

Appendix

1 Gold standard vs REC-LS substitute choices

In the following table, we present examples where the gold standard substitute choice differs from the REC-LS suggestion. These examples show that the suggestions from the REC-LS system are not only valid alternatives to the gold standard choices, but are also arguably simpler.

<i>Examples where gold standard substitution choice is different from the REC-LS substitution</i>	Gold Standard Substitutions	REC-LS Substitutions
Usually portrayed _{0.80} as being bald , with long whiskers , he is said to be an incarnation of the Southern Polestar .	portrayed → shown	portrayed → presented
The island chain forms part of the Hebrides , separated _{0.87} from the Scottish mainland and from the Inner Hebrides by the stormy waters of the Minch, the Little Minch and the Sea of the Hebrides .	separated → demarcated	separated → split
It is therefore pointless _{0.71} to email our contact addresses asking for permission to reproduce articles or images , even if rules at your company or school or organization mandate that you ask web site operators before copying their content .	pointless → useless	pointless → silly
He is also renowned _{0.90} for his work in mechanics , fluid dynamics , optics , and astronomy .',	renowned → known	renowned → noted

Table 1: Examples from Wikilarge and Wikismall

2 Gold standard vs REC-LS simplification choices

The following table shows examples where the gold standard does not contain simplifications that the REC-LS system performs. These words are identified as complex by our CWI labeller. Each word is additionally annotated with its respective complexity score. It is worth noting that despite the fact that these simplifications are not included in the gold standard, many of them are appropriate and valid.

<i>Examples where word was not simplified in gold standard but was simplified by the REC-LS system:</i>	Original	REC-LS Substitution
When the biological value of nicotinic acid was realized, it was thought appropriate to choose a name to distance it from nicotine , in order to avoid the perception _{0.89} that vitamins or niacin-rich food contains nicotine , or that cigarettes contain vitamins .	perception	feeling
When the properties of nicotinic acid were discovered, it was thought prudent to choose a name to dissociate _{0.94} it from nicotine , in order to avoid the perception that vitamins or niacin-rich food contain nicotine .	dissociate	distance
Sustainability is an issue of major transcendence in the contemporary world, which needs to find new options for economic , social , and cultural development ; seeking an improved utilization of natural resources for the benefit of the planet and its inhabitants _{0.88} .	inhabitants	residents
Sustainability is an issue of major transcendence in the contemporary _{0.90} world, which needs to find new options for economic , social , and cultural development ; seeking an improved utilization of natural resources for the benefit of the planet and its residents .	contemporary	modern
Sustainability is an issue of major transcendence in the modern world, which needs to find new options for economic _{0.70} , and social growth ; seeking an increased utilization of natural resources for the benefit of the planet and its residents .	economic	financial
When a band of captive chimpanzees _{0.96} was transferred from the Netherlands to Scotland 's Edinburgh Zoo in 2010 , they used different grunts to refer to apples .	chimpanzees	monkeys

Table 2: Examples from Wikismall and Newsela

3 DRESS-LS, P&S and REC-LS simplification choices

In the following table, we present examples from all three systems. Both P&S and DRESS-LS sometimes perform simplifications that are not context appropriate, e.g. *butterflies* → *birds/moths*. In such cases, if the REC-LS system is not able to produce a substitute that fits the context, the word will not be simplified. As there is no complex word identification step used with the P&S system, simple words are further simplified, which can lead to meaning distortion and ungrammatical substitutions, e.g. *caught* → *spotted*.

<i>Examples where lexical substitution differs between REC-LS and DRESS-LS</i>	DRESS-LS Substitutions	P&S Substitutions	REC-LS Substitutions
He ends up being whipped to death for not revealing the location of escaped _{0.69} slaves	escaped → dead	escaped → hid	no simplification
There are dragons , eagles , butterflies _{0.72} and ingenious mechanical kites , including one that features a frog seeking to devour a praying mantis .	butterflies → birds	butterflies → moths	no simplification
The lava _{0.69} flow moves in fits and starts, moving for a few days , then pausing for weeks .	lava → ice	lava → volcano	no simplification
Archaeologists digging on the grounds _{0.77} have found objects from the 19th century , such as flatirons and milk bottles .	grounds → zebras	no simplification	grounds → site
The company opened twice as many Canadian outlets _{0.65} as McDonald 's " Wendy 's confirms Tim Hortons IPO by March " , Ottawa Business Journal , December 1 , 2005 , and system-wide sales also equalled those of McDonald 's Canadian operations as of 2002 .	outlets → shops	outlets → stores	outlets → stores
These letters usually coincide with the German districts, in few cases an urban district and the surrounding non-urban district share the same letter _{0.05} code .	no simplification	letter → correspondence	no simplification
Much of the fish and sealife caught up in the nets are thrown _{0.45} back dead .	no simplification	thrown → spotted	no simplification

Table 3: Examples from Newsela, Wikilarge and Wikismall