Cite as:

Curry, N. & Clark, T. (2020). Corpus-informed evaluation of examination criteria for MENA. In L. McCallum, and Coombe, C. (eds.), *The assessment of L2 Written English across the MENA Region: A Synthesis of Practice* (pp. 359-392). Palgrave Macmillan.

Spelling Errors in the Preliminary English B1 Exam: Corpus-informed Evaluation of Examination Criteria for MENA Contexts

Niall Curry
Coventry University
Tony Clark
Cambridge Assessment

1. Introduction

This chapter presents the analysis of written learner language in different MENA contexts in order to add to the literature on corpus linguistics for language testing and assessment. To do so, first this chapter presents relevant research on corpus linguistics, testing and assessment, and learner corpus research. Building on these broader linguistic contexts, research on spelling and assessment in Arabic-speaking contexts is then discussed. The novelty in this research pertains to the study of spelling errors in the error-annotated component of the *Cambridge Learner Corpus*, which is a 30-million-word corpus of learner language from the Cambridge English language examinations. This study focuses on comparing spelling errors from 5 subcorpora of the *Cambridge Learner Corpus, which* capture written examinations of Arabic speakers and a larger reference subcorpus of non-Arabic speakers. More specifically, these include subcorpora from Libya and the United Arab Emirates, which are compared with other texts produced by other Arabic speakers, and all other speakers from the B1 preliminary exam subcorpora, in order to test for variation across a variety of Arabic speaking contexts and different first language backgrounds.

In the context of the Preliminary English B1 exam, spelling errors are contrasted in order to illustrate that while similarities do occur in exam takers across and beyond MENA contexts, spelling errors are particularly salient among Arabic speakers. Spelling errors were chosen as they occur frequently in each subcorpus and are under investigated in Arabic speaking users of English (Abu-Rabia & Taha, 2006), at least from a corpus linguistic perspective. To address this gap, two case studies on spelling errors are presented, focussing on misspelled verbs and misspelled words that begin with 'wh/w'. The findings are then considered in terms of the assessment criteria for the Preliminary English B1 exam with the aim to produce and deliver guidance on how assessment criteria could be localised for MENA contexts.

2. Literature Review

2.1 Corpus Linguistics and Language Pedagogy

Corpus linguistics has made and continues to make strong contributions to many areas of language pedagogy. These contributions can be seen as indirect—informing materials used to teach/learn language— and direct—allowing users to directly access and analyse corpus data. This chapter is concerned with a lesser-studied area of the long established indirect applications of corpus linguistics to

language pedagogy *i.e.* language testing and assessment. Looking back to the 'corpus revolution' of the 1980s (Rundell & Stock, 1992), corpus linguistics made significant contributions to lexicography and nowadays it is commonplace for reference material for language learning, such as dictionaries and grammars, to be corpus-informed (Hunston, 2002). In fact, everyday teachers and learners, who may never have heard of a corpus, are using corpus-informed material. Lexicographers, for example, have profited extensively from corpora (Römer, 2011) which, according to Hunston (2002), can help lexicographers decide how to present entries, basing decisions on frequency, collocation and phraseology, variation, lexis/grammar, and authenticity, among other things. Similarly, other reference materials, such as grammars are also often informed by corpora and an extensive study of such grammars can be seen in McEnery and Xiao (2005), where non-corpus-informed grammars are seen to contain biases. The *Longman Grammar of Spoken and Written English* (Biber 2012) or the *Cambridge Grammar of English* (Carter & McCarthy, 2013) are of particular note as, due to corpus intervention, they arguably give more accurate and authentic examples, contextualised examples, and real descriptions, based, not on rationalistic thinking, but empirical research.

Beyond reference material, a corpus approach to syllabus development is also a noteworthy indirect application. Evidently, for corpus-based syllabi to be effective, corpus-based materials are a prerequisite (Mishan, 2005). Corpus-based materials such as the Touchstone series of English language textbooks (McCarthy et al., 2004-2006) are a good example of the indirect use of corpora in informing language teaching materials. They are based on the *Cambridge Reference Corpus* and the *Cambridge Learner Corpus* and 'present vocabulary, grammar and functions students encounter most often in real life' (McEnery & Xiao, 2011, p.367).

A further indirect application of corpus linguistics can be seen in its incorporation into teacher training where it has exhibited extremely results. In such cases, 'corpus evidence [can] address teachers' questions' (Tsui, 2006, p.57) and raise awareness. Researchers like O'Keeffe and Farr (2012) have seen corpora as tools to improve teachers' knowledge, efficacy, and insight and in so doing develop teaching expertise. Knowledge, for example, may be acquired through increased awareness of how corpora work. Others such as Frankenberg-Garcia (2012) highlight how task-based approaches to teaching teachers about corpus linguistics can raise awareness, for example. This improved awareness could help improve teaching efficacy as seen in studies such as Abdelkader et al.'s (2010) study, which highlights how corpora can help improve teaching efficiency in online Arabic learning environments. Finally, insight can be invoked through 'the study of naturally occurring classroom discourse [that can] provide student teachers with a valuable resource' (Amador Moreno et al., 2006 p.99).

Most pertinent here however, is the indirect application of language corpora to language testing. The field of language testing and assessment 'is concerned with the development of valid and reliable assessments that measure language ability through specific tasks for particular purposes' (Barker et al., 2015). The role of corpora in language testing and assessment could be considered a relatively emerging field (Callies & Götz, 2015; Park, 2014) where corpora are now being used 'to make inferences about language ability and capacity for' learners to use a language (Chapelle & Plakans, 2013, p.241). This movement allows us to move away from assessment based on more dated criteria towards a data-driven approach to language assessment (Callies & Götz, 2015). Some corpora, such as *the Cambridge Learner Corpus* (2016), *the Michigan Corpus of Academic Spoken English* (MICASE) (Simpson 2002) or *The Marburg Corpus of Learner English* (MILE) (Kreyer, 2015) have been developed with language testing in mind. This has been a reaction to the movement to incorporate corpora in syllabi and materials development, where

researchers believe that if we use corpora to inform our teaching, we must respond to this in our assessments (O'Keeffe et al., 2007). In language testing, corpora have allowed us to build a collection of exam scripts, which can help inform learner corpus research, as seen in Barker et al.'s (2015) work on language testing and assessment, and learner corpora. Their work is exemplary of how the role of corpora in language assessment has emerged, owing to previously theoretically weak approaches to the definition of the Common European Framework of Reference for Languages (CEFR) bandscale descriptors. Traditionally such descriptors were often based on teachers' perception of language and language skills that constituted a specific level. Corpora offered a more empirical and evidence-informed approach, based on examples of learner production. Other studies in the field have similarly used learner corpus research to develop test material, as seen in Qin et al. (2016) who look at corpora as a means to develop speech testing, or Banerjee et al. (2015) who look at corpora as a means to effectively assess writing and standardise testing. Barker et al. (2015) discuss the role of *the Cambridge Learner Corpus* for informing language testing and assessment, by showing what learners can do at particular levels. They discuss how the corpus can be used to inform test design through demographic contextual metadata, and through task rating with error analysis and corpus annotation. Building on this and the many other studies referred to here, this chapter presents the analysis of spelling errors in written examinations in 5 subcorpora of the *Cambridge Learner Corpus* in order to inform assessment criteria specific to MENA contexts.

2.2 Spelling and Arabic-speaking Contexts

Before the results are explored and discussed in detail, there are several points about the differences between English and varieties of Arabic to be noted, if their implications are to be understood in context. Firstly, all 28 Arabic letters represent consonants—the majority of the phonemes also appear in English, and most English phonemes feature in Arabic too (Allaith & Joshi, 2011). It should also be clarified that the language can be split into spoken Arabic and Standard Arabic; the latter has both spoken and written forms. Error analysis regarding spelling challenges among Arabic-speaking learners of English has indicated that phonological differences between Arabic and English can be used to predict spelling errors (Allaith & Joshi, 2011), and such errors can be categorised as insertion, omission, substitution, or transposition errors. This means that learners might add a letter, omit a letter, change a letter for another, or rearrange the order of letters in words, respectively (Al-Zuoud & Kabilan, 2013). Therefore, the role of the L1 in predicting errors is well established (Randall, 2005) and in their development of the BUiD corpus, Randall and Groom (2009) found that many of the errors produced by Arabic speaking learners of English were categorised according to a range of orthographical and phonological possible causes. As such, the phenomenon of L1 interference for influencing spelling errors is deep-rooted; however, from another perspective, so too is the target language phonological interference, where the sounds of English impact on the accuracy of spelling among learners, and in the case of Alhaisoni, Al-Zuoud, and Gaudel (2015), to an even greater degree than L1 interference.

As the current study focuses on learners from Arabic-speaking contexts, these distinctions should be borne in mind. The research identified and described here offers a valuable insight into common spelling errors for these national groups, but it is not to be assumed that this can necessarily be transferred to other Arabic-speaking countries. Further investigation is required if that is to be conclusively established. This important caveat highlights the complexities of contemporary assessment in the MENA region. As will be described, research is beginning to help practitioners better understand the implications of linguistic variance. If test development is to be adequately informed, robust research

is required to support decision making on aspects such as marking criteria. This research can potentially help assessment bodies to accommodate speakers of different first languages and from regional or national varieties of these first languages, and, in so doing, ensure appropriate test fairness. The Arabic-speaking world (and English education and exam practices within that) provides an excellent case study of the challenges this would entail. Before addressing these issues, the corpora used in this study must be described.

3. Data

3.1 Corpus Data: the Cambridge Learner Corpus

The *Cambridge Learner Corpus* is a 55-million-word corpus of keyed students' written responses to Cambridge English exams. This paper is based on the 30-million word error-annotated component of that corpus. The construction of this corpus has been ongoing for over 20 years and a detailed description of its construction is presented in Nicholls (2003). Here, owing to limitations of space, a brief description of the corpus, its error annotation, and the subcorpora used herein are presented.

This corpus has been jointly constructed by Cambridge University Press and Cambridge Assessment and, to-date, amasses 55 million words of learner English. This learner English is taken from transcriptions of students' examination, which have been produced by over 200,000 students from 173 countries around the world. Alongside the text captured, the corpus also contains metadata, such as students' first language, nationality, exam, CEFR level of exam, CEFR level of student performance, year of taking exam, educational levels, age, years of English study, gender, and whether or not the student passed or failed. The error-coded component of the corpus currently contains 30 million words. It is on this error tagged part of the corpus that this analysis is based.

For error tagging, Hawkins and Buttery (2010), Hawkins and Filipovic (2012), and Nicholls (2003) outline in detail the annotation process and a taxonomy of errors tags in the Cambridge Learner Corpus, which are manually added by trained linguists. The corpus is also part-of-speech tagged by the Robust Accurate Statistical Parser (RASP; Briscoe *et al*., 2006), using the CLAWS2 tagset.

3.2 Subcorpora from the B1 Preliminary Exam

This study used the exam scripts from subcorpora containing the B1 Preliminary exam, formerly known as Cambridge English: Preliminary Exam Test (PET), in which learner performance was at B1 CEFR level. This excludes learners who scored below B1 and above B1 when taking this exam. The B1 Preliminary exam was chosen as the area of focus for a number of reasons. First, in the context of Arabic speakers, B1 Preliminary level texts in the *Cambridge Learner Corpus* account for the largest number of Arabic speaker texts in the corpus. As such, it makes sense to focus specifically on this level. Second, in order to avoid including too many variables in the subcorpora chosen, the decision was made to focus only on the B1 Preliminary level. This makes for a clearer comparison across varieties of Arabic and other first languages. Later in this section, Figs. 1, 2, 3, and 4 show that there are several metadata that account for variations within the subcorpora of the B1 Preliminary exam. Further examinations and levels, for example, would increase the differences in the subcorpora studied, which would afford less clarity for comparison.

Within the *Cambridge Learner Corpus,* specifically, this study used a subcorpus of B1 Preliminary exam by Arabic speakers from Libya (PALB1), B1 Preliminary exam by Arabic speakers from the United Arab Emirates (PAUAEB1), B1 Preliminary exam by other Arabic speakers in the Cambridge Learner Corpus

(PAOB1)[1], B1 Preliminary exam by all Arabic speakers in the Cambridge Learner Corpus (PAB1)[2], and all other B1 Preliminary exam by non-Arabic speakers (PNAB1)[3]. Table 1 shows the size of each subcorpus studied. It was deemed important to also look across regions as a variable, as this can allow for considerations of varietal differences as well as contextualisation of education within geo-political boundaries.

Table 1 Size of subcorpora of B1 Preliminary for Libyan Arabic, UAE Arabic, remaining Arabic speakers, all Arabic speakers and non-Arabic speakers

Corpus	Tokens
PALB1	10251
PAUAEB1	10542
PAOB1	17779
PAB1	38572
PNAB1	1633214

While at first glance some of these subcorpora may appear small, they are extremely specialised owing to their explicit focus on subcorpora containing the B1 Preliminary exam from Arabic speakers. Following Aston (2001), small specialised corpora can boast advantages and can be heavily patterned. As such, they can produce valuable insights into language use in the respective areas of language they represent. Within these specialised subcorpora, there are rich metadata, which allow for a better understanding of the exam context. These metadata are now briefly presented, starting with Fig. 1, which shows the years in which the exam texts studied in each subcorpus were written.

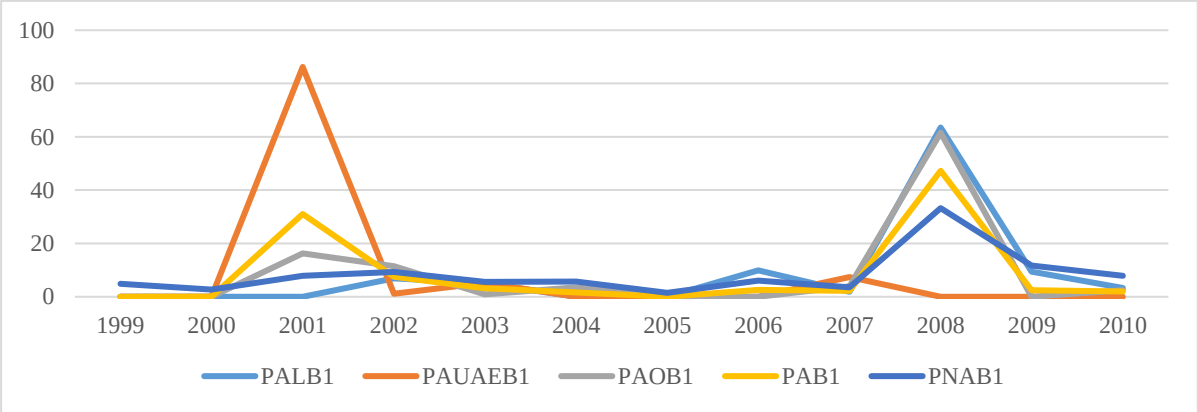

Fig. 1 ***Years of Exam Scripts per Subcorpus in Percentage***

Overall, most subcorpora peak in 2001 and 2008 with a steadier distribution in between and following these years. In terms of the ages of the exam takers in each subcorpus, Fig. 2 shows the range where most exam takers are aged between 10 and 30 in each subcorpus. While age does appear to be an important variable, the main difference is the absence of younger exam takers in PALB1 and PAUAEB1 when compared to the other subcorpora.

[1] See Appendix A and Fig 7 for a list of countries represented in PAOB1.
[2] See Appendix B and Fig 8 for a list of countries represented in PAB1.
[3] See Appendix C and Fig 9 for a list of first languages represented in PNAB1

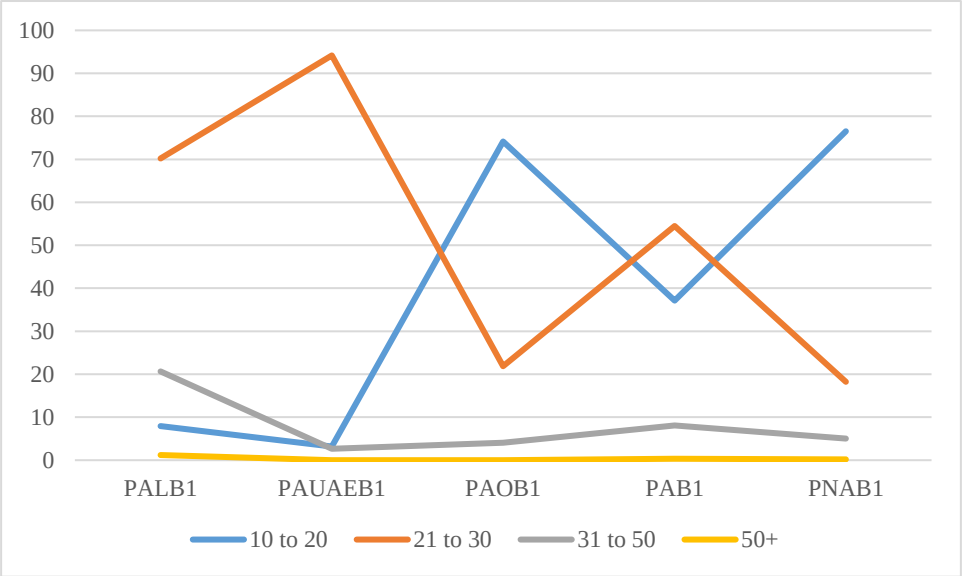

Fig. 2 ***Ages of Exam Takers per Subcorpus in Percentage***

In terms of types of writing, Fig. 3 shows the different types of writing completed across the subcorpora, where the majority of subcorpora here use informative writing. These ensures that the data are reasonably comparable.

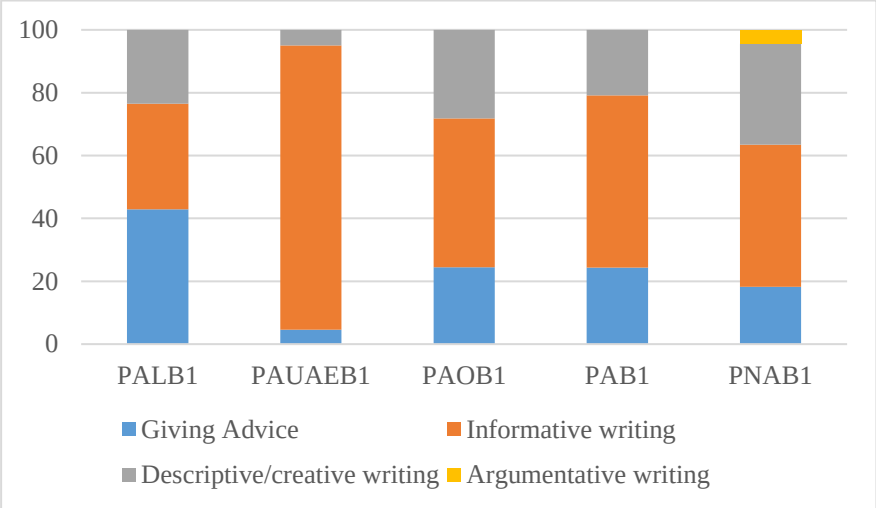

Fig. 3 ***Type of Writing per Subcorpus in Percentage***

In terms of task type, most tasks completed were letter writing type tasks, as can be seen in Fig. 4. Letter writing tasks are identified by the need for students to write a letter, often to a friend, as part of a writing task. This may contain elements of texts that are informative or instructional, for example. However, the key difference is the purpose of the task. Letter tasks serve to demonstrate students' ability to write in such a genre, while informative or instructional texts are tasks that require students to comment on topics or deliver advice and information in essay like formats.

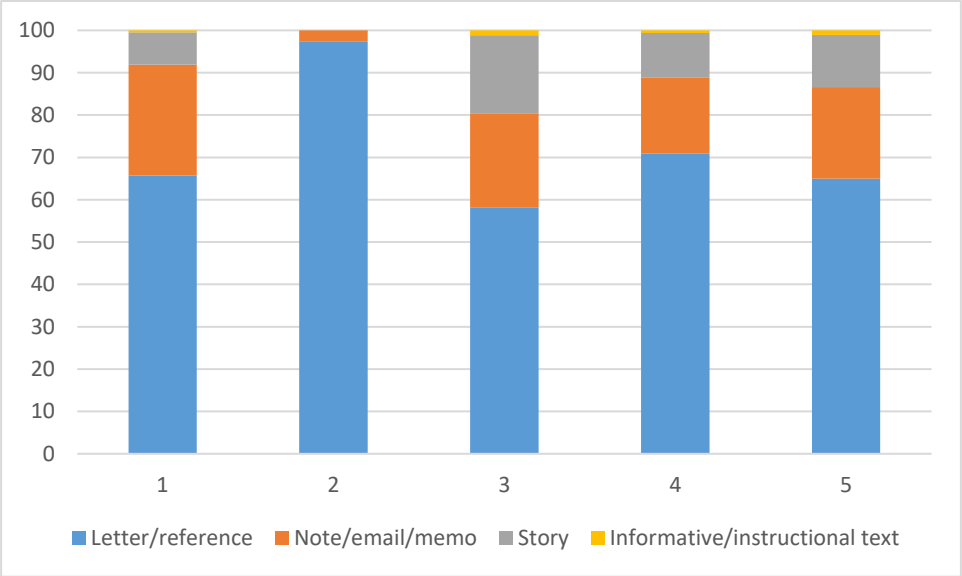

Fig. 4 ***Task Type per Subcorpus in Percentage***

Overall, these metadata serve to further illustrate the context in which the exam texts in each subcorpus were written. Next, the analysis of errors is presented, focusing on the identification of errors, the analysis of spelling errors, and the presentation of two case studies on misspelled words.

4. Analysis, Results, and Discussion

4.1 Corpus Analysis: Errors Identified

Using Sketch Engine, each corpus was searched for all errors by searching '#.*'. What was retrieved was used to identify highly occurring errors, which were then analysed using both quantitative and qualitative approaches. Table 2 shows the top 5 errors in each subcorpus of the *Cambridge Learner Corpus* in words per 10,000 words. What emerges is the shared nature of errors that occur most frequently in each subcorpus. This is especially interesting given the varied metadata presented in Figs. 1, 2, 3, and 4.

Errors Identified									
PALB1		**PAUAEB1**		**PAOB1**		**PAB1**		**PNAB1**	
Error	Freq. per 10k	Error	Freq. per 10k	Error	Freq. per 10k	Error	Freq. per 10k	Error	Freq. per 10k
#RP - replace punctuation	256	#RP - replace punctuation	157	#MP - missing punctuation	160	#RP - replace punctuation	140	#RP - replace punctuation	91
#MP - missing punctuation	106	#S.* - spelling error	135	#S.* - spelling error	104	#S.* - spelling error	113	#TV - incorrect tense of verb	59
#TV - incorrect	91	#TV - incorrect	98	#TV - incorrect	88	#MP - missing	81	#S.* - spelling error	67

tense of verb		tense of verb		tense of verb		punctuation			
#S.* - spelling error	82	#R - replace error	94	#MP - missing punctuation	81	#TV - incorrect tense of verb	65	#MP - missing punctuation	47
#UP - unnecessary punctuation	67	#MP - missing punctuation	93	#RV - replace verb	60	#RV - replace verb	53	#RV - replace verb	47

Table 2 ***Top Five Errors in each Subcorpus***

While frequency is certainly a guiding factor, it is important to consider these data, presented in Table 2, in the wider context. Many of these errors are of note and are worthy of consideration. However, the rationale for the focus on spelling errors in this chapter is based both on their frequency as an error type as well as spelling errors being an area of error research that has remained somewhat under-investigated in studies on Arabic-speaking users of English, from a corpus linguistic perspective. Moreover, while syntax and lexical choice are likely to be influenced by the varied task types presented in Figs. 3 and 4, spelling is less likely to be impacted by the task type. As such, only spelling is considered here, which minimises the impact of differing corpus construction on the findings. For spelling, distribution across the corpora was checked in order to test for significance, and dispersion tests were applied to bolster the statistical analysis and ensure that outliers did not skew the data. This is important as through these tests, we can determine that the corpus did not contain texts that overly skewed the data, and we could determine that spelling errors were normally distributed across all the texts in the data. Subsequently, spelling errors were analysed in terms of word class, and most common errors.

4.2 Spelling Errors: Corpus Findings

Focusing on spelling, Table 3 shows the number of errors in each subcorpus per 10,000 words.

Corpus	Spelling Errors per 10,000 words
PALB1	82
PAUAEB1	135
PAOB1	104
PAB1	113
PNAB1	67

Table 3 ***Spelling Errors per Corpus per 10,000 words***

Comparing the distribution of errors across these corpora, Fig. 5 shows the 95% confidence intervals for the distribution of errors across the 5 subcorpora and Fig. 6 shows a boxplot of distribution. The value of these dispersion tests is discussed next.

Focusing on Fig. 5, spelling errors appear to be distributed similarly in the Libyan and Emirati corpora and there is no significant difference in their dispersion within their texts. This is understandable

given that spelling errors occur in 66.67% of texts in PALB1 and 69.09% of texts in PAUAEB1, and that their dispersion measured at a 0.88 and a 0.9 Julian's D score, respectively, 'which normalizes the coefficient of variation into the range (0,1)' (Coats 2019, p.24). This means that there is an even dispersion of errors across the two subcorpora and that they are not overly skewed by outliers. Focusing on the red line in Fig. 6, both subcorpora also show a similar mean of 2.36 and 2.45 spelling errors respectively per text in every 10,000 words. However, when considering spelling errors in PAOB1, PAB1, and PNAB1, there is a notable difference where the significance of the spelling error drops for each corpus. This is especially interesting given that both PALB1 and PAUAEB1 contain texts written by exam takers who are largely of the age range 21-30, while PAOB1 and PNAB1 contain texts written by exam takers who are largely of the age range 10-20.

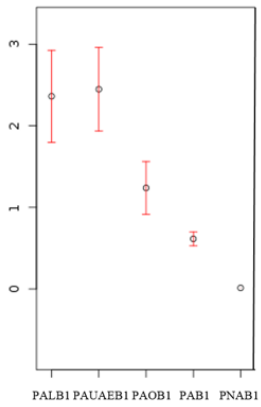

Fig. 5: ***95% confidence interval for spelling subcorpora***

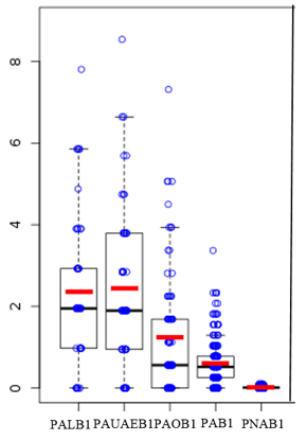

Fig 6: ***Boxplot of spelling error distribution errors across across subcorpora***

In essence, this means that spelling errors in both PALB1 and PAUAEB1 represent a significant number of errors and spelling for these learners is potentially a greater challenge. Spelling errors in PAOB1 appear less significant and only occur in 44.58% of texts, at a mean value of 1.25 errors per text in every 10,000 words and with a normal distribution of 0.87 Julian's D. In PAB1, spelling errors are even less significant with a range of 24.04% of texts, a mean value of 0.61 errors per text in every 10,000 words, and a Julian's D score of 0.93 representing a normal distribution. Most noticeable here is PNAB1 which represents all Preliminary B1 exams completed by learners with languages other than English. PNAB1, comparably, does not see spelling errors as particularly significant despite their normal distribution of 0.99 Julian's D score. Therefore, learners in Libya and the United Arab Emirates appear to encounter greater challenges with spelling at B1 level.

When looking at the specific spelling errors identified, Table 4 shows the top 5 errors per part-of-speech (PoS) per 10,000 words.

Parts of Speech in which Spelling Errors Occur									
PALB1		**PAUAEB1**		**PAOB1**		**PAB1**		**PNAB1**	
PoS	Freq. per 10k	PoS	Freq. per 10k	PoS	Freq. per 10k	PoS	Freq. per 10k	PoS	Freq. per 10k
Nouns	41	Nouns	70	Nouns	42	Nouns	50	Nouns	1.2

Adverbs	16	Adjectives	21	Verbs	52	Adjectives	33	Adjectives	0.8
Adjectives	15	Adverbs	19	Adjectives	19	Verbs	26	Verbs	0.6
Verbs	14	Verbs	17	Adverbs	4	Adverbs	6	Adverbs	0.2
Preposition	11	Preposition	2	Preposition	4	Preposition	5	Preposition	0.1
Total	97	Total	127	Total	121	Total	120	Total	2.9

Table 4 ***Top Five Parts of Speech in which Spelling Errors Occur in each Subcorpus per 10,000 words***

What emerges of note here is the consistency in the top 5 spelling mistakes being some ranking of nouns, adverbs, adjectives, verbs and prepositions, with nouns as the most commonly misspelled words in each subcorpus. Errors ranked from 2 to 4 are either adverbs, adjectives, or verbs, with prepositions at number 5 in each subcorpus. Although other errors did occur, as in most corpus studies, we begin to see steep drop off on data as the list continues. Overall, consistent with the findings discussed in Figs. 5 and 6, PNAB1 contains relatively fewer examples of errors when compared to the 4 Arabic subcorpora. This further supports the view that spelling is a significant challenge for Arabic speakers. Given the infrequency of spelling errors in PNAB1, Table 5 shows the 10 most commonly misspelled words in PALB1, PAUAEB1, PAOB1, and PAB1 only. Note that in Appendix 4, Table 5 contains the top 10 spelling errors in PNAB1 in words per 10,000. Table 6 shows that the same spelling mistakes are much less frequent in all other B1 Preliminary examination taken by non-Arabic speakers when compared to Table 5. This is unsurprising given that Figs. 5 and 6 show significant differences in the presence of spelling mistakes between Arabic and non-Arabic speakers.

Words in which Spelling Errors Occur							
PALB1		**PAUAEB1**		**PAOB1**		**PAB1**	
Word	Freq. per 10k	Word	Freq. per 10k	Word	Freq. per 10k	Word	Freq. per 10k
whether	3	centres	7	centre	2.8	centre	2.6
programme	3	here	5	teacher	2.2	centres	2.3
centre	3	museum	4	because	2.2	because	1.8
because	3	famous	4	tomorrow	1.7	beautiful	1.8
town	2	weather	3	programme	1.7	received	1.6
together	2	restaurant	3	beautiful	1.7	programme	1.6
recommend	2	mountains	3	wearing	1.1	famous	1.6
received	2	beautiful	3	was	1.1	museum	1.3
later	2	Centre	3	then	1.1	here	1.3

Table 5 ***Top 10 Most Commonly Misspelled Words in each Arabic Subcorpus per 10,000 words***

Unlike the grouping of PoS in Table 4, Table 5 shows greater heterogeneity. There are fewer individual words that are frequently misspelled. Among those misspellings that recur are words like centre or centres, programme, whether, weather, museum, because, and received. Taking a more qualitative view, two errors are presented in greater detail. First, spelling errors in verbs are discussed and, second, errors in words beginning with 'wh/w' are considered in more detail.

4.3 Case Study #1: Misspelled Verbs

In terms of verbs, in order to capture all spelling errors in verbs, the following search was queried: [tag="V.*"] within <corr type="#S.*"/>. This shows all verbs, which are part of a correction tag where the error made was related to spelling. Normalised to 10,000 words, Table 6 shows the most commonly misspelled verbs in the 4 Arabic speaker subcorpora.

Spelling Errors in Verbs							
PALB1		**PAUAEB1**		**PAOB1**		**PAB1**	
Word	Freq. per 10k	Word	Freq. per 10k	Word	Freq. per 10k	Word	Freq. per 10k
recommend	2	centres	3	wearing	1	received	2
received	2	recommend	2	was	1	recommend	1
realised	1	received	2	received	1	centres	1
qualified	1	organised	2	postpone	1	wearing	1
prepare	1	climb	2	believe	1	was	1
prefer	1	waiting	1	've	1	postpone	1
jogging	1	varied	1	write	1	organised	1
is	1	riding	1	were	1	introduced	1
introduced	1	read	1	visit	1	climb	1
introduce	1	describe	1	understand	1	believe	1

Table 6 ***Top 10 Most Commonly Misspelled Verbs in each Arabic Subcorpus per 10,000 words***

Initial reviews of the errors appear to show a range of words that are misspelled with no evident relationship between them. Taking a more qualitative view and following Al-Zuoud and Kabilan's (2013) taxonomy for categorising written errors as either insertion, omission, substitution, and transposition, related spelling errors, some interesting trends were identified. While the likes of insertion errors did occur, there was only one example among these top 10 misspelled verbs:

> Well, if you prefere | prefer to go with your friends, I think you should talk to your parents about that. (PALB1)

Here the 'e' is added to prefer. However, more interesting are omission, substitution, and transposition related spelling errors, which will now be briefly discussed. In terms of omission there were a number of examples with verbs like 'prepare', 'was', 'postpone', and 'have' in its contracted form 've'.

> I heard that you should prepar | prepare yourself before the Final exam. (PALB1)

> I'm really sorry to tell you, but it is necessary to pospouon | postpone our meeting for a few hours. (PAOB1)

> Dear Mary, I'v | I've just arrived here on my holiday in India. (PAOB1)

In the case of these errors by omission, 'prepare', 'postpone', and 'have' in its contracted form 've' contain a letter which is not pronounced i.e. 'prepare' in standard British English /prɪˈpeər/[4] and standard American English /prɪˈper/ both end with the sound /r/ where the grapheme 'e' is not pronounced. Similarly, the contraction 'I've' ends in standard British English as /-v/ or /-əv/ and in standard American English as /-v/ /-əv/, again with an 'e' not needing to be pronounced. For postpone, in the case of omission, it is the letter 't' that is problematic, where, among the standard British English pronunciations, /pəʊstˈpəʊn/ or /pəstˈpəʊn/ and, the standard American pronunciation, /poʊstˈpoʊn/ there is always a /t/ which is usually dropped in connected speech (Alameen & Levis, 2015). Therefore, following Allaith and Joshi (2011) and Abu-Rabia and Taha (2006) the pronunciation of these words appears to interfere with their spelling among Arabic speakers. For the other examples of omission, there is spelling confusion where vowels are dropped, likely owing to the use of a schwa /ə/:

> And I have orgnised | organised a lot of interesting things to do for the rest of the holiday. (PAUAEB1)

> In the morning I will vist |visit the centre and museum. (PAOB1)

Or, as in the following examples were two vowels letters are used, often to represent a diphthong:

> I'm wating | waiting for you to write to me. (PAUAEB1)

> Suddenly, she saw a person waring | wearing the same hat that the man she was looking for was waring | wearing. (PAOB1)

> About the weather, you won't belive | believe how the weather is here. (PAOB1)

Therefore, vowels sounds that are reduced or combined seem to pose challenges for Arabic learners' spelling of the words in which they occur.

For substitution, a number of spelling errors occur, where consonant letters are replaced with other consonants, as in the following examples:

> So, why don't you go there with your friends? The food there id | is delicious as you know and the people are very friendly. (PALB1)

> Next week, I have decided to climp | climb some mountains and take some pictures from the top. (PAUAEB1)

> I hope you understant | understand. (PAOB1)

The errors with letters 'd', 'p', and 't' are consistent with Arabic learner errors reported in Allaith and Joshi (2011) who found that 'the absence of some phonemes from Arabic has a negative effect on its speakers' spelling of novel English phonemes'. Further substitution is evident in schwa /ə/ related challenges, where an attempt to replicate graphically the sound of the schwa was made:

[4] Note that the Cambridge Dictionary was used to capture phonetic transcriptions used in the chapter. This dictionary is available at: https://dictionary.cambridge.org/ [last accessed 01 August 2019]

We intreduced | introduced ourselves and then we started our lesson. (PALB1)

Finally, for transposition, a number of diphthong related spelling errors appear to occur in the subcorpora:

They interveiwed | interviewed our head-teacher and some good students. (PALB1)

I relaised | realised that after I was on the train. (PALB1)

Dear friend, I was happy to have recieved | received your letter in which you explained to me about what your parents want. (PAUAEB1)

I 'm really sorry to tell you this, but it is necessary to pospouon | postpone our meeting for a few hours. (PAOB1)

I went running to my friend to warn him but he didn't beleive | believe me and thought I was joking. (PAOB1)

Among these errors, again, there is a clear consistency surrounding errors with vowel sounds, where the vowels are inverted or in a different position to the corrected form. This again supports the relationship between phonological challenges and spelling errors among Arabic speaking learners of English (Al-Zuoud & Kabilan, 2013).

4.4 Case Study #2: Spelling Errors in Words Beginning with 'W/Wh'

Drawing on the most commonly misspelled words in Table 5, the word 'whether' has been misspelled in multiple ways, such as 'wether' and 'wheither'. The following examples illustrate this in practice:

Thanks a lot for your letter. I haven't heard from you for ages. You wanted me to give you advice about wheither | whether you should go with your family or with your friends. (PALB1)

I hope that you will have good time on your holiday wether | whether you go with your friends or with your parents. (PALB1)

PAUAEB1 also shows issues with the word 'weather', which is misspelled like whether with 'wether'

I bought some clothes for the cold wether | weather. PAUAEB1

Interestingly a further 'wh/w' spelling errors occur with 'where' in PAUAEB1, *e.g.*:

I would like to tell you more about the place were | where we are staying.

While of course these words differ, they do share phonological traits, which have been found to be a primary cause for spelling errors among Arabic learners of English (Abu-Rabia & Taha, 2006). As 'whether' and 'weather' are homophones in standard British English /ˈweð.ər/[5] and standard American English

[5] Note that the Cambridge Dictionary was used to capture phonetic transcriptions used in the chapter. This dictionary is available at: https://dictionary.cambridge.org/ [last accessed 01 August 2019]

/weð.ɚ/, it would appear that the 'wh/w' words are problematic for Arabic learners. Arguably, with the 'wh' cluster at the beginning of 'whether' sounding the same as the 'w' at the beginning of 'weather', the errors appear to be linked to the relationship between phoneme and grapheme. The same can be seen with 'were' and 'where' where, although the pronunciations do differ, the initial sound for each word is /w/ followed by vowel sounds such as /eə/, /e/, /ɜː/, /ə/, or /ɚ/. As such, the 'h' in the 'wh' words is dropped. This omission of the 'h' grapheme is also consistent with Al-Zuoud and Kabilan (2013) who found that the majority of spelling errors among Arabic learners were omission errors. It is also of note that through conducting key word analyses of the subcorpora of PALB1, PAUAEB1, PAOB1, and PAB1 with PNAB1, none of the 'wh/w' are key to Arabic speakers. Therefore, Arabic speaking learners of English do not use these words in a noticeably more frequent manner than non-Arabic first language speaking learners and it is not through increased usage that these omission errors were made.

4.5 Summary of Corpus Findings

Overall, this corpus analysis has shown that spelling errors are significant among Arabic first language speakers of English, and especially among those in Libya and the United Arab Emirates. Moreover, in these contexts, it appears that learners aged between 21 and 30 produce many of these errors, which is a finding not shared by the learners represented in the other subcorpora. Therefore, it is difficult to identify whether regional variety of Arabic, teaching and learning cultures, or age and learner profile render spelling for learners in this region particularly problematic. Naturally, this is beyond the scope of a corpus-based analysis. However, this does raise interesting questions that are worthy of academic pursuit.

In terms of the specific findings, a range of open-class words are commonly misspelled *i.e.* nouns, adjectives, adverbs, and verbs. The two case studies of verbs and common words beginning with 'wh/w' take a more qualitative perspective and bolster the view that:

> Arab students, in general, have spelling problems because of the between English and Arabic sound systems (such as, the number and quality of vowels and diphthongs, consonant clusters in word initial, medial and final positions and the Arabic diacritic system is different for the English sound system (Al-Zuoud & Kabilan, 2013, p.173).

Among the errors identified, there is evidence of insertion, omission, substitution, and transposition errors, as defined earlier by Al-Zuoud and Kabilan (2013). These errors are associated with diphthongs, the use of the schwa, and silent letters in words. Of course, further analyses could be undertaken in in any one of these areas to gain a deeper view of phonology-related spelling errors in Arabic first language speakers of English. For now, what has been presented is sufficient to consider the implications of these findings on assessment.

5. Implications for Assessment: Reflections on Key Findings from the Corpus Analysis

In order to understand the significance of these results, it should be recalled that the identified error patterns contribute to a modest but emerging body of existing work on the challenges posed by spelling for Arabic first language speakers who are learning the English language. Owing to limitations of access to corpus metadata in the Cambridge Learner Corpus, further details on tagging, inter-rater agreement, and participant data are not available. This due to the nature of the texts, as international and authentic examinations, as being protected under policies such as GDPR. That being said, in spite of these

limitations, the size and specialised nature of the corpus offers much in way of data that can be of value to the community of researchers who study learner errors, such as spelling.

As noted earlier, phonological errors appear to be the most common source of spelling difficulties for the wider Arabic first language group (Abu-Rabia & Taha, 2006), and certain phonemes in particular may diverge from established English usage (Allaith & Joshi, 2011). Our research adds to the above findings, demonstrating, from a corpus linguistic perspective, the range of spelling errors that occur in the data. It should now be determined what this means from an educational perspective and what possible action could be taken to better support English language testing in Arabic speaking contexts.

5.1 Research-led Investigation and Informing Testing and Assessment in a Global Context

Two potential approaches to impacting assessment will now be discussed. These pertain to altering exam practices to fit local contexts and further targeted test preparation support, based on corpus studies. It is important to note that these approaches should not be considered to be mutually exclusive.

The first of these involves modifying English language examinations in order to better accommodate such groups as L1 Arabic speakers—a notion which should at least be considered in the design of future tests. This becomes particularly apparent when the finding that these learners differ from others in their spelling of certain phonemes is considered in more depth; the question may then be asked whether or not differing from others is necessarily 'incorrect'. In each of the above studies, Arabic students' spelling is defined against some form of established yardstick—be it L1 English usage, their non-Arabic L2 counterparts or other Arabic speakers. In the contemporary educational landscape, this may not be an entirely fair comparison, especially as English is now a dominant language of instruction in a large number of tertiary study institutions in the region (Hopkyns, 2015). This development is not limited to the Arab world, and even those who question its merits can do little to deny its prominence (Bjorkman, 2011). Defining what 'English' *is* becomes problematic however, particularly how to accommodate international variations and decide what or who dictates acceptable academic usage. If Arabic students use alternative spellings for some of the phonemic features listed above, and the meaning of the remains clear, it should not necessarily be classed as an error. In case studies 1 and 2 described above, the spelling errors would not be described as incomprehensible or obfuscating meaning.

The continued need to share the English language and to accommodate international variety is evident, but how to do this is less clear. The term 'international English' may be used to refer to this shared concept, but such an expression leads to further confusion surrounding what international means; establishing *which* English variations are to be included under the term is far from straightforward (Flowerdew, 2012). Consequently, this presents a significant challenge for global assessment, if regional variations of usage are to be accommodated. This does not mean that it should be ignored, and it is advisable that the results of the above studies feed into discussions on how to approach spelling variation more inclusively in future.

5.2 Using Findings to Inform Assessment Practices: Recommendations

In relation to assessment practices, localised assessment variations may provide a viable solution, if practical matters such as cost and defining a manageable scope can be addressed. Future considerations might include:

- retraining examiners of candidates from certain regions not to penalise spelling or phonological errors that have been identified as attributable to L1 interference. This could be trialled on a small cohort

in the first instance, and the implications monitored. Of course, this level of change would need to be directed at larger assessment bodies in order to practically effect change;

- altering existing writing band descriptors to reduce or mitigate any potentially punitive implications of areas that have been categorised as repeatedly problematic for specific learners, through research-led investigation; and
- collating established banks of writing scripts from candidates and analysing them to inform test content or marking practices. The potential benefits of this would not be limited to addressing spelling errors; however, this study demonstrates how consideration of linguistic and cultural contexts in international examinations is important, if an equitable examination experience is sought.

Each of the above notions should be considered, both as new tests are developed, and current models revisited and updated. Of course, the complexities of doing so should not be underestimated. One issue of which the current study reminds us is that the needs of an entire region cannot be addressed using a blanket approach. The challenges that Libyan and Emirati learners face at B1 level may differ from Saudis at B2, for example. However, 30 such differences (and the admittedly somewhat daunting task of accurately documenting them) should not mean that the status quo is preserved, as this would effectively accommodate nobody in particular. Building a sufficient database of learner errors to ensure that the above suggested action can confidently be taken would neither be a quick nor a straightforward process. However, corpora such as the *Cambridge Learner Corpus* and *Trinity Lancaster Corpus* (Gablasova et al., 2019) are noteworthy in this respect.

In the shorter term, it is perhaps more manageable to conceive of other means of helping first language Arabic learners to avoid having their exam scores reduced through these recurring spelling errors (if it is to be conceded that this is an accurate term, in the current climate at least). One approach is to intervene at the test preparation stage, and better equip candidates to meet acceptable spelling standards before taking the exam in question. Doing so may also require further reflection on the possible other contributing causes of the spelling difficulties described, and whether more could be done in response. It has been suggested that English education in the region has a tendency to focus on listening, speaking, writing, reading, development of lexis, and grammar—meaning that spelling receives relatively less attention (Al-Jarf, 2010). This, combined with the difficulties of phonology (Allaith & Joshi, 2011) and first language or 'phonological' (Alhaisoni *et al*., 2015) already outlined, may be exacerbating the issue. It may be more straightforward to address this, than alterations to global assessment policy. If local test preparation courses were able to incorporate sessions that specifically dealt with spelling, and brought to candidates' attention that certain phonemes tend to cause Arab speakers problems in spelling (specifically identifying each one and explaining why it is different in English may help lower the frequency of errors).

6. Conclusions

In conclusion, this chapter has essentially raised a series of pertinent questions, and provided several answers, which serve as a starting point for future research. By adding to existing corpus linguistic work on Arabic speakers and spelling, it has become clear that there does appear to be a pattern for Arabic speakers as a group struggling with spelling. The notion that first language can predict spelling errors (Randall, 2005) is supported in this dataset; Libyan and Emirati students appear to be no exception. Understanding more about what these problems were in particular for these cohorts was the purpose of this research.

To summarise, it is now apparent that a series of open class lexical items are often misspelled, including nouns, adjectives, adverbs, and verbs. Furthermore, it has been documented that insertion, omission, substitution, and transposition were also recurring problems. The implications of this for assessment practices in the region were then discussed. It is evident that further work needs to be done on developing the existing knowledge bank of common spelling errors for the various groups of Arabic speakers, and that if this is done, there could be legitimate scope for adjusting marking approaches to accommodate these patterns of variation.

Promoting fairness for test takers would be at the centre of these suggested discussions, and although the considerable practical challenges of doing so should not be underestimated, testing formats that can be adapted to local contexts are a highly appealing prospect. This may be considered an indirect application of corpus linguistics to assessment practices (particularly in the MENA region, but which should also be looked at elsewhere), two fields which should overlap more than they currently do. Finally, the implications of the study for the short-term have also been noted. The above findings can be used to improve test preparation practices in the countries featured and help test takers prepare for the examination using a pedagogical approach that accommodates their needs.

Acknowledgements

We would like to thank Cambridge University Press and Cambridge Assessment for use of the Cambridge Learner Corpus.

Appendix A: Countries in PAOB1

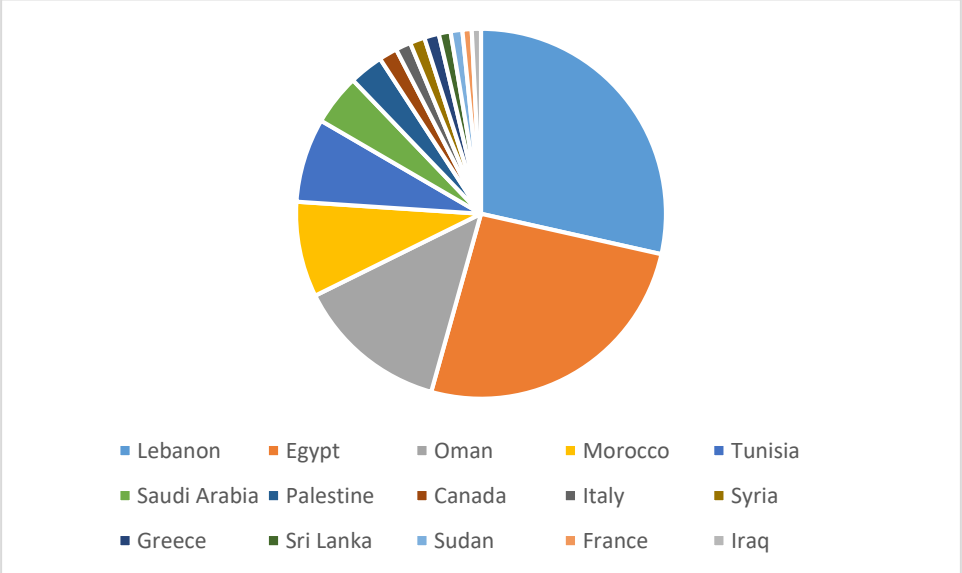

Fig 7 ***Countries Represented in PAOB1***

Appendix B: Countries in PAB1

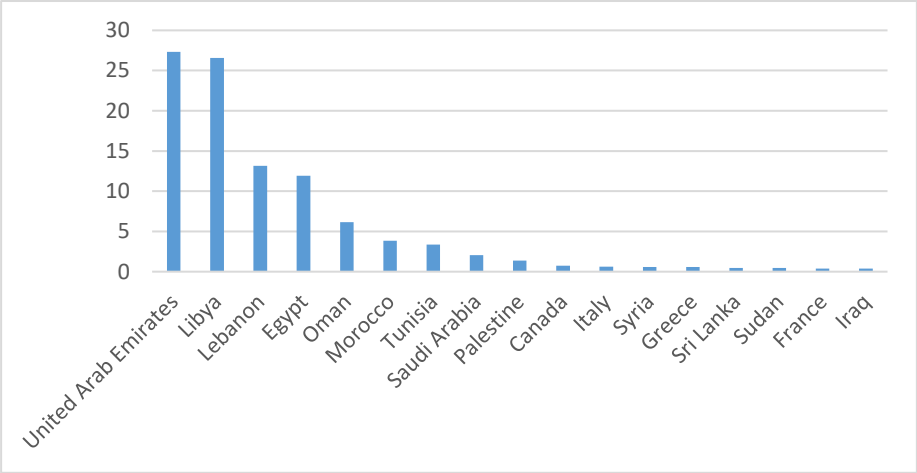

Fig. 8 ***Countries Represented in PAB1***

Appendix C: First Languages

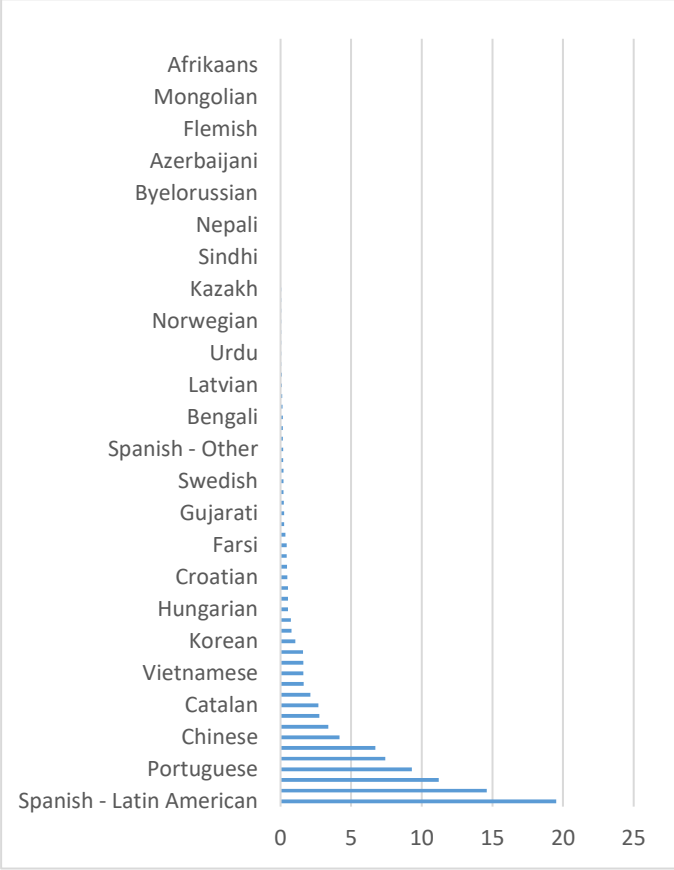

Fig. 9: ***First Languages Represented in PNAB1***

Appendix D: Misspelled words

Top 10 errors in PNAB1	
Word	**Freq. per 10K words**

too	0.88782
centre	0.863328
programme	0.765362
which	0.74087
received	0.716379
beautiful	0.704133
believe	0.600044
decision	0.557184
favourite	0.514323
thought	0.42248

Table 14.7: ***Top 10 Misspelled Words in PNAB1 per 10,000 words***

References

Abdelkader, A., Boumiza, D.S., & Braham, R. (2010). An online Arabic learning environment based on IMS-QTI. In Proceedings—*10th IEEE International Conference on Advanced Learning Technologies* (pp. 116–118). ICALT 2010. .

Abu-Rabia, S., & Taha, H. (2006). Phonological errors predominate in Arabic spelling across grades 1–9. *Journal of Psycholinguistic Research*, 35(2), 167–188.

Alameen, G., & Levis, J.M. (2015). Connected speech. In M. Reed & J.M. Levis (Eds.), *The handbook of English pronunciation* (pp. 159–174) Sussex: Wiley Blackwell.

Alhaisoni, E.M., Al-Zuoud, K.M., & Gaudel, D.R. (2015). Analysis of Spelling Errors of Saudi Beginner Learners of English Enrolled in an Intensive English Language Program. *English Language Teaching*, 8(3), 185–192.

Al-Jarf, R. (2010). Spelling corpora in EFL. *Sino-Us English Teaching, 7* (1), 6–15.

Allaith, Z.A., & Joshi, R.M. (2011). Spelling performance of English consonants among students whose first language is Arabic. *Reading and Writing*, 24(9), 1089–1110.

Al-Zuoud, K.M., & Kabilan, M.K. (2013). Investigating Jordanian EFL |Students' Spelling Errors at Tertiary Level. *International Journal of Linguistics*, 5(3), 164.

Amador Moreno Carolina, P., Chambers, A.,. & Riordan, S.T. (2006). Integrating a corpus of classroom discourse in language teacher education: the case of discourse markers. *ReCALL*, 18 (1), 83–104.

Aston, G. (2001). Learning with corpora: an overview. In G. Aston (Ed.), *Learning with corpora* (pp. 7–45). Bologna: CLUEB.

Banerjee, J., Yan, X., Chapman, M., & Elliott, H. (2015). Keeping up with the times: Revising and refreshing a rating scale. *Assessing Writing*, 26, 5–19.

Barker, F., Salamoura, A., & Saville, N. (2015). Learner corpora and language testing. In S. Granger, G. Gilquin, & F. Meunier (Eds.), *The Cambridge Handbook of Learner Corpus Research* (pp. 511–534).Cambridge: Cambridge University Press.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (2012). *Longman grammar of spoken and written English*. Harlow: Pearson Education.

Bjorkman, B. (2011). English as a Lingua Franca: Implications for EAP. *Iberica*, *22,* 79–100.

Briscoe, E., Carroll, J., & Watson, R. (2006). The second release of the RASP system. In *Proceedings of the coling/acl 2006 interactive presentation sessions* (pp. 77–80). Sydney, Australia.

Callies, M., & Götz, S. (2015). Learner corpora in language testing and assessment: prospects and challenges. In M. Callies & S. Götz, (Eds.), *Learner corpora in language testing and assessment* (pp. 1–12). Amsterdam: John Benjamins.

Carter, R., & McCarthy, M. (2013). *Cambridge grammar of English guide to spoken and written English grammar and usage*. Cambridge: Cambridge University Press.

Chapelle, C.A., & Plakans, L. (2013). Assessment and testing: overview. In C.A. Chapelle (Ed.), *The encylopedia of applied linguistics* (pp. 241–244). New Jersey: Wiley-Blackwell.

Coats, S. (2019). Language choice and gender in a Nordic social media corpus. *Nordic Journal of Linguistics*, 42(1), 31-55. doi:10.1017/S0332586519000039

Flowerdew, J. (2012). Corpora in Language Teaching from the Perspective of English as an International Language. In L. Alsagoff, S.L. Mckay, G. Hu, & W.A. Renandya (Eds.), *Principles and practices for teaching english as an international language* (pp. 226–243). Abingdon: Routledge.

Frankenberg-Garcia, A. (2012). Raising teachers' awareness of corpora. *Language Teaching*, 45(4), 475-489.

Gablasova, D., Brezina, V., & McEnery, T. (2019). The Trinity Lancaster Corpus: Development, description and application. *International Journal of Learner Corpus Research*, 5(2), 126-158.

Hawkins, J. A., & Buttery, P. (2010). Criterial features in learner corpora: Theory and illustrations. *English Profile Journal*, 1(1), 1-23. doi: 10.1017/S2041536210000103

Hawkins, J. A., & Filipovic, L. (2012). *Criterial features in L2 English: Specifying the reference levels of the Common European Framework*. Cambridge: Cambridge University Press.

Hunston, S. (2002). *Corpora in applied linguistics.* Cambridge: Cambridge University Press

Hopkyns, S. (2015). A conflict of desires: english as a global language and its effects on cultural identity in the united arab emirates. In R. Al-Mahrooqi and C. Denman, (Eds,). *English education in the Arab world* (pp. 6-37)*,* Newcastle: Cambridge Scholars Publishing.

Kreyer, R. (2015). The Marburg corpus of Intermediate learner English (MILE). In M. Callies and S. Götz, (Eds,). *Learner corpora in language testing and assessment* (pp. 13-34). Amsterdam: John Benjamins.

McCarthy, M., McCarten, J. & Sandiford, H. (2004-2006). *Touchstone* (four levels). Cambridge: Cambridge University Press.

McEnery, A. & Xiao, R. (2005). Help or help to: What Do Corpora Have to Say? *English Studies*, 86(2), 161-187. doi: http://dx.doi.org/10.1080/0013838042000339880.

McEnery, T. & Xiao, R. (2011). What corpora can offer in language teaching and learning. In E. Hinkel, ed., *Handbook of research in second language teaching and learning* (PP. 364-380). London: Routledge.

Mishan, F. (2005). *Designing authenticity into language learning materials.* Bristol: Intellect.

O'Keeffe, A., McCarthy, M. & Carter, R. (2007). *From corpus to classroom : language use and language teaching*. Cambridge: Cambridge University Press.

O'Keeffe, A. & Farr, F. (2012). Using language corpora in inital teacher education: pedagogic issues and practical applications. In D. Biber and R. Reppen, (Eds,). *Corpus linguistics Vol. 4* (335-362). Los Angeles: Sage.

Nicholls, D. (2003). The Cambridge Learner Corpus: Error coding and analysis for lexicography and ELT. *In Proceedings of the Corpus Linguistics 2003 conference* (pp. 572-581). Corpus Linguistics 16.

Park, K. (2014). Corpora and Language Assessment: The State of the Art. *Language Assessment Quarterly*, 11(1), 27-44, DOI: 10.1080/15434303.2013.872647.

Qin, M., Du, X., Tao, J. & Qiu, X. (2016). A study on the optimal English speech level for Chinese listeners in classrooms. *Applied Acoustics*, 104, 50-56. doi: http://dx.doi.org/10.1016/j.apacoust.2015.10.017.

Randall, M. (2005). Orthographic knowledge and first language reading: evidence from single word dictation. In B. Bassetti and V. Cook, (Eds,). *Second Language Writing Systems* (pp. 122-147). Clevedon: Multilingual Matters.

Randall, M. & Groom, N. (2009). The BUiD Arab learner corpus: a resource for studying the acquisition of L2 English spelling. In *Proceedings of the Corpus Linguistics Conference*, Liverpool, UK.

Römer, U. (2011). Corpus Research Applications in Second Language Teaching. *Annual Review of Applied Linguistics*, 31, 205-225.

Rundell, M., & Stock, P. (1992). The corpus revolution. *English Today*, 8(3), 21-32. doi:10.1017/S0266078400006520

Simpson, R.C., Briggs, S.L., Ovens, J. & Swales., J.M. (2002). *The Michigan Corpus of Academic Spoken English*, Ann Arbor: The Regents of the University of Michigan.

Tsui, A.B.M. (2006). What teachers have always wanted to know – and how corpora can help. In J.M. Sinclair, (Ed,). *How to use corpora in language teaching* (pp. 39-61). Amsterdam: John Benjamins.