

KoKo-Korpus: Übersicht über die Annotationen im Bereich Orthographie

Alle Fehlerkategorien basieren auf den amtlichen deutschen Rechtschreibregeln (laut Duden Band 1, 24. Auflage, 2006)

Tagname (in ANNIS): Level 1	Tagname (in ANNIS): Level 2	Fehlerkategorie	Beispiele	Beschreibung
orth_err="orth_error"	orth_err="01 lcp instead of cap: polite form"	01 Klein- statt Großschreibung: Höflichkeitsform	<i>ihnen</i> statt <i>Ihnen</i> , <i>sie</i> statt <i>Sie</i> in der Anrede	Annotiert werden fehlerhafte Kleinschreibung von Höflichkeitsformen bei Anreden.
orth_err="orth_error"	orth_err="02 lcp instead of cap: other cases"	02 Klein- statt Großschreibung: anderes	<ul style="list-style-type: none"> • <i>haus</i> statt <i>Haus</i>, <i>das beweisen</i> statt <i>das Beweisen</i> • <i>das Jugendlichsein</i> statt <i>das Jugendlich-Sein</i> 	Annotiert werden fehlerhafte Kleinschreibungen <ul style="list-style-type: none"> • von Substantiven und Substantivierungen • auch nach Bindestrichen, sofern es keinen Fehler in der Getrennt- und Zusammenschreibung gibt (s. Kategorie 05).
orth_err="orth_error"	orth_err="03 cap instead of lcp"	03 Groß- statt Kleinschreibung	<ul style="list-style-type: none"> • <i>Schön</i> statt <i>schön</i>, <i>Jetzt</i> statt <i>jetzt</i>, <i>Laufen</i> statt <i>laufen</i>, <i>Er</i> statt <i>er</i>, <i>Mit</i> statt <i>mit</i>, etc. • <i>das Immer-Der-Kleinste-Sein</i> statt <i>das Immer-der-Kleinste-Sein</i> 	Annotiert werden fehlerhafte Großschreibungen <ul style="list-style-type: none"> • von Adjektiven, Adverbien, Verben, Pronomen, Präpositionen, etc. • Nach Bindestrichen, sofern es keinen Fehler in der Getrennt- und Zusammenschreibung gibt (s. Kategorie 05).
orth_err="orth_error"	orth_err="04 lcp/cap behind punctuation marks"	04 Groß-/Kleinschreibung nach Satzzeichen	<p>: <i>es solle noch regnen.</i> statt : <i>Es solle noch regnen.</i></p> <p>. <i>sie war sich sicher.</i> statt . <i>Sie war sich sicher.</i></p>	Annotiert wird die Nicht-Beachtung der Groß- und Kleinschreibung nach Punkten, Kommata, Doppelpunkten, Anführungszeichen und Semikola.
orth_err="orth_error"	orth_err="05 lcp/cap in combination with sep instead of tog/tog instead of sep"	05 Klein-/Großschreibung bei Getrennt-/Zusammenschreibung	<i>krankenhaus reif</i> statt <i>krankenhausreif</i> (auf das 1. Element bezogen)	Annotiert wird die fälschliche Groß- bzw. Kleinschreibung bei den ursprünglichen Wortarten eines Kompositums, wenn es fälschlicherweise nicht als

				solches erkannt wurde.
orth_err="orth_error"	orth_err="06 sep instead of tog: compounds"	06 Getrennt- statt Zusammenschreibung: Komposita	<i>Disko Besuche</i> statt <i>Diskobesuche</i> oder <i>Diskobesuche</i>	Annotiert wird die fälschliche Getrenntschreibung von Nomen-, Verb- und Adjektiv-Komposita.
orth_err="orth_error"	orth_err="07 sep instead of tog: other cases"	07 Getrennt- statt Zusammenschreibung: anderes	<ul style="list-style-type: none"> • <i>das selbe</i> statt <i>dasselbe</i>, <i>in sofern</i> statt <i>insofern</i> <i>tag ein</i> statt <i>tagein</i> • <i>12 jährige</i> statt <i>12-jährige</i> 	Annotiert wird die fälschliche Getrenntschreibung von <ul style="list-style-type: none"> • Pronomen, Adverbien, Konjunktionen, Partikel, Präpositionen, • auch fehlende Bindestriche, die eine Zusammenschreibung anzeigen.
orth_err="orth_error"	orth_err="08 tog instead of sep: minimal phraseologism"	08 Zusammen- statt Getrenntschreibung: minimale Phraseologismen	<i>desweiteren</i> statt <i>des Weiteren</i> , <i>vor allem</i> statt <i>vor allem</i>	Annotiert wird die fälschliche Zusammenschreibung von minimalen Phraseologismen (mit eigenem Lexikoneintrag).
orth_err="orth_error"	orth_err="09 tog instead of sep: other cases"	09 Zusammen- statt Getrenntschreibung: anderes	<i>Sogut</i> statt <i>so gut</i> <i>nunmal</i> statt <i>nun mal</i>	Annotiert wird die fälschliche Zusammenschreibung von zwei Wörtern, die keine minimalen Phraseologismen darstellen.
orth_err="orth_error"	orth_err="10 missing or unnecessary hyphen"	10 Fehlende oder überflüssige Bindestriche	<i>süß-sauer</i> statt <i>süßsauer</i>	Annotiert werden überflüssige oder fehlende Bindestriche innerhalb eines Kompositums
orth_err="orth_error"	orth_err="11 omissions: of double consonants"	11 Fehlen eines Graphems: bei Doppelkonsonant (außer an Silbengrenze)	<i>das</i> statt <i>dass</i> , <i>den</i> statt <i>denn</i> , <i>kan</i> statt <i>kann</i>	Annotiert werden fehlende Doppelkonsonanten, davon ausgeschlossen sind Doppelkonsonanten, die eine Silbengrenze darstellen (vgl. Kategorie 14).
orth_err="orth_error"	orth_err="12 omissions: of a vowel: (i) instead of (ie)"	12 Fehlen eines Graphems: Vokal: <i> statt <ie>	<i>wider</i> statt <i>wieder</i> , <i>widerum</i> statt <i>wiederum</i> , <i>sowiso</i> statt <i>sowieso</i>	Annotiert wird die falsche Verwendung von <i> statt <ie>.
orth_err="orth_error"	orth_err="13 omissions: of graphemecluster"	13 Fehlen von zwei oder mehr aufeinanderfolgenden Graphemen	<i>Juglicher</i> statt <i>Jugendlicher</i>	Annotiert werden Wörter, in denen zwei oder mehr Grapheme in direkter Reihenfolge fehlen.
orth_err="orth_error"	orth_err="14 omissions: of one grapheme"	14 Fehlen eines Graphems (auch Doppelkonsonant an Silbengrenze)	<ul style="list-style-type: none"> • <i>ziehn</i> statt <i>ziehen</i> • <i>Enttäuschung</i> statt <i>Enttäuschung</i>, <i>Motorad</i> statt 	Annotiert werden alle Wörter, <ul style="list-style-type: none"> • in denen ein einziges Graphem fehlt, • auch bei

			<i>Motorrad</i>	Doppelkonsonanten, die eine Silbengrenze darstellen.
orth_err="orth_error"	orth_err="15 insertions: of double consonants"	15 Einfügen eines Graphems: Doppelkonsonant	<i>dass</i> statt <i>das</i> , <i>hatt</i> statt <i>hat</i>	Annotiert werden alle Wörter, bei denen fälschlicherweise Doppelkonsonanten verwendet werden.
orth_err="orth_error"	orth_err="16 insertions: of a vowel: (ie) instead of (i)"	16 Einfügen eines Graphems: Vokal <ie> statt <i>	<i>wieder</i> statt <i>wider</i>	Annotiert werden alle Wörter, in denen <ie> statt <i> verwendet wird.
orth_err="orth_error"	orth_err="17 insertions: other cases"	17 Einfügen von ein oder mehreren aufeinanderfolgenden Graphemen: anderes	<i>stürzt</i> statt <i>stürzt</i> , <i>zerstöhren</i> statt <i>zerstören</i>	Annotiert werden Wörter, in denen fälschlicherweise Grapheme, auch mehreren, eingefügt wurden.
orth_err="orth_error"	orth_err="18 transpositions: of vowels"	18 Verwechseln/Vertauschen: Vokal: anderes	<i>absolut</i> statt <i>absolut</i>	Annotiert werden Wörter in denen fälscherweise ein Vokal mit einem anderen vertauscht/verwechselt wurde.
orth_err="orth_error"	orth_err="19 transpositions: of (e/ä) / (ä/e)"	19 Vertauschen: Vokal: <e/ä> / <ä/e>	<i>prädigen</i> statt <i>predigen</i> , <i>Souverenität</i> statt <i>Souveränität</i>	Annotiert werden Wörter, in denen <ä> und <e> vertauscht/verwechselt wurden.
orth_err="orth_error"	orth_err="20 transpositions: of (i/e) / (e/i)"	20 Vertauschen: Vokal: <i/e> / <e/i>	<i>priveligiert</i> statt <i>privilegiert</i>	Annotiert werden alle Wörter, in denen <i> und <e> vertauscht/verwechselt wurden.
orth_err="orth_error"	orth_err="21 transpositions: of vowel and umlaut"	21 Vertauschen: Vokal und Umlaut	<i>wachst</i> statt <i>wächst</i>	Annotiert werden Wörter, in denen das der Umlaut statt des Grundvokals und vice versa steht.
orth_err="orth_error"	orth_err="22 transpositions: of consonants: Fortis/Lenis"	22 Vertauschen: Konsonant: Fortis/Lenis	<i>bedeuted</i> statt <i>bedeutet</i> <i>hüpsch</i> statt <i>hübsch</i>	Annotiert werden alle Wörter, in denen <d/t>, <p/b> <g/k> vertauscht/verwechselt wurden.
orth_err="orth_error"	orth_err="23 transpositions: of consonants: other cases"	23 Vertauschen: Konsonant: anderes	<i>weis</i> statt <i>weiß</i> <i>provieren</i> statt <i>probieren</i>	Annotiert werden alle Wörter, in denen andere Konsonanten, z.B. <ß/s>, <ph/f>, <c/k>, <g/ch>, <v/f/w>, <v/b> vertauscht/verwechselt wurden.
orth_err="orth_error"	orth_err="24 transpositions: of (ss)	24 Vertauschen: <ss> statt <ß>	<i>weiss</i> statt <i>weiß</i>	Annotiert werden alle Wörter, in denen <ss> und <ß>

	and (ß)"			vertauscht/verwechselt wurden.
orth_err="orth_error"	orth_err="25 incorrect positioning of two graphemes"	25 Falsche Position eines Graphems innerhalb einer Silbe	<i>Tehma</i> statt <i>Thema</i>	Annotiert werden alle Wörter, in denen ein Graphem an die falsche Position innerhalb einer Silbe verschoben wurde.
orth_err="orth_error"	orth_err="26 missing or incorrect use of apostrophes"	26 Fehlender oder falscher Apostroph	<i>Disko's</i> statt <i>Diskos</i>	Annotiert werden alle Wörter, in denen fälschlicherweise Apostrophe verwendet oder ausgelassen werden.
orth_err="orth_error"	orth_err="27 abbreviations"	27 Fehler bei Abkürzung	<i>SmS</i> statt <i>SMS</i> , <i>bzw</i> statt <i>bzw.</i> , <i>e.t.c.</i> statt <i>etc.</i>	Annotiert werden alle Fehler, die Abkürzungen betreffen, dazu gehört auch die Groß- und Kleinschreibung innerhalb von Abkürzungen
orth_err="orth_error"	orth_err="28 Eigennamen"	28 Fehler bei Eigennamen	<i>Enzenberg</i> , <i>Erzberger</i> , <i>Enzbacher</i> statt <i>Enzensberger</i>	Annotiert werden alle Fehler bei der Verwendung von Eigennamen.
orth_err="orth_error"	orth_err="Unknown"	anderes	<i>Souverinät</i> statt <i>Souveränität</i> , <i>Schangsen</i> statt <i>Chancen</i>	Annotiert werden Fehler, die nicht eindeutig einer der oben genannten Kategorien zuweisbar sind oder mehr als zwei Kategorien betreffen würden.