Norms and varieties of English and TESOL teacher agency

M. Obaidul Hamid
The University of Queensland, Brisbane, m.hamid@uq.edu.au

Lingyan Zhu
Zhejiang University of Finance and Economics, PRC, lyzhu1980@hotmail.com

Richard B. Baldauf Jr.
The University of Queensland, r.baldauf@uq.edu.au
Norms and varieties of English and TESOL teacher agency

M. Obaidul Hamid
The University of Queensland

Lingyan Zhu
Zhejiang University of Finance and Economics, PRC

Richard B. Baldauf
The University of Queensland

Abstract: The growing recognition of the plurality of English underlying the World Englishes (WE) paradigm has problematised the conventional second language acquisition (SLA) views of errors. If English use in emerging English-speaking contexts is to be judged by local norms, as argued by WE scholars, applying exocentric norms in these contexts can be inappropriate. On the other hand, despite the significant growth of WE, varieties of new Englishes have yet to develop widely acceptable endocentric norms. These developments have raised a critical question: How can TESOL teachers distinguish between errors in the SLA sense and varietal features in the WE sense? Framed around language management theory (LMT) and teacher agency, this paper investigates how a group of global TESOL practitioners in an Australian university evaluated usages of English as a second language, what criteria they used and what implications their judgments and decision-making processes have for TESOL pedagogy and WE research.

Introduction

Regardless of what language purists (e.g. Honey, 1997) have to say about non-native Englishes, the world Englishes (WE) paradigm that has brought together varieties of Englishes (see Bolton, 2006) has been widely recognised in the past three decades. Kachru (1988) argues that WE makes three basic claims:

1) there exists a repertoire of models for English, not just the native speaker (NS) varieties;

2) localised innovations in English which have their pragmatic bases deserve recognition; and

3) the ownership of English lies with all those who use it.

Recent interpretations of WE have led to its more comprehensive characterisation. For example, Proshina (2012) identified seven characteristics of WE: diversity, plurality, equality, functionality, inclusivity, appropriateness and variability. Essentially, it is argued that the global spread of English has generated multiple norms at the local, national and regional levels in response to socio-cultural and pragmatic needs of speech communities. These non-native speaker (NNS) norms should be treated as equal to native speaker norms and should serve as criteria in judging the appropriateness of English use in Outer Circle contexts (Kachru, 1983, 1990).

These WE arguments challenge the traditional second language acquisition (SLA)
paradigm which set native-like competence as the English learning goal for all learners (Sridhar & Sridhar, 1986). As Kachru (1985, p. 30) explained:

In my view, the global diffusion of English has taken an interesting turn: the native speakers of this language seem to have lost the exclusive prerogative to control its standardization [...] What we need now are new paradigms and perspectives for linguistic creativity in multilingual situations across cultures.

Second language (L2) errors can be taken as an example to illustrate divergences between SLA and WE views of linguistic competence. In SLA, deviations from NS norms, which are believed to result from limited linguistic competence, are called errors. However, these may not necessarily be errors in WE, particularly when an idiosyncratic form appears systematic and is shared by a speech community.

It must be acknowledged that the orthodox view of SLA as represented above has given way to more complex and diversified understandings in recent years. For instance, a constructive interpretation of errors—errors as signposts to learning—has informed error analysis from the beginning (Corder, 1981). Similarly, discussion of native-like competence as L2 learning goal, as pointed out by Sridhar and Sridhar (1986), needs to refer to the concept of multi-competence, as expounded by Cook (2008). Likewise, principles of instructed second language acquisition (Ellis, 2005, 2008) point to a comprehensive view of L2 learning goal which may not necessarily correspond to NS norms. More crucially, the problematisation of ‘native speaker’ has suggested more refined categories including ‘native user’, for instance, to refer to users of English, whether native or non-native (Davies, 2013).

Despite these welcome developments, a narrowly-focused view of SLA with its emphasis on Inner Circle norms has maintained its dominance in the pedagogical domain. For example, education authorities in Singapore have strictly pursued Standard English (SE), which is labelled ‘good English’, at the expense of varieties spoken by people (Rubdy, 2001, 2007). Similarly, Hong Kong has imposed unjustifiable Inner Circle-based standards in assessing English language proficiency of local English teachers (Glenwright, 2005). Exocentric norms informed by traditional SLA have also guided English language curriculum and pedagogy in many Asian countries (Kaplan et al, 2011; Kirkpatrick, 2002).

The dominance of traditional SLA in TESOL pedagogy on the one hand and the increasing recognition of WE on the other raise a critical question: How can we distinguish between errors in the SLA sense and varietal features in the WE sense (see Hamid & Baldauf, 2013; Bamgbose, 1998; Davies, Hamp-Lyons & Kemp, 2003; D. Li, 2010; Mollin, 2006; van Rooy, 2011)? As Bamgbose noted:

The main question that arises with innovation is the need to decide when an observed feature of language use is indeed an innovation and when it is simply an error. An innovation is seen as an acceptable variant, while an error is simply a mistake or uneducated usage. (1998, pp. 1-2)

The distinction is essential in TESOL pedagogy because TESOL teachers, who enact language policies as ‘performative action’ in the micro context (Lo Bianco, 2010), are expected to correct errors and nurture innovations. Failure to draw clear boundaries between them may have academic and social consequences (see Hsu, 2012; Kenkel & Tucker, 1989).

One—probably the most efficient—way of drawing boundaries between errors and innovations is by looking through a large corpus of L2 English data, as some researchers have done (e.g. D. Li, 2010; Okunrinmeta, 2011), to discover patterns that point to some kind of standard. However, as we discuss in more detail later, this approach has limitations. First, it
assumes that L2 English has gone through the whole process of institutionalisation and codification, which, in reality, is still a long way away. It may be reminded that Australian English has completed this process only recently. Second, TESOL teachers have to provide judgments on L2 English as part of everyday teaching practice; they cannot suspend their judgments until local norms are established and made available. In other words, since teachers are already involved in the process of judging L2 English as part of their practice, it is important to investigate the processes and criteria for distinguishing between errors and innovations from their perspectives. However, there has been inadequate research on teachers’ judgments of L2 use from the SLA-WE contrastive perspectives. For instance, we do not know 1) whether there is a consensus among teachers in their understandings of errors and varietal features; 2) what criteria they use in judging the status of L2 features; and 3) whether these criteria are different from those suggested in the literature. The present study seeks to explore these issues by generating insights into the following questions:

- What kind of acceptability judgments do TESOL teachers offer on select features of L2?

- What criteria and decision-making processes underlie teacher judgments of the status of L2 use as errors or varietal features?

Errors and Innovations: Language Management Theory

Although error research has drawn considerable attention in the literature, error identification seems to have been taken for granted, without explicating the process (Hamid, 2007; Hamid & Doan, forthcoming). Language management theory (LMT) (Jernudd & Neustupný, 1987; Jernudd & Nekvapil, 2012), which focuses on the management of language utterances distinguishing them from the generation of utterances, provides a framework for understanding how teachers may distinguish between errors and varietal features. Although LMT has a major focus on linguistics, it “considerably widens its scope to include all behaviours toward language” (Nekvapil, 2009, p. 3). Language management can occur at individual and institutional levels, leading to a distinction being made between “simple management” and “organized management.” Simple management can take place in four phases, as explained by Jernudd and Nekvapil (2012, p. 33):

(a) the speaker notes a deviation from the expected course of communication;

(b) the speaker can but need not evaluate the deviation; if it is evaluated negatively, the language management model refers to this as an inadequacy, if positively, a gratification;

(c) the speaker can but need not think of an adjustment to the inadequacy;

(d) the speaker can but need not implement this adjustment.

Following this language management framework, we can investigate how TESOL teachers judge L2 usages (whether they note deviations or not), whether they evaluate them positively or negatively – “this phase is essential for LMT, for it provides a clear opportunity to define a language problem” (Nekvapil, 2009, p. 3) – whether they suggest adjustment (correction) and how that adjustment might be implemented. However, LMT’s notion of evaluation of a “deviation from the expected course of communication” or norms needs further explication when applied to the WE context. LMT takes norms against which deviations are to be noted.
as a contextual given, without problematising that context. However, in the WE context, the issue of evaluation of a deviation from norms is more problematic to assess and criteria have been suggested in the literature to distinguish between errors and varietal features. These are discussed in the following section.

Errors or Varietal Features? Theoretical Criteria

Kachru (1992) made a distinction between errors and innovations which he labelled “mistakes” and “deviations” respectively, arguing that the latter are produced in new contexts of English use and are systematic within a variety. While systematicity can helpfully explain features of new Englishes (see Mollin, 2006; van Rooy, 2011), Bamgbose (1998) has proposed a more elaborate set of five criteria: 1) demographic strength or number of users using the new form; 2) its geographic spread; 3) authority of influential people or media behind the usage; 4) its codification; and 5) its acceptability. Of these, the second criterion, as D. Li (2010) argued, is hard to measure. Therefore, he recommended doing a Google/Yahoo search which he added as the sixth criterion.

Although the gate-keeping potential of these criteria cannot be underestimated, their use involves conceptual and practical challenges for TESOL teachers. For instance, it is argued that a varietal feature is social rather than individual (Kachru, 1992; Mollin, 2006; van Rooy, 2011). However, socially conventionalised forms must emerge as innovations at the individual level in the first place before these can be codified as varietal features passing the test of time and social acceptability (Hamid & Baldauf, 2013). Because codification is a lengthy process, there must a significant temporal gap between the first emergence of an innovative form (as individual creativity) and its ultimate acceptance as a varietal feature. The majority of the features of national varieties of NNS English are passing through this transitional phase at present, making it difficult for people to judge their status (as errors or varietal features) with certainty and confidence. Teachers, nevertheless, have to judge their acceptability and status as part of their professional responsibility.

Scholars have also suggested some other criteria. For instance, Smith and Nelson (1985; Nelson, 2012; Smith, 2009) have argued for the concept of intelligibility as the goal of communication which can be used as a criterion to determine the status of idiosyncratic forms. Some researchers have verified some of these criteria drawing on corpus data. For instance, van Rooy (2011), through an analysis of three grammatical features of African English, concluded that grammatical stability and acceptability can be used to identify conventionalised innovations. In an attempt to distinguish between Izon (Nigerian)-induced errors and permissible local variations manifested in the syntax of Nigerian English, Okunrinmeta (2011) made a rigorous acceptability test applying widespread usage, appropriateness, grammaticality and intelligibility as yardsticks.

However, one problem with intelligibility is that the same L2 usage may appear intelligible to some and unintelligible to others. Melchers and Shaw (2003) have classified intelligibility into international intelligibility, national intelligibility and local intelligibility which is useful because this suggests that although intelligibility at a local level may not be a problem, it can be problematic at a global level. Furthermore, although intelligibility can be relevant to grammatical deviations, we cannot necessarily expect that lexical innovations would be intelligible to people who have not previously encountered the innovation.

In addition to conceptual problems, there are also practical ones. For example, the existence of the criteria does not guarantee that TESOL practitioners will be familiar with them. Moreover, being familiar with such criteria is not a guarantee that teachers will apply them for the simple reason that curricular and institutional contexts may have policy
prescriptions for specific norms (e.g. British English in Singapore). Furthermore, teachers may also have their own ideologies about standards which may vary from the norms. This calls for understanding teacher attitudes and judgments, i.e., recognising their agency.

**Error-Innovation Distinctions: Teacher Agency**

Traditional error research has not recognised teacher agency in error detection and analysis. However, we argue that teachers including those teaching content subjects because they are teachers of languages in the first place (Bullock, 1975) need to be seen as policy actors who are also gate-keepers of language standards. This reconceptualisation is a response to recent recognition of teacher agency in related fields including language policy and planning (e.g. Baldauf, 2005, 2006; Menken & Garcia, 2010), TESOL pedagogy (e.g. Canagarajah, 1999; Lin, 1999) and teacher education (e.g. Campbell, 2012; Feryok, 2012). In the context of WE in particular, teachers need to be seen as mediators of norms and their permissible variability given the growing recognition of multiple norms (e.g. Young & Walsh, 2010). In their role, teachers decide on the acceptability of English use in a complex institutional context which is permeated by the recognition of world Englishes, policy and curricular prescriptions, students’ language variety backgrounds, the purpose for learning English, and their own language ideologies.

**Research on Teacher Judgment of Errors Versus Varietal Features**

Research on errors and innovations is at an early stage at present. A few studies that have highlighted distinctions between them (e.g. Kenkel & Tucker, 1989; D. Li, 2010; Okunrinmeta, 2011; van Rooy, 2011) have done so mainly with reference to L2 corpus data. Research that focuses on error/feature distinction based on people’s judgments includes French (2005) in Japan, Buregeya (2006) in Kenya, O’Hara-Davies (2010) in Brunei, Kaushik (2011) in India and Hamid and Baldauf (2013) in Bangladesh. Although these studies attest that teachers have been asked for judgments about error-feature issues, it is not clear what criteria the judges applied and what kind of decision-making processes is involved. The present study seeks to expand this research by focusing on the criteria and decision-making process involving TESOL teachers in a teacher-education class.

**Methodology**

**Context and Participants**

The present study was located in a postgraduate critical TESOL education class in a major Australian university. The data were the outcomes of a classroom activity related to the topic of varieties of Englishes and their implications for TESOL teachers. Seventeen TESOL teachers (out of 25 enrolled in the course), who were present on the day, participated in the activity. They constituted a global community of TESOL teachers representing the polities of Vietnam, Australia, Malaysia, Saudi Arabia, China, Hong Kong, Taiwan, South Korea and Indonesia (in numerical order). Three of them were speakers of Australian English; the others were proficient speakers of English as a second language. The first author was the course lecturer.
Corpus, Task and Data

The classroom activity was divided into two parts and was carried out in two phases. In the first phase, the participants commented on the intelligibility of seven features of L2 English (see the Appendix) individually using three response categories of “Yes”, “No” and “Not Sure”. Using the same set of responses, they were next asked to decide whether they would consider those features as errors. In the second phase, the participants were divided into three groups (six in two groups and five in one group) and were asked to discuss the usages and decide on their status as errors or varietal features.

The seven examples of L2 English are related to grammar and lexis (see the Appendix) and were selected from the internet and English use in the academic context. While the lexical items mark linguistic creativity, the grammar items deviated from corresponding norms of SE. Only a small corpus of L2 data was used deliberately to enable in-depth discussion among the participants and to satisfy the constraints of time for the classroom activity. Each of the discussions lasted 16 to 17 minutes and was recorded and transcribed verbatim which produced a text of about 7000 words.

Findings

Analysis of Teachers’ Individual Responses

Table 1 presents the teachers’ individual judgments on the intelligibility (Do you understand the meaning?) and error status (Would you call it an error?) of the seven items. As can be seen from the table, all 17 teachers found four of the seven items (i3, i4, i5 and i6) intelligible. Only one teacher was Not Sure about the meaning of i1 and i7 each while for i2 three teachers found it unintelligible, and another three were Not Sure.

Compared to the intelligibility judgment, the error judgment produced notably different responses. First, teacher judgments were mixed, as not a single item was judged as an error or a varietal feature by all participants. Based on their responses, the seven items can be divided into three groups. In one group (grammaticality) we can include i4, i5 and i6 which were judged errors by over 55% of the judges. Two of these items (i4 and i5) are related to countability and the third one (i6) refers to phrasal verb particle (preposition). In the second group we can include i3 and i7 both of which had an error rating of below 30%. The variety ratings for both items were higher, 35% for i3 and 47% for i7. Like i5 and i6, i3 refers to countability (a research) while i7 refers to subject (pronoun) drop. This group is also related to grammaticality, although the items received somewhat different judgments from the teachers. The third group includes i1 and i2 which are related to lexis. Both items received very low error rating (about 12%) and high variety rating (53%), with a considerable proportion of teachers (about 35%) being Not Sure.

<table>
<thead>
<tr>
<th>Teachers</th>
<th>Responses (intelligibility)</th>
<th>Responses (error)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>i1</td>
<td>i2</td>
</tr>
<tr>
<td>T1</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>T2</td>
<td>√</td>
<td>✗</td>
</tr>
<tr>
<td>T3</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>T4</td>
<td>√</td>
<td>✗</td>
</tr>
<tr>
<td>T5</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>T6</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>T7</td>
<td>√</td>
<td>?</td>
</tr>
<tr>
<td>T8</td>
<td>√</td>
<td>√</td>
</tr>
</tbody>
</table>
Teachers’ individual responses to the two tasks allow for making some observations. First, lexical innovations tend not to be judged as errors and therefore appear to be more acceptable as varietal features. On the other hand, L2 items involving grammaticality tend to be judged more as errors than as varieties. However, specific grammatical items may be exceptional. For instance, although i3, i4, and i5 violated the same grammatical norms, i3 (a research) had lower error rating than the other two. We find a clue to this differential judgment in the following excerpt:

**Excerpt 1, Group 1**

PF1: This is the same problem like a research, feedbacks, and equipments. They are all the same.

PF3: But this is rather interesting actually. One is OK for me to say a research but it’s not acceptable for me to say feedbacks. You know, why, it does sound weird.

PF1: I don’t think it is acceptable to say feedbacks, to others…

PF3: I know, but for me, I don’t know. You know what I am saying…,

When PF3’s intuitive judgment, having to do with her own language experience as a NS, could not persuade others as well as herself about the judgment, she referred to a non-linguistic variable (“weird”). She then expressed the need for some criteria that would guide teachers in their judgment which points to the ambiguity involving errors and innovations and its consequences for English teachers.

The data indicate that intelligibility is not strongly related to acceptability and is probably unreliable for distinguishing between errors and varietal features. Interestingly, unintelligibility may not be attributed solely to NNS use. For instance, i2 (“dental holiday”), the only item to have low intelligibility rating, is probably an NS innovation. A Google search of the phrase generates 7 million results, attesting to the extent of its spread in the virtual space. Even providing a detailed context of the phrase (see the Appendix) did not enhance its intelligibility. Compared to this, the Korean English innovation (“handphone”) received higher intelligibility rating from the teachers.

### Analysis of Group Discussions

Our analysis of the group discussion data was guided by the stages of LMT previously discussed as well as Scollon’s (1998) framework of speaker roles in discourse: receptors (reading/noticing L2 items), interpreters (their understandings and explanations) and judges (providing judgments). We present our findings under four criteria that the participants...
Conformity to SE Norms

As the first step of LMT specifies, all three groups of teachers noticed the idiosyncrasies underlying the L2 items. Noticing may have been enhanced by the nature of the task that asked them to judge only a small sample of language behaviours. However, some items received more attention than others. For instance, all three groups devoted relatively less amount of time to the discussion of items 3 to 6, indicating a shared view that these four items were less disputable. Let us consider Excerpt 2:

Excerpt 2, Group 2
PM2 : The difference is that No 3, 4, 5, 6 are kind of grammatical…
PF6 : I think [whether] you accept it or not depends on the context.
PM2 : Yes, that does, but there is still difference, like 1 and 2 are different kinds of problems to 3 to 6.
PF1 : Yeah, that is making nouns.
PM2 : So that is like the vocabularies. But 3 to 6 are kind of, different I don’t know exactly, but they are different.
PF5 : 3, 4, 5, 6 are more grammatical. I think they are structures, what I call structures than content.
PM2 : Because there is a leeway.
PF1 : Because I guess grammar is harder to change than adding or subtracting words from vocabulary.
PM2 : So 1 and 2 are kind of vocabulary which seems to be more malleable.
PM4 : Isn’t it acceptable?
PM2 : Yes, you can change it more easily than you can change the structure.

In the extract, the participants played different but complementary roles, which contributed to the collective evaluation and judgment. Their use of various strategies is noteworthy. For instance, as interpreter, PM2 identified the nature of the problem in a tentative manner (“kind of grammatical”). This tentativeness continued in the next three turns, until PF5 added more certainty (“are more grammatical”), adding another descriptor for the items (“structures”). By now the classification (“grammar/structure” versus “vocabulary/content”), another strategy, was established. So PF1 and PM2 were able to refer to the relative permeability of grammar and vocabulary (“harder to change” versus “malleable”). At this stage, a group decision was implied, but not explicitly articulated. So, PM4, as coordinator asked for confirmation of the decision. PM2 confirmed that grammar items cannot be changed easily (therefore are errors) while vocabulary items were acceptable. So, the group relied on the authority and stability of grammar in deciding on the status of the L2 items.

Group 3 also relied on grammar as a yardstick, although the decision-making process was somewhat different, as can be seen in Excerpt 3:

Excerpt 3, Group 3
PF3 : […] For me, errors [are] related to grammar, but for vocabulary, you can make a sense of this. It’s a variety. But for grammar, such as feedbacks, a keen interest to, I call them errors.
PF2 : What about No. 7?
PF1 : So you use grammar, but…
PF3 : I call it (No. 7) an error, because I think it is related to grammar.
PF1 : But can I just [point to something]… on the radio, on the TV, everywhere, people don’t use adverbs. Adverbs have like disappeared. And that’s grammar. Even here like “She drives really careful”, and a lot…But now it’s just all like for the last nearly five years, I just…you never hear adverb any more, that’s here.
PF3 : Language change[s].
PF1 : Yeah, but I agree with you like most of my choices will depend on this. I’m more open for vocabulary than grammar.
PF3 : That’s for the vocabulary we can understand the meaning of a compound, because based on components, but for grammars, I call them errors. I will tell my students that they are errors.

As judge, PF3 pointed to a general rule: errors are related to grammar. As for vocabulary items, she argued that these are intelligible. From her point of view, grammatical problems were errors while vocabulary issues were varietal features. Knowing that this rule may not apply to all cases, PF2 drew PF3’s attention to i7. Similarly, PF1 pointed to the supposed disappearance of adverbs from spoken English as a counterexample of the infallibility of grammar. To this, PF3 responded by introducing another rule: that languages change. PF1 then accepted PF3’s interpretation and judgment, which was probably accepted by PF2 as well. Thus, the excerpt shows that group discussion has the potential to influence participants’ understanding.

Intelligibility

Neologisms are outcomes of the rapid social, economic and technological developments. In Excerpt 4, Group 3 considered intelligibility to be of utmost importance in judging the acceptability of neologisms.

Excerpt 4, Group 3
PF1 : The first one, handphone driving, understand the meaning? Accept it as English?
PF2 : Not sure, maybe it’s a new phrase.
PF1 : I think…
PF4 : Innovation, we discussed it before.
PF1 : […] like Australians, we say mobilephone, in America, cellphone,
PF4 : And in Korea, they call it handphone.
PF1 : So it already like so many different versions. It’s technology, so it’s not like…
PF3 : So what is the criterion to distinguish between errors and varieties? What is the criterion?
PF1 : For me, here, I said because it is an innovation, like a new technology, if it’s new…
PF2 : I don’t think it’s an error. I just think it’s an innovation, a new phrase, and also we can understand it very well, so I don’t think it’s an error.
PF1 : And also because different versions already exist, so you can’t say: “No, mobilephone is… does it?” like because of mobilephones, cellphones, pocketphones, handphones, …
PF3 : I can understand handphone, but not sure about handphone driving, if there is
no context.

PF5 : It is easy to lead to misunderstand without context.
PF1 : But for me, the handphone driving, I can…(PF4: You can imagine it) I can imagine it.
PF2 : Yeah, I can imagine it. I can guess the meaning of this phrase.
PF1 : I’ve never heard of this before, but like…it makes sense, like in English to put it, make it shorter, like driving while talking…like someone says handphone driving, or mobilephone driving. Oh, that’s OK.
PF2 : I think most people can get this meaning from this phrase,…Let’s move to the second one.

The interaction opens with PF1, as initiator, inviting group members’ comment on the item’s (“handphone”) intelligibility and acceptance. PF2 indicated her inability (“not sure”) adding that this was a “new phrase” to her. PF1 took the turn to give her opinion, but it was taken over by PF4, who confidently judged the item as innovation, referring to a previous discussion, probably outside the activity. PF1 then interpreted the phrase, referring to comparable forms (mobile phones in Australia and cellphones in the US). PF4 aptly noted that “handphone” was another form innovated by Koreans. Although, at this point it was probably agreed that “handphone” was an innovation, the group needed to establish the criteria for the judgment, as can be seen from PF3’s invitation (line 8). PF1 repeated her previous reference to technology-induced innovation while PF2 observed that the new phrase was intelligible. To this PF1 added that because there had already been different versions, there was nothing wrong in adding another version called “handphone”. PF3 and PF5 referred to the importance of context, particularly for “handphone driving”, without which, they argued, the form might be unintelligible. In response, PF1 and PF2 pointed out that they could visualise (“imagine”) the expression “handphone driving”. PF2 provided the group judgment and concluded that the meaning was intelligible to most people. Therefore, intelligibility was the group’s main consideration in judging “handphone” and “handphone driving” as innovations.

Context

The teachers made frequent reference to “context” in considering the acceptability of the items, particularly those related to grammar (items 3 to 7). By “context”, they referred to two situations in particular:

1) Everyday usage versus language teaching practice

The teachers noted that judgments on L2 features would vary depending on whether the context was everyday usage or the pedagogical domain. They drew a distinction between mundane usage and pedagogical sphere, suggesting that the latter should not compromise standards. Let us consider Excerpt 5.

Excerpt 5, Group 3

PF1 : Yeah, I would say in my usage I would accept it [item 6] easily, because it is kind of, like I know there is another, people don’t use adverbs, like…
PF6 : How about your teaching? When you are teaching, and your students make those kinds of errors? Will you correct it?
PF2 : You tell them it’s ok?
PF1 : It depends on, depends on the whole thing. I wouldn’t focus on them, because I
don’t really teach to the exam. Like TOFEL, you know, proficiency exam, I wouldn’t mark someone a low proficiency because they said “a keen interest to send” rather than “a keen interest in sending”. I wouldn’t even think about it.

PF4 : What if this appears in their writing?
PF1 : Yeah, we depend on the whole thing, but…
PF4 : For me, I can easily accept this in oral English, but…
PF2 : But not in our writing, in our grammar. So in our context, in our classroom, we cannot teach students like this, because it will lead them to make mistakes in the exam…
PF1 : But the exams you are teaching to, are they a kind of grammatical exams or proficiency exams?
PF2 : I just teach the junior high school students, for the final, for the entrance examinations for the senior high school, there will be this kind of grammar.
PF1 : And the sentences they have to make or change…I suppose in that situation…maybe I would…But if I am over proficiency…, I wouldn’t like [to] send off alarms: “She is wrong.”….
PF2 : Pardon?
PF1 : … [whether] you call it an error, it depends where you are teaching.
PF2 : Yeah, it depends on the context.

Focusing on i6 (“interest to”), the teachers established a general rule (i.e. context-dependence), suggesting different aspects of context. PF1 opened the conversation as judge, pointing out that she would consider i6 acceptable. However, her judgment did not specify the context. So PF6 and PF1 inquired whether PF2 was referring to the pedagogical domain. PF1 pointed out that she would not worry about the form even in the classroom context unless she was teaching to the exam where grammatical accuracy was a requirement. PF4 drew her attention to the context of writing in particular (line 10). While PF1 probably pondered on this context, PF4 provided her own judgment saying that although she would be happy to accept it in oral communication, she would disapprove of it in writing. But she could not finish her turn which was taken by PF2 who, referring to her own context of preparing students for entrance exam, asserted that she would not encourage students to use this form because they would be penalised. On PF1’s request for clarification, PF2 specified that the exam in question was grammar-based, not proficiency-focused. PF1 then stated that she would not offer a negative judgment on the item in a proficiency-focused context. Both PF1 and PF2 reiterated that judgments would depend on the context.

Based on the extract we can make several observations. First, instead of treating context as singular, various possibilities including everyday usage versus pedagogy, oral versus writing, exam-oriented versus non-exam-oriented and grammar-focused versus proficiency-focused are suggested. Second, the context as a criterion appears to be a consensual outcome. Although we are not sure whether PF1 was initially unwilling to make any allowance for the context of writing in the pedagogical domain, in the end she also supported the criterion put forward by PF2 and PF4. This suggests that group discussion has the potential to promote certain criteria required for consensual judgments.

2) Spoken English versus written English

In Excerpt 5, spoken versus written communication was indicated, but this specific context was given more explicit attention in Excerpt 6 by Group 1. The teachers indicated that issues of standards and varieties can be treated more leniently in oral than in written communication. Davies (2013) argues that innovations characterise mainly spoken English; written communication in English usually maintains universal standards. The teachers stated
that as long as spoken communication was carried on smoothly and intelligibly, it was not necessary to pay much attention to idiosyncrasies in use.

**Excerpt 6, Group 1**

PF2 : …And I mean, when you look at it, you will understand what the person says, it’s just different, it’s just switched, it’s like a variety of abnormal grammatical order, it’s just switched, so I don’t know…

PF1 : I think it’s just the subject is missed.

PM3 : It is acceptable in spoken English.

PF2 : Yes, not in written English. But it is very much, when you speak, you say, “yes, teaches at school locally.” It’s fine. But when you write, it’s not acceptable.

In Excerpt 6, PF2 noted the problem (“abnormal grammatical order”), explained why it occurred (“just switched”) and provided her evaluation. However, she was unable to provide a judgment. PF1 explained the nature of the problem more specifically (“subject is missed”), but she did not provide a judgment either. PM3 then came up with her unhesitant judgment: “acceptable in spoken English”. PF2 agreed with the judgment and, at the same time, warned that the form will not be acceptable in written communication. The excerpt indicates that collaborative language management work may lead to consensual judgment about the status of idiosyncratic forms.

**Gate-keeping Authorities**

Context in the sense of polity is referred to in Excerpt 7 by Group 1 in which the members resolved error/feature issues by drawing on the authority of textbooks and exams:

**Excerpt 7, Group 1**

PF1 : But then so the other ones, like No 3: “I have done a research regarding textbook evaluation”. I said that it is an error.

PF2 : Yes, that is an error, you don’t say I have done a research…

PM3 : Do you think the first language plays a role in this because I see it acceptable because in my first language we do a research, and something. So when I say I understand what persons are trying to say that he has done one research…

PF1 : Then it’s like… So, OK, if you are teaching in your country, but then whatever, what English do you teach, if you are to go to Vietnam, for example, would you teach… would you say a research?...

PM3 : But I’m not trying to say that we should teach English as an international language, more than a foreign language.

PF1 : Yes, as a foreign language, but then it was like, maybe there are some people who don’t say a research.

PM3 : What if we taught like in a context where Arabic, because of our first language for example, it would be acceptable there but would it be considered as English?

PF4 : […] if you think about the other varieties, if we say that many varieties and many countries used that kind of language […], you can tell your students that in the spoken language in many countries, they use this one, and this is kind of accepted, but [it is unacceptable] if it is only one or two [examples]…

PF1 : I don’t know. There are so many things that come into it, because like I do that, when if a student said “a research” in the classroom, then I will probably say:
“Yes, lots of people say that, but in a textbook or an exam it would be wrong.” And so as soon as they hear textbook or exam, it means that the fact that I’ve said “Yes, lots of people said that” just goes out of the window completely…

In Extract 7, PF1 and PF2 judged the violation of countability (“a research”) as an error. However, PM3 questioned their judgment based on his own understanding of its use in his own context. To him, this was acceptable for two reasons: 1) in his first language research is used as a count noun; and 2) the form is intelligible. However, PF1 challenged PM3’s judgment by highlighting the problems of accepting usages based on specific national or L1 context. PM3 retreated somewhat from his initial position of accepting the form. Instead, he presented this as problematic to the group: if a form is acceptable in one country, can it be accepted as (standard) English? The teachers’ responses to the problem indicate uncertainty, tentativeness, and, to some extent, their struggles. While, in general, they were seen to have been guided by sensibilities towards diversity and plurality of English (Proshina, 2012), they might have been equally pressed by a concern that De Klerk (1999, p. 315) aptly pointed out: “When does a substratal feature assert itself sufficiently to overcome the fear that if deviations are allowed, the rules will be abandoned and chaos will ensue?” For instance, in providing her tentative judgment, PF4 referred to geographic spread and acceptability, as previously discussed. She clarified that if there were many cases of a usage in multiple polities, it is “kind of accepted”. The tentativeness, despite all conditionalities, is to be noted. PF1’s immediate reaction is uncertainty: “I don’t know. There are so many things that come into it.” Underlying her explanation is a sense of the lack of power and agency and the futility of teacher judgment “which goes out of the window completely” in the face of what is prescribed by traditional authorities (textbook or exam) which serve as gate-keepers (Jenkins, 2007).

The discussion in Excerpt 7 points to the complexity of issues of errors and innovations. The teachers’ sharing of their perspectives does not indicate what can be done about the acceptability of local innovations. It also indicates their helplessness compared to the power of conventional authorities. Nevertheless, it can be argued that the collective engagement would have a positive impact on the participants. For instance, although the initial judgment of PF1 and PF2 prevailed at the end, presumably they understood that the judgment was not unquestionable. Similarly, PM3 would probably have a more complex understanding that would guide him in dealing with features of local English. As a whole, relying on traditional authorities however reluctantly appears to be the consensual judgment on the specific language problem. However, an awareness of the complexity and the need for critical engagement with issues of errors and innovations appears to be the more important outcome of the discussion.

Discussion

The growing recognition of WE has made it imperative to draw a line between errors and innovations. Although relying on large corpuses of L2 English data may be an efficient way to address this need, for a number of conceptual and practical reasons, as we have discussed in the paper, this approach appears problematic. Under the circumstances, TESOL teachers have to judge L2 English data by default, as part of their professional practice. In this paper we have investigated teacher responses to features of new Englishes in a TESOL teacher-education class. These usages were seen from two contrastive perspectives—as deviations from SE (errors) or as features of WE varieties (innovations). Although SE cannot be seen as monolithic at the expense of the variability that it permits, our focus was on the
academic context where curriculum and instruction typically emphasise uniformity rather than flexibility in use. Similarly, although the recent views of SLA, L2 errors and L2 learning goals have undergone desirable changes, L2 pedagogy in many parts of the world are still guided by traditional views of SLA and errors.

Our analyses of the teacher data have shown that individually teachers provided mixed judgements on L2 features, rejecting some items as errors while accepting others as varietal features. A clear pattern underlying their judgments referred to the type of features: lexical issues received approval while grammatical deviations were given negative judgment. Intelligibility did not appear to be strongly related to this judgment.

The teachers’ collective judgements emerging from the group task reflected similar patterns: They observed that grammatical features needed to conform to SE, but lexical innovations deserved recognition as varietal features. A similar distinction between grammatical and lexical items characterised teachers’ judgments of L2 items in a Bangladeshi study (Hamid & Baldauf, 2013). The teachers’ acceptance of lexical items as varietal features rather than errors reflects an awareness of the continuous addition of these items to the lexicon of English following developments in science and technology and the spread of English globally. The astuteness of their judgment recognizes the openness of lexical resources of English and the mundaneness of linguistic creativity.

The analysis of the discourse of the groups of teachers indicates that they were not guided by exclusively WE or SLA; their collective stance reflected characteristics of both: the need to conform to SE and, at the same time, the importance of recognising language change and innovations. In other words, they represent a critical middle ground which mediates the parameters of norms, of linguistic creativity and pedagogical intervention. This is understood from Excerpt 8:

**Excerpt 8, Group 1**

PF1 : I think it would be really difficult for the students if we just say: “Yeah, everything is OK” … So, but if we just say: “Yeah, everything is all right,” then what’s the classroom for, what’s the lesson for?

PF1 suggests that an absolute form of “liberal linguistics” (“everything is ok”) may deny the necessity of teacher instruction. While teacher instruction is acknowledged, it is implied that teachers cannot perform “orthodox linguistics” which may deny linguistic creativity and the diversity of English in the context of its globalisation and the localisation.

This assertive position testifies to the relevance of teacher agency in the micro-context (Baldauf, 2006; Liddicoat & Baldauf, 2008). Although some of the teachers discussed their powerlessness in the face of institutional and curricular prescriptions, their practice, as can be surmised from the group discussions, may not be that of passive technicians (Kumaravadivelu, 2003); rather, this reflects the exercise of their agency. In other words, these teachers can be seen as actors with “agentive space” (Hornberger & Johnson, 2007), who demonstrate an awareness of the changing face of English on the one hand and their role as gate-keepers of language norms on the other. Thus, their position corroborates the relative autonomy of the micro-context (Lo Bianco, 2010) which has also been highlighted by classroom-based L2 research (e.g. Canagarajah, 1999; Lin, 1999; Martin, 2005).

However, emphasising the teachers’ agency and their articulation of a specific set of criteria including context, exocentric norms, intelligibility and gate-keeping authorities cannot be seen as undermining the complexity of teacher judgments of new Englishes which are mediated by teacher ideologies and institutional and contextual parameters. Our analysis of the group discussion data showed that in discussing various criteria, the teachers raised three major issues that would influence teacher judgment of features of new Englishes: 1)
whether the new feature is coined by native or non-native speakers; 2) whether the judges are native or non-native speakers of English; and 3) what kind of language attitudes and ideologies teachers/judges have. For instance, teacher ideologies were pointed out in Excerpt 9 in which PF2 thought that teacher judgment depended on how NNSs perceived their own varieties of English and what kind of attitudes they had towards L1 and L2 varieties.

**Excerpt 9, Group 1**

PF2 : ...If they [teachers] think, if they perceive their own varieties to be wrong, they want to conform to speak the native variety of English, they would do it [judge the features as errors]. So it’s also the question, just a perception of how they want to speak. So it doesn’t matter whether the teacher, you know, draws a line, saying that, you know, even when they introduce the non-native variety […] it’s how they perceive, and how they judge their own variety…

These issues point to the individuality of teachers and subjective judgments which result in a lack of consensus among them (e.g. Hamid & Baldauf, 2013). However, from our study we can also observe that teacher group discussion and collective judgment tasks have the potential to establish this consensus. In our analysis we have highlighted several occasions where it was clear that there was a positive impact of the group discussion on individual teachers which led them to a) change their initial understanding and judgment; b) develop a more critical attitude; and c) gain in confidence. Therefore, collaborative discussion on errors and innovations may hold the potential for teacher professional learning and development which can be organised by individual institutions and professional bodies. It will be important to follow this line of inquiry in a more focused way to understand when, how and under what circumstances individual teachers give up their personal judgments and accept collective judgments.

Our focused and in-depth analysis of the teacher data would allow us to draw out implications for WE and TESOL pedagogy. As we have discussed in the paper, the debate on standards and varieties of English with regard to the distinctions between errors and innovations illustrates the concept of micro-level agency (Baldauf, 2006) versus macro level standards and provides justification for seeing TESOL teachers as language policy actors (M. Li, 2010; Menken & Garcia, 2010; Zhao, 2011). Traditionally, national language policy actors have proscribed local innovations in English (e.g. Rubdy, 2001, 2007), either explicitly or implicitly. However, it appears from our work that collectively teachers are in a position to exercise their agency and provide appropriate judgments on the issues of errors and innovations, taking goals of communication, curricular and educational requirements and the sociolinguistic reality of English into consideration. Given these insights, it would be important to investigate how teachers perform language norms and their permissible variations through their pedagogical practice in the classroom.

Based on the teachers’ perspectives on errors and innovations, we can comment on WE research as well. As we have discussed in the paper, scholars have suggested macro-level criteria for drawing a distinction between errors and innovations. While these criteria may provide useful references for teachers, it cannot be taken for granted that these will provide a sufficient basis for teacher judgment. Given that scholars/researchers and teachers work in different contexts under differential circumstances, there can be divergence between their understandings and judgments (Zhu & Hamid, 2013). Therefore, WE researchers need to engage with practicing TESOL teachers and recognise their agency if the WE arguments are to impact education and society through teacher mediation.

Finally, the changing norms and varieties of English in the globalising world should have important input into teacher education programs. The inclusion of WE in teacher education
courses may contribute to teachers’ knowledge base which, in turn, may empower them in judging their students’ language use. At the same time, it would be important for teachers to collectively deal with examples of their students’ language use at the institutional level in the form of professional development workshops.

Endnotes

1. The first author keeps a record of innovative English uses by his postgraduate students from which some of these examples were selected.
2. We indicate gender identity of the participant by adding F (female) and M (male) to P (participant).
3. The participants in the groups were found to play other roles as well such as initiator and coordinator.

References


Englishes, 30(1), 141-150. http://dx.doi.org/10.1111/j.1467-971X.2010.01693.x


debates (pp. 17-25). Hampshire: Palgrave Macmillan.
[http://dx.doi.org/10.1111/j.1467-971X.1985.tb00423.x](http://dx.doi.org/10.1111/j.1467-971X.1985.tb00423.x)
[http://dx.doi.org/10.1111/j.1467-971X.1986.tb00636.x](http://dx.doi.org/10.1111/j.1467-971X.1986.tb00636.x)
[http://dx.doi.org/10.1075/scl.44.10roo](http://dx.doi.org/10.1075/scl.44.10roo)
[http://dx.doi.org/10.1080/07908311003797627](http://dx.doi.org/10.1080/07908311003797627)

**Appendix**

1. Handphone driving is considered to be a serious offence equivalent to reckless or dangerous driving because there are many accidents involved during the course of handphone usage.
2. What is a dental holiday?  
   Due to increased prices in dental care, more and more people are travelling abroad for their dental work. During this time they are not just getting a needed procedure done, they are getting a holiday adventure as well.
3. I have done a research regarding textbook evaluation.
4. Please find my draft chapter [attached]. Really appreciate your feedbacks.
5. The new equipments made in France will be the only items on sale this week.
6. The Prime Minister of Sri Lanka has shown keen interest to send his agricultural scientists to interact with Pakistani scientists.
7. A: Gautam was there with his wife shopping.  
   B: Doesn’t his wife work now somewhere?  
   A: Yes, teaches at a school here locally.