Bionomina, **1**: 11–14 (2010)
www.mapress.com/bionomina/

Correspondence

ISSN 1179-7649 (print edition)
BIONOMINA
ISSN 1179-7657 (online edition)

Being obsessive-compulsive about terminology and nomenclature is not a vice, but a virtue

Jens H. KUHN[1] and Victoria WAHL-JENSEN
Integrated Research Facility at Fort Detrick, National Institute of Allergy and Infectious Diseases (NIAID), National Institutes of Health (NIH), B-8200 Research Plaza, Fort Detrick, Frederick, Maryland 21702, USA. <kuhnjens@mail.nih.gov>.

At first glance, discussions among linguists seem utterly absurd to those without linguistic training. Scientific debates on the Chinese languages are a good example. For decades, linguists have been arguing over how many Chinese languages exist. Should Chinese characters be classified as ideographs or pictographs, and are they phonetic or rather syllabic in nature (DeFrancis 1984)? Reading linguistic treatises is reminiscent of reading any specialized scientific article—they are difficult to understand if one is not among the experts. It is only natural that non-experts ask sooner or later whether all of these linguistic discussions are of any practical use, and how much money, often derived from taxes, is spent on them. Of course these discussions are of practical use and money is well spent.

For instance, assume a European or an American has just begun studying a Chinese language. Most likely, he (or she) will realize that one cannot say studying "Chinese". A more specific description is required, such as "Mandarin". This descriptor will most likely confuse the student's peers, as they probably do not know that Cantonese is as different from Mandarin as French is from Spanish, or English from Dutch (and, to be fair, many Chinese may not know the vast differences between the Romance or Germanic languages, either). At the same time, the student's Chinese friends and colleagues know what is meant with "Mandarin". Consequently, whereas communication among one group of people has been made more difficult, communication among another group of people has been simplified.

The student then begins the arduous memorization process for the large number of Chinese characters and, before he realizes it, becomes a member of the linguistic discussion mentioned earlier. He may feel that there is a logical connection among the characters that, once figured out, would make memorization more efficient—but he cannot quite pinpoint the connection and will continue to rely on memorization. He will then have to choose among dictionaries that use different character sorting criteria and will wonder why dictionaries are not organized in some other, better way. Of course, one task of linguists is indeed to achieve a better, systematic and precise description of a given language, such as Mandarin. Even the still imperfect dictionaries sort characters into classes and groups that are inherently logical on a small scale. A student may recognize parts of a new character from already known characters, and the learning progress accelerates. One must conclude that classification and grouping seems to be a very useful endeavor for learning and communicating, even if it is not quite perfect and albeit discussions on improvements continue.

In the life sciences, similar processes are at work. The terminology of subspecialty fields compares somewhat to languages or at least dialects. There are specialized vocabularies composed of scientific terms, specialized rules of engagement, such as writing style and types of oral presentations, and numerous misunderstandings when members of two or more fields suddenly interface—"interdisciplinary" can indeed mean the need for interpreters. Even the evolution of language can be observed within subspecialty fields: the

1. Corresponding author. Phone: +1-301-631-7245. Fax: +1-301-619-5029. The authors performed this work as employees of Tunnell Consulting, Inc., a subcontractor to Battelle Memorial Institute's under its prime contract with NIAID, under Contract No. HHSN272200200016I.

meaning of certain words is different in different fields, which one sometimes could suspect is a way of setting groups apart rather than bringing them together. Newcomers to a scientific subspecialty field are often as confused as German students trying to learn Mandarin or Chinese students trying to learn English. "Linguists" specializing in the language of science, and systems to standardize scientific "vocabulary" and "grammar" are therefore of utmost necessity so that international and interdisciplinary communication can be simplified and comprehension of complex subjects be improved. Collaboration then becomes easier, public debates less confusing, and policy measures more useful. Such systems indeed exist and are called terminologies or nomenclatures.

Many scientists, especially those working in the applied biosciences, actively ignore discussions or publications addressing terminology or nomenclature. The word "nomenclature" alone seems to be repellent. Worse, many laboratory scientists appear to look down on experts who spend time on committees trying to standardize terminology. It is often considered a waste of time and of no relevance to "real" science. Curiously, the same scientists can be offended quickly in peer review or during seminars when their object of affection, say a protein or a bacterium, is not named *correctly*. "Correctly" in this case means what is correct in these scientists' own minds and often pertains to the fact that a certain nomenclature had been common usage for a longer period of time, usually the time since the scientist in question was taught that particular nomenclature during his education. Nomenclature changes seem to be generally acceptable if they are felt to come from within a group considered to be the "experts" on a particular subject, but objectionable if they are proposed by an Other. "Interdisciplinary" unfortunately means that small expert groups frequently have to communicate and come to consensus with Others.

It is an interesting question, for instance, whether the protein "bone marrow stromal antigen 2 (BST-2)" should be continued to be referred to as that, or, as recently suggested, as "tetherin" (Neil *et al*. 2008). Should an old designation prevail to honor the person who discovered the item and named it, and also to avoid confusion? Should designations be altered to reflect new developments and findings, as in the case of BST-2/ tetherin, which recently was found to function as a virus-tethering factor and hence part of the cellular innate immune system (Neil *et al*. 2008)? Perhaps the name change is almost mandatory, because a more functional name leads to less confusion? And what if there is an alternative nomenclature system (CD317)? Should there even be alternative nomenclatures?

Almost all scientists are involved with and use nomenclatures every day. But many do so without consciously knowing it. It seems odd that scientists could oppose rationally conceived changes. Does it matter when virologists state erroneously that the mortality of Machupo virus infection is ≈ 20 %? It does matter, at least to epidemiologists. A mortality of 20 % would ring alarm bells, asking for swift intervention, because according to the typical epidemiological definition of mortality (number of dead per number of *healthy* individuals) this would mean a large number of Machupo virus infections in a given area. But if, instead of mortality, lethality or case-fatality rate was meant (number of dead per number of *sick* individuals) then maybe the disease may be classified as rather insignificant to public health and as "exotic", which Machupo virus infections indeed are. It is possible that scientists confuse the two terms erroneously (not knowing any better), but they could also do it mistakenly (theoretically knowing better and hence bordering on deliberate). If Machupo virus research funding is dependent on Machupo virus case numbers then such a mix-up of terms can have far-reaching consequences.

Not following a standardized nomenclature (or not having one) may lead to some bizarre misunderstandings. Assume the headline of a newspaper in the Democratic Republic of the Congo states "*Ebola hampers industrial development*". Microbiologist may immediately think of another disease outbreak due to filovirids, which most likely caused terrible losses in one of the poorest countries on earth. But the article title did not state "*Ebola* virus *hampers industrial development*". This small difference can be important. "Ebola" is a river (Gonzalez 1995). It may be that common flooding of the river turned large areas into mud; or that the river is in the way of building long transects. Or, of course, "Ebola virus" was indeed meant—and not the river. Only precise language, governed by rules, conveys true meaning. Some scientists tend to forget that they are not only doing work for the greater good (or actually for their own median long-term impact factor, as one former colleague phrased it probably more truthfully and eloquently). Their results

need to be communicated to the scientifically less educated public and also to politicians. Research today is rarely performed anymore for the mere accumulation of knowledge, but for the solution of problems via applications. It is therefore crucial for all stake holders to be able to communicate. The US hit comedy series *Big Bang Theory* nicely demonstrates the difficulties in communication between non-scientists and scientists. Penny, who works as a waitress, wants to, but does not, understand physicists Dr. Leonard Hofstadter and Dr. Sheldon Cooper (Anonymous 2007–2010). Leonard and Sheldon (may) understand each other, but their communication is equivalent to that of a small tribe in a large country speaking a minority language. A minority language may be cohesive for the minority group, but will hardly achieve true productivity and solve problems between this group and the majority (it should go without saying that both parties need to learn *each other's* languages – majority does not automatically have the high ground). *Big Bang Theory* exemplifies this nicely: Leonard is in love with Penny, but just as Penny does not understand his language, Leonard does not understand hers.

It is, theoretically, up to editors and peer reviewers to enforce correct nomenclature in scientific publications. The virologists among them should know that a virus species can only be eradicated by inducing collective human amnesia—because a virus species is a theoretical construct and not a physical entity and can therefore not be eradicated by killing its members (Kuhn & Jahrling 2010). Zoologists may disagree, as their species concept is less clearly defined. Peer reviewers and editors should cringe when experts in subspecialty fields use language that is only used among their few peers but not by the larger community, or when experts openly defy set standards. A sentence such as "*Vesicular stomatitis virus* causes increased serum glutamic oxalic transaminase (sGOT) levels in the livers of cynomolgus macaques without appearance of clinical symptoms" should not be acceptable. First, only virus species are italicized and only viruses can cause disease, but species are never viruses (Kuhn & Jahrling 2010). Second, the virus in question has been officially called vesicular stomatitis Indiana virus since 2005 (Fauquet *et al*. 2005). The enzyme is called aspartate transaminase (Anonymous 2010) and its "level" does not increase, but its concentration. Third, cynomolgus macaques are called crab-eating macaques by zoologists, undoubtedly the authorities when it comes to animals (Wilson & Reeder 2005). Finally, animals cannot have clinical symptoms (as these are subjective descriptions provided by the affected individual) but instead show clinical signs. One can, of course, object to any one of those corrections. There may be reasons to do so. For instance, virologists usually think of what was once called the "Indiana serotype" of "vesicular stomatitis virus (VSV)" when they hear the phrase "vesicular stomatitis virus." Maybe it would be better to make this old name the official virus name and rename vesicular stomatitis New Jersey virus "New Jersey virus" and vesicular stomatitis Alagoas virus "Alagoas virus." But one has to admit that the problem is the recognition that there are at least three viruses known that are unique but appear to be "vesicular stomatitis virus." Ignoring new developments and sticking to old terms just because one has always used them is rarely smart. How many people are left who argue that HIV-1 should still be called "HTLV-III"? It should not be so difficult to research terminology during the drafting of an article. Scientists constantly keep up with the newest scientific developments, so why not with changes in nomenclature as well?

Sure enough, sometimes nomenclature and taxonomy experts go overboard and terminology is introduced that is outright silly. Vesicular stomatitis Indiana virus (VSIV) may be a good example since the virology community has never used that name or that abbreviation. On the other hand, the virology community has ample options to be part of the taxonomic naming and classification process by voting on proposals, participating in study groups, or writing opinion articles. So who is at fault for ensuing chaos? The committee that suggested a name change, although it had rational reasons to do so (in that case the existence of three "VSVs")? The virology community because it is ignoring the change? Or both parties because they obviously do not communicate effectively?

When terms get renamed just for the sake of renaming them then outrage at nomenclature experts is justified. But it is a two-way street. Nomenclature without discourse with the scientific community working in laboratories is useless—but science without nomenclature cannot be performed, either.

We all make numerous errors and this article probably contains its fair share. As scientists, we are supposed to be curious and willing to learn. We should be willing to listen to those who may know better so

we will not make the same mistakes in the future and our proceedings will become more precise. We should welcome improved nomenclature and battle nonsensical terminology, even if that means that we have to adjust our ways a little bit. Otherwise we will still wonder ten years from now how creationists get away with stating (correctly) that "even scientists say evolution is only a theory"! Maybe scientists should communicate better and teach the population at large what the difference is between a hypothesis and a theory, and that "theory" in science is often almost the same as "fact". Maybe a reader of this article stumbled over "≈ 20 %" and wondered why this was not written "~ 20 %"? Maybe a virologist reading this article wondered why "filovirid" was chosen rather than the more common "filovirus"? It is to be hoped that the reaction of those readers is not outright dismissal but rather curiosity and willingness to explore. "Filovirid" may sound strange at first but perhaps this word has its merits, as explained by Vetten & Haenni (2006)? Scientific writing is difficult. What again is the difference between analogous and homologous proteins? Is it really possible to perform a heterologous challenge in a vaccination experiment, or would it not be better to use the word "heterotypic"? What is the difference between an antibody and an immunoglobulin? And is there really a difference between "mistake" and "error" as we alluded to in this article?

The readers of this article probably know many things we authors do not. We will not be offended when we get corrected.

References

Anonymous (2007–2010) *The Big Bang Theory*. TV-Series, New York (CBS Broadcasting).

Anonymous [International Union of Biochemistry and Molecular Biology (IUBMB)] (2010) *EC 2.6 Transferring Nitrogenous Groups.* London (University of London). <http://www.chem.qmul.ac.uk/iubmb/enzyme/EC2/0601a.html#1; last accessed November 23, 2010>.

DeFrancis, J. (1984) *The Chinese language*. Honolulu (University of Hawai'i Press): 1–344.

Fauquet, C. M., Mayo, M. A., Maniloff, J., Desselberger & U., Ball, L. A. (ed.) (2005) *Virus taxonomy. Eighth report of the International Committee on Taxonomy of Viruses.* San Diego (Academic Press): 1–1162.

Gonzalez, J.-P. (1995) Ebola, une rivière tranquille au cœur de l'Afrique. *Cahiers Santé*, **5** (3): 145–146.

Kuhn, J. H. & Jahrling, P. B. (2010) Clarification and guidance on the proper usage of virus and virus species names. *Archives of Virology*, **155** (4): 445–453.

Neil, S. J., Zang, T. & Bieniasz, P. D. (2008) Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature*, **451** (7177): 425–430.

Vetten, H.-J. & Haenni, A.-L. (2006) Taxon-specific suffixes for vernacular names. *Archives of Virolology*, **151** (6): 1249–1250.

Wilson, D. E. & Reeder, D. M. (ed.) (2005) *Mammal species of the world. A taxonomic and geographic reference*. Third edition. Baltimore (Johns Hopkins University Press): 1– 2000.

■ *Accepted by Alain Dubois: 23 November 2010. – Published: 24 December 2010.*