The Language Learning Journal

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/rllj20

Text classification and gradation in Arabic textbooks

Salwa Mohamed

To cite this article: Salwa Mohamed (2024) Text classification and gradation in Arabic textbooks, The Language Learning Journal, 52:6, 629-649, DOI: 10.1080/09571736.2023.2213695

To link to this article: https://doi.org/10.1080/09571736.2023.2213695

Published online: 26 May 2023.

THE LANGUAGE LEARNING JOURNAL
2024, VOL. 52, NO. 6, 629–649
https://doi.org/10.1080/09571736.2023.2213695

Routledge
Taylor & Francis Group

Text classification and gradation in Arabic textbooks

Salwa Mohamed

Department of Languages, Information & Communications, Faculty of Arts and Humanities, Manchester Metropolitan University, Manchester, UK

ABSTRACT

Text classification and text gradation are important for language teachers. Profiling and readability studies examine textual and linguistic features that determine text difficulty. Arabic, as an under-resourced language, suffers from a lack of such studies which results in material developers and textbook writers relying on their intuitions and experiences in selecting appropriate texts. At the classroom level, teachers often supplement or adapt the texts they use to suit their students' developmental stages. However, they often do this based on their intuition and experience as well. This study aims to sensitise material developers and particularly teachers to a range of textual, linguistic and contextual features that affect text difficulty so, they can make informed decisions about the suitability of the texts they use and how to adapt and/or supplement them. Although this is most important for teachers of under-resourced languages such as Arabic; the study will be useful for teachers of resource-rich languages as well since research shows that even with systematically classified texts, discrepancies exist between texts and their assumed level.

ARTICLE HISTORY
Received 1 August 2022
Accepted 9 May 2023

KEYWORDS
Text classification; text gradation; text difficulty; text readability; Arabic as a foreign language; CEFR

Introduction

Text classification and gradation are very important in pedagogical contexts to increase the effectiveness of learning and aid students' progression. In first language learning (L1) contexts, texts are graded using statistical formulas that measure texts' difficulty/readability level for target readers (usually school graders). Such 'readability studies' examine the linguistic features of a text to determine its 'ease of comprehension' score (Graesser et al. 2004; Crossley et al. 2007). In L2 teaching, texts may be levelled according to proficiency criteria or frameworks such as the American Council on the Teaching of Foreign Languages (ACTFL) (2012), the Interagency Language Roundtable levels of proficiency (ILR) (2012) and the Common European Framework of Reference for languages (CEFR) (Council of Europe [CoE] 2001). There have, however, been concerns about the consequences and limitations of such frameworks (see Byrnes 2012 for a detailed discussion). However, these frameworks have managed to provide a more detailed and less ambiguous concept of levelling than the proficiency level categories which are frequently used in SLA research, e.g. beginners, intermediate, advanced or higher and lower levels (Hulstijn, Alderson, and Schoonen 2010) and a more principled basis for classifying teaching materials (Sung et al. 2015) and assessing learners' performance (Pilan, Volodina, and Zesch 2016). The CEFR, in particular, has gained wide popularity as a proficiency framework with its well-established six-level proficiency scale, describing

CONTACT Salwa Mohamed s.a.mohamed@mmu.ac.uk

learners' ability (what they can do with the target language and how well) in terms of receptive and productive skills. The CEFR acknowledges the limitations and imprecision of levelling language ability but defends the enterprise because 'for practical purposes it is useful to set up a scale of defined levels to segment the learning process for the purposes of curriculum design, qualifying examinations, etc.' (CoE 2001, 17).

It is intuitively obvious that teachers and educators should select texts that are suitable for their learners' proficiency/developmental level. Textbooks constitute the main source of L2 input texts for many language courses and students. However, little research has addressed the consistency and reliability of these texts' classification into proficiency levels (Barrot 2015; Hong et al. 2020). Research shows that texts display varied linguistic features and different levels of difficulty, even in textbooks and materials that claim to be systematically classified into the same proficiency level (Francois and Fairon 2012; Hong et al. 2020). There are a number of reasons for these discrepancies besides the aforementioned note that the whole process of levelling is approximate. Text difficulty is affected by a hybrid of linguistic as well as cognitive and textual factors beyond the surface lexical and syntactic features. Additionally, text comprehension may depend on factors external to the text such as learners' prior knowledge and skills (Graesser, McNamara, and Kulikowich 2011).

In terms of the CEFR, as the main scale adopted in this study, it is an action-oriented framework which focuses on communicative competencies, strategies and activities/tasks that are evident in productive skills. Receptive skills, on the other hand, suffer from a lack of clarity (Alderson et al. 2004). Furthermore, in their pursuit to be language-independent, the CEFR descriptors that are related to linguistic competence, i.e. vocabulary range, grammatical accuracy and phonological control, are unspecified, which can make it difficult to differentiate between adjacent levels (Martyniuk and Noijons 2007; Zheng, Zhang, and Yan 2016). This requires further specification to identify detailed linguistic features for different CEFR levels in a given language (Hulstijn, Alderson, and Schoonen 2010). Such work is known as profiling studies; these are important in guiding teachers, curriculum designers and material developers to a suitable diet of linguistic features for each proficiency level. Recent L2 readability studies seek to link their automated text linguistic analyses to L2 proficiency scales in order to increase their pedagogical value (see Sung et al. 2015 for link with the CEFR and Saddiki, Bouzoubaa, and Cavalli-Sforza 2015 for link with the ILR).

In the field of TAFL, there is a lack of profiling studies that translate the CEFR reference level descriptors into Arabic-specific linguistic features that correspond to different proficiency levels, and a lack of L2 readability studies that assess text difficulty, using statistical, automated measures. Texts are usually classified and graded based on expert judgement, that is, by textbook authors and/or publishers. This results in wide discrepancies between texts that claim to be at the same difficulty level, e.g. beginner, intermediate, advanced (See Mohamed 2021b).

Background and aims of the study

In the absence of profiling studies in Arabic in reference to the CEFR or any regularising proficiency scales, in a previous study, I collated a list of salient features for the different CEFR proficiency levels. Based on that, I created a CEFR-aligned Arabic curriculum (CEFR-AC) with exemplary linguistic features (vocabulary, grammar, discourse markers, etc.) and general textual characteristics for different levels (Mohamed 2021a); the purpose was to facilitate proficiency-based teaching.

In an endeavour to select/create proficiency-appropriate teaching materials, I analysed four commonly used AFL textbooks to gauge their suitability for implementing the developed CEFR-AC (Mohamed 2021b). Since these textbooks were classified into generic proficiency levels, i.e. beginners, intermediate, etc., each series was divided into CEFR-labelled sections based on the calculation of guided learning hours (GLHs) (Cambridge University 2018) to facilitate a fair and consistent comparison. The following excerpt from Mohamed (2021b, 63) explains how this was calculated (see reference for more details):

> The teaching hours assigned by the authors of each textbook were identified, then an equivalent number of hours was added for guided self-study. The total was then divided by the number of units in the textbook to get the average teaching hours (ATHs) per unit. Finally, the GLHs for each CEFR level were divided by the ATHs per unit to calculate how many units in a textbook series counted for a specific CEFR level.

Then the reading texts in each series were analysed, using an adapted version of the Dutch CEFR Grid (Alderson et al. 2004) to assess their difficulty level and judge their actual CEFR proficiency level accordingly. The Dutch CEFR Grid has a range of textual features that were developed in relation to the CEFR 's specifications and guidelines to make up for the ambiguity of the descriptors for receptive skills and was used by the CEFR Manual (CoE 2009) in calibrating assessment materials, including reading texts, to the CEFR. However, the Grid has its limitations, as will be discussed in the methodology section, hence the use of an adapted version. The results showed that although these textbooks claimed similar generic proficiency levels, their reading content showed substantial discrepancies in interpreting and representing proficiency across levels, as was evidenced in the variation in text difficulty at the same CEFR-labelled sections, especially at the intermediate (B1 and B2) and higher (C1 and C2) levels.

This paper, in turn, follows Mohamed (2021b) and investigates each series independently with the purpose of:

a) ascertaining the extent to which the reading texts are accurately classified and consistently graded within each textbook series. This is assessed by comparing the judged proficiency level against the pre-calculated CEFR level they correspond to within each series.
b) sensitising teachers to the role certain textual, linguistic and contextual features play in determining text difficulty. This is done by qualitatively (i) analysing salient textual and linguistic features, as per the adapted Dutch CEFR Grid, at the text level using sample texts from each series and (ii) reflecting on textual, linguistic and contextual features beyond the text level that affect difficulty in some of these texts.

Rationale and significance

Profiling studies and readability models are important in translating the under-specified CEFR levels into concrete, language-specific linguistic features. The field of TAFL, as mentioned earlier, suffers from a shortage of studies in both areas, which leaves teachers and material writers without much guidance on selecting or developing proficiency-appropriate materials. This study is an attempt to raise the awareness of teachers of Arabic and similar under-resourced languages to the role such features play in determining text difficulty and, by implication, its receptive proficiency level. The value of this goes beyond the contexts of under-resourced languages as readability research shows that even with resource-rich languages, where texts are classified according to proficiency scales such as the CEFR, texts can show a varied range of difficulty.

On the other hand, despite the important contributions profiling studies and readability models can make to improving the effectiveness of texts, there are limitations for this work when texts are utilised at the classroom level. Profiling studies, aiming to specify linguistic features for different proficiency levels, tend to focus on language forms producing lists of vocabulary and grammar individually or contextualised at the sentence level. This does not necessarily capture contextual and extra-textual factors that may as well contribute to defining text difficulty. Examining proficiency at the textual level, rather than at the surface linguistic level, sheds light on these contextual and extra-textual features and makes them more visible. This would encourage/promote a deeper and more critical interpretation of the CEFR descriptors rather than a narrow focus on surface forms. It would also have the added value of applying the target proficiency criteria flexibly in different contexts with different practices.

Similarly, readability models focus on form, quantifying specific linguistic features and providing a comprehension score that arguably determines text difficulty. Texts' qualitative characteristics, however, cannot be judged merely based on quantitative formulas or on separate individual features. As Graesser, McNamara, and Kulikowich (2011, 223) note:

> Computers obviously cannot identify and scale texts on all levels of language, discourse, and meaning. Available computer systems cannot fully comprehend the deep metaphors, literary devices, and historical contexts of Shakespeare's plays, for example. Some characteristics of texts require humans to provide informed, deep, critical analyses.

Besides, at the classroom level, teachers are more concerned with students' comprehension of the text than with the text readability or proficiency level per se (Pilan 2013). Comprehension, as (Graesser, McNamara, and Kulikowich 2011, 223) put it, 'involves much more than an analysis of text characteristics alone because prior knowledge, inference mechanisms, and skills of readers are also critically important'. In Pilan's (2013) study, teachers' evaluation was an important factor in pointing out additional factors for improving readability and comprehensibility of texts/sentences e.g. background context, suitability of and familiarity with the topic (Pilan 2013).

Hence, this study aims to raise teachers' awareness of the role different linguistic, textual and contextual features play in determining text difficulty (and comprehensibility) so that they can conduct an indicative global assessment of available reading texts. This has implications for improving teachers' ability to select proficiency-appropriate materials, comparing different textbooks/texts, adjusting levels of difficulty to facilitate comprehension, creating and customising additional materials to specific students, increasing the efficacy of teaching and helping students progress along the language proficiency spectrum (Barrot 2015; Hong et al. 2020; Saddiki, Bouzoubaa, and Cavalli-Sforza 2015; Uchida and Negishi 2018).

Methodology

The current paper builds on the work in Mohamed (2021b) and assesses the reading texts' difficulty in four popular Arabic textbook series generically classified as (Beginner, Intermediate, Advanced, etc.). The same quantitative method of analysis used in Mohamed (2021b) will be employed to ascertain the extent to which the texts are classified and graded consistently within each series. The analysis will present the distribution of texts' judged CEFR level, using an adapted version of the Dutch CEFR Grid, against their pre-calculated proficiency level based on calculated GLHs (see Mohamed 2021b). Qualitative analyses will, in turn, examine salient textual, linguistic and contextual features that affect text difficulty (a) at the text level and (b) beyond the text level. The purpose of this section is to raise teachers' awareness of the role of these features rather than employing statistical measurements of target linguistic and textual characteristics, which are typical in readability studies (e.g. Sung et al. 2015; Uchida and Negishi 2018; Uri and Abd Aziz 2018).

The following is a list of the textbook series that are analysed in this study (for more details about the series and the reason for choosing them, refer to Mohamed 2021b):

- *Mastering Arabic* (henceforth M.A), book 1-third edition, book 2-first edition (Wightwick and Gaafar 2009, 2015).
- *Arabiyyat Al-naas* (henceforth A.AL) (Younes, Weatherspoon, and Foster 2014 [book 1]; Younes and Al-Masri 2014 [book 2]; Younes and Chami 2014 [book 3]).
- *Al-Kitaab fii taallum al-arabiyya* (henceforth AL-K), second edition (Brustad, Al-Batal, and Al-Tonsi 2004a, 2004b, 2006, 2007).
- *Ahlan wa Sahlan* (henceforth AwS) (Alosh 2006, 2010a, 2010b)

In this section, I will first give an overview of the main tool of analysis (the adapted Dutch CEFR Grid) and the textual and linguistic features included in it, quantitative analysis and qualitative analyses.

The adapted Dutch CEFR grid

The original Dutch CEFR Grid for reading texts contained 9 categories, with categories 1–7 focused on textual features. The only two linguistic categories, 8 and 9, referred in a very generic sense to

vocabulary and grammar respectively; thus, the former ranged from frequent to extended and the latter ranged from simple to complex (see Appendix 1). The problem with the Grid as it is was that it focused more on textual than linguistic features; besides, the linguistic categories were not specified enough to enable accurate assessment of text difficulty (Natova 2019).

To work around these shortcomings, I adapted the Grid by adding two more linguistic categories; that is, category 10 (discourse markers), ranging from simple connecting words to complex and logical markers and category 11 (salient stylistic features), including elements such as modals, determiners, collocations, etc. (see Appendix 2). To further enable more accurate and informed assessment of texts, the categories of the adapted Grid were corroborated with more salient features from the CEFR descriptors, consulting language-specific exemplars from the CEFR-AC (Mohamed 2021b). Additional input from readability studies was added for wider and deeper understanding of the role of the different categories/features in influencing text difficulty.

Textual features

Textual features in the adapted Dutch CEFR Grid include, in order: text source, authenticity, discourse type, domain, topic, nature of content and text length.

Text source impacts the language complexity, so for example everyday signs and menus would have simpler, more accessible language and more salient intention than editorials, literature or technical texts. In the same line, texts that belong to the personal domain are likely to be easier than those which belong to the occupational or educational domain. Discourse type refers to the text genre, which is believed to affect 'structural complexity' (Barrot 2015, 73). From a readability perspective, descriptive and narrative texts are generally thought to be easier than expository and argumentative texts (Graesser, McNamara, and Kulikowich 2011).

Topic and nature of content relate to content progression, thus lower-level learners are expected to cover topics that relate to everyday basic immediate needs with concrete content whereas higher level learners deal with literary, technical and academic topics with more abstract content (Uri and Abd Aziz 2018). In practice, though, texts do not necessarily belong strictly to one domain/type or the other, so in many cases weighing different features is necessary for categorising a text. Alderson et al. (2004) and Zheng, Zhang, and Yan (2016) agree that length of text correlates with text difficulty. This is also confirmed by readability studies (Barrot 2015; Sung et al. 2015).

Authenticity is a very important feature in determining text difficulty and a debatable characteristic in L2 teaching materials, in general (See Crossley and McNamara 2008 and Crossley et al. 2007 for an overview of this debate). Authenticity has multiple dimensions such as authenticity of text, authenticity of task, including purpose, activities, interaction, suitability and relevance of the language to the target students (Guariento and Morley 2001; Beatty 2015). However, In the context of the Dutch CEFR Grid, authenticity refers to the use of genuine texts written for native speakers as opposed to the use of pedagogic or simplified texts. Although exposing students to authentic texts is important for preparing learners for real-life language use, the use of simplified texts provides ideal comprehensible input that aids understanding and long-term language acquisition and matches the linguistic and communicative needs of learners, especially at the beginning and intermediate levels (Crossley et al. 2007; Guariento and Morley 2001; Sung et al. 2015).

Linguistic features

L2 research on proficiency profiles and L2 text readability intersect in their concern with aspects of the language that specifically characterise different proficiency/difficulty levels and that distinguish one level from another. These usually include formal features such as vocabulary, grammar, discourse markers and salient stylistic features. It is important to represent the right features in input texts as they will reinforce students' output (Barrot 2015; Pilan, Volodina, and Zesch 2016).

An important linguistic feature that is agreed to play an important role in determining text difficulty and influencing comprehensibility is vocabulary (Grabe and Stoller 2011; Graesser, McNamara, and Kulikowich 2011; Pilan 2013; Saddiki, Bouzoubaa, and Cavalli-Sforza 2015). Students'

knowledge of vocabulary determines how much time they take to read, process and understand a text (Graesser, McNamara, and Kulikowich 2011). At lower levels, words need to be simple and concrete; repetition of key words can also be helpful as this reduces the time and burden of processing the content and increases understanding (Graesser et al. 2004; Uri and Abd Aziz 2018). In other words, texts aimed at lower levels should be characterised by high-frequency, simple, familiar everyday vocabulary, whereas higher level texts are characterised by less frequent and diverse vocabulary (Barrot 2015; Graesser, McNamara, and Kulikowich 2011; Hong et al. 2020).

Grammar is another important feature that affects text difficulty as it is associated with readers' working memory load when processing sentences. Shorter sentences are syntactically easier as they contain 'fewer words per noun phrase, fewer words before the main verb of the main clause, and fewer logic-based words' (Graesser, McNamara, and Kulikowich 2011, 226). They also maintain that parallel sentences tend to have comparable syntactic structure, which makes processing and understanding the text easier. Conversely, longer sentences are usually complex in structure (Saddiki, Bouzoubaa, and Cavalli-Sforza 2015) and contain multiple clauses (subordinate and/or relative) (Pilan and Volodina 2018). Having more than one modifier or many words before the main verb, they exert a lot of pressure on learners' working memory, delaying or hindering processing the text and understanding the content (Graesser, McNamara, and Kulikowich 2011; Uri and Abd Aziz 2018). Likewise, difficult word order can be an obstacle for learners while trying to understand a text (Pilan 2013). Thus, sentences at lower levels should be characterised by simple and familiar structures and short sentences with one modifier (Graesser, McNamara, and Kulikowich 2011; Uri and Abd Aziz 2018), whereas complex structures and subordination should be left to higher levels (Barrot 2015; Salamoura and Saville 2010).

Discourse markers and salient stylistic features, such as connectives, conjunctions, modals, determiners, intensifiers, collocations, etc., improve the cohesion of text and, consequently, aid comprehension (Hong et al. 2020; Uri and Abd Aziz 2018). In higher level texts, sentences have complex relationships, showing more discontinuity and transition and so are more demanding to process and understand (Graesser, McNamara, and Kulikowich 2011; Hong et al. 2020). Thus, as texts get longer and more complex, more conjunctions and discourse markers are needed to 'clarify and stitch together the actions, goals, events, and states conveyed in the text' and so facilitate comprehension (Graesser, McNamara, and Kulikowich 2011, 227).

Quantitative analysis

This section presents an overview of the distribution of texts' judged CEFR levels against their corresponding pre-calculated proficiency levels in each series and so, gives an indication of how far texts are/are not accurately classified according to their relevant proficiency levels and whether they are graded consistently within each level.

M.A. series has two books with a total of 79 reading texts (see Table 1). Based on GLHs, 33 (42%) texts added up to A1 level (96 h), 33 (42%) to A2 level (124 h) and the remaining 13 (16%) to B1 level (60 h with a shortage of about 140 h). Thus, in theory, this series caters for A1 and A2 levels. Based on text analysis, (53%) of the texts were judged at A1 level (including the overlapping A1/A2 level), (40%) at A2 level (including the overlapping A2/B1 level) and (6%) at B1. In terms of consistency

Table 1. Quantitative distribution of judged against pre-calculated levels in M.A. series.

Texts' pre-calculated levels based on GLHs		Distribution of texts' judged levels based on text analysis										
		A1	A1/A2	A2	A2/B1	B1	B1/B2	B2	B2/C1	C1	C1/C2	C2
A1	33 (42%)	31 (94%)	2 (6%)	–	–	–						
A2	33 (42%)	6 (18%)	3 (9%)	23 (70%)	1 (3%)	–						
B1	13 (16%)			1 (8%)	7 (54%)	5 (38%)						
Total no. of texts	**79**	**37 (47%) 53%**	**5 (6%)**	**24 (30%) 40%**	**8 (10%)**	**5 (6%) 6%**						

of grading, most texts at the pre-calculated A1 section were judged at their corresponding level ([94%] and [100%] when combining the adjacent overlapping level). At the pre-calculated A2 section, a majority of texts (73% combined) were judged at their corresponding level and the remaining (27%) were judged at A1 (18%) and A1/A2 (9%). At the pre-calculated B1 section, only (38%) were judged at B1 and (62%) were judged at A2 (8%) and A2/B1 (54%). The fact that majority of the texts at the pre-calculated B1 section were judged at A2 and A2/B1 levels concurred with the fact that these texts came at the end of book 2, majority of which were calculated as A2, and it seemed logical that they gradually moved up towards B1 level. Accordingly, the texts' judged levels overall matched the proficiency levels expected from them, excluding the B1 level, where there were not enough texts to yield reliable results. The gradation of texts within each level also showed consistency within the two main levels covered (A1 and A2).

A.AL series has three books with a total of 210 reading texts (see Table 2). Based on GLHs, 24 (11%) added up to A1 level (97 h), 22 (10%) to A2 level (118 h), 73 (35%) to B1 level (204 h) and 91 (43%) to B2 level (240 h). Thus, in theory, this series caters mainly for the intermediate levels (B1 and B2), and especially the higher intermediate (B2). Based on text analysis, (15%) were judged at A1, (21%) at A2, (40%) at B1 and (24%) at B2. In other words, although based on judged texts the series catered for the intermediate levels as was pre-calculated, the emphasis was actually more on the lower intermediate than the higher intermediate level. Besides, more texts were judged at A1 and A2 than was calculated (The former was double the latter). A closer look to the distribution of judged levels within each pre-calculated level shows more irregularity. For example, the pre-calculated A2 section has only 27% (this jumps to 63% if we add the adjacent overlapping level) corresponding texts while the rest are judged at either the B1 level (14%) or the A1 level (23%). At B1 and B2 levels, under half the texts were judged at their corresponding levels, and when combined with the adjacent overlapping levels, they showed a small majority of 57% and 52% agreement with their corresponding levels respectively. This shows that texts' judged levels did not always match the proficiency levels expected from them. The gradation of texts within each level was not consistent especially at intermediate levels, which the series was expected to cater for; a considerable number of texts at B1 and B2 lingered at the lower levels, i.e. were easier than would be expected. Besides, in terms of quality, most of the texts at the intermediate levels were fragmented and generally shorter than expected, which justifies the disproportionately higher number of texts compared to the lower levels.

AL-K series has three books plus a booklet for script and sounds with a total of 127 reading texts (see Table 3). Based on GLHs, 18 (14%) texts added up to A1 level (90 h), 21 (17%) to A2 level (120 h), 28 (22%) to B1 level (210 h), 23 (18%) to B2 level (240), and 37 (29%) to C1 level (240 h) (it has to be noted that this is the only series that had enough GLHs to count for C1 level). Thus, in theory, this series caters mostly for the C1 level. Based on text analysis, (10%) were judged at A1, (19%) at A2, (12%) at B1, (29%) at B2, (18%) at C1 and (12%) at C2. If we compare the number of pre-calculated texts and judged texts per whole band, they seem to be close, i.e. band A: pre-calculated (31%), judged (29%); band B: pre-calculated (40%), judged (41%); band C: pre-calculated (29%), judged (30%). However, we find discrepancies within band B, where the judged texts at B1 dropped to 12% compared to the pre-calculated 22%, and at B2 increased from the pre-calculated (18%) to (29%). Also, at band C, almost (40%) of the texts pre-calculated at C1 were judged as C2. A closer look shows that, at the intermediate and higher levels, the B1 section had more texts judged at B2 than the B2 section, and the B2 section has more texts judged at C1 than the C1 section, then almost (40%) of the C1 section was judged at C2. Thus, again in this series, texts' judged levels did not always match the proficiency levels expected from them and the gradation of texts within each level was not consistent either. Contrary to the A.AL series, the reading texts in this series tended to race ahead of the proficiency level expected of them, i.e. texts were more difficult than would be expected.

AwS series has two books plus a booklet for script and sounds with a total of 107 reading texts (see Table 4). Based on GLHs, 21 (20%) texts added up to A1 level (100 h), 25 (23%) to A2 level (125 h),

Table 2. Quantitative distribution of judged against pre-calculated levels in A.AL Series.

Texts' pre-calculated levels based on GLHs		Distribution of texts' judged levels based on text analysis										
		A1	A1/A2	A2	A2/B1	B1	B1/B2	B2	B2/C1	C1	C1/C2	C2
A1	**24 (11%)**	20 (83%)	3 (13%)	1 (4%)	–	–						
A2	**22 (10%)**	1 (5%)	4 (18%)	6 (27%)	8 (36%)	3 (14%)	–	–				
B1	**73 (35%)**	3 (4%)	—	11 (15%)	15 (21%)	34 (47%)	7 (10%)	3 (4%)				
B2	**91 (43%)**	1 (1%)	–	–	2 (2%)	28 (31%)	11 (13%)	44 (48%)	4 (4%)	1 (1%)	–	–
Total no. of texts	**210**	**25 (12%)**	**7 (3%)**	**18 (9%)**	**25 (12%)**	**65 (31%)**	**18 (9%)**	**47 (22%)**	**4 (2%)**	**1 (0.5%)**		
		15%		**21%**		**40%**		**24%**		**0.5%**		

Table 3. Quantitative distribution of judged against pre-calculated levels in AL-K Series.

Texts' pre-calculated levels based on GLHs		Distribution of texts' judged levels based on text analysis										
		A1	A1/A2	A2	A2/B1	B1	B1/B2	B2	B2/C1	C1	C1/C2	C2
A1	18 (14%)	10 (56%)	1 (6%)	2(11%)	4(22%)	1(6%)						
A2	21 (17%)	–	2 (10%)	9 (43%)	4 (19%)	3 (14%)	1 (5%)	1 (5%)	1 (5%)			
B1	28 (22%)	–	–	2 (7%)	2 (7%)	5 (18%)	3 (11%)	9 (32%)	5 (18%)	1 (4%)	1 (4%)	
B2	23 (18%)	–	–	–	1 (4%)	1 (4%)	–	5 (22%)	3 (13%)	5 (22%)	5 (22%)	3 (13%)
C1	37 (29%)						1 (3%)	7 (19%)	6 (16%)	8 (22%)	3 (8%)	12 (32%)
Total no. of texts	**127**	**10 (8%)**	**3 (2%)**	**13 (10%)**	**11 (9%)**	**10 (8%)**	**5 (4%)**	**22 (17%)**	**15 (12%)**	**14 (11%)**	**9 (7%)**	**15 (12%)**
		10%		**19%**		**12%**		**29%**		**18%**		**12%**

Table 4. Quantitative distribution of judged against pre-calculated levels in AwS Series.

Texts' pre-calculated levels based on GLHs		Distribution of texts' judged levels based on text analysis										
		A1	A1/A2	A2	A2/B1	B1	B1/B2	B2	B2/C1	C1	C1/C2	C2
A1	21 (20%)	20 (95%)	1 (5%)	–	–	–						
A2	25 (23%)	10 (40%)	5 (20%)	8 (32%)	1 (4%)	1 (4%)	–	–				
B1	38 (36%)	1 (3%)	8 (21%)	13 (34%)	7 (18%)	9 (24%)	–	–	–	–	–	–
B2	23 (21%)			6 (26%)	3 (13%)	5 (22%)	–	3 (13%)	2 (9%)	1 (4%)	2 (9%)	1 (4%)
Total no. of texts	**107**	**31 (29%)**	**14 (13%)**	**27 (25%)**	**11 (10%)**	**15 (14%)**		**3 (3%)**	**2 (2%)**	**1 (1%)**	**2 (2%)**	**1 (1%)**
		42%		**35%**		**14%**		**5%**		**3%**		**1%**

38 (36%) to B1 level (208 h), and 23 (21%) to B2 level (166 h with a shortage of av. 60 h). Thus, in theory, this series caters more for the intermediate levels, especially the lower intermediate, and least for A1. However, based on text analysis, (42%) of the texts were actually judged at A1, (35%) at A2, (14%) at B1 and only (5%) at B2. Hence, unlike what was pre-calculated, the series actually catered for the lower levels with a majority of (77%) (A1 and A2 combined); texts judged at A1 were more than double what was pre-calculated based on GLHs. On the contrary, the number of texts judged at the B band dropped two-thirds from the number based on GLHs at this level (from [57%] for B1 and B2 combined to [19%]). Thus, this series showed the most discrepancy between the pre-calculated and the judged proficiency levels, indicating lack of consistency in grading texts between and within levels. Similar to A-AL, and contrary to AL-K, as the level went up more texts were judged at lower levels. Texts lingered by two levels down in a considerable number of cases, e.g. at the B1 section, about quarter of the texts (24%) were judged at A1, and at the B2 section, a considerable percentage of (39%) was judged at A2, i.e. were much easier than would be expected.

This discrepancy between different textbook series and lack of consistency in the classification and/or gradation of reading texts within each series seems to result, in a great part, from the lack of recognised standards that specify what linguistic or textual features characterise different proficiency levels and the reliance on intuition and personal experience in developing and selecting material. At the classroom level, most teachers find that they have to adapt and/or supplement these materials to equitably meet their students' needs and target proficiency level. However, as discussed earlier, even with language resources that are aligned with recognised proficiency frameworks such as the CEFR, discrepancies (more or less) do happen. This highlights the need for pedagogical work that, not only specifies the CEFR generic descriptors into language-specific vocabulary and grammar lists per proficiency level, but that also provides exemplary classified texts and guidelines on what textual, linguistic and contextual features contribute to classifying them into their respective levels. The next section will try to do this by qualitatively analysing a sample of the above texts in terms of the textual, linguistic and contextual features that determine their difficulty and, in turn, their difficulty and receptive proficiency level.

Qualitative analysis

This section will provide (a) a micro analysis of textual and linguistic features that affected text difficulty at text level, using a sample of the texts analysed above; (b) a critical macro analysis of textual and contextual factors, beyond the text level, that generally contributed to text difficulty.

Micro qualitative analysis at text level

In this section, I will analyse sample texts from each series and highlight salient textual and linguistic features which contributed to determining text difficulty and so, proficiency level of the text. (See Appendix 3 for an English literal translation of the texts discussed below [Figures 1–8])

M.A. showed the most accuracy in terms of text classification and consistency in grading texts within each pre-calculated proficiency level. Figure 1 (book 1, unit 12, p. 156, pre-calculated section A1, text no. 12) illustrates textual characteristics of level A1, which include simplified descriptive content; a topic of immediate and concrete need in the personal domain; that is food/drinks – likes/dislikes. Linguistically, the text shows a commitment to the use of simple, high-frequency vocabulary and short phrases, employing everyday expressions. Other linguistic salient features include repetition of some key words, use of simple linear connectors such as 'and', 'but' and 'also' and simple grammatical structures (simple nominal sentences and simple present tense). This reduces the burden off students at this lower level and aids processing and inference.

Figure 2 (book 2, unit 3, p. 51, pre-calculated section A2, text no. 13) illustrates textual characteristics of level A2, i.e. although the text falls within the occupational domain, it addresses a familiar predictable topic of interest; that is, job advertisements; it has concrete content with limited

Figure 1. Shows a text from M.A. book 1 that was pre-calculated as A1 and judged as A1.

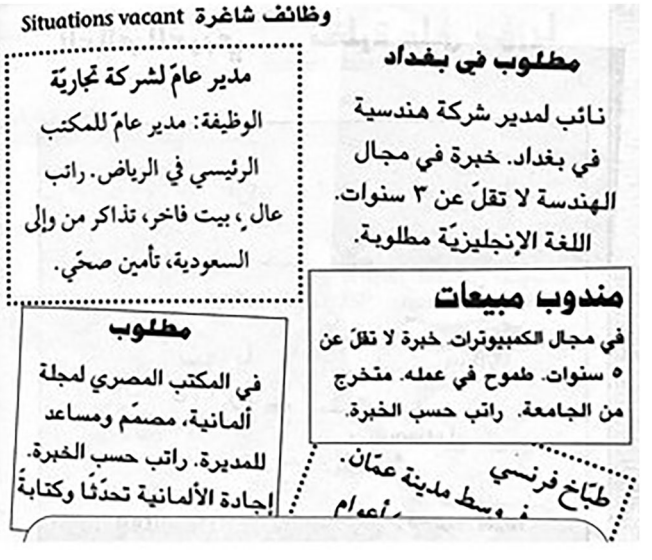
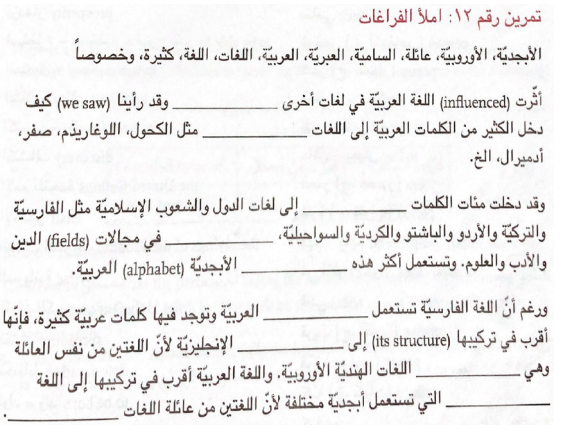


Figure 2. Shows a text from M.A. book 2 that was pre-calculated as A2 and judged as A2.

information. Linguistically, it employs simple and short phrases with descriptive and concrete vocabulary that is familiar and predictable in similar contexts. Although some higher-level connectors and adverbs are used, e.g. 'according to', 'no less than', they are used to link simple short phrases and groups of words in concise advertisement style, and some key words are repeated. Grammatical structure is simple (mainly adjectival phrases).

A.AL showed discrepancies in text classification and inconsistency in the grading of texts, especially at B1 and B2 with under half the texts judged at their corresponding levels and a considerable number of texts lingering at the lower levels. Figure 3 (book 2, unit 3, p.62, pre-calculated section B1, text no. 9) illustrates the textual characteristics of level B1. The text is expository in



تمرين رقم ١٢: املأ الفراغات

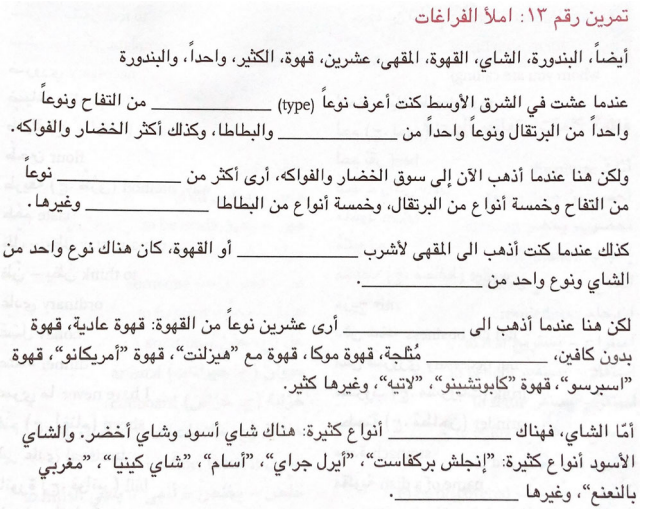
الأبجديّة، الأوروبيّة، عائلة، الساميّة، العبريّة، العربيّة، اللغات، اللغة، كثيرة، وخصوصاً

أثّرت (influenced) اللغة العربيّة في لغات أخرى _____ وقد رأينا (we saw) كيف دخل الكثير من الكلمات العربيّة إلى اللغات _____ مثل الكحول، اللوغاريذم، صفر، أدميرال، الخ.

وقد دخلت مئات الكلمات _____ إلى لغات الدول والشعوب الإسلاميّة مثل الفارسيّة والتركيّة والأردو والباشتو والكرديّة والسواحيليّة، _____ في مجالات (fields) الدين والأدب والعلوم. وتستعمل أكثر هذه _____ الأبجديّة (alphabet) العربيّة.

ورغم أنّ اللغة الفارسيّة تستعمل _____ العربيّة وتوجد فيها كلمات عربيّة كثيرة، فإنّها أقرب في تركيبها (its structure) إلى _____ الإنجليزيّة لأنّ اللغتين من نفس العائلة وهي _____ اللغات الهنديّة الأوروبيّة. واللغة العربيّة أقرب في تركيبها إلى اللغة _____ التي تستعمل أبجديّة مختلفة لأنّ اللغتين من عائلة اللغات _____.

Figure 3. Shows a text from A.AL book 2 that was pre-calculated as B1 and judged as B1.



تمرين رقم ١٣: املأ الفراغات

أيضاً، البندورة، الشاي، القهوة، المقهى، عشرين، قهوة، الكثير، واحداً، والبندورة

عندما عشت في الشرق الأوسط كنت أعرف نوعاً (type) _____ من التفاح ونوعاً واحداً من البرتقال ونوعاً واحداً من _____ والبطاطا، وكذلك أكثر الخضار والفواكه.

ولكن هنا عندما أذهب الآن إلى سوق الخضار والفواكه، أرى أكثر من _____ نوعاً من التفاح وخمسة أنواع من البرتقال، وخمسة أنواع من البطاطا _____ وغيرها.

كذلك عندما كنت أذهب الى المقهى لأشرب _____ أو القهوة، كان هناك نوع واحد من الشاي ونوع واحد من _____.

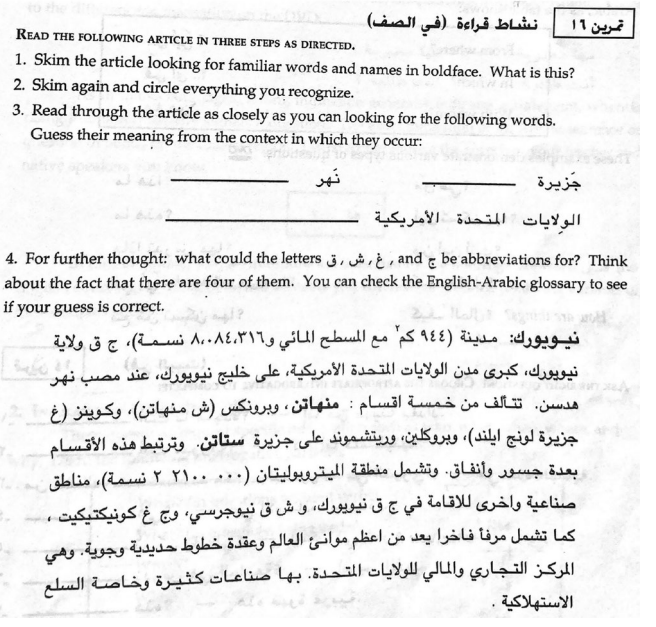
لكن هنا عندما أذهب الى _____ أرى عشرين نوعاً من القهوة: قهوة عادية، قهوة بدون كافين، _____ مُثلجة، قهوة موكا، قهوة مع "هيزلنت"، قهوة "أمريكانو"، قهوة "اسبرسو"، قهوة "كابوتشينو"، "لاتيه"، وغيرها كثير.

أمّا الشاي، فهناك _____ أنواع كثيرة: هناك شاي أسود وشاي أخضر. والشاي الأسود أنواع كثيرة: "إنجلش بركفاست"، "أيرل جراي"، "أسام"، "شاي كينيا"، "مغربي بالنعنع"، وغيرها _____.

Figure 4. Shows a text from A.AL book 2 that was pre-calculated as B1 but judged as A2.

genre and covers a topic in the educational domain; that is, the influence of the Arabic language on other languages. Although it uses mainly concrete language, it has some idiomatic uses (e.g. language family). Linguistically, the text links sentences into connected sequence of points in a non-complex manner and uses some cohesive devices like 'such as' and 'despite' to connect ideas. Although the text is quite short, sentences are increasing in length and complexity than sentences at the A2 level, e.g. 'despite the fact that the Persian language uses the Arabic alphabet and has many Arabic words, it is closer in structure to the English language because the two languages are from the same family, which is the Indo-European languages'. A number of advanced grammatical structures were also employed, e.g. the use of 'قد' which compares to the perfect tense in English, passive verbs, comparative, and relative pronouns.

Figure 4 shows a text from the following lesson to the above one (book 2, unit 4, p. 82, pre-calculated section B1, text no. 13); it is of similar nature to the previous text (i.e. both texts occur at the end of their units for further reading practice and require students to fill in gaps from a list of words). However, this text shows the characteristics of the A2 level and was judged as A2; it covers a topic of



تمرين ١٦ نشاط قراءة (في الصف)

READ THE FOLLOWING ARTICLE IN THREE STEPS AS DIRECTED.

1. Skim the article looking for familiar words and names in boldface. What is this?
2. Skim again and circle everything you recognize.
3. Read through the article as closely as you can looking for the following words. Guess their meaning from the context in which they occur:

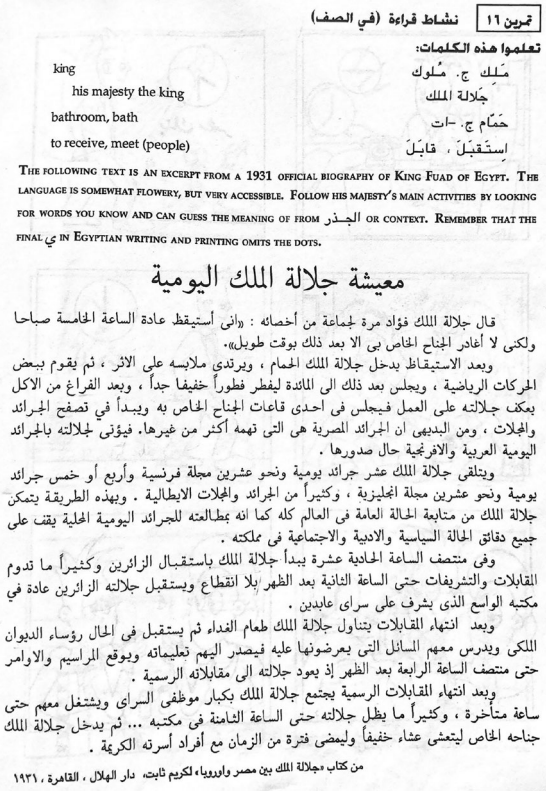
جَزيرة _____ نَهر _____

الولايات المتحدة الأمريكية _____

4. For further thought: what could the letters ق , ش , غ , and ج be abbreviations for? Think about the fact that there are four of them. You can check the English-Arabic glossary to see if your guess is correct.

نيـويورك: مـدينة (٩٤٤ كم٢ مع المسطح المائي و٨،٠٨٤،٣١٦ نسـمـة)، ج ق ولاية نيويورك، كبرى مدن الولايات المتحدة الأمريكية، على خليج نيويورك، عند مصب نهر هدسن. تتـألف من خـمسـة اقسـام : **منهاتن**، وبرونكس (ش منهاتن)، وكوينز (غ جزيرة لونج ايلند)، وبروكلين، وريتشموند على جزيرة **ستاتن**. وترتبط هذه الأقسـام بعدة جسور وأنفـاق. وتشمل منطقة الميتروبوليتان (٠٠٠ ٢١٠٠ ٢ نسمة)، مناطق صناعية واخرى للاقامة في ج ق نيويورك، و ش ق نيوجرسي، وج غ كونيكتيكيت ، كما تشمل مرفأ فاخرا يعد من اعظم موانئ العالم وعقدة خطوط حديدية وجوية. وهي المركـز التجـاري والمالي للولايات المتحـدة. بهـا صناعـات كثـيـرة وخاصـة السلع الاستهلاكية .

Figure 5. Shows a text from AL-K book 1 that was pre-calculated as A1 but judged as B1.



تمرين ١٦ نشاط قراءة (في الصف)

تعلموا هذه الكلمات:

king	مَـلِك ج. مُـلوك
his majesty the king	جَلالة الملك
bathroom, bath	حَمّام ج. -ات
to receive, meet (people)	اِستَـقبَـلَ ، قابَـلَ

THE FOLLOWING TEXT IS AN EXCERPT FROM A 1931 OFFICIAL BIOGRAPHY OF KING FUAD OF EGYPT. THE LANGUAGE IS SOMEWHAT FLOWERY, BUT VERY ACCESSIBLE. FOLLOW HIS MAJESTY'S MAIN ACTIVITIES BY LOOKING FOR WORDS YOU KNOW AND CAN GUESS THE MEANING OF FROM الجـذر OR CONTEXT. REMEMBER THAT THE FINAL ي IN EGYPTIAN WRITING AND PRINTING OMITS THE DOTS.

معيشة جلالة الملك اليومية

قال جلالة الملك فؤاد مرة لجماعة من أخصائه : «انى أستيقظ عادة الساعة الخامسة صباحا ولكنى لا أغادر الجناح الخاص بى الا بعد ذلك بوقت طويل».

وبعد الاستيقاظ يدخل جلالة الملك الحمام ، ويرتدى ملابسه على الاثر ، ثم يقوم ببعض الحركات الرياضية ، ويجلس بعد ذلك الى المائدة ليفطر فطوراً خفيفا جداً ، وبعد الفراغ من الاكل يعكف جلالته على العمل فيجلس فى احدى قاعات الجناح الخاص به ويبدأ في تصفح الجرائد والمجلات ، ومن البديهى ان الجرائد المصرية هى التى تهمه أكثر من غيرها. فيؤتى لجلالته بالجرائد اليومية العربية والافرنجية حال صدورها .

ويتلقى جلالة الملك عشر جرائد يومية ونحو عشرين مجلة فرنسية وأربع أو خمس جرائد يومية ونحو عشرين مجلة انجليزية ، وكثيراً من الجرائد والمجلات الايطالية . وبهذه الطريقة يتمكن جلالة الملك من متابعة الحالة العامة فى العالم كله كما انه بمطالعته للجرائد اليومية المحلية يقف على جميع دقائق الحالة السياسية والادبية والاجتماعية فى مملكته .

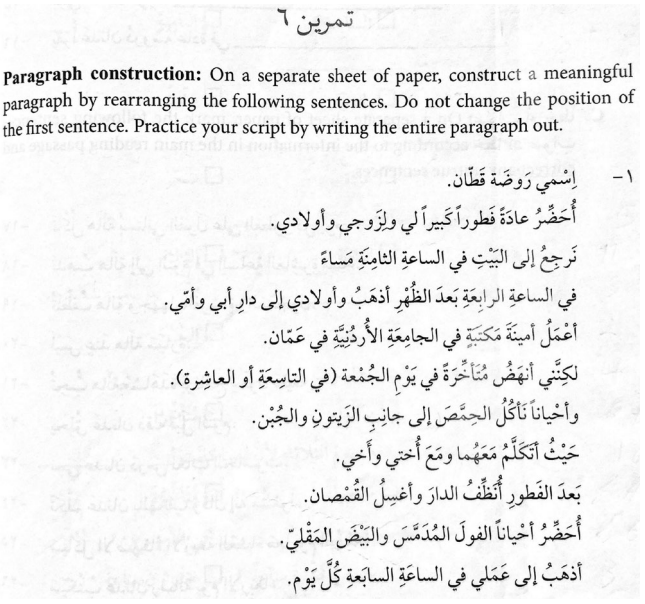
وفى منتصف الساعة الحادية عشرة يبدأ جلالة الملك باستقبال الزائرين وكثيراً ما تدوم المقابلات والتشريفات حتى الساعة الثانية بعد الظهر بلا انقطاع ويستقبل جلالته الزائرين عادة في مكتبه الواسع الذى يشرف على سراى عابدين .

وبعد انتهاء المقابلات يتناول جلالة الملك طعام الغداء ثم يستقبل فى الحال رؤساء الديوان الملكى ويدرس معهم المسائل التى يعرضونها عليه فيصدر اليهم تعليماته ويوقع المراسيم والاوامر حتى منتصف الساعة الرابعة بعد الظهر إذ يعود جلالته الى مقابلاته الرسمية .

وبعد انتهاء المقابلات الرسمية يجتمع جلالة الملك بكبار موظفى السراى ويشتغل معهم حتى ساعة متأخرة ، وكثيراً ما يظل جلالته حتى الساعة الثامنة فى مكتبه ... ثم يدخل جلالة الملك جناحه الخاص ليتعشى عشاء خفيفاً وليمضى فترة من الزمان مع أفراد أسرته الكريمة .

من كتاب «جلالة الملك بين مصر واوروبا» لكريم ثابت، دار الهلال ، القاهرة ، ١٩٣١

Figure 6. Shows a text from AL-K book 1 that was pre-calculated as A2 but judged as B2.

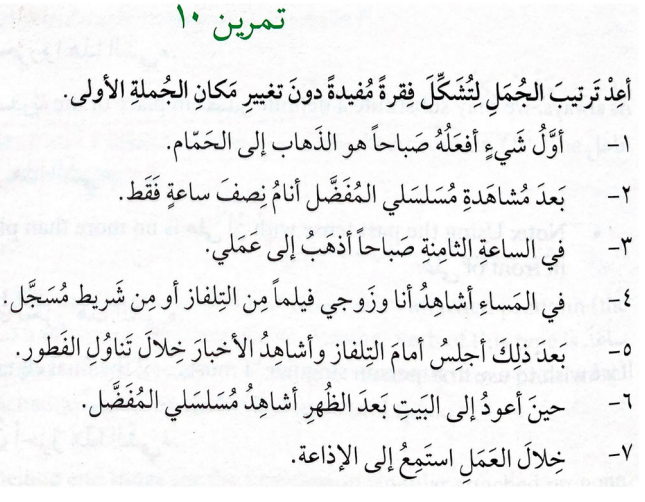


تمرين ٦

Paragraph construction: On a separate sheet of paper, construct a meaningful paragraph by rearranging the following sentences. Do not change the position of the first sentence. Practice your script by writing the entire paragraph out.

١- اِسْمي رَوضَة قَطّان.
أُحَضِّرُ عادَةً فَطوراً كَبيراً لي ولِزَوجي وأولادي.
نَرجِعُ إلى البَيْتِ في الساعةِ الثامِنَةِ مَساءً
في الساعةِ الرابِعَةِ بَعدَ الظُهْرِ أذهَبُ وأولادي إلى دارِ أبي وأمّي.
أعْمَلُ أمينَةَ مَكتَبَةٍ في الجامِعَةِ الأُردُنِيَّةِ في عَمّان.
لكِنَّني أنهَضُ مُتَأخِّرَةً في يَوْمِ الجُمُعة (في التاسِعَةِ أو العاشِرة).
وأحْياناً نَأكُلُ الحِمَّصَ إلى جانِبِ الزَيتونِ والجُبْن.
حَيْثُ أتَكَلَّمُ مَعَهُما ومَعَ أُختي وأَخي.
بَعدَ الفَطورِ أُنَظِّفُ الدارَ وأغسِلُ القُمْصان.
أُحَضِّرُ أحْياناً الفولَ المُدَمَّسَ والبَيْضَ المَقْليّ.
أذهَبُ إلى عَمَلي في الساعَةِ السابَعةِ كُلَّ يَوْم.

Figure 7. Shows a text from AwS book 1 that was pre-calculated as A2 but judged as A1.



تمرين ١٠

أعِدْ تَرتيبَ الجُمَلِ لِتُشَكِّلَ فِقرةً مُفيدةً دونَ تغييرِ مَكانِ الجُملةِ الأولى.

١- أوَّلُ شَيءٍ أفعَلُهُ صَباحاً هو الذَهاب إلى الحَمّام.
٢- بَعدَ مُشاهَدةِ مُسَلسَلي المُفَضَّل أنامُ نِصفَ ساعةٍ فَقَط.
٣- في الساعةِ الثامِنةِ صَباحاً أذهَبُ إلى عَمَلي.
٤- في المَساءِ أشاهِدُ أنا وزَوجي فيلماً مِن التِلفاز أو مِن شَريط مُسَجَّل.
٥- بَعدَ ذلِكَ أجلِسُ أمامَ التِلفاز وأشاهِدُ الأخبارَ خِلالَ تَناوُلِ الفَطور.
٦- حينَ أعودُ إلى البَيتِ بَعدَ الظُهرِ أشاهِدُ مُسَلسَلي المُفَضَّل.
٧- خِلالَ العَمَلِ استَمِعُ إلى الإذاعة.

Figure 8. Shows a text from AwS book 1 that was pre-calculated as B1 but judged as A1.

immediate and concrete need in the personal domain; that is food/drinks; it uses simple, descriptive and concrete language encountered in familiar and predictable everyday situations. Despite the employment of few high-level connectors for A2 such as 'etc.' and 'as for', it mainly uses linear and simple connectors to link simple phrases. Key words are extensively repeated, and it uses a high number of familiar loan words, e.g. 'caffeine', 'espresso', 'cappuccino'. Sentences are short with simple grammatical structures, mainly simple past and simple present tenses.

AI-K also showed discrepancies in text classification and inconsistency in the grading of texts, with reading texts racing ahead of the proficiency level expected of them. Figure 5 (book 1, unit 1, p.16, pre-calculated section A1, text no. 3) actually shows the characteristics of the B1 level and was

judged as B1. It is an authentic text taken from an encyclopaedia; it covers a topic that belongs to the public domain; that is, location of New York city. It is partly expository and partly descriptive. Although the text is short and contains few proper nouns (regions and rivers), the text displays a high range of advanced low-frequency language, idiomatic expressions and even technical vocabulary, e.g. 'gulf', 'subways', 'harbour', 'body of water', 'river mouth', 'a network of air and railway lines', etc. Sentences are long, can be complex and use advanced structures, e.g.

> It comprises the Metropolitan area (2,2100,000 people), industrial areas, and other residential areas in southeast New York, southeast New Jersey, and southwest Connecticut, just as it comprises a luxurious harbour which is considered one of the greatest harbours in the world and a network of air and railway lines. (Brustad, Al-Batal, and Al-Tonsi 2004b, 16)

More advanced cohesive devices that mark the relationships between sentences and ideas are used such as, 'several', 'just as' 'particularly'. It is worth mentioning here that the text does not belong to the main topic of its corresponding lesson (introducing oneself and family) and is aimed at developing reading strategies (more on this point will be discussed in the following section).

Figure 6 (book 1, unit 10, p. 182, pre-calculated section A2, text no. 11) actually shows the characteristics of the B2 level and was judged as B2. The text thematically relates to the personal domain, covers a familiar topic; that is, daily routine, is mainly descriptive/narrative with a mix of concrete and abstract content. However, it is authentic and uses extensive vocabulary, many of which are abstract, low frequency, e.g. 'political affairs', 'his majesty', 'ceremonies'. Additionally, there is an extensive use of classical, archaic, collocational and highly idiomatic language, e.g. 'commit himself to work', 'browsing the newspapers', 'become acquainted with minute details of the political situation', 'meetings and ceremonies', 'signing decrees and orders'. Sentences are long and complex and use advanced structures, e.g. ' … it is self-evident that the Egyptian newspapers are of interest to him more than others, and so the Arabic and French daily newspapers are brought to His Majesty as soon as they are released' and 'This way, His Majesty the king is able to follow the general situation in the whole world, and also by reading the local daily newspapers, he becomes acquainted with the minute details of the political, literary and social situation in his kingdom'. An extensive range of advanced and highly formal discourse markers, cohesive devices, adverbs and organisational patterns are employed to convey finer shades of meaning, e.g. 'in consequence', 'It is self-evident/ obviously', 'as soon as', 'without interruption/ ceaselessly ', 'immediately', 'since', 'a period of time'.

AwS series showed the most discrepancies in text classification and grading, with texts sometimes lagging two levels behind. Figure 7 (book 1, unit 13, p.250, pre-calculated section A2, text no. 11) actually displays the characteristics of level A1. It is a simple narrative text designed for pedagogical purposes. It covers a basic everyday topic in the personal domain, that is daily routine and has concrete content. Linguistically, the text uses simple short sentences and a range of basic, high-frequency vocabulary, e.g. 'I go to my work at seven o'clock every day', 'I work as a librarian in the Jordanian University in Amman'. Sentence structure is simple, mainly present tense, and simple linear connectors such as 'but', 'after', 'where' and adverbs of frequency such as 'usually', 'sometimes' are employed.

Figure 8 (book 1, unit 22, p. 487, pre-calculated section B1, text no. 13) shows the characteristics of level A1. The text is very similar to the text in Figure 7; it is narrative, designed for pedagogical purposes and covers the same topic, daily routine, but even shorter. Linguistically, it has almost the same range of vocabulary and structures as the previous text, e.g. 'at eight o'clock in the morning, I go to my work'; 'in the evening, my husband and I watch a film on TV or a video tape'. Sentence structure is simple, mainly present tense, and simple connectors, e.g. 'after that', 'when', 'during', etc. are used.

Macro qualitative analysis beyond text level

In this section, I will critically reflect on some of the wider factors beyond text level that affected text difficulty in general. I will particularly focus on two series; that is, AL-K and AwS as they showed the

most discrepancies in opposite ways, i.e. texts in AL-K exhibited textual and linguistic features that made them too difficult for the proficiency levels expected from them, whereas texts in AwS exhibited textual and linguistic features that made them too easy for the proficiency levels expected from them.

As for AL-K, the following are some of the factors that increased text difficulty for the proficiency level expected from them:

a) Absence of basic topics and social functions at lower levels. In book 1, some basic and familiar themes typical at lower levels, e.g. health and ailments, booking a room/table/ticket, shopping, bank dealings, just to mention a few, were missing. In a few cases, basic topics were introduced too late in the book. For example, whereas abstract topics such as 'making hard decisions', 'immigration', 'identity issues' were introduced in units 12, 13 and 15 in book 1 respectively, topics such as 'houses', 'body parts', 'personal descriptions' and 'colours' were introduced in units 14 and 18. Usually there were no reading texts featuring these topics; they were presented either as subsidiary oral activities, additional exercises to supplement a broader topic or as trivia in the culture section rather than as independent topics worthy of teaching. This was reflected in the complete absence of frequent basic vocabulary in some cases or the sporadic listing of vocabulary that belonged to the same topic across a number of lessons in other cases. In other words, vocabulary lists sometimes lacked focus or development within a specific thematic area. This creates gaps in students' lexical knowledge which is important 'in building a mental representation of a text and therefore impact[s] its understanding' (Francois and Fairon 2012, 470).
b) Lack of commonality and shared context between some of the reading texts and the wider theme/context of the lesson. In book 1, in particular, the main thematic topic of each lesson focussed on the dramatic development of the main characters' life story through video clips. Although these covered a range of themes related to personal and family life, the reading texts were not necessarily related to the main video/story nor to the same thematic area. An example of that was in lesson 1 where the content of the main video/topic of the lesson was introducing oneself (name, study, nationality, family members, work); however, one of three reading exercises at the end of the lesson focused on the census of the city of New York (see Figure 5). This absence of a wider context and lack of familiarity with the situation/topic and ideas involved in the reading text creates a gap in students' background knowledge which increases the difficulty of texts and hinders comprehension (Uri and Abd Aziz 2018). As the authors of the series explain, the main purpose of these texts is to develop reading strategies and processing skills, e.g. skimming, inferencing and guessing word meanings, recognising their roots and using the context without the need to understand the entirety of the text (Brustad, Al-Batal, and Al-Tonsi 2004b, xiii). However, this comes at the expense of 'meaningful reading practice' (Renandya 2015, 81) and does not necessarily contribute to the student's language skills and proficiency. Besides, at lower levels, students are usually focused on developing their language skills that they might not benefit from any training on reading strategies.
c) Introduction of authentic complex texts from the very early stages. This increased the difficulty of the texts and placed them at a much higher proficiency level than expected (these texts were usually footnoted with citations from the original sources). Although text authenticity has a positive effect on students' learning, the use of too complex demanding texts for learners' current proficiency level can be linguistically, cognitively and culturally overwhelming and demoralising for students, which can negatively affect their progression (Crossley et al. 2007; Renandya 2015). Tomlinson (2012, 162) explains that 'an authentic text is one that is produced in order to communicate rather than to teach'. Simplified texts, on the other hand, provide ideal comprehensible input at lower levels, focusing on the language learners need in order to achieve appropriate linguistic and communicative goals according to their own level (Crossley and McNamara 2008). Gutierrez Bermudez (2014) and Athanasiou et al. (2016) agree that authenticity is not just about

real-life materials, but also real-life situations and the vocabulary needed in such contexts. Athanasiou et al. (2016, 300, citing Dudley-Evans and St. John 1998) add that authenticity also refers to 'the way learners interact with the material to acquire knowledge'. The text in Figure 6 is an example of such an 'authentic' text with unrealistic range of archaic, obsolete vocabulary that are irrelevant to students' needs and real-life situations. As Guariento and Morley (2001) point out, texts do not just provide receptive input; they also stimulate output. Providing learners with too complex texts for the sake of authenticity may come at the expense of task and learner output authenticity. Beatty (2015), Guariento and Morley (2001) and Sung et al. (2015) agree that what actually matters is the suitability of the material and language for students' level of proficiency and learning purposes. This coincides with what teachers in Athanasiou et al. (2016) reported in that authentic texts were produced in a different context for a different purpose and would have to be adapted to suit students' proficiency.

As for AwS series, the following factors contributed to texts being too easy for the proficiency level expected from them:

a) Focus on topics within the personal domain; this was particularly the case with book 1, which covered levels A1, A2 and just over half B1 sections. 16 out of 24 units focused repeatedly on topics such as introductions, family, study, daily routine, food, travel. Even when the main topic of the unit was different, there was revisiting to one of those basic topics, e.g. unit 15 was about festivals; however, there was a text revisiting study and family, unit 20 was about sports; however, there was a text revisiting food. The final unit in the book (unit 24) expansively revisited family, work and travel. Although there were some uses of advanced structures/expressions every now and then, these were limited and were mainly used to introduce certain grammatical forms. Most texts remained confined to the personal domain within familiar everyday contexts and the language and structure were generally simple, pre-packaged and repeated. In book 2, which covered the second half of the B1 section and all of the B2 section (18 units), the content up to unit 6 remained confined to the personal domain. In addition to this, texts were confined to descriptive and narrative genres with mainly concrete content throughout books 1 and 2; only in the last four units of book 2 there were few expository texts and some abstract content in units 13-18. Thus, texts did not progress in complexity and fell short of the proficiency level expected from them.
b) Repeated content and linguistic features under different headings, e.g. different persons, cities/countries, days, etc.; this was more the case with book 1 again. An example of this was in lesson 21 'The geography of the Arab World and the USA'; although the text was about 740 words, the same content/information was repeated under different cities/regions. Similarly, a few texts followed diary style and were sometimes duplicated, i.e. the first half narrated the diary of X, then the second half narrated the diary of Y in almost the same range of vocabulary, content, structure and small details, e.g. lessons 8, 9, 13, 14, 17, 19, 20, 23. Additionally, the texts used a lot of listing, e.g. in lesson 20, we read 'I now eat a lot of vegetables such as beans, peas, lettuce, ladies' fingers, aubergine, courgette, green peppers, etc.', then in the next paragraph 'but there are lots of other fruits such as orange, grapes, bananas, berries, watermelon, mango and pears'. Repetition is a desirable feature in early levels to help learners consolidate new language (Beatty 2015). This is particularly true in the case of Arabic due to its specific orthographic features, among which is the fact that Arabic is a cursive language where letters must be joined (from both sides with majority of the letters and from the right side only for few letters) to form words. Besides, many letters have the same basic form (distinguished only by dots above or below the letter), and they change their base shape depending on their position in the word. Research shows that this slows reading and word processing at early levels even for native speakers, let alone for non-native speaking readers of Arabic (Daniels and Share 2018; Yassin,

Share, and Shalhoub-Awwad 2020). However, *repetition is a characteristic of early levels and should be reduced as the proficiency level goes up*.

c) Lack of authenticity; all texts in book 1 clearly had pedagogic input. As for book 2, it was not always straightforward to figure out text authenticity as, even when authentic texts were used, e.g. known stories, historical expository articles, etc., original sources were not referred to except in few cases. However, some texts were completely pedagogic (units 1-5, 9, 12), some seemed mainly adapted or having a mix of pedagogic and authentic/adapted content (units 6, 7, 10, 15) and some seemed to be mainly authentic (8, 11, 13, 14, 16-18). There was a specific type of reading exercise that recurred at the end of several lessons in book 1 (7 exercises) and book 2 (15 exercises), which required students to re-arrange sentences to form a coherent paragraph. This exercise used particularly pedagogic input and simple language throughout and focused on the same personal/everyday topics as in (a) above. These were regularly judged as A1 and A2 except the last two in units 16 and 17 in book 2 which were judged as B1 and an earlier one in lesson 13 which was judged as A2/B1. This overuse of simplified and repeated input can lead to students losing interest and motivation, which affects the readability of texts (Pilan 2013; Uri and Abd Aziz 2018).

Conclusion

In this paper, I built on my previous work on aligning Arabic teaching with the CEFR to assess the consistency of text classification and gradation of the input reading texts within four widely used Arabic textbook series. The textbooks series were divided into pre-calculated CEFR-labelled sections, then an adapted version of the Dutch CEFR Grid (reinforced with additional input from the CEFR descriptors, CEFR-AC and readability studies) was used to judge the actual proficiency level of texts. Quantitative analysis, comparing the judged level and the pre-calculated level of texts revealed discrepancies between different textbook series and a lack of consistency in the classification and gradation of texts within each series, with the exception of M.A. Qualitative analysis showed a range of textual, linguistic and contextual features that affected text difficulty both at the micro (text level) and the macro (beyond text) levels. Many of the texts exhibited textual and linguistic features that were either too easy or too difficult for their corresponding pre-calculated levels.

In practice, this means that teachers will need to adapt and/or supplement these texts to meet their students' needs and target proficiency level. However, as Tomlinson (2012) says, there is little guidance in the literature on how to do this in a methodical and principled way. Providing pre-service and in-service training for teachers on developing their own materials and adapting existing materials is important for language teachers in general, but more so for teachers of under-resourced languages such as Arabic where existing materials tend to be based on expert judgements and experience rather than on unified, recognised criteria. Part of this training is to be able to evaluate existing material; the current study aimed at facilitating this by raising teachers' and professionals' awareness of the role certain textual, linguistic and contextual features play in determining text difficulty and proficiency level. This is hoped to encourage teachers to critically assess the content of published teaching materials in relation to their contexts and their students' developmental levels and learning needs, and to guide them on what features to focus on when adapting existing materials or developing their own.

Recommendations for further research

It has to be noted that this is an exploratory study that aimed mainly at raising teachers' awareness; more in-depth quantitative and qualitative research is needed to validate the assumptions made in this study. More qualitative studies that focus on smaller numbers of both reading and audio texts in Arabic L2 teaching materials, and analyse them more systematically and conclusively, can further the

approach and assumptions in the current study. Quantitative, computational studies in Arabic L2 text readability are, in turn, needed to statistically validate qualitative analyses and findings. These will be most valuable when combined with regulating proficiency scales such as the CEFR for the classification and advancement of Arabic teaching materials.

Future studies can feed into detailed linguistic profiles (vocabulary and grammar) for Arabic proficiency levels. However, focusing on text rather than on language forms only is important to develop a more comprehensive and contextualised view of proficiency on both the receptive and productive levels.

Acknowledgement

The author would like to thank the reviewers for their valuable feedback.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Salwa Mohamed http://orcid.org/0000-0002-3047-9097

References

Alderson, J. Charles, N. Figueras, H. Kuijper, G. Nold, S. Takala, and C. Tardieu. 2004. The development of specifications for item development and classification within the common European framework of reference for languages: learning, teaching, assessment: final report of the Dutch CEF construct project. Paper presented at workshop on research into and with the CEFR, University of Amsterdam.

Alosh, M. 2006. *Ahlan wa Sahlan: Functional Modern Standard Arabic for Intermediate Learners*. New Haven: Yale University Press.

Alosh, M. 2010a. *Ahlan wa Sahlan: Functional Modern Standard Arabic for Beginners, Letters and Sounds of the Arabic Language*, 2nd ed. New Haven: Yale University Press.

Alosh, M. 2010b. *Ahlan wa Sahlan: Functional Modern Standard Arabic for Beginners*. 2nd ed. New Haven: Yale University Press.

American Council on the Teaching of Foreign Languages [ACTFL]. 2012. *ACTFL Proficiency Guidelines*. 3rd ed. Alexandria: ACTFL INC.

Athanasiou, A., E. K. Constantinou, M. Neophytou, A. Nicolaou, S. P. Sophocleous, and C. Yerou. 2016. Aligning ESP courses with the common European framework of reference for languages. *Language Learning in Higher Education* 6, no. 2: 297–316.

Barrot, J. S. 2015. Comparing the linguistic complexity in receptive and productive modes. *Journal of Language Studies* 15, no. 2: 65–81.

Beatty, K. 2015. Language, task and situation: authenticity in the classroom. *Journal of Language and Education* 1, no. 1: 27–37.

Brustad, K., M. Al-Batal, and A. Al-Tonsi. 2004a. *Alif baa: Introduction to Arabic Letters and Sounds*. 2nd ed. Washington, DC: Georgetown University Press.

Brustad, K., M. Al-Batal, and A. Al-Tonsi. 2004b. *Al-Kitaab fii taallum al-arabiyya: A Text for Arabic, Part 1*. 2nd ed. Washington, DC: Georgetown University Press.

Brustad, K., M. Al-Batal, and A. Al-Tonsi. 2006. *Al-Kitaab fii taallum al-arabiyya: A Text for Arabic, Part 2*. 2nd ed. Washington, DC: Georgetown University Press.

Brustad, K., M. Al-Batal, and A. Al-Tonsi. 2007. *Al-Kitaab fii taallum al-arabiyya: A Text for Arabic, Part 3*. 2nd ed. Washington, DC: Georgetown University Press.

Byrnes, H. 2012. Of frameworks and the goals of collegiate foreign language education: Critical reflections. *Applied Linguistics Review* 3, no. 1: 1–24.

Cambridge University. 2018. *How Long Does It Take to Learn a Foreign Language?* Cambridge: Cambridge University Press.

Council of Europe. 2001. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.

Council of Europe. 2009. Relating language examinations to the common European framework of reference for languages: learning, teaching, assessment. A manual. http://www.coe.int/t/dg4/linguistic/manuel1_en.asp. Accessed 15 March 2020.

Crossley, S. A., D. F. Dufty, P. M. McCarthy, and D. S. McNamara. 2007. Toward a new readability: a mixed model approach. In *Proceedings of the 29th Annual Cognitive Science Society*, eds. D. S. McNamara, and J. G. Trafton, 197–202. Austin, TX: Cognitive Science Society.

Crossley, S. A., and D. S. McNamara. 2008. Assessing L2 Reading texts at the intermediate level: an approximate replication of Crossley, Louwerse, McCarthy & McNamara (2007). *Language Teaching* 41, no. 3: 409–429.

Daniels, P. T., and D. L. Share. 2018. Writing system variation and its consequences for Reading and dyslexia. *Scientific Studies of Reading* 22: 101–116.

Dudley-Evans, T., and M. J. St. John. 1998. *Developments in English for Specific Purposes: A Multi-Disciplinary Approach* Cambridge: Cambridge University Press.

Francois, T., and C. Fairon. 2012. An AI readability formula for French as a foreign language. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, Jeju Island, Korea, 12–14 July 466–477.

Grabe, W., and F. L Stoller. 2011. *Teaching and Researching Reading*. 2nd ed. London: Routledge.

Graesser, A. C., D. S. McNamara, and J. Kulikowich. 2011. Coh-Metrix: providing multilevel analyses of text characteristics. *Educational Researcher* 40: 223–234.

Graesser, A. C., D. S. McNamara, M. M. Louwerse, and Z. Cai. 2004. Coh-Metrix: analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers* 36: 193–202.

Guariento, W., and J. Morley. 2001. Text and task authenticity in the EFL classroom. *ELT Journal* 55, no. 4: 347–353.

Gutierrez Bermudez, J. F. 2014. An exercise in course-book evaluation: strengths, weaknesses, and recommendations regarding new *English File*: *Elementary*. *Latin American Journal of Content and Language Integrated Learning* 7, no. 1: 98–111.

Hong, J. F., C. Y. Peng, H. C. Tseng, and Y. T. Sung. 2020. Linguistic feature analysis of CEFR labeling reliability and validity in language textbooks. *Journal of Technology and Chinese Language Teaching* 11, no. 1: 57–83.

Hulstijn, J.H., J. C. Alderson, and R. Schoonen. 2010. Developmental stages in second-language acquisition and levels of second-language proficiency: are there links between them? In *Communicative Proficiency and Linguistic Development: Intersections Between SLA and Language Testing Research*, eds. I. Bartning, M. Martin, and I. Vedder, 11–20. European Second Language Association. www.lulu.com.

Interagency Language Roundtable. 2012. Levels of proficiency. https://www.govtilr.org/index.htm.

Martyniuk, W., and J. Noijons. 2007. Executive summary of results of a survey on the use of the CEFR at national level in the Council of Europe member states. *Council of Europe, Language Policy Division*. Accessed 30 April 2020. https://www.coe.int/en/web/common-european-framework-reference-languages/documentsMohamed 2021a.

Mohamed, S. 2021a. The development of an Arabic curriculum framework based on a compilation of salient features from CEFR level descriptors. *The Language Learning Journal*. https://www.tandfonline.com/doi/full/10.108009571736.2021.1923781.

Mohamed, S. 2021b. An analysis and levelling of reading texts across Arabic textbooks based on the CEFR proficiency levels. In *Trends and Developments for the Future of Language Education in Higher Education*, ed. C. Xiang, 57–83. Pennsylvania: IGI Global.

Natova, I. 2019. Estimating CEFR reading comprehension text complexity. *The Language Learning Journal*. https://www.tandfonline.com/doi/full/10.108009571736.2019.1665088.

Pilan, I. 2013. NLP-based approaches to sentence readability for second language learning purposes. Master's dissertation, Gothenburg: University of Gothenburg. file:///C:/Users/55096616/Downloads/NLP_based_Approaches_to_Sentence_Readabi.pdf.

Pilan, I., and E. Volodina. 2018. Investigating the importance of linguistic complexity features across different datasets related to language learning. In *Proceedings of the Workshop on Linguistic Complexity and Natural Language Processing*, Santa Fe, New Mexico, USA, 49–58.

Pilan, I., E. Volodina, and T. Zesch. 2016. Predicting proficiency levels in learner writings by transferring a linguistic complexity model from expert-written coursebooks. In *Proceedings of the 26th International Conference on Computational Linguistics*, 2101–2111.

Renandya, W. A. 2015. Reading in a foreign language: what else is important besides skills and strategies? In *Developing Indegenous Models of English Language Teaching and Assessment*, eds. F. A. Hamied, I. B. P. Yadnya, and I. G. A. G. Sociowati, 81–94. Bali, Indonesia: Udayana University Press.

Saddiki, H., K. Bouzoubaa, and V. Cavalli-Sforza. 2015. Text readability for Arabic as a foreign language: what performance to expect from simple predictors? In *IEEE/ACS 12th International Conference of Computer Systems and Applications (AICCSA)*, IEEE, 1-8.

Salamoura, A., and N. Saville. 2010. Exemplifying the CEFR: criterial features of written learner English from the English profile programme. In *Communicative Proficiency and Linguistic Development: Intersections Between SLA and Language Testing Research*, eds. I. Bartning, M. Martin, and I. Vedder, 101–132. European Second Language Association. www.lulu.com.

Sung, Y. T., W. C. Lin, S. B. Dyson, K. E. Chang, and Y. C. Chen. 2015. Leveling L2 texts through readability: combining multilevel linguistic features with the CEFR. *The Modern Language Journal* 99, no. 2: 371–391.
Tomlinson, B. 2012. Materials development for language learning and teaching. *Language Teaching* 45, no. 2: 143–179.
Uchida, S., and M. Negishi. 2018. Assigning CEFR-J Levels to English Texts Based on Textual Features. In *Fourth Asia Pacific Corpus Linguistics Conference (APCLC 2018)*, Takamatsu: Japan, 463–467.
Uri, N. F. M., and M. S. Abd Aziz. 2018. Assessing readability of a national exam reading texts in Malaysia. *European Journal of English Language Teaching* 4, no. 1: 149–164.
Wightwick, J., and M. Gaafar. 2009. *Mastering Arabic 2*. London: Palgrave Macmillan.
Wightwick, J., and M. Gaafar. 2015. *Mastering Arabic 1*. 3rd ed. London: Palgrave Macmillan.
Yassin, R., D. L. Share, and Y. Shalhoub-Awwad. 2020. Learning to spell in Arabic: the impact of script-specific visual-orthographic features. *Frontiers in Psychology* 11: 1–10.
Younes, M., and H. Al-Masri. 2014. *Arabiyyat Al-Naas Part Two*. New York: Routledge.
Younes, M., and Y. Chami. 2014. *Arabiyyat Al-Naas Part Two*. New York: Routledge.
Younes, M., M. Weatherspoon, and M.S. Foster. 2014. *Arabiyyat al-Naas Part One*. New York: Routledge.
Zheng, Y., Y. Zhang, and Y. Yan. 2016. Investigating the practice of the Common European Framework of Reference for languages (CEFR) outside Europe: a case study on the assessment of writing in English in China. British Council. Accessed 19 March 2020. https://www.teachingenglish.org.uk/article/investigating-practice-cefr-outside-europe-a-case-study-assessment-writing-english-china.

Appendices

Appendix 1. Copy of the original Dutch CEFR Grid (Alderson et al. 2004)

	Descriptive categories
1	Text Source
2	Authenticity
3	Discourse type
4	Domain
5	Topic
6	Nature of Content
7	Text Length
8	Vocabulary (frequent > extended)
9	Grammar (simple > complex)
	Comprehensible by learner at CEF level

Appendix 2. Copy of the adapted Dutch CEFR Grid

	Descriptive categories
1	Text Source
2	Authenticity
3	Discourse type
4	Domain
5	Topic
6	Nature of Content
7	Text length
8	Vocabulary
9	Grammar
10	Discourse markers (connectives, logical markers, signposting markers, etc.)
11	Salient stylistic features (modals; determiners; adverbs; intensifiers; collocations, etc.)
	Comprehensible by learner at CEF level

Appendix 3. Translation of the sample texts analysed (Figures 1–8)

Figure 1 My name is Jamal and I like chicken and oven made meat. I like rice and potatoes as well but I do not like cheese or milk. My favourite dish is grilled chicken with rice.My name is Karima and I do not like meat or chicken, I prefer fish and vegetables. My favourite dish is fish fried in oil with tomato salad.I am Mido and my favourite dish is pizza! I also like fried chicken and fried potatoes (chips) and cola but I do not like vegetables or fruits.

Figure 2 **General manager for a trade company** – Post: General manager for the main office in Riyad. High salary, luxurious house, tickets from and to Saudi, health insurance.**Required in Baghdad** – Deputy manager for an engineering company in Baghdad. Experience in the field of engineering not less than 3 years. English language is required.**Required** – in the Egyptian office of a German magazine designer and manager assistant. Salary according to experience. Mastering German spoken and written.**Sales representative** – in the field of computers. Experience for no less than 5 years. Ambitious in his work. University graduate. Salary according to experience.**French Chef** - … centre of Amman city …

Figure 3 Exercise no. 2: Fill in the gaps [with one of the following words]alphabet, European, family, semitic, Hebrew, Arabic, languages, language, many, especiallyThe Arabic language influenced _____ other languages and we have seen how a lot of Arabic words entered the _____ languages such as Alcohol, algorithms, cipher, admiral, etc.Hundreds of _____ words entered the languages of Islamic countries and nations such as Persian, Turkish, Urdu, Pashto, Kurdish and Swahili, _____ in the fields of religion, literature and science. Most of these _____ use the Arabic alphabet.Although the Persian language uses the Arabic _____ and has many Arabic words, its structure is closer to the English _____ since the two languages are from the same family which is the Indo-European _____ The Arabic language is closer in its structure to the _____ language which uses a different alphabet as the two languages are from the _____ languages family.

Figure 4 Exercise no. 13: Fill in the gapsAlso, Tomatoes, tea, the coffee, café, twenty, coffee, much, one, and tomatoesWhen I lived in the Middle East I knew _____ type of apples and one type of oranges and one type of _____ and potatoes and so most vegetables and fruits.But here when I go to the vegetables and fruits market, I see more than _____ types of apples and five types of oranges, and five types of potatoes _____ and so on.Likewise when I went to the café to drink _____ or the coffee, there was one type of tea and one type of _____.But here when I go to _____ I see twenty types of coffee: normal coffee, decaffeinated coffee, iced _____, mocha, hazelnut coffee, americano, espresso, cappuccino, latte and much more.As for tea, there is _____ many types: there is black tea and green tea. And the black tea has many types: English breakfast, earl grey, Assam, Kenya tea, Moroccan mint tea, and _____ more.

Figure 5 Exercise 16 – In-class reading activityNew York: city (944 square kilometres including water body and 8,084,316 people), southeast of New York state the largest city in the United States of America, on New York Bay, at the mouth of the Hudson River. It consists of five sections: Manhattan, Bronx (north of Manhattan), Queens (west of Long Island), Brooklyn, and Richmond on Staten Island, and these sections are connected by several bridges and tunnels. It comprises the Metropolitan area (2,2100,000 people), industrial areas, and other residential areas in southeast New York, southeast New Jersey, and southwest Connecticut, just as it comprises a luxurious harbour which is considered one of the greatest harbours in the world and a network of air and railway lines. It is the commercial and financial centre for the United States. It has many industries, especially consumer goods.

Figure 6 **Exercise 16 – In-class reading activity**The daily life of his majesty the kingOnce, his majesty king Fouad said to a group of his specialists: I usually wake up at five o'clock in the morning but I do not leave my wing until long after that.After waking up, his majesty the king enters the bathroom, and gets dressed straight after that. Then he does some exercises. After that he sits at the table to have a very light breakfast. After finishing food his majesty commits himself to work so he sits in one of the halls of his private wing and starts browsing newspapers and magazines. It is self-evident that the Egyptian newspapers are of interest to him more than others, and so the Arabic and French daily newspapers are brought to His Majesty as soon as they are released.His majesty the king receives ten daily newspapers and about twenty French magazines and four or five daily newspapers and about twenty English magazines, as well as many Italian newspapers and magazines. This way his majesty the king is able to follow the general situation in the whole world and also by reading the local daily newspapers he becomes acquainted with the minute details of the political, literary and social situation in his kingdom.At half past eleven his majesty the king starts receiving visitors and it is often that meetings and ceremonies continue ceaselessly until two o'clock in the afternoon and his majesty usually receives visitors in his spacious office which overlooks Abdeen palace.After the end of meetings his majesty the king takes dinner then immediately meets the heads of the royal court and studies with them the issues they present to him then he announces/gives his instructions, signs decrees and orders until half past four in the afternoon. when his majesty returns to his official meetings.After the end of official meetings his majesty the king assembles with the senior officials of the palace and works with them until late hours. It is often that his majesty remains in his office until eight o'clock … then his majesty the king enters his private wing to have a light supper and spend a period of time with members of his honourable family.From: Thabet, Karim. 1931. *His Majesty the King between Egypt and Europe*. Cairo: Dar Al-Hilal.

Figure 7 **Exercise 6**1- My name is Rawda Qattan.I usually prepare a big breakfast for me and my husband and my children.We return to the house at eight o'clock in the evening.At four o'clock in the afternoon I and my children go to my father and mother's house.I work [as] a librarian in the Jordanian University in Amman.But I wake up late on Friday (at nine or ten o'clock).And sometimes we eat Hummus besides olives and cheese.When I chat with them [two] and with my sister and brother.After breakfast I clean the house and wash the shirts.Sometimes I prepare Foul Modammas [processed fava beans] and fried eggs.I go to my work at seven o'clock every day.

Figure 8 **Exercise 10 – Rearrange the sentences to form a meaningful paragraph without changing the position of the first sentence.**1- The first thing I do in the morning is going to the bathroom.2- After watching my favourite show, I sleep for half an hour only.3- At eight o'clock in the morning, I go to my work.4- In the evening, I and my husband watch a film on TV or on a recorded tape.5- After that I sit in front of the TV and watch the news while eating breakfast.6- When I return home in the afternoon, I watch my favourite soap opera.7- During work I listen to the radio.