RaY

Research at the University of York St John

For more information please contact RaY at
ray@yorksj.ac.uk

This is a pre-copyedited, author-produced PDF of an article accepted for publication in

ELT Journal following peer review. The published version is available online at:

http://eltj.oxfordjournals.org/content/early/2014/03/18/elt.ccu016.

Moving beyond accuracy:

from tests of English to tests of *Englishing*

Christopher J. Hall

York St John University

Abstract

This article examines how English is conceptualized in the domain of testing, and particularly the tendency to identify the concepts of 'standard English' and 'native English' with 'the language itself'. I argue that such a monolithic view is inconsistent with the diversity of Englishes attestable across both native and non-native users and uses, and that this undermines the inclusion of accuracy criteria in English language tests. Adopting an alternative 'plurilithic' orientation, I challenge the traditional view on both cognitive and social grounds, arguing that the Englishes encountered and appropriated by non-native speakers will inevitably be qualitatively different from 'standard English' models, and that the effectiveness of the resources learners do develop should be assessed, where appropriate, independently of linguistic criteria. I conclude that a shift is required from tests of English to tests of *Englishing*—from testing how people *use* the language to testing what they can *do* with it.

In the past decade or so, teachers have become more aware of the global diversity of English and the implications of this for ELT. But the notion of multiple Englishes has yet to be granted legitimacy by many, and even for those who recognize the inevitable variation in forms which comes with global usage, it is still widely believed that accuracy should be a central learning objective. In this article I examine how English is conceptualized in testing, questioning in particular the tendency to identify the language uniquely with a single 'standard' variety. In doing so, I am motivated by the precept articulated by Alderson and Banerjee (2002: 80):

> Central to testing is an understanding of what language is, and what it takes to learn and use language, which then becomes the basis for establishing ways of assessing people's abilities.

Admittedly, great strides have been taken over the past half century to understand 'what language is' such that 'people's abilities' can be adequately assessed. Especially notable is the expansion of assessment design and practice from a unique concentration on accurate *forms* to the broader *functional* notion of communicative competence (reflected in the 'can-do' statements of the Common European Framework of Reference, for example). And yet accuracy is still a major component in most tests, all the way from formative classroom assessments to big standardized tests. For accuracy in production to be assessed, a unitary set of norms is needed as a yardstick, so that errors can be counted and correctness judged. This requires a

monolithic ontology of English: a belief in the existence of a single definitive inventory of correct forms.

Such a monolithic view does appear to prevail, albeit implicitly, in the mindsets of many people, including teachers and linguists (Hall 2005, 2013). The belief is a tenacious one, accounting for a range of phenomena in language ideology, including the expectation of conformity with codified norms in English-medium higher education and the aspirations of millions of NNSs to learn 'the best' English. Of course the general public, including liberal academics, will not be abandoning their insistence on 'standards' in English any time soon, and as Taylor (2006: 51-2) points out, it would be foolish to take no account of learners' views. So although my aim here is to invite teachers and testers to question the monolithic position, I should declare from the outset that I am not arguing for 'standard' varieties to be abandoned as targets in all learning contexts. Indeed, I recognize the need to test conformity with such varieties under many circumstances (for example in some EAP contexts), as well as to 'make allowance for individual aspirations to [NS] norms' (Hu 2012: 138). Moreover, I acknowledge that form-focused instruction, informed by 'standard' English norms, facilitates learning. What is not helpful, I maintain, is the presentation of these norms as the only ones for successful English usage.

Given the evidently contentious nature of the issues I raise and the somewhat radical orientation I adopt, and considering my own lack of expertise in testing theory and practice, my goal here is to encourage reflection and debate, rather than presume to offer any blueprint for action.

'Standard English' ≢ 'the language itself'

Two major problems with the monolithic view will be raised in this paper. One is social, the other cognitive, and they both compromise the construct validity of current English language tests. The first, essentially social, issue stems from the fact that 'NS English' is generally equated with the notion of 'Standard English' (SE), which in turn is implicitly viewed as constituting 'the language' itself. Yet SE does not adequately reflect the linguistic competence and performance of most NS and NNS users and uses. Given the prevalence of unstandardized forms in NS usage, and of variance and innovation among NNSs, it is clear that SE will not be relevant for all contexts of learning and use. So, I will argue, SE is inauthentic as the sole reference point for assessing the adequacy of NNS forms.

'Standard English' is unattainable by most adults

The second problem with the monolithic position is that 'native-like' SE is not actually attainable for most adult learners in educational contexts, because of the way learners experience the language and the way their minds process it. Learners will invariably develop sets of linguistic resources that differ qualitatively from those of L1 users, even if they are functionally equivalent to the 'target' variety. This suggests that monolithic NS norms might be inherently unrealistic as attainable learning outcomes, and that an insistence on them in teaching implicitly invokes a deficit model of learning. If so, it follows that tests cannot reliably measure learning trajectories as stages on a continuum from 'no competence' to 'native-like competence' in SE.

In the light of these problems with the monolithic approach, I propose moving towards what we might call a 'plurilithic' alternative. I make the case that, where appropriate, we need to take the focus off language forms altogether, and move towards a position proposed 30 years ago by Troike (1983: 214):

> Since language often appears to be learnt best when it is not the direct object of instruction, it may be that language is tested best when it is not the immediate object of assessment.

The arguments for the position presented below draw on sociolinguistic and psycholinguistic approaches to language learning and use, and on the experiences of both NSs and NNSs. First, however, I will say a little more about monolithic and 'plurilithic' views.

Monolithic and 'plurilithic' approaches to English

Human language is normally used unconsciously and its development, storage, and processing are not directly observable to users. Given the impossibility of telepathy, this is the optimal evolutionary design, providing as it does a seemingly instantaneous but altogether physical medium for communication. Cognitively, then, we operate under what I have called 'the language spell' (Hall 2005), oblivious to the psychological nature of the formal resources through which we share meanings and construct our identities. On the sociocultural plane, a side-product of the language spell is a relatively recent (500-year-old) public fixation on language as a national attribute, in which metalinguistic awareness has been formed by ideological processes. As a result,

people tend to adopt uncritically a belief in 'the' English language, identified with SE, the dialect of institutionalized power. This view of SE as 'official', the language of the state, has been fostered over the centuries by educational authorities, leading to a 'doctrine of correctness'.

Such factors have resulted in a deep-seated belief in the essentially monolithic nature of English, in which 'the' language is governed by a single set of 'correct' grammatical rules. Its educated native speakers are the (albeit imperfect) exponents and guardians of its idealized form. It is this form to which its users are expected to aspire, in acquisition and use. The proficiency of both native acquirers and non-native learners is judged, either directly or indirectly, on their ability to conform to monolithic SE norms.

An alternative view, informed by sociolinguistics and associated with several minority viewpoints within general and applied linguistics, is that English and all named languages are essentially fictions (albeit necessary ones). In Pennycook (2007), for example, the belief in English as a 'real world object' is characterized as a myth, and he invites readers to separate 'doing things with the language, in terms, perhaps, of Englishing' from 'the institutional entity English' (p. 111). Hall (2013) appropriates some of Pennycook's critical claims, as well as his later-coined term *plurilithic*, in order to expose the 'myth' of monolithic English from a cognitive, as well as a social, perspective. This view sees English, and all other 'named languages', not as singular reified entities (Greek *mono* 'single' + *lithos* 'rock'), codified in books and dictionaries, but as dynamic sets of overlapping phonological, grammatical, and lexical resources, stored in millions of individual minds, which interact in multiple communities and

cultural practices. In other words, rather than a planet comprising a single terrestrial mass, English resembles a galaxy of millions of discrete objects (a plurality of *lithoi*), bound together by the gravitational pull of effective communication.

Pennycook (2007: 112) argues that recognition of the plurilithic nature of English should lead language teachers to reassess what and how they teach. But what about testing? Nowhere are the monolithic myth and plurilithic reality of English more evident yet less acknowledged than in testing, an activity which perhaps more than any other dictates what is taught. I think we can identify both cognitive and social corollaries of the monolithic position which compromise the practical and ethical value of traditional approaches to testing, and which jointly signal the need to reorient it along the plurilithic lines Pennycook proposes for pedagogy.

'Standard English' ≢ 'NS English'

Let us start at the beginning, examining the concepts of 'standard English' and 'NS English' from a plurilithic perspective, before moving on to NNS Englishes. First, note that the monolithic concepts of 'correctness' and 'accuracy' assumed in most tests are associated uniquely with SE, and this is implicitly equated with NS usage. Unfortunately, much of the NS usage that L2 learners will experience outside the classroom does not abide by SE norms. This represents a problem of authenticity for tests, and also raises the more general issue of the circumstances under which 'correctness' and 'accuracy' can be judged to be appropriate assessment criteria. Consider the following 'incorrect' options from a random selection of four online multiple choice grammar tests for L2 English learners:

1 You look cold. Will I close the window for you?

2 Will we go to the cinema?

3 Will I carry this bag for you?

4 You have a terrible fever! Will I call a doctor?

In these examples, the 'mistake' is the use of *will* instead of *shall* in first person questions intended to obtain advice or confirmation of a proposal. This may seem self-evident and uncontroversial to many readers. And yet for very significant numbers of NSs, the 'mistake' is part of the linguistic resources they have naturally acquired, which they retrieve from long-term memory in their everyday interactions, and which identify them as NSs from a particular locality. Consider the following utterances, for example:

5 Will I bring along a train set the next time?

6 Will we count how many hedgehogs there are?

7 Will I phone us a taxi?

8 Cathy, will I put the kettle on for you?

These utterances were spoken by four different adults from Belfast, Northern Ireland, in interactions involving pre-school children and their mothers.[1] The utterances, and many others like them, demonstrate that in the mental lexico-grammars of members of this quite typical NS community, modal *will* is used where other NSs use *shall*.

Unsurprisingly, all the children in the study naturally acquire the forms used by the adults around them. Here are examples from four of the children, taken from the same set of interactions in which (5) - (8) appeared:

9   Will I show you what I can do?

10   Will we count how many bees there are?

11   Will I open them for you?

12   Will we make this again?

According to a strong monolithic orientation, what these children acquired and use naturally is incorrect. In the 'proper' version of 'the language', the correct form is, unquestionably, the *shall* option licensed by SE and considered as 'NS usage' in tests for NNSs. Of course, it cannot be denied that SE norms are more *socially valued* than unstandardized ones for many purposes. Indubitably, all children living in officially Anglophone countries need the opportunity to develop competence in SE, because its resources are those institutionalized in formal literacy and oracy practices, and they are needed for socialization into many aspects of public life (cf. Widdowson, 1993). Accordingly, after they have acquired their own English, most NS children become familiar with SE norms to a greater or lesser degree by engaging in the processes of literacy development and *Second Dialect Acquisition* (SDA; Siegel 2010). These occur principally via formal education.

Data on what happened to these Belfast children after they started school are unavailable, but it is certain that they would have been exposed to, and were perhaps explicitly taught, the (arbitrarily) 'correct' SE versions of (9)-(12). Like them, most NSs

develop two overlapping sets of norms, corresponding to a naturally acquired local English and a learnt repertoire of SE forms. Whatever metalinguistic awareness they develop will be associated with, or referenced to, SE norms—hence the perpetuation of the 'doctrine of correctness'. The limited data within the SDA literature on the learning of 'standard' dialects in classroom contexts suggest that native mastery is rare (Siegel 2010). It appears that learnt lexico-grammar (whether SE or L2) is stored in *declarative memory*, in a neuroanatomically distinct set of brain structures, as explicit knowledge *about* language. This knowledge can be deployed only deliberately, with more or less confidence, most frequently in formal writing. If used often enough, the forms learnt about may also be incorporated into *procedural memory*, in separate brain regions, where they can be activated automatically for spontaneous oral interaction (Paradis 2009). The significance of this for NNS Englishes will become apparent in the next section. For the moment, we will simply conclude that few, if any, NSs know and use only SE; many use it only with conscious effort; and most have unstandardized rules and structures as part of their natural linguistic repertoires, which they deploy in unmonitored usage.

Challenges for testing

It follows from the foregoing that 'grammatical accuracy' is a spurious criterion if used in tests which aim to assess learner approximations to authentic NS competence across a broad range of usage contexts. Furthermore, as I now go on to demonstrate, the monolithic position is also challenged when we recognize the social and cognitive

realities of NNSs' learning experiences, and when proficiency is viewed independently of the native speaker.

The cognitive challenge

As we have seen, the monolithic position rests, in part, on the assumption that there is a single attainable target set of forms (those licensed by SE norms), which is made available to learners through teachers, textbooks, and tests. I will shortly be challenging this assumption, but for the moment let us suppose that learners are indeed exposed to a single version of English, namely 'standard' (British or American) English. On the traditional view, the learner's task is essentially to 'clone' in their own minds the phonological, lexico-grammatical, and semantic system underlying the variety. Proficiency is defined, in part, in terms of the accuracy with which the system has been copied, and how faithfully linguistic output conforms to the system in performance. Output that fails to match the external system is judged 'errorful'. This reflects, as I have already suggested, a deficit view of language learning, in which most learners fail to one degree or other.

Most actual language learning is not the result of the step-by-step internal reproduction of an externally existing 'target' system. It is true that substantial amounts of English knowledge develop through instructionally-regulated explicit learning. But for the majority of learners, English resources also inevitably develop through natural acquisition, resulting in the dynamic construction of a unique and constantly shifting mental system, created on the basis of unique, localized experiences of use (just as in the case of the Belfast children). Here the implicit

objective is *Englishing*, not acquiring English, and the driving force is successful communication, not the prospect of assessment. This view is consistent with Usage-Based approaches to language acquisition and storage in memory (cf. Eskildsen 2009), according to which individuals develop linguistic resources through interaction with other users. They do this by gradually abstracting away from experienced and initially unanalysed form-meaning pairings to more general constructional patterns (or rules), on the basis of frequency of use in particular sociocultural practices.

Inevitably, the patterns underlying Englishing that develop in learners' minds through usage events will not be identical to the single SE model taught and modelled by teachers, textbooks, and tests. Learners in the globalizing world are increasingly exposed to a great deal of NNS and unstandardized English: in interaction with peers and teachers inside the classroom; and increasingly *outside*, through participation in online activities like gaming, social networking, special interest forums, and the like, and through everyday life in multilingual environments such as migrant communities, tourist zones, and major urban centres. The knowledge of English acquired this way, through implicit procedures rather than explicit 'noticing', leads to fluency, but not necessarily 'accuracy' in SE. And it's worth noting that teachers can't stop it happening.

Learning, then, is not confined to internally reproducing the external target system as declarative knowledge, but also involves constructing a unique, personal, procedural system which will inevitably differ from the taught system in important ways. For assessment, the implication is clear: testing conformity with the external system is inappropriate unless accuracy in SE is a requirement for future contexts of use (for

example in EAP or for certain jobs in NS contexts). Moreover, testing accuracy in SE is also counter-productive: the more that assessment expectations set learners up to reproduce an external system of declarative *facts*, the harder it is for them to develop automatized *procedures* for effective use (which is presumably the prime objective of ELT). Ultimately, *proficiency* can most usefully be understood as 'ability for Englishing', and Englishing can optimally be done through the implicit regularities that learners themselves detect and build on, rather than the explicit regulations of SE underpinning the accuracy criterion.

The social challenge

According to a strong monolithic orientation, 'the English language' belongs to its native speakers, and non-native speakers can only 'borrow' it. This is the false premise behind the common identification of all L2 English as English as a *Foreign* Language, effectively challenged by research on English as a Lingua Franca (ELF; cf. Jenkins 2006). High stakes tests like TOEFL, IELTS, and the Pearson Test of English use NS competence as a yardstick, glossing over two important facts:

- Most NSs use, identify with, and are identified with, unstandardized forms (as we have seen); furthermore, many NSs are poor communicators in some contexts.

- Most interaction in English these days is between NNSs, who often have different first languages, resulting in a great diversity of forms being exploited, often expertly, for mutual understanding (ELF).

Consider, for example, the following 'can-do' descriptor for the highest proficiency C2 level of spoken interaction on the CEFR:

> *I can take part effortlessly in any conversation or discussion and have a good familiarity with idiomatic expressions and colloquialisms.*

The assumed interlocutor is the NS, who uses culturally-bound language forms that users in ELF contexts would not need or would sensibly avoid, unless they were planning to use their English primarily or exclusively with NSs (Jenkins 2006: 47-48).

A plurilithic stance privileges no single group of speakers or set of forms. Instead, it recognizes the unique interaction between specific groups of English users' shared or unshared experiences and identities, as they negotiate languaging events using unequal linguistic (and other semiotic) resources, at specific points in time, for specific purposes. From such a perspective, the kind of proficiency to be developed and tested must be locally tuned, sensitive to users' individual needs and identities.

A plurilithic perspective on testing

Recognition of the plurilithic reality of English necessitates a radical rethinking of the nature and purpose of English testing. Consistent with a plurilithic perspective would be the assessment of a learner's *Englishing*: what they do with the language in specific situations. As suggested decades ago by Troike (1983) and argued since by many others, what should be assessed (where possible and appropriate) is learners' performance on communicative tasks in English that are appropriate to their own goals and contexts, rather than their knowledge and use of 'English itself'.

Of course, this is likely to be considered impractical for many contexts in the short term, especially given the vested interests of the big testing organizations and examination boards, and the strict conventions of academic writing. But in many local classroom contexts it may provide a more valid and relevant approach to assessing the kind of Englishing ability that learners may actually need. The idea is (imperfectly) realized in at least one international test: the Test of Interactive English (TIE), as described in this journal a few years ago by McGinley (2006). TIE is based on self-selected projects that the learner engages in during classroom-based learning, such as following a news story and investigating a topic of interest. At test, candidates speak and write about these activities, as well as performing a spontaneous problem-solving task with another candidate. As the test handbook explains (TIE 2012: 4): 'The aim is to get students to bring their own interests, opinions and experiences into the classroom and use them in order to perform their best at the test'.

But monolithic assumptions are too entrenched to be resisted entirely. Although TIE candidates 'are assessed on what they "can do" rather on what they are expected to be "able to do"' (TIE ibid.: 4), the assessment criteria include 'the accurate and appropriate use of the candidate's grammatical resources' (TIE ibid.: 15), in both spoken and written components. In a sample marked paper, for example, the phrase *variety of options <u>to choose</u>* is flagged as an accuracy problem (TIE ibid.: 24), presumably because a final preposition *from* is required in SE, even though its lack has no negative effect on understanding. It is clear from examples such as this that what is being tested is not, in fact, 'the accurate and appropriate use of *the candidate's* grammatical resources' (emphasis mine), but rather, those of the external model, SE.

There is, therefore, a long way to go before we see tests which focus on Englishing rather than English. But in the short to medium term, ELT professionals convinced by the arguments presented here can press for at least two fundamental shifts in testing orientation, each consistent with the plurilithic reality of English:

- Following Jenkins (2006), tests should include NNS as well as NS Englishes for the assessment of receptive skills, and should reward the successful use of communicative strategies over conformity with SE forms in the assessment of productive skills.

- Following Troike (1983), tests should move to more explicit recognition of the non-linguistic factors involved in contextualized Englishing, such as personality and knowledge of content (for example professional, academic, and cultural).

Ultimately, the best tests will not be tests of any kind of English, whether SE, 'Interactive English', English as an International Language, or ELF. In future, as English becomes viewed as a set of basic skills like numeracy (Graddol 2006), learners will take tests which require Englishing, but are not tests of English.

Conclusion

Testing how people *use* English, rather than what they can *do with* English, is problematic, because it assumes a fixed notion of what English is, such that *'it'* can be used. In other words, testing English, rather than Englishing, inevitably requires acceptance of a monolithic view of the language. It denies the plurilithic reality of English as it is differently known across NS and NNS users, and as each of these

cognitive representations of English changes continuously through learning, use, and mixing with other languages.

As Jenkins (2006) points out, testing regimes must change if ELT practices are to have any chance of adapting to the new realities of globalized English and ELF. I would argue that an understanding of the plurilithic nature of English, at both social and cognitive levels (Pennycook 2007; Hall 2013), is essential if testers and teachers are to fully embrace the challenge. This shift in approach won't be achieved only through adjustments to current testing practices (Taylor 2006). Instead, it will require testers, teachers, and other stakeholders to reappraise their fundamental beliefs about what English is and how it is learnt and used.

Note

1. The data are taken from transcriptions made by Alison Henry and John Wilson in their work on the acquisition of Belfast English, available in the CHILDES database (http://childes.psy.cmu.edu).

[4000 words]

References

Alderson, J. C. and J. Banerjee. 2002. 'Language testing and assessment (Part 2)'. *Language Teaching* 35, 79–113.

Eskildsen, S.W. 2009. 'Constructing another language — Usage-based linguistics in second language acquisition'. *Applied Linguistics* 30/3, 335-57.

Graddol, D. 2006. *English Next*. Online. Retrieved 8 August 2013 from:

http://www.britishcouncil.org/learning-research-englishnext.htm.

Hall, C.J. 2005. *An Introduction to Language and Linguistics. Breaking the Language Spell*. London: Continuum.

Hall, C.J. 2013. 'Cognitive contributions to plurilithic views of English and other languages'. *Applied Linguistics* 34, 211-31.

Hu, G. 2012. 'Assessing English as an International Language' in L. Alsagoff, S. McKay, G. Hu and W. Renandya (eds.). *Principles and Practices for Teaching English as an International Language*. New York, NY: Routledge.

Jenkins, J. 2006. 'The spread of EIL: a testing time for testers'. *ELT Journal* 60/1: 42–50.

McGinley, K. 2006. 'The 'Test of Interactive English'—from conception to implementation'. *ELT Journal* 60/4: 374-81.

Paradis, M. 2009. *Declarative and Procedural Determinants of Second Languages.* Amsterdam: John Benjamins.

Pennycook, A. 2007. 'The myth of English as an International Language' in S. Makoni and A. Pennycook (eds.). *Disinventing and Reconstituting Languages*. Clevedon: Multilingual Matters.

Siegel, J. 2010. *Second Dialect Acquisition.* Cambridge: Cambridge University Press.

Taylor, L. 2006. 'The changing landscape of English: implications for language assessment'. *ELT Journal* 60/1, 51-60.

TIE. 2012. *Test of Interactive English. Handbook.* Dublin: ALCS. Retrieved 18 July 2013 from: http://www.tieexams.com.

Troike, R. C. 1983. 'Can language be tested?'. *Journal of Education* 165/2: 209–16.

Widdowson, H. G. 1993. 'Proper words in proper places'. *ELT Journal* 47/4, 317-329.

The Author

Chris Hall is Associate Professor and Reader in Applied Linguistics at York St John University in the UK. He received his PhD in linguistics from USC (Los Angeles, USA) in 1987, and taught EFL and applied linguistics at the University of the Americas (Puebla, Mexico) between 1990 and 2007. He is the author or co-author of several books, including *An Introduction to Language and Linguistics: Breaking the Language Spell* (2005), *Mapping Applied Linguistics* (2011), and *Introducing Language in Use* (2nd edn, 2014). His British Council-funded online resource *Changing Englishes: An Interactive Course for Teachers* was launched in 2013. Email: c.hall@yorksj.ac.uk