children and about half the Asian and mixed-race children’), and a dearly ironic usage:

but we always pulled our weight when it came to glossy pictures for the annual report. Who could forget our three young people (mixed race, mixed gender) leaning over a computer print-out and smiling while one of them points to some key figure? Or, our two middle-aged executives (one dynamic, one thoughtful) studying a report carefully?

Ten from British books refer to South Africa, Nazi Germany, South America, and the USA. The twenty from US National Public Radio are all about South Africa or the USA, as are the eight from *The Economist*. The other sub-corpora, reflect very similar patterns of usage. Basically, ‘mixed-race’ is primarily used about South Africa and the USA. The term is gradually increasing in a British context, partly because of American influence, and partly because the older term ‘half-caste’ is clearly seen as pejorative:

When I said half-caste, she said the correct term was mixed race.
(*Today*)

The British references to ‘mixed-race’ are usually to children, especially in the context of adoption. There are twenty-three occurrences of ‘interracial’, with a similar spread. Six refer to the USA, five to South Africa. There are actually ten occurrences that refer to Britain, but three are historical (one about the British in India), two are about adoption and child care, and one is from a ‘new age’ text exhorting peace and harmony. There is also one reference to North American Indians, and one to former Yugoslavia.

In my comments on the newspaper article, I suggested that in Britain we might more readily talk about ‘mixed marriages’. There are thirty-two occurrences of ‘mixed marriage(s)’. However, I had overlooked the historical dimension: most references to ‘mixed marriages’ are still concerned with inter-denominational and inter-faith relationships rather than interracial ones: marriages between Catholics and Protestants, Christians and Jews. So of the thirty-two, twelve are for this usage, eight are for marriages between blacks and whites in Britain, and the rest of the occurrences are very varied, referring to Nazi Germany, former Yugoslavia, the West Indies, South Africa, Russia, and France. Two refer to Shakespeare’s *Othello*, and one jocularly to pig breeding.

This exercise has pointed up one limitation of COBUILD's current corpus tools: we need to know not only of whom a term is used, but also who uses it, and specific textual sources are not easy to pinpoint with the existing software.

The dictionaries' use of 'race' (as well as religion, language, culture, immigration, nation, society and other concepts) in their definitions of 'ethnic' is reflected by the frequent juxtaposition of the two terms in the corpus: ninety-four concordance lines have both, interestingly thirty have 'ethnic' before 'racial' ('ethnic and racial differences, ethnic or racial enmity', etc.), whereas sixty-five reverse the order ('racial or ethnic discrimination, racial and ethnic minorities', etc.). So the less frequent word is usually placed first.

The use of 'ethnic' rather than 'racial' in academic contexts has been noted above in the collocation with 'studies'. In both British and American sources, 'Ethnic Studies' refers mainly to an academic discipline in universities.

CED gave the examples 'the ethnic dances of Bosnia', 'the ethnic look' and 'ethnic food'. The 5,128 lines for 'ethnic' in the corpus reveal only one line for 'ethnic dance' ('Asian women are proud of the sari, . . . it is not merely a novelty piece to be taken out of the wardrobe only for national holidays and ethnic dances.'), three lines for 'ethnic look', and ten for 'ethnic food':

. . . lunch, and could it not be	ethnic food. He ate ethnic food . . .
. . . not be ethnic food. He ate	ethnic food all the time . . .
. . . is plenty of chi-chi and	ethnic food in Chicago. But the . . .
. . . has a range of books on	'ethnic' food and explained why . . .
. . . In the last 10 years	ethnic food sales soared by 400 . . .
. . . its way to being the new	ethnic food everyone turns to. . . .
. . . originally seen as sort of	ethnic food becomes more and . . .
. . . photograph. Scour markets,	ethnic food shops and . . .
. . . and – perhaps the least	ethnic food ever created – . . .
. . . in October and featuring	ethnic food, parades, and . . .

LDEL gave 'ethnic look' but also 'ethnic music'. This proves to be a stronger collocation in the corpus: there are eighteen lines for it. AHD offered 'enclaves, neighborhoods, theater'. Looking only at the American data in the corpus, out of 969 occurrences of 'ethnic', three are for 'enclaves', eight for 'neighborhoods', and one for 'theater'.

LDOCE gave 'ethnic art/traditions' and 'This music would sound more ethnic if you played it on steel drums'. There are no instances of 'ethnic art' in the corpus (but there is one for 'ethnic arts and crafts') and two occurrences of 'ethnic traditions'. The second example is very unusual: 'steel drums' and 'sound' are not found in the proximity of 'ethnic', it is normally used before a noun, and it is rarely qualified by 'more': there are two lines (both from American sources):

I mean, who is more ethnic and who is more American, Colin Powell or Frank Sinatra?

For more ethnic types of preparations, like baccala salad (page 101), you may wish to shorten the soaking period.

CCELD gives 'composition' and 'music': the first occurs eight times, the second twenty-one times. OALD adds 'clothes' and 'restaurant': there are no lines for 'ethnic clothes' (but there is one for 'ethnic-inspired clothes'), but there are eleven for 'restaurant', mainly from American books.

For 'racial', CED optimistically gives the example of 'racial harmony' (25 occurrences, against 202 for 'discrimination', plus all the other negative collocations noted earlier). LDOCE offers 'pride' (4 in corpus) and 'customs' (none), before the better-attested 'violence' (49), 'discrimination', 'harmony' and 'segregation' (66). CCELD cites 'discrimination', 'prejudice' (66), 'harassment' (63), 'reconciliation' (6), 'barriers' (8), 'inequality' (10) and 'memory' (5). OALD gives 'feature' (one line for 'features'), 'type' (3 – but 17 for 'stereotypes'), 'difference' (6 – but 14 for the plural form), then the more frequent 'conflict' (20), 'harmony', 'hatred' (45), the rarer 'pride' and, finally, the commonest collocates 'discrimination'.

LDOCE's collocates for 'tribal' are 'dance' (6), 'chief' (11 – but 44 for the plural), 'warfare' (11) and 'divisions' (1). CCELD offers 'leaders' (the commonest collocate, 45) and 'war' (15). OALD gives 'loyalties' (1), 'dances' (3), 'gods' (2) and 'wars' (2). My criticism of LDOCE's emphasis on 'chiefs' is not backed by the corpus evidence, but note that it is the plural form that is common (44), not the singular form (11) used in both definition and example. Also, 'leader' (22) and 'leaders' (45) taken together are commoner than 'chief' (11) and 'chiefs' (44).

At any rate, the pejorative connotations of 'tribal' are as clear in the corpus as in the newspapers and the dictionaries. 'Primitive' co-occurs with it six times, but never with 'ethnic' (which is four times as frequent) or 'racial' (which is twice as frequent). Similarly, 'tribal' attracts a higher proportion of 'insurgents' and 'rebels' than the other terms: 12 each, whereas 'racial' only has 1 and 2 respectively, and 'ethnic' only 15 and 19 (whereas proportionately, it should attract 48 of each).

The main area in which dictionaries other than CCELD (which was after all based on a corpus, though much smaller than the one being used in this chapter) seem to repeatedly miss the target is in their examples. Often, their examples are in the appropriate semantic environment, but fail to specify the most frequent lexical realisations. Sometimes, they select the appropriate word, but use the less common form (as with several instances cited above of using a singular when the plural is better-attested). Occasionally, the syntactic framework as well as the lexis is extremely marked (as, for instance, the example 'This music would sound more ethnic if you played it on steel drums.').

Conclusions

At the beginning of this chapter, I suggested that language not only affords us a means of understanding a language and its users, but also might cause us to unwittingly adopt their attitudes and opinions. The English word 'tribal' clearly has pejorative connotations, and if we continue to use it, and apply it only to certain groups of human beings, we are merely recycling the prejudices that the English-speaking culture has developed with regard to those groups. As English develops into a truly international language, one of the rites of passage must surely be for it to divest itself of these culture-bound terms. After all, for similarly organised small groups in the British Isles, we use the term 'clan'.

In his *Keywords*, subtitled 'A vocabulary of culture and society', Raymond Williams (1983) includes 'ethnic' and 'racial' in the headword list, but not 'tribal'. Of 'ethnic' he says:

Ethnic has been in English since mC14. . . . It was widely used in the sense of heathen, pagan or Gentile, until C19, when this sense was generally superseded by the sense of RACIAL (q.v.) characteristic. **Ethnics** came to be used in the United States as what was described in 1961 as 'a polite term for Jews, Italians, and other lesser breeds'. . . . The scientific uses are now specialized areas within anthropology. . . . Meanwhile in mC20 **ethnic** reappeared, probably with effect from the earlier American use of **ethnics**, in a sense close to FOLK (q.v.), as an available contemporary style, most commonly in dress, music, and food.

(Williams, 1983: 119)

At the end of his lengthy article on 'race', he concludes:

It is clear that the very vagueness of **race** in its modern social and political senses is one of the reasons for its loose and damaging influence. . . . Physical, cultural and socio-economic differences are taken up, projected and generalized, and so confused that different kinds of variation are made to stand for or imply each other. The prejudice and cruelty that then often follow, or that are rationalized by the confusions, are not only evil in themselves; they have also profoundly complicated, and in certain areas placed under threat, the necessary language of the (non-prejudicial) recognition of human diversity and its actual communities.

(Williams, 1983: 250)

One aspect which has barely been touched upon in this chapter paper, except in the broadest sense, is the identity of the user. Who actually sponsors

the usages of ‘ethnic’, ‘racial’ and ‘tribal’? We need to know not only which particular text the term occurs in, but also the identity of the speaker or writer of the individual sentence in which the term occurs, whether authorial or fictional, and the degree to which the context reflects a deliberate choice, whether sincere or ironic. How many users of the terms might themselves be described as ‘ethnic’, ‘racial’ or ‘tribal’, and by whom?

The problems raised in this chapter are not new. The focus merely shifts from one term to another. Published on the eve of the Second World War, the frontispiece to a short text by Huxley (1939) reads:

The vague term ‘race’ has been much misused in modern pseudo-scientific writings and nationalist propaganda.

Huxley defines ‘nation’ in terms very similar to the dictionary entries for ‘ethnic’ and ‘racial’ quoted earlier:

Very many human activities, aspirations, and emotions have contributed, either naturally or artificially, to build up the great synthesis that we term a ‘nation’: language, religion, art, law, even food, gesture, table manners, clothing, and sport all play their part.
(Huxley, 1939: 1)

In a similar work of the same period by Walker (1940) we read:

Thirdly, there are the 6,600,000 Bantu, the Natives, who range from a still tribal majority to the few who have become thoroughly westernized and even hold professional qualifications.
(Walker, 1940: 23)

Evidence more recent than the COBUILD corpus shows that the problems are becoming more acute, not less:

And the whole affair depended on the unspeakability of certain words. As Claudia Brodsky Lacour argues, Thomas’s race was the essential factor in his nomination, but the word ‘race’ could not be mentioned, and opposition to Thomas would automatically be understood as ‘racism’: ‘The word “racism” and not the thing . . . was the object of concern.’

(Wendy Steiner, ‘The witch, the judge and the Pepsi’,
Independent on Sunday, 21 February 1993: 28)

The terminology used with reference to adoptive children in Britain continues to be problematic:

As the (natural) mother of two children of mixed race, my heart went out to the Norfolk couple whose application for adoption was turned down because they were 'racially naive' . . . a white single parent bringing up two mixed race children. . . . Because my ethnicity is in question . . . I do indeed understand racism.

(Louise Gosnell, *Guardian*, 14 July 1993: 18)

Not 'ethnicity' but 'racism'; the term 'ethnicism' does not occur in the COBUILD corpus either.

'Tribe' and 'tribal' still strike me as raising a problem of a different dimension.

Tribes, states and empires are all agents of war. Purely tribal wars, however, are rare in the contemporary world. Ethnic and international conflicts are not reversions to the primitive, since they occur specifically in modern state and imperial structures. . . . The extreme violence of the post-Yugoslav conflict is therefore best understood by reference neither to tribal hatreds nor to a struggle between the primitive and the civilised.

(Michael Freeman, 'Death toll from the war in our midst',
Times Higher Education Supplement, 5 March 1993: 23)

A new publication is listed in the *Bookseller* as:

Alexander Stuart

Tribes

Short, sharp, frightening tale of football hooliganism from the author of *The War Zone*.

(Sarah Broadhurst, 'Paperback Preview',
Bookseller, 26 March 1993)

Yet journalists continue to see the 'humorous' side:

Essex, USA

The first in a series on British tribes looks at holidaymakers who find everything they like about America at Disney World

. . . I have always liked America, but I could never like it the way the Essex tribe do, with a fierce, transferred patriotism that admits no fault in the adopted nation and embraces even its prejudices.

(Martyn Harris, *Daily Telegraph*, 14 August 1993: 1)

References

AHD *American Heritage Dictionary*, 1982

CCELD *Collins COBUILD English Language Dictionary*, 1987

CORPUS LINGUISTICS

CED *Collins English Dictionary*, 1991

CTCD *Chambers 20th Century Dictionary*, 1983

LDEL *Longman Dictionary of the English Language*, 1991

LDOCE *Longman Dictionary of Contemporary English*, 1987

OALD *Oxford Advanced Learner's Dictionary*, 1990

W9 *Webster's Ninth New Collegiate Dictionary*, 1983

Huxley, J. (1939) *Race in Europe, Oxford Pamphlets on World Affairs no. 5*, Oxford: Clarendon Press.

Sinclair, J. McH. (ed.) (1987) *Looking Up*, London: Collins ELT.

Walker, E. A. (1940) *South Africa, Oxford Pamphlets on World Affairs no. 32*, Oxford: Clarendon Press.

Williams, R. (1983) *Keywords*, London: Fontana.