by Stephen S Hale

Data Are, Datta Is

We database managers
   are known to be fastidious,
We block data violations
   both obvious and insidious.
So there's nothing that makes us
   feel more slandered
Than questioning our search for
   data standards.

We believe in data standards
   with all our heart.
Without them, our databases
   would fall apart.
In writing copious documentation
   we do not slumber
But we can't agree on *data's* pronunciation
   or number.

Many of us say ta
Pronounce the word as DAY-tuh.
Others think it ought ta
Be pronounced as DAH-tuh.
To others, that don't matta,
They pronounce it as DAT-uh.

Habits are hard to erase,
But standards will never pass
If some of us load a database
While others load a dattabass.

What's more, *data*, the plural of *datum*,
Takes a plural verb
   (other forms, we hate 'em).
Good grammar we do not mar
When we say the data are.
But we flunk our grammar quiz
When we say the data is.

And, in the past tense,
We can always infer
That the data were.
It would make our ears buzz
To hear that the data was.

It gives some of us a righteous frown
To hear *data* used as a collective noun.
In this new singular situation
Data is akin to *information*.
But if *data* has taken on a singular form
Then, for the plural, will *datas* become the
   norm?

If we don't follow Latin rules verbatim
At least, here is an ultimatum:

Please do not use *are* with *datum*!
No doubt it would be foolish to go so far
To wonder whether *metadata* is or are.
And, what is worse (this will really grate
   'em),
Is there such a thing as a *metadatum*?

It's really quite simple, you see.
Subject and verb in number must agree.
Else, we sound like grammatical rubes:
"Standards is going down the tubes."

In scientific circles, if we slip and say
   the data is,
Our status and tenure we will be trimmin'.
Just like using a title other than *Ms*
At a meeting of the National Organization
   for Women.

For language and data we do need standards.
(We can't have Frenchmen speaking
   Chinese to Spaniards.)
And we couldn't enjoy a game of Scrabble
If we were a talking Tower of Babel.

But very strict grammar would be a bummer.
We like to say "Let's *database* those numbers."
By this usage we are not perturbed,
We believe any noun can be verbed.

But I'm beginning to fear,
Although we can design complex databases
   with numerous parameters,
It's becoming painfully clear
That we can't write consistent poetic quatrains
   in iambic pentameters.

English standards ease people communication
And data standards ease database integration.
But let's not pretend we can reach
   data nirvana
When we use whatever English grammar
   we wanna.

Our search for data standards is hopeless,
   you see,
Until on the word *data* itself we can agree.
We will never have data standards
   harmony and bliss
While some say DAY-tuh are
   and others say DAT-uh is.

STEPHEN S HALE *is an ecologist with the Atlantic Ecology Division, US Environmental Protection Agency, in Narragansett, RI.*