There may be regular guys but there are no regular native speakers: Lexis and native-speaker-like competence

Alan Durant
Middlesex University, London

1. Introduction

Sometimes the main themes of an academic’s work emerge early. In articles leading up to her 1986 Lund University thesis - and then in the thesis itself, published as Lexis in composition: A performance analysis of Swedish learners’ written English (Linnarud, 1986) - Moira Linnarud examined the acquisition and use of vocabulary. In Lexis in composition, she presented a quantitative analysis of lexical differences between student compositions in English by native speakers and by Swedish speakers, showing how native-speaker compositions display more ambitious, varied and dense lexis than comparable pieces of non-native speaker writing.

To structure her comparison, Moira Linnarud correlated lexical patterning in her two parallel corpora with overall evaluations of the compositions made by three different types of language stake-holder: Swedish university teachers; native-speaker university teachers; and other native speakers not professionally involved with English. Prominent among her conclusions was the insight that written compositions are valued more for predictable combinations of lexical items than for one-off interesting words, no matter how apposite those words may be individually. On this basis, she urged - as has become widely accepted since - that vocabulary development should be concerned less with single items than with collocations (e.g. ‘adjective plus noun’ pairings, which often differ from language to language) and with idioms (phrases whose overall meaning cannot be worked out from the meaning of their individual words); it should also foreground indicators of register and context of use. Inflected in a variety of ways subsequently, Moira Linnarud’s interest in the expressive capability of non-native speakers’ English vocabulary has continued throughout her teaching and writing; and the topic of lexical collocation, revolutionized by computer corpus techniques that were only emerging in the mid 1980s, remains as fascinating now as it was then, bringing together a cluster of profound questions about verbal discourse and playing a significant but still elusive role in advanced-level second language acquisition.
Choosing student compositions as data for investigating non-native speaker lexis may have slightly obscured the significance of Moira Linnarud’s 1980s research. Taking students as your case study can look like working with whatever data or informants come most easily to hand. But two decades later, focus on the writing of advanced non-native speakers, rather than on beginners or intermediate learners, stands out more impressively. For a variety of reasons, the processes by which advanced learners converge on native-speaker-like competence have received less attention than more clear-cut challenges facing learners at earlier stages; and the closer the convergence between non-native and native speaker, the more difficult diagnostic judgements are and the less systematic pedagogic strategies will be for enhancing performance. Yet as Moira Linnarud’s own evidence showed, significant differences in lexical patterning remain a major factor in evaluations by both native-speaker and non-native speaker readers, as well as sometimes a cause of low self-confidence among non-native writers themselves. So the nature and basis of the differences are important to understand.

2. Judging words

Questions worth exploring about how lexical choices contribute to overall discourse effect can be tackled in a number of ways. Corpus-based research into the distribution and collocation of linguistic items is now perhaps pre-eminent. A different, but complementary approach is to focus on the basis of intuition-led evaluation of lexis expressed in terms of correctness or appropriacy. Perhaps because there are many dimensions of lexical effect, and because successive vocabulary choices interact with one another, continuing mystique surrounds native-speaker intuitions in this area, reflecting a presumption that native speakers can introspect reliably concerning the relevant lexical features. Faced with all the difficulties that lexis presents, many learners wish for a level of verbal self-assurance that they imagine goes with ‘native-speaker-like’ competence. There are, however, complications with native-speaker judgements in matters of lexis; and it is appeals to native-speaker intuitions that I would like to use this opportunity to reflect on.

When someone asks a native speaker about lexical meaning or effect, what sort of response do they expect? What special set of insights, for example, were the native-speaker teachers in Moira Linnarud’s 1986 comparison thought to have access to in their evaluation of written compositions, such that they should constitute two out of the three evaluating groups in the design of the
study? Presumably what is appealed to is some notion of native-speaker competence in matters of lexis, somewhat along the lines of native-speaker grammatical competence. It is hardly an exaggeration to say that non-native speakers often consider native speakers to be authoritative in judgements of lexical meaning or effect far beyond the extent to which they would accept that any piece of language can have fixed rather than fuzzy meanings.

For the person called on to display native-speaker competence who is nevertheless not a walking dictionary, the question immediately arises where that knowledge will come from. Over the last few weeks, I have been asked to comment on the scope of the meaning of *slipstream*, figurative use of *glazed*, why the verb *reckon* is unsuitable as a synonym for *think* in an essay, what the difference is between *at the moment* and *at the minute*, and many other cases of usage and nuance. In responding ‘as a native speaker’, rather than as a teacher consulting dictionaries, concordance lines, or other resources, I have wondered how far my replies report generalised native-speaker intuition and how far I merely reflect my own idiosyncratic language experience and preferences. When recently my eleven year-old twin daughters informed me, seriously and in agreement with one another, that the difference between *umbrella* and *brolly* in English is that umbrellas have a U-shaped handle - hence *um-brella* - whereas brollies have a straight handle, I could recognise that countless, hopefully smaller-scale misunderstandings may be buried in my own adult vocabulary. At the same time, there is undoubtedly a complex web of intuitions about which I can be fairly confident, intuitions that might be expressed (if I care to organise them formally) as judgements of synonymy, hyponymy, entailment, pleonasm, contradiction, semantic anomaly and relative improbability.

As significant in this context as personal, sub-Wittgensteinian meditation is the fact that, when native speakers talk together about word choice or meaning, interesting differences of viewpoint almost invariably arise. Consensus among native speakers on usage and meaning cannot be taken for granted to anything like the extent that may be the case as regards sentence structure. There is far more variation; indeed substantial interpretive variation may be common rather than exceptional.

Not in all cases, though, and not to the same extent with different dimensions of lexical effect. It may therefore be interesting to list briefly some of the main dimensions of verbal effect, signalling where variation is more and less likely. Differences between the dimensions may suggest that native-speaker competence is only part of what is going on in vocabulary judgements. It may even suggest - encouragingly, I hope, for non-native speakers - that competence
is less an achieved, steady state of knowing than something more approachable when less idealised. Arguably the more non-native speakers distinguish between researched information about lexis and faith in native-speaker authority, the more comfortable such speakers may feel communicating even with substantial areas of lexical uncertainty - and the more willing they may be to recognise that the vocabulary of a language, as Dr Johnson famously discovered, is not something that can easily be fixed and possessed.

3. Asking native speakers

Native-speaker intuitions make themselves felt in reading or listening somewhere along a continuum, from a sense of the particular appositeness or vividness of an expression, through discomfort or amusement with anomalous usage, to a clear sense that something is not possible as an expression of the language. Sometimes an intuition is focused on a particular lexical item, sometimes it is spread across a phrase, sentence, or longer stretch of discourse. In relation to non-native speaker queries, such intuitions can be made more accessible by linking them to the kinds of question that elicit them. Here, I think, are the main types of question that non-native speakers put to native speakers about lexis:

1. Is that a word of English? Is there such an expression?

A non-native speaker asks such a question when she or he half-remembers a word or idiom, or if the word-form being queried is close to a similar expression in the speaker’s own first language. In most cases, the native speaker will simply know. But in some cases (e.g. with possible technical terms, idioms, jargon or obscure Scrabble words) the native speaker may not know if such an expression exists or not. On most estimates of the size of an individual speaker’s vocabulary compared with the number of words in English - allowing for much discussed difficulties in making such estimates - a typical native speaker seems unlikely to know more than 20 per cent of the publicly recorded lexicon, and may well know less than 10 per cent. So an answer in the negative about whether a possible word exists will not necessarily be definitive. Even if the proportion of words of the language likely to be recognised sounds low, as it does to me, nobody could reasonably expect a native speaker to know all the words in her or his language - and given that most words are polysemous, certainly not all of the current senses of all of those words. And the speaker’s
difficulties obviously increase significantly if trying to identify words from an earlier period of the language (of which the native speaker was not a native speaker), or worse still, attempting to perform such an earlier state of the language in speech or writing.

Knowing lots of words - both passively and actively - is a useful skill, and native speakers have plenty of opportunity to do so. But this is not the nature of their competence as native speakers. Competence is not a store of learnt knowledge of the public language but a mental representation of the underlying systems of that language. When native speakers are asked whether a particular expression exists, the question runs a quick check in someone’s handy, on-board dictionary. But the dictionary is always a ‘shorter’ edition. Where doubt persists, more standard reference resources will need to be consulted - resources that are produced on the basis of large amounts of collected evidence, and that are equally available to native speaker and non-native speaker alike.

2. Is this the right form of the word? What preposition does it take? Is that the right plural/past tense, etc.?

Questions of this kind bring together grammatical and semantic competence (a meeting of domains reflected in the sub-field of lexicogrammar), and present even advanced non-native speakers with persistent difficulties. Native-speaker intuitions, by contrast, are likely to be fairly reliable, except in seemingly marginal cases where social dialect differences are at issue or where usage is unclear (e.g. different to, different from, different than). Typical questions concern variant forms of an expression, where there may be doubt about alternative possible inflections (learned/learnt, spin/span/spun/*spinned). Other cases, especially with phrasal verbs, involve selection between closed-set word classes, such as prepositions (e.g. *they knuckled in to their work, presumably prompted by the sound similarity between in and down, as well as by resemblance to got stuck in to the work).

Some cases turn out to involve significantly divergent meanings alongside the issue of choice of form, as for example with the cluster of senses associated with carried on, carried off, and carried away. For the native-speaker informant, however, even apparently simple cases may turn out to be equivocal. In Usage and Abusage, for example, Eric Partridge (1947/1994) characterises fill in and fill out, and the meaning of indicate, and the contrast between anticipate and expect - as well as very many other expressions - in ways that square so poorly with my own intuitions that I can be convinced only by considering myself not to be speaking the same dialect (partly true, in time, place and class), or if I conclude
that I have nurtured misguided intuitions about my own and only language for most of my life (quite possible, but with implications worth exploring as regards the degree of trust commonly placed in native-speaker lexical intuitions).

Although inflectional and syntax errors with lexemes are difficult for non-native speakers to avoid, they may not be especially significant from the point of view of communication. It is usually easy in context to recognise mistakes of this kind, because the formulations produced are localised and easily corrected. There is no expression *knuckle in*, so we assume the idiom *knuckle down* was intended, especially if *knuckle* is co-selected with *work*. What is decisive is whether an expression appears only grammatically ill-formed (and so in need of local correction), or whether it appears semantically anomalous, possibly causing discourse to veer off-course and making it difficult to understand or construe in terms of register.

3. *I know this word, but what does it mean?*

A non-native-speaker asks this type of question to check intuitions about meaning and use against a native speaker’s more extensive experience of the target language. At advanced levels of proficiency, particular importance will be attached to clarifying different ways of saying broadly the same thing. In straightforward matters of denotation, non-native speakers routinely make links back into their own first language, mapping networks of concepts and vocabulary and noting relevant differences, including lexical gaps. Dictionaries help, by supplementing experience of unfamiliar words with definitions, either analytically (‘genus/superordinate plus differentiae’) or by means of loose synonyms with various tags and restrictions on usage specified. Whatever unique value is to be found in a native speaker’s intuitions lies in explicating cases where alternative words denote roughly the same thing - they may even have been presented as synonyms in a dictionary definition - but where the alternatives display subtle shades of meaning. Questions of this kind address a difficulty we might call the dilemma of the thesaurus-user.

The main issues appear to be ones of conceptual or descriptive meaning: the questions address ways of giving names to ideas and things, and of specifying the set of things that can be designated by any particular expression. A non-native speaker might query, for example, whether some expression would be correct if used to describe something. Is this large entrance more precisely a ‘door’, a ‘gate’, a ‘portico’, or a ‘portal’? What qualities are ascribed if someone is characterised as ‘parsimonious, not stingy’, or is not so much
‘escorted’ as ‘frogmarched’? Some, but not all, of the relevant properties of meaning here can be grasped in terms of sense relations: that is, in terms of how the words of a given semantic field are mapped by relationships between them (relationships, such as synonymy or hyponymy, that help explain for instance why the pain must have been agony has a very different effect from the agony must have been pain.) But sense relations are only part of any explanation of word meaning, because such relations always need to be plugged back into the language user’s encyclopaedic knowledge.

A grackle, for example, is a type of horse’s bridle (so the word grackle is a hyponym of the word bridle); and a bridle, in turn, is part of a horse’s tack, or riding equipment (so the word bridle is a meronym of the word tack); one result of these relations is that a sentence such as ‘There were grackles, bridles and tack all over the place’ should seem to a native speaker to be semantically anomalous. I suspect that native-speaker intuitions may not be consistent on this. However that may be, it is not an immutable fact, or given, that a grackle is a type, rather than a part, of a bridle; a bridle is a composite device, including a browband, throatlash, bit and other accessories, and usage about part and whole is imprecise and inconsistent. Deciding between different models of sense relations in this semantic sub-field requires knowledge not only about words but also about horses. In this way, and across most or possibly all semantic fields, questions of sense relations are inextricably bound up with the function of denotation to designate whatever is properly picked out by any given word (as the history of the compilation of the OED as ‘the meaning of everything’ (Winchester, 2003) entertainingly shows).

Difficulty with the interplay of sense relations and denotation does not arise only with technical or obscure expressions, either, or only with words that have obviously varying regional or social use. It can also be encountered in apparently simple, everyday words. When a ‘complimentary meal’ in a restaurant turns out not to include drinks, the question arises whether the word meal is standardly used to denote a combination of something to eat and something to drink, not just something to eat - or only in some circumstances (circumstances that will not be easy to specify). In this case, what seems needed in order to understand the use of the word meal is not only knowledge about words and about cultural conventions to do with eating, but also about speaker intention in a given context (perhaps especially in contexts of commercial or promotional speech).

Interestingly, close linkage between verbal and experiential knowledge does have a spin-off: that it is not all one-way traffic between native speaker
and non-native speaker. In many situations, native speakers can be guided as to word meaning by non-native speakers whose world-knowledge enables them to configure words in a particular semantic field that the native speaker does not know much about. More commonly for the non-native speaker, though, a question about word meaning will arise in a specific context and call for selective interpretation of an expression in and for that context. Loose synonyms suggested by a native speaker will function like entries in a glossary accompanying a text, rather than enquiries into the play of affinities and disaffinities between the queried word and others in the language, as might be undertaken in lexical semantics. It is true that extra information about an expression’s meaning potential might provide a better guide to future use of the word. But understandably, because any appeal to native-speaker intuition is always motivated and about something, the answer sought is usually a matter of local interpretation rather than a general account of the word’s meaning and use.

4. Does this word have any resonances or associations I ought to know about?

To ask this question is to be interested in the connotations of an expression: the culturally-attached meanings that accumulate around it. A textbook example would be knowing that champagne, which denotes a type of white wine produced in a particular region of France, also signifies luxury, wealth, or celebration. Connotations shift attention from the language system further into its relations with socio-cultural knowledge; the connotations of a word are meanings that are superfluous to the word’s denotative capability - they are its extra meanings or associations.

What most distinguishes connotation as a kind of meaning is that the social attributes embedded in connotations are specific to social groups of many different sorts and sizes, and even to individuals (because of their personal memories or experiences). Some connotations become generalised across a whole language community, and may be relatively stable over time (as in the case of champagne); but the connotations of a word will often be different depending on where you look from, and when. The compound expression village life, for example, may connote traditional pastoral tranquillity and values; or it may connote dullness, claustrophobia or threat. As social experience of village life changes, or if village life is seen or publicly represented from new or different perspectives, connotations of the expression will alter. This is true even of the textbook examples, for instance when champagne takes on
connotations of decadence, or mindless or ritualistic self-indulgence (as in they opened the inevitable champagne) - even leaving aside the imputation of hypocrisy associated with the word in its oxymoronic, compound expression champagne socialist.

Connotations are attributes derived from perceptions of the referent. As a result, they are often thought to hold across different media (e.g. the connotations for something will be the same either for a word or a picture of it). If this is the case, then connotations might in theory also be relatively stable from one language to another, and not something the non-native speaker needs to learn afresh. But because within any given language area connotations vary from social group to social group, and because differences between languages are often also differences of lived environments, it is doubtful how far this is actually the case.

What makes connotations particularly important is that they charge up language, generating much of the power that words have to instil attitude or inspire emotion; and when the power of a connotation colours the meaning of a sentence in which it occurs, connotation merges into what is called semantic prosody, or the capability of a word to create strong expectations about what follows it, or to establish its own attitudinally-marked collocational context. More puzzlingly, the powerful reactions that words trigger - which have been described so far as highly culturally specific connotations - have been suggested by some empirical research to be reducible to a small number of core factors of attitude, factors which can be distilled into three principal scales: good/bad, strong/weak, active/passive. Osgood’s semantic differential (Osgood et al., 1957) is now generally neglected because of the behaviourist paradigm in which the relevant research was framed; but it remains an important, cautionary statement of the non-conceptual, emotive force-field of lexical effect.

Faced with evidence of social as much as linguistic power inherent in lexis, we must ask how reliable the single native-speaker’s intuitions will be. Clearly the value of such intuitions will vary, depending not only on the multiple determinants of the informant’s social identity, but also according to the particular native speaker’s ability to relate her or his own associations and beliefs to views and attitudes attributed to others.

5. Can I use this word 24/7? Or is it only good for formal essay writing, etc.? 

Alongside a word’s denotative meaning and connotations, other effects are created depending on the expression’s relation to the circumstances of its use.
Those circumstances include both its collocation, or pattern of co-occurrence with other words, and also the type of situation or context in which it is used. Knowing these dimensions of meaning is essential to active lexical knowledge: to knowing when to use a word and when not to. Such aspects are typically indicated in dictionaries as usage constraints, either in terms of time (e.g. ‘archaic’, ‘obsolete’), of dialect (e.g. ‘Southern English’, ‘US’), of status (e.g. ‘colloquial’, ‘slang’), or of speaker attitude (e.g. ‘dismissive’, ‘offensive’).

Textbook cases of register (such as general salt compared with technical sodium chloride, or horse compared with steed, nag or gee-gee) fit established scales quite neatly. But even a stylistic series like bring to an end, finish, conclude, complete, and terminate is more troubling as regards discrete intuitions, and this can lead to difficulties for non-native speakers where so many situational variables are in play. How, for example, could the non-native speaker who recently asked about use of reckon in his essay have grasped the word’s register properties more successfully? He had identified a relationship between reckon and reckoning, understanding historical connections between the word and the notion of calculation; my suggestion that reckon is a more colloquial alternative to think, rather than a marker of formal estimation or computation, inevitably seemed counter-intuitive. It is easy to say the speaker could have noticed collocations and situations of use that mark this verb as informal and mostly spoken; but pointing up difficulties with his own initial hypothesis about reckon seems to depend as much on noticing the casual nature of the thought processes typically being reported following reckon - i.e. on an interpretive judgement - as on the word’s collocations or patterns of use.

A corresponding difficulty with scales and contrasts may apply in the acquisition of colloquial words, slang and so-called taboo words. These expressions are easily recognised in context as attitudinally loaded; but they are likely to seem, at least initially, to be members of a single class rather than elements within a network of graduated sensitivity and potential offensiveness (though some system of scales is clearly suggested by public-opinion research on swearwords that periodically results in a bad language top-ten). Appropriacy judgements model not just how words reflect situations, but also a dynamic sense of how vocabulary choices act in, contribute to, or provoke situations. Dictionary guidance as to usage digests collective behaviour, extracting regularities and norms by distilling from very varied social experience and different ways of behaving. By contrast, each native-speaker informant is a located individual, with individually shaped horizons of recognition, empathy and tolerance as regards usage.
6. Is this something you would write or say yourself? What sort of people use it?

As in the previous section, this aspect of lexis is concerned with words less as means of denoting something (i.e. for their ideational properties) than as contributions to the interpersonal dimension of communication. Sometimes this area of lexical effect is called ‘social’ meaning, but it also overlaps with the ‘tenor’ aspect of traditional classifications of register. Words communicate information about the speaker’s origins (as well as emotional state and attitude towards the addressee and topic); and effects of this kind are as important for non-native speakers as for native speakers, if the non-native speaker is not to appear devoid of mood or character, or to display misleading indicators of personality (e.g. when some non-native speaker discourse appears oddly conservative because of vocabulary items drawn from nineteenth-century literary sources that are sometimes used without context as educational materials).

Again the textbook illustrations seem straightforward. Using the word *outwith*, for example, is likely (if noticed) to suggest that the speaker may be Scottish, speaking or writing either formally or informally, whereas *dreich* may still indicate Scottishness but now of an informal or colloquial variety that would (for instance) be a marked choice in writing. Many words have demonstrably regional distributions, though what amounts to the map for each word - and so the condition of its specific effect when the expression is used off-map - may not be recognised by many native speakers of the regional dialect themselves. Social markers, such as those of the U and non-U classifications of the 1950s, or updated fifty years to reflect the complex social relations of contemporary English-using societies, function similarly, but the mapping in this case is of social fields and classes rather than regions, and the speaker-identities are a combination of permanent and more transient roles, positions and aspirations.

It is hardly news that the facts of language use overspill such basic classifications. People and situations are not as separate as our distinction between dialect and register suggests. Speakers move in unpredictable ways between regions and circumstances; some actively command many registers, others fewer; and some speakers code-switch between regional or social dialects in ways that form a major part of their pattern of register variation and deployment of cultural capital. Language users, that is to say, have differently activated repertoires within what, from the point of view of English as a whole, is an abstract system of choices. Like social choice in politics, though, it is not
only the systemic possibility of choice that matters but also the on-the-ground realities of available alternatives, in this case for any given speaker. Using and responding to English involves tuning into, learning and engaging with prevailing, but always subtly changing associative meanings that reflect and possibly reproduce social relations - which raises the question how reliable, in such circumstances, appropriacy judgements made by any individual native speaker will be.

Asking a native speaker is not like consulting a reference book or corpus of language data. With a speaker’s intuitions, it is not the language as a whole, or as represented by a wide-ranging sample, that is being consulted, but one particular speaker’s repertoire and awareness of her or his varied and changing language environment. Idiosyncrasies of personal history may be pervasive within what appears to be the common wisdom of the language community - perhaps especially with expatriates, for whom social markers may be the first aspects of the native language to slip from currency. Only with data from many native speakers do variant models combine into an overall, polyphonic image of a public language. An individual’s judgment that an expression ‘is not something I would normally say’ may bring with it a complex mix of dialectal and register baggage.

4. Conclusions

Users of a language do not generally invite comment in the particularised ways indicated above. They seek broad-brush reactions to utterances or writing, reassurance about ‘oddities’, or general advice on how to revise something. Each of the different question-types above picks out a specific aspect of lexis; but in any given utterance or text the different properties of each word function together, and successive words interact in complex patterns of co-selection. ‘Knowing a word’ has many dimensions, and so does using it.

Extending second-language vocabulary is for this reason a multi-dimensional process, involving not only acquisition of systematic, linguistic knowledge but also an accumulation of socio-cultural awareness that will vary from one English-using country to another. The shades of nuance and style at stake include, but also go beyond, formulae which might assign ready-made roles to interactants in typical transactions or interactions; and lexical awareness evolves into not so much a dispassionate grasp of abstract verbal alternatives as a process of becoming located within a language ‘world’. When in the ‘Preface’ to Keywords (Williams, 1983), Raymond Williams describes how the book was
written to address the sense he had formed, at a moment of significant cultural change, that many of his fellow users of English ‘just don’t speak the same language’ (Williams, 1983:11), he draws attention to a common word stock with varying, sometimes conflicting uses made of it to articulate core aspects of social, including political experience. Much in our lexis is not in this sense a matter of key words, but the same general principle of potentially discrepant social meanings applies. When a native speaker explicates lexical meaning or effect, she or he is responding to denotation and sense relations interlaced with features of connotation, register, and perceived speaker intention. The experience brought to bear on these lexical effects is inevitably selective, partial and to some extent partisan; so it is important that, for all the richness of native-speaker insights into their own usage, trust in native-speaker intuitions should be kept within limits that reflect this. Meaning or effect judgements put forward in a given context are interpretive rather than semantic; intuition plays the role of practical guide rather than golden key.

Developing educational materials to stimulate and support advanced vocabulary work is a pedagogic challenge that requires us to combine intuitions as to meaning and effect with distributional facts gathered and represented in other ways, especially in dictionaries and corpus files. How to link observed facts about lexis with native-speaker intuition and theoretical enquiry is something nobody would claim to have perfected. And that in itself is reason enough why studies along the lines of, and inspired by, Moira Linnarud’s commitment throughout her career to non-native speaker lexis will have a continuing role to play, both in research and in the classroom.

References