Linguistic Annotation Workshop (LAW-XIX-2025)

Guidelines for Fine-grained Sentence-level Arabic Readability Annotation

Nizar Habash,† Hanada Taha-Thomure,‡ Khalid N. Elmadani,†
Zeina Zeino,‡ Abdallah Abushmaes††

†New York University Abu Dhabi
‡Zayed University ††Abu Dhabi Arabic Language Centre

Introduction

- Readability refers to how easy or difficult to read and understand a piece of text.

Arabic	Translation
كُرَة	Ball
غُرْفَةُ النَّوْمِ	The bedroom
سُلوكي مَسْؤولِيَّتي	My behavior is my responsibility
كانت الحديقة واسعة، تطل على شاطئ النيل،	The garden was spacious, overlooking the Nile shore.
تعريف أصول الفقه	Definition of the Principles of Islamic Jurisprudence
بين طعن القَنا وخَفْق البُنودِ	Between the thrusts of lances and the fluttering of ensigns

Introduction

- Readability refers to how easy or difficult to read and understand a piece of text.

Grade	Arabic	Translation
KG	كُرَة	Ball
1st	غُرْفَةُ النَّوْمِ	The bedroom
2nd	سُلوكي مَسْؤولِيَّتي	My behavior is my responsibility
4th	كانت الحديقة واسعة، تطل على شاطئ النيل،	The garden was spacious, overlooking the Nile shore.
8th	تعريف أصول الفقه	Definition of the Principles of Islamic Jurisprudence
Uni	بين طعن القَنا وخَفْق البُنودِ	Between the thrusts of lances and the fluttering of ensigns

Introduction

- Readability refers to how easy or difficult to read and understand a piece of text.

RL	Grade	Arabic	Translation
1	**KG**	كُرَة	Ball
3	**1st**	غُرْفَةُ النَّوْمِ	The bedroom
6	**2nd**	سُلوكي مَسْؤولِيَّتي	My behavior is my responsibility
10	**4th**	كانت الحديقة واسعة، تطل على شاطئ النيل،	The garden was spacious, overlooking the Nile shore.
14	**8th**	تعريف أصول الفقه	Definition of the Principles of Islamic Jurisprudence
17	**Uni**	بين طعن القَنا وخَفْق البُنودِ	Between the thrusts of lances and the fluttering of ensigns

Introduction

- **Why Readability?**
  - Helps aligning texts with students' reading abilities.
  - Impacts understanding, retention, reading speed and engagement.
  - Developing readability models is crucial for improving literacy, language learning, and academic performance.

- **BAREC**: the **B**alanced **A**rabic **R**eadability **E**valuation **C**orpus
  - Over 69K sentence (1M words) annotated into **19 readability** levels inspired by the Taha/Arabi21 book-leveling approach.

Why 19 levels?

Taha-Thomure (2017)'s 19-level Arabic text leveling framework

- Inspired by Fountas and Pinnell (2006)'s 26 levels.
- 11 of the 19 levels covering up to 4th grade to support teachers in matching books to students' reading abilities.
- Ten qualitative and quantitative criteria: from text genre, and sentence structure, to book production quality, and word count.
- The Arab Thought Foundation adopted this framework and funded the leveling of over 9,000 children's books (**Arabi21**)
- *We mapped and restricted the features to apply on sentences.*

Arabic Orthographic Ambiguity

- Arabic script uses optional diacritical marks
  - 1.5% of newspaper words have some diacritical marks
  - Standard Arabic has 6.8 diacritizations and 2.7 lemmas/word

ولعين

وَلِعِينَ وَلِعَينِ وَلَعِينٌ

Infatuated (m.pl) # and for an eye/spring # and cursed

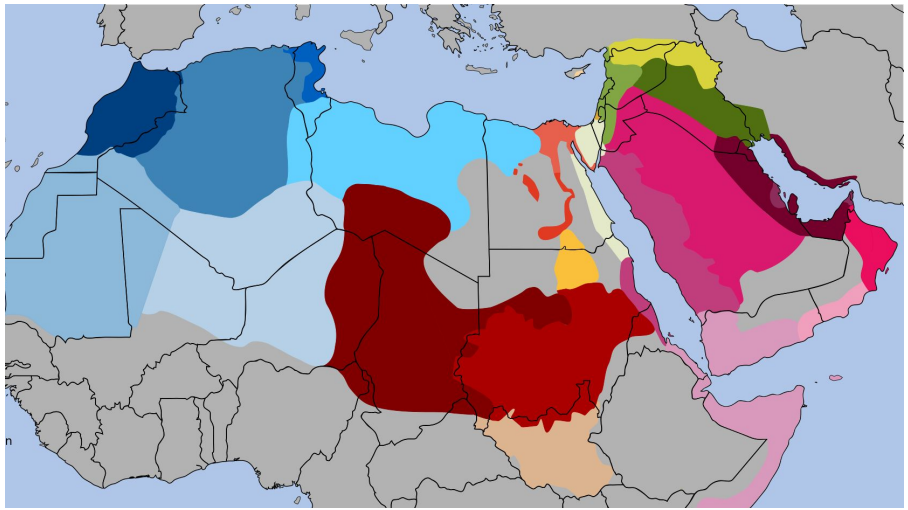
Arabic Morphological Richness

- Arabic has a very rich inflectional+cliticizational system
  - For example, Arabic verbs have 5,400+ forms
  - Whereas English verbs have 6!

قال، قالت، قالا، قالوا، قلتَ،	وسنقولها
قلتِ، قلتما، قلتم، قلتن،	/wasanaqūluhā/
يقولُ، يقولَ، يقل، تقولُ، تقولَ،	و+ س+ ن+ قول + ها
تقل، تقولين، تقولي،	wa+sa+na+qūl+u+hā
... فقال، فقالت، فقالا	and+will+we+say+it
... وسأقولها، **وسنقولها**، ...	*And we will say it*

Arabic Dialectal Variation

- Classical Arabic
  - Quranic Arabic, Historical texts
- Modern Standard Arabic
  - Official language
  - Language of news & media
  - Standard writing & grammar
  - The National Language
- Dialectal Arabic
  - Predominantly spoken
  - No official standardization
  - The Mother Tongue
  - Increasing use on social media

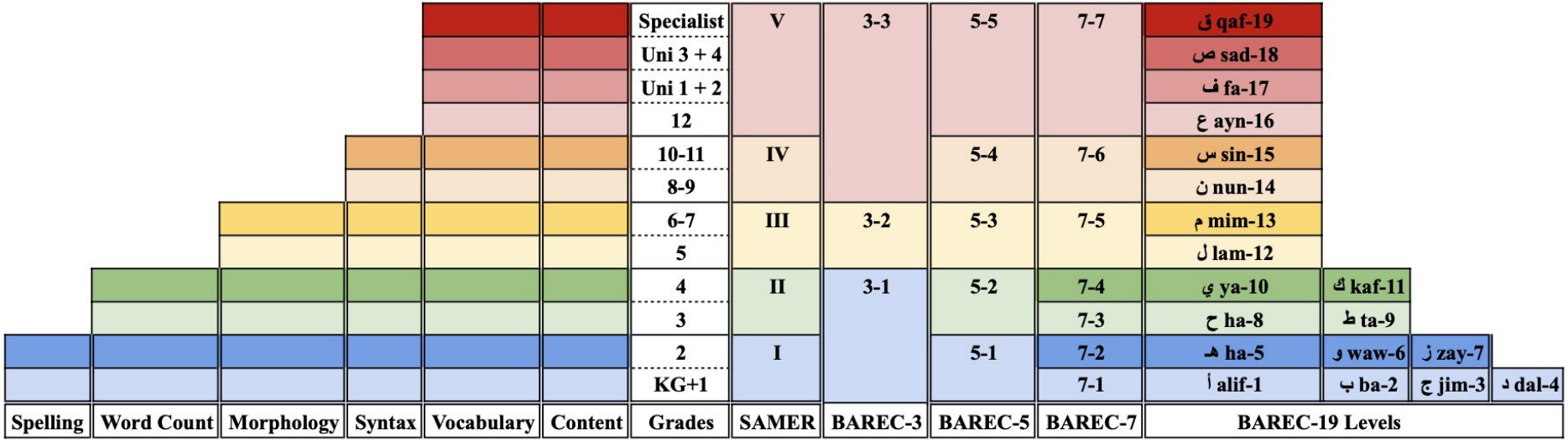


Annotation Guidelines

- **Textual Features**: Levels are assessed in six key dimensions.
  a. **Spelling**: Word length & syllable count.
  b. **Word Count**: The number of unique words.
  c. **Morphology**: Inflection & derivation.
  d. **Syntax**: Sentence structure & syntactic relations.
  e. **Vocabulary**: The complexity of word choices.
  f. **Content**: The required prior knowledge and abstraction levels.

BAREC Pyramid

The BAREC Pyramid illustrates the relationship across BAREC levels and linguistic dimensions, three collapsed variants (3 levels, 5 levels and 7 levels), and educational grades.



BAREC Annotation Guidelines

مستوى بارق	صف	ACTFL	عدد كلمات	تهجئة وإملاء	تصريف واشتقاق	تراكيب نحوية	مفردات	فكرة ومحتوى
أ	روضة-1	مبتدئ أدنى	1	• كلمات من مقطع واحد أو مقطعين	• الفعل المضارع المفرد	• كلمة واحدة	• اسم جنس • اسم علم (متداول بسيط تركيبيا) • ضمير منفصل • مفردات متطابقة مع العامية - سامر I • الأرقام (العربية أو الهندية) 1-10	• فكرة مباشرة وصريحة وحسية. • لا رمزية في النص.
ب	1	مبتدئ أدنى	≤2	• كلمات من 3 مقاطع		• جملة اسمية (هو يلعب) • إضافة حقيقية (باب البيت) • صفة وموصوف (باب كبير)	• فعل • صفة • مفردات متشابهة مع العامية - سامر I • العدد الأصلي بالأحرف • الأسماء الخمسة: أبو، أخو	
ج		مبتدئ متوسط	≤4	• كلمات من 3 مقاطع	• سوابق: ال التعريف • سوابق: واو العطف • لواحق: ضمير المتكلم المفرد المتصل	• بدل كل: (صديقي أحمد) • بدل إشارة: (هذا البيت)	• مفردات فصيحة شائعة - سامر I • اسم الإشارة المفرد • الأرقام (العربية أو الهندية) 11-100	
د		مبتدئ متوسط	≤6	• كلمات تستخدم مد الألف (أ)	• الفعل المضارع الجمع • سوابق: حروف جر متصلة • ظرف منون	• جملة فعلية بدون مفعول به • جار ومجرور	• حروف الجر	
ه	2	مبتدئ أعلى	≤8	• كلمات من 4 مقاطع	• لواحق: ضمير متصل مفرد أو جمع • المثنى (في الأسماء والصفات) • جمع المؤنث السالم	• جملة فعلية مع مفعول به واحد اسم • جمل معطوفة • أدوات استفهام أساسية: ماذا، متى، من، أين، ما، كيف • صيغة التعجب "ما أفعل"	• العدد الترتيبي • الأرقام (العربية أو الهندية) 101-1,000 • اسم اشارة مثنى، جمع	• المحتوى من حياة القارئ. • لا رمزية في النص.
و		مبتدئ أعلى	≤9	• كلمات من 5 مقاطع	• الفعل الماضي المفرد والجمع • جمع مذكر سالم	• جملة فيها فعلين (مثلا جملة فعلية مفعولها أن المصدرية)	• مفردات فصيحة - سامر I	
ز		متوسط أدنى	≤10	• كلمات من +6 مقاطع • أفعال/أسماء معتلة الآخر	• الفعل الماضي المثنى •الفعل المضارع المثنى • فعل الأمر المفرد • لواحق: ضمير المثنى المتصل • جمع التكسير • واو القسم (والله)	• مفعول فيه (ظروف زمان ومكان) • حال • أداة الاستفهام هل	• مفردات فصيحة شائعة - سامر II	• بعض الرمزية أو عدم التصريح المباشر بكل المقصود في الجملة

BAREC Annotation Guidelines

BAREC Level	Grade	ACTFL	Word Count	Spelling/Pronunciation	Morphology	Syntax	Vocabulary	Idea/Content
1-alif	Pre1-1	Novice Low	1	• One-syllable and two-syllable words	• Singular imperfective verb	• One word	• Common noun • Proper noun (frequent and simple) • Personal pronouns (non-clitics) • Vocabulary identical to dialectal form - SAMER I • Numbers (Arabic or Indo-Arabic) 1-10	• Direct, explicit, and concrete idea. • No symbolism in the text.
2-ba	1	Novice Low	≤2	• Three-syllable words			• Verb • Adjective • Vocabulary similar to dialectal form - SAMER I • Spelled cardinal numbers • The five nouns: *Abw (father)*, *Axw (brother)*	
3-jim		Novice Mid	≤4		• Prtoclitic: Definite article *Al+* • Proclitic: Conjunction *wa+* • Enclitic: First Person Singular pronoun	• Apposition (full) • Demonstratives	• Common MSA vocabulary - SAMER I • Singular demonstrative pronoun • Numbers: 11-100	
4-dal		Novice Mid	≤6	• Words with an elongated Alif (e.g. /ʔāsif/)	• Plural imperfective verb • Prepositional proclitics • Nunated adverbials	• Verbal sentence w/o direct object • Preposition and object	• Prepositions	
5-ha	2	Novice High	≤8	• Four-syllable words	• Enclitic: Singular and Plural pronouns • Dual (in nouns and adjectives) • Sound feminine plural	• Verbal sentence with one nominal direct object • Conjoined sentences • Basic interrogative particles: what, when, who, where, how • Exclamatory form: how <comparative adjective>	• Ordinal numbers • Numbers: 101-1,000 • Dual and plural demonstrative pronoun	• Content is from the reader's life. • No symbolism in the text.
6-waw		Novice High	≤9	• Five-syllable words	• Singular and plural perfective verb • Sound masculine plural	• Sentence with two verbs (e.g., a verbal sentence a clausal direct object introduced with *Masdar 'an [~to/that]*)	• MSA vocabulary - SAMER I	

BAREC Annotation Guidelines

BAREC Level	Grade	ACTFL	Word Count	Spelling/Pronunciation	Morphology	Syntax	Vocabulary	Idea/Content
1-alif	Pre1-1	Novice Low	1	• One-syllable and two-syllable words	• Singular imperfective verb	• One word	• Common noun • Proper noun (frequent and simple) • Personal pronouns (non-clitics) • Vocabulary identical to dialectal form - SAMER I • Numbers (Arabic or Indo-Arabic) 1-10	• Direct, explicit, and concrete idea. • No symbolism in the text.
2-ba	1	Novice Low	≤2	• Three-syllable words			• Verb • Adjective • Vocabulary similar to dialectal form - SAMER I • Spelled cardinal numbers • The five nouns: *Abw (father)*, *Axw (brother)*	
3-jim		Novice Mid	≤4		• Prtoclitic: Definite article *Al+* • Proclitic: Conjunction *wa+* • Enclitic: First Person Singular pronoun	• Apposition (full) • Demonstratives	• Common MSA vocabulary - SAMER I • Singular demonstrative pronoun • Numbers: 11-100	
4-dal		Novice Mid	≤6	• Words with an elongated Alif (e.g. /ʔāsif/)	• Plural imperfective verb • Prepositional proclitics • Nunated adverbials	• Verbal sentence w/o direct object • Preposition and object	• Prepositions	
5-ha	2	Novice High	≤8	• Four-syllable words	• Enclitic: Singular and Plural pronouns • Dual (in nouns and adjectives) • Sound feminine plural	• Verbal sentence with one nominal direct object • Conjoined sentences • Basic interrogative particles: what, when, who, where, how • Exclamatory form: how <comparative adjective>	• Ordinal numbers • Numbers: 101-1,000 • Dual and plural demonstrative pronoun	• Content is from the reader's life. • No symbolism in the text.
6-waw		Novice High	≤9	• Five-syllable words	• Singular and plural perfective verb • Sound masculine plural	• Sentence with two verbs (e.g., a verbal sentence a clausal direct object introduced with *Masdar 'an [~to/that]*)	• MSA vocabulary - SAMER I	

BAREC Annotation Guidelines

BAREC Level	Grade	ACTFL	Word Count	Spelling/Pronunciation	Morphology	Syntax	Vocabulary	Idea/Content
7-zay		Intermediate Low	≤10	• Six-syllable or more words • Verbs/nouns with weak final letters	• Dual perfective verb • Dual imperfective verb • Singular imperative verb • Enclitics: dual pronoun • Broken plurals • Waw of oath	• Adverbial accusative (time and place adverbs) • Circumstantial accusative • Interrogative particle *hal*	• High frequency MSA vocabulary - SAMER II	• Some symbolism, or not everything is stated directly in the sentence.
8-ha	3	Intermediate Low	≤11		• Plural imperative verb • Feminine plural suffix (*nun*) in nouns and verbs • Other proclitics: future *sa+*, continuation *wa+*, conjunction *fa+* • Conjunctions (e.g., then, until, or, whether, but, as for)	• Absolute object (emphasizing the verb) • Object of purpose • Object of accompaniment • Verbal sentence with two direct objects	• MSA vocabulary - SAMER I and II • Negation particles • Numbers: 1,001-1,000,000	• Some symbolism that requires the reader to seek help to understand the idea.
9-ta		Intermediate Mid	≤12		• Dual imperative verb • Interrogative Hamza • Ba of oath • Oath: The particle of oath, the object of the oath, and the answer to the oat	• Vocative	• Vocabulary describing positive and negative emotional and mood states like joy, happiness, anger, regret, sorrow	• Some symbolism at the event level in the sentence that the reader understands through prior knowledge.
10-ya	4	Intermediate Mid	≤15		• Passive voice	• *Inna* and its sisters (particles introducing a subject) • *Kana* and its sisters (past tense verbs) • Preposed predicate, postponed subject • Chain of narration • *rubba* preposition construction • Relative clauses • Circumstantial and object clauses	• Singular relative pronouns • Verbal particles *qad* and *laqad* • Preposition-Conjunctions: *mimma*, *fima*...	
11-kaf		Intermediate High	≤20		• Acting derivatives (e.g., the active participle)	• Nominal sentence with a nominal predicate • False idafa (tall in stature)	• Dual and plural relative pronouns	• A degree of symbolism and a need for prior knowledge to understand the meaning of the sentence.
12-lam	5	Advanced Low			• Diminutive form	• Parentheticals (explanation, blessing) • Exception • Exclusivity • Apposition (e.g., partitive or containing) • Specification (*tamyiyz* construction)	• MSA vocabulary - Samer III • Frozen Verbs (e.g., *Āmiyn* Amen) • Numbers: > 1,000,000 • Five Nouns: Dhu (possession nominal) • Interjections: *bala*, *Ajal*, etc.	

BAREC Annotation Guidelines

BAREC Level	Grade	ACTFL	Word Count	Spelling/Pronunciation	Morphology	Syntax	Vocabulary	Idea/Content
13-mim	6-7	Advanced Mid			• Energetic mood (emphatic *nun*) • Ta of oath	• Conditional sentences • Jussive particle *lamma* (not yet)	• Words describing deep psychological states like depression, loss, psychological alertness • Use of coined, uncommon words • Abbreviations (e.g., LLC)	• Symbolic ideas and deeper meanings, especially in terms of the psychological dimension of characters/events. • Local cultural expressions that may not be understood by those outside the
14-nun	8-9	Advanced High				• Semantic emphasis • Praise and dispraise • *Masdar 'an* clause as a subject • Exclamatory form: <comparative adjective> *bih min*	• MSA vocabulary - SAMER IV • General legal, scientific, religious, political vocabulary, etc. • Five Nouns: *fw*, *Hmw*	
15-sin	10-11	Superior Low				• Uncommon constructions that are ambiguous and need diacritization for clarification	• Specialized vocabulary that requires understanding the concept/idea to comprehend it • Shortening in proper names (e.g., *fatim* for *fatima*)	• Symbolic, abstract, scientific, or poetic ideas that require prior linguistic and cognitive knowledge to understand.
16-ayn	12	Superior Mid					• MSA vocabulary - SAMER V • Specialized and highly elevated Arabic vocabulary. • Vocabulary mostly distant from dialects.	
17-fa	University Year 1-2	Superior High					• Scientific and heritage vocabulary not in use today, but familiar to a novice specialist	
18-sad	University Year 3-4	Distinguished					• Scientific and heritage vocabulary not in use today, but familiar to a specialist	
19-qaf	Specialist	Distinguished+					• Scientific and heritage vocabulary not in use today, but familiar to the advanced researcher specialist	
Difficulty	This tag is used when there is difficulty in assessing the level. It is preferred to use this tag so that the team can find a solution (for example, by adjusting the criteria or adding explanatory details).							
Problem	Generally, we use this tag for sentences containing:		• Spelling mistakes (e.g., Hamzas, Ta Marbuta, Alif maqsura/Ya) • Errors in diacritics • Linguistic awkwardness (illiteracy, colloquialism, poor translation from a foreign language) • Inappropriate topics (racism, bias, bullying, pornography, etc.) • Sentences and phrases mostly written in languages other than Arabic or in non-Arabic script		However, in the following cases, we provide the level and add a note in the comments column: • Error in Hamzat al-Wasl/Hamzat al-Qat' >> (أ) • Offensive words >> (ع) • Error in diacritics at the beginning of the sentence >> (ت) • Dotted Yaa missing at the end of the word >> (ي)		

Annotation Guidelines

- **Textual Features**: Levels are assessed in six key dimensions.
  a. **Spelling**: Word length & syllable count.
  b. **Word Count**: The number of unique words.
  c. **Morphology**: Inflection & derivation.
  d. **Syntax**: Sentence structure & syntactic relations.
  e. **Vocabulary**: The complexity of word choices.
  f. **Content**: The required prior knowledge and abstraction levels.

RL	Arabic Sentence/Phrase	Translation	Reasoning
1-alif	**أَرْنَب**	**Rabbit**	*One bisyllabic familiar noun*
2-ba	**ملعبٌ واسعٌ**	**A large playground**	*Noun-adjective*
3-jim	أنا أحب **اللون** الأحمر.	I love **the** color red.	*Definite article*
4-dal	الشمس تشرق **في الصباح الباكر**.	The sun rises early **in the morning**.	*Prepositional phrase*
5-ha-	القطة تستريح على السرير **وتستمتع بأشعة الشمس الدافئة**.	The cat rests on the bed **and enjoys the warm sunshine**.	*A conjoined sentence*
6-waw	سُلوكي **مَسْؤوليَّتي**	My behavior is **my responsibility**	*Five syllable word*

Annotation Process

- **Annotation Team**: Five annotators, all of whom are experienced Arabic language educators with experience in Arabi21 project.
- They receive folders containing annotation sets for individual annotation.
- Each set contains 100 randomly selected sentences.
- An average of 3 hours per set (~2min/sentence).
- **Result**: 1,039,371 words - 69,441 sentences - 1,922 documents.
- 19 shared sets for measuring inter-annotator agreement (IAA).

Annotation Team

	A0^{P}	**A1**	**A2**	**A3**	**A4**	**A5**^{L}
Native Language	Arabic	Arabic	Arabic	Arabic	Arabic	Arabic
Other Language	En, Fr	En	En, Fr	En, Fr	En, Fr	En, Fr
Nationality	Syrian	Lebanese	Lebanese	Lebanese	Lebanese	Lebanese
Residence	USA	Lebanon	Lebanon	Lebanon	UAE	Lebanon
Gender	Female	Female	Female	Female	Female	Female
Background	Muslim	Muslim	Muslim	Muslim	Christian	Muslim
Degree	MA	BA	BA	MA	MA	B MA
Major	Applied Ling.	Arabic Lit.	Geography	Arabic Lit.	Arabic Lit.	Arabic Lit.
Experience	CT, LA, RA	PT, LA	PT, LA	CT, LA	CT, LA	CT, LA, RA
School	Private	-	-	Public&Private	Private	Public
Level	University	Elementary	Elementary	Secondary	Secondary	Secondary
Students	L2	L1	L1	L1	L1	L1
Years	16	16	22	22	8	25

Annotation Interface

Notes ملاحظات	Idea/Content فكرة / محتوى	Vocabulary مفردات	Syntax تراكيب نحوية	Morphology تصريف واشتقاق	Spelling/Pronunciation تهجئة/املاء	Word Count عدد الكلمات	Level المستوى	Length عدد الكلمات	Sentence/Phrase الجملة \ العبارة
	• المحتوى من حياة القارئ. • لا رمزية في النص.	• مفردات فصيحة - سامر ١	• جملة فيها فعلين (مثلا جملة فعلية مفعولها أن المصدرية)	• الفعل الماضي المفرد والجمع • جمع مذكر سالم	• كلمات من ٥ مقاطع (بدون حساب حركات الإعراب)	٩ هو أعلى عدد كلمات مطبعية غير متكررة بدون علامات الترقيم	6-waw و (صف 2)	1	خَبَّرَ
	• بعض الرمزية أو عدم التصريح المباشر بكل المقصود في الجملة	• مفردات فصيحة شائعة - سامر ٢	• مفعول فيه (ظروف زمان ومكان) • حال • أداة الاستفهام هل	• الفعل الماضي المثنى •الفعل المضارع المثنى • فعل الأمر المفرد • جمع التكسير • واو القسم (والله)	• كلمات من ٦ مقاطع أو أكثر (بدون حساب حركات الإعراب) • أفعال/أسماء معتلة الآخر	١٠ هو أعلى عدد كلمات مطبعية غير متكررة بدون علامات الترقيم	7-zay ز (صف 2)	2	جودي بقربي
	• بعض الرمزية يحتاج معها القارئ إلى مساعدة من يشرح له المقصود من الفكرة	• مفردات فصيحة - سامر ١ و سامر ٢ • أحرف النفي • الأرقام (العربية أو الهندية) 1,000,000-1,001	• المفعول المطلق • المفعول لأجله • المفعول معه • جملة فعلية تتعدى إلى مفعولين	• فعل الأمر الجمع • نون النسوة في الأسماء والأفعال (انتظرنَ دورهنَّ) • سوابق أخرى: سين الاستقبال ، واو الاستئناف، فاء العطف • (ثم ، حتى ، أو ، أم ، لكن ، أمّا)		١١ هو أعلى عدد كلمات مطبعية غير متكررة بدون علامات الترقيم	8-ha ح (صف 3)	4	بيروت في يوليو ١٩٦٦
	• هناك درجة من الرمزية وحاجة للمعرفة السابقة كي يُفهم المقصود من الجملة	• أسماء الوصل المثنى والجمع • متلازمات لفظية مثل شارد الذهن، وارف الظلال	• جملة أسمية خبرها جملة أسمية (فيها مبتدآن) • إضافة خيالية (لفظية) طويل القامة	• المشتقات على أنواعها (تركز على المشتقات العاملة لاسيما اسم الفاعل واسم المفعول)		٢٠ هو أعلى عدد كلمات مطبعية غير متكررة بدون علامات الترقيم	11-kaf ك (صف 4)	4	كتابةُ خطَّةٍ لمشروعِ الوحدةِ
	• المحتوى من حياة القارئ. • لا رمزية في النص.	• مفردات فصيحة - سامر ١	• جملة فيها فعلين (مثلا جملة فعلية مفعولها أن المصدرية)	• الفعل الماضي المفرد والجمع • جمع مذكر سالم	• كلمات من ٥ مقاطع (بدون حساب حركات الإعراب)	٩ هو أعلى عدد كلمات مطبعية غير متكررة بدون علامات الترقيم	6-waw و (صف 2)	4	اجْتَمَعَ الأَهْلُ في العِيدِ.
	• هناك درجة من الرمزية وحاجة للمعرفة السابقة كي يُفهم المقصود من الجملة	• مفردات فصيحة - سامر ٣ • اسم الفعل : إيهِ - صَهْ - آمين - حَيَّ - هَاؤُم - هاكَ - هَيَّا - هيتَ - هَلُمَّ إلى - مَهْ - رويدكَ - • الأرقام (العربية أو الهندية > 1,000,000 • ذو • (بَل - بلى - أجل)	• جمل اعتراضية (تفسير - دعاء...) • استثناء • حصر • البدل (غير حالات ما بعد اسم الإشارة) • تمييز	• التصغير		لا حد لعدد الكلمات المطبعية	12-lam ل (صف 5)	4	وَلا يُخَالِطُنا عَجْزٌ وَلا خَوَرُ

Inter Annotator Agreement

Stage	#Sets	Distance	Acc^{19}	$\pm 1Acc^{19}$	QWK
Pilot 3	1	1.69	37.5%	58.5%	79.3%
Phase 1	2	1.38	48.4%	64.4%	80.2%
Phase 2A	6	1.21	49.4%	67.4%	72.4%
Phase 2B	10	0.80	67.6%	78.3%	78.8%
Overall / Macro	19	1.04	58.2%	72.3%	76.9%
Phase 2 / Macro	16	0.96	60.8%	74.2%	76.4%
Phase 2 / Micro	16	0.95	61.1%	74.4%	81.8%

Average pairwise inter-annotator agreement (IAA) across different annotation stages.

Inter Annotator Agreement

Sentence (Arabic)	A1	A2	A3	A4	A5	UL	MM	Comments
أبي.. أبي.. *Dad .. Dad .. [lit. my father .. my father ..]*	2	2	2	3	3	**3**	*1*	First person singular pronoun is level 3.
اِحْتِضانُ الْأُمِ لَهُم. *The mother's embrace for them.*	9	12	5	5	5	**5**	*7*	Disagreement over احتضان 'embrace': standard or dialect aligned.
أشعر بالتعب والجوع.. *I feel tired and hungry..*	9	9	9	9	4	**9**	*5*	Vocabulary describing emotions (level 9).
يتم ضمان حيادية الإدارة بموجب القانون. *Administrative neutrality is guaranteed by law.*	12	12	12	14	12	**12**	*2*	Disagreement over حيادية 'neutrality': general advanced or specialized.

Table 8: Examples of Annotator Disagreements with Unified Levels (UL) and Max-Min Differences (MM)

Selected examples of IAA errors.

Corpus Selection

- 1,922 documents, from 30 different resources.
- Manually categorized into **3 domains**
  - **Arts & Humanities**
  - **Social Sciences**
  - **STEM**
- And **3 readership groups**
  - **Foundational**: up to grade 4
  - **Advanced**: individuals with average adult reading abilities
  - **Specialized**: from grade 9

	#Documents		#Sentences		#Words	
Arts & Humanities	1,367	71%	50,442	73%	652,995	63%
Social Sciences	375	20%	14,319	21%	275,731	27%
STEM	180	9%	4,680	7%	110,645	11%
Foundational	633	33%	27,781	40%	314,068	30%
Advanced	731	38%	22,696	33%	381,660	37%
Specialized	558	29%	18,964	27%	343,643	33%
	1,922	100%	69,441	100%	1,039,371	100%

Corpus Observations

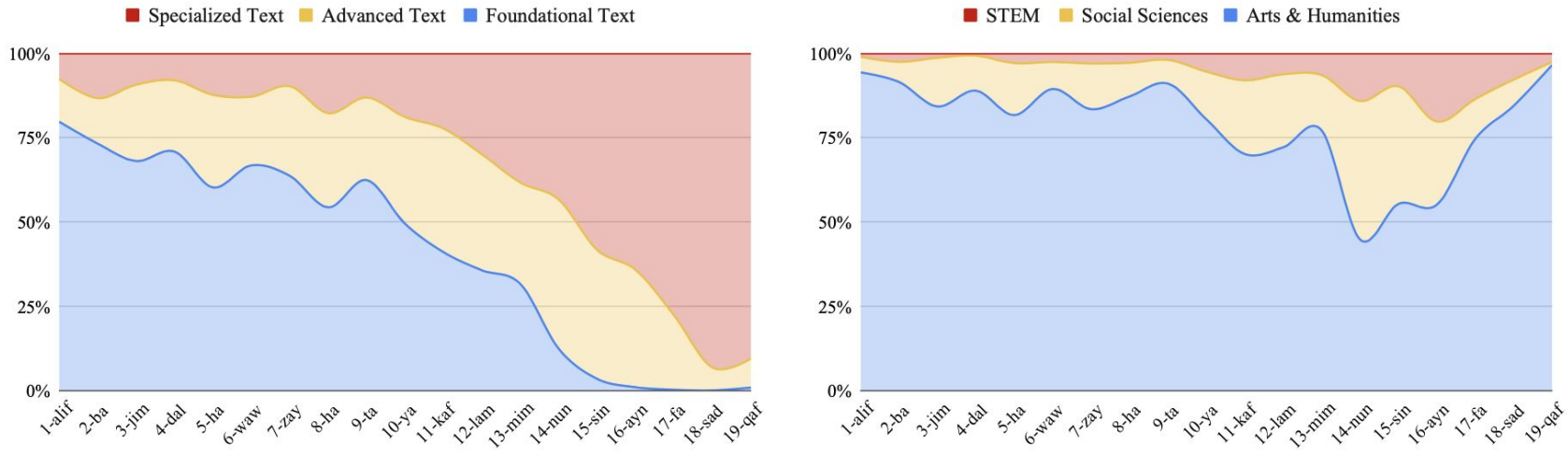


Figure 4: The relative distribution of readership groups and domains across **BAREC** levels.

Automatic Assessment Evaluation

- We define Readability Assessment as an ordinal classification task.
- Metrics: Distance, Accuracy, ±1 Accuracy, Quadratic Weighted Kappa
- Fine-tuned AraBERTv02 (Antoun et al., 2020)

Train	Distance	Acc^{19}	$\pm1\ Acc^{19}$	QWK	Acc^{7}	Acc^{5}	Acc^{3}
12.5%	1.35	45.0%	61.3%	77.2%	56.8%	63.0%	71.3%
25.0%	1.33	46.9%	63.0%	77.6%	58.8%	64.3%	72.3%
50.0%	1.16	52.4%	68.1%	80.7%	62.9%	67.6%	74.0%
100.0%	1.09	55.8%	69.4%	81.0%	64.9%	69.1%	74.7%

Table 9: Performance at different training data sizes across multiple evaluation metrics.

A Large and Balanced Corpus for Fine-grained Arabic Readability Assessment

Khalid Elmadani | Nizar Habash | Hanada Taha

Findings of the Association for Computational Linguistics: ACL 2025

BAREC Analyzer

BAREC Analyzer determines the readability level of Arabic texts using artificial intelligence.

سمير وحبّ الكتب
سمير طفلٌ صغير، يحبّ القراءة كثيرًا. كلّ يوم يقرأ كتابًا جديدًا قبل النوم. في المدرسة، يختار سمير كتبًا صعبة ليفهم كلمات جديدة. أصبح يكتب قصصًا قصيرة ويقرأها لأصدقائه في الصف. بفضل حبّه للكتب، صار يفكّر بذكاء ويحلّ المسائل بسهولة.

ANALYZE

Results:

Readability Levels	Expert				Proficient		Advanced		Intermediate				Beginner							
	ق 19	ص 18	ف 17	ع 16	س 15	ن 14	م 13	ل 12	ك 11	ي 10	ط 9	ح 8	ز 7	و 6	هـ 5	د 4	ج 3	ب 2	أ 1

Level 10 Ya	Text Level
سمير وحبّ الكتب	Level 7 Zay
سمير طفلٌ صغير، يحبّ القراءة كثيرًا.	Level 3 Jim
كلّ يوم يقرأ كتابًا جديدًا قبل النوم.	Level 10 Ya
في المدرسة، يختار سمير كتبًا صعبة ليفهم كلمات جديدة.	Level 8 Ha
أصبح يكتب قصصًا قصيرة ويقرأها لأصدقائه في الصف.	Level 10 Ya
بفضل حبّه للكتب، صار يفكّر بذكاء ويحلّ المسائل بسهولة.	Level 10 Ya

BAREC Shared Task 2025

Arabic Readability Assessment

The Third Arabic Natural Language Processing Conference (ArabicNLP 2025) @ EMNLP 2025

Suzhou, China

Conclusion & Future Work

- We introduced **BAREC**, the largest manually annotated dataset for Arabic Readability Assessment.
- We achieve a high level of agreement between annotators.
- We plan to expand the corpus, enhancing its size and diversity to cover additional genres and topics.
- We also aim to add annotations related to vocabulary leveling and syntactic treebanks to study less-explored genres in syntax.
- We look forward to readability leveling being used for other NLP tasks and as an analytical tool - e.g. educational systems, summarization, reasoning, etc.

Thank You

شكرًا

Website

Data

Questions?