

Error tagset for Arabic learner corpora جدول لوسم الأخطاء اللغوية في المدونات العربية

Abdullah Alfaifi and Eric Atwell

Version 3

عبدالله الفيفي وإيريك أتويل

النسخة الثالثة

Error Category مجال الخطأ	Error Type نوع الخطأ	Tag الرمز
1. Orthography الإملاء 'imlā'	1.1. Hamza (ء، أ، إ، و، ي، ن)	<OH>
	1.2. Confusion in Hā' and Tā' Mutadarrifaṭīn (ت، هـ، عـ، ث)	<OT>
	1.3. Confusion in 'alif and Yā' Mutadarrifaṭīn (ي، ا، ي، ي)	<OA>
	1.4. Confusion in 'alif Fāriqa (اكتبوا)	<OW>
	1.5. Confusion between Nūn (ن) and Tanwīn (و، و، و)	<ON>
	1.6. Shortening the long vowels (اوي → و، و، و)	<OS>
	1.7. Lengthening the short vowels (اوي → و، و، و)	<OG>
	1.8. Wrong order of word characters (الخطأ في ترتيب الحروف داخل الكلمة)	<OC>
	1.9. Replacement in word character(s) (استبدال حرف أو أحرف من الكلمة)	<OR>
	1.10. Character(s) redundant (زيادة حرف أو أكثر)	<OD>
	1.11. Character(s) missing (نقص حرف أو أكثر)	<OM>
	1.12. Other orthographical errors (أخطاء إملائية أخرى)	<OO>
2. Morphology الصرف 'ssarf'	2.1. Word inflection (الخطأ في اختيار بنية الكلمة المناسبة)	<MI>
	2.2. Verb tense (الخطأ في زمن الفعل)	<MT>
	2.3. Other morphological errors (أخطاء صرفية أخرى)	<MO>
3. Syntax النحو 'nnaḥw'	3.1. Case (الخطأ في الإعراب)	<XC>
	3.2. Definiteness (الخطأ في التعريف والتكثير)	<XF>
	3.3. Gender (الخطأ في الجنس (التذكير والتأنيث))	<XG>
	3.4. Number (singular, dual and plural) (الخطأ في العدد (الإفراد والتثنية والجمع))	<XN>
	3.5. Word redundant (كلمة زائدة)	<XT>
	3.6. Word missing (كلمة ناقصة)	<XM>
	3.7. Other syntactic errors (أخطاء نحوية أخرى)	<XO>
4. Semantics الدلالة 'ddalāla'	4.1. Word selection (الخطأ في اختيار الكلمة المناسبة)	<SW>
	4.2. Fasl wa wasl (confusion in use/non-use of conjunctions) (الخطأ في الفصل والوصل (الخطأ في استخدام أدوات العطف))	<SF>
	4.3. Other semantic errors (أخطاء دلالية أخرى)	<SO>
5. Punctuation علامات الترقيم 'alāmāt 't-tarqīm'	5.1. Punctuation confusion (علامة ترقيم خاطئة)	<PC>
	5.2. Punctuation redundant (علامة ترقيم زائدة)	<PT>
	5.3. Punctuation missing (علامة ترقيم مفقودة)	<PM>
	5.4. Other errors in punctuation (أخطاء أخرى في علامات الترقيم)	<PO>

6 Error categories and 29 Error types

29 نوعاً موزعة على 5 مجالات عامة