Controlled language for text simplification

Mike Unwalla

mike@techscribe.co.uk
Do you object to simple text?

Loss of technical accuracy?

- It has been observed that a domestic feline with a propensity to slothfulness and of an inherently obese nature was seated on the mat.
- The lazy fat cat sat on the mat.

Questions:

- Is it necessary for a scientific text to be linguistically complex? (I mean the grammatical structures and the non-technical vocabulary. I do not mean the technical content.)
- Do you have objections to simple text in scientific documents? If yes, why?
Phrasal verbs cause confusion

Phrasal verbs cause problems for non-native readers (Thrush, 2001).

“At moderate fields and temperatures a lattice of triangular symmetry is observed, **crossing over** to a lattice of square symmetry with increasing field and temperature.” (https://arxiv.org/abs/1403.1767)
One word can have different parts of speech

As (conjunction/preposition)

- Push the switch **as** you move the lever. > Conjunction = when/while
- Be careful, **as** the material is poisonous. > Conjunction = because
- Do not use the pipe **as** a support. > Preposition
STE checker example

Correct: If the oil is dirty, clean the filter.
Incorrect: Do not oil the lever.

STE 7, rule 1.6: oil
lubricate
More...

Starting check in English (British)...
The structure of a grammar rule

<rule id="DEMO_POS_OIL" name="Noun approved: oil">

<pattern>

<token>oil</token>
</pattern>

<message>Make sure that '<match no="1"/>' is a noun.</message>

<example type="incorrect"><marker>Oil</marker> the valve.</example>

<example type="correct">Put <marker>oil</marker> on the valve.</example>

</rule>
Ambiguous parts of speech

Ambiguous POS in standard English (the examples are not STE)

- Let's stop getting this **biweekly**. Noun or adverb?
- ... purchased from Ukraine in 2006 and Il-76 **transport** aircraft. Noun in a noun cluster or a verb? In BrE, we can treat company names as collective nouns.
- ... with the former **being** manufactured by Acme Industries and the latter... Noun or present participle? Our world knowledge tells us 'probably a verb'.