TAXONOMY AND TRANSPARENCY IN INTERNATIONAL PHARMACEUTICAL NOMENCLATURE

Rachel Bryan

The language of medicine, although highly specialised, has a broad usership comprising multiple strata of the population with varying levels of knowledge for multiple purposes. This usership includes general practitioners, consultants, nurses, pharmacists, patients, parents and caregivers. No single person holds a comprehensive knowledge of every area, and so there is great variation in understanding of the terminology and the degree to which its use is specialised. Medication names such as *morphine*, *Benadryl*, *paracetamol* and *adrenaline* surround us in our daily lives, and are an important and under-researched area of terminology.

In antiquity, medications were named after the gods, e.g. morphine after Morpheus, the god of dreams and anandamide after Sanskrit *ananda*, 'bliss, delight' (OED). In the present day, pharmaceutical substances are named within a complex system of nomenclature which is managed by multiple government bodies. As illustrated in Figure 1, a pharmaceutical substance such as salbutamol (an asthma medication) will have three types of name.

Chemical name	(RS)-4-[2-(tert-butylamino)-1-hydroxyethyl]-2-(hydroxymethyl)phenol
	(formula: $C_{13}H_{21}NO_3$)
Generic names	international nonproprietary name: salbutamol
	British Approved Name: salbutamol
	United States Approved Name: albutamol
Proprietary names	Ventolin, Aerolin, Ventorlin, Asthalin, Proventil, ProAir

Figure 1. The pharmacopoeial monograph[1] for salbutamol

[1] A pharmacopoeial monograph is a single document describing the name(s) and chemical formula of a pharmaceutical substance.

- One chemical name, based upon the chemical formula of the substances, indicating the position of hydroxy groups, the length of the carbon chain, and so on. This name is designated by the International Union of Chemistry, and is published multilingually. There are some interesting translation problems in this area, but they are beyond the scope of this chapter.
- At least one brand, or proprietary, name, chosen by the manufacturer that originally created the substance. This name is commercially driven, capitalised and legally bound to not imply any therapeutic benefit. It is typically laconic and euphonious. Once out of patent (up to 20 years in the EU), a substance can be marketed by other companies and so will be assigned more brand names.
- At least one generic, or nonproprietary name. On a global level, it will be assigned an International Nonproprietary Name (INN) by the World Health Organization (WHO), and in each country in which it is approved for use, it will be assigned a national generic name, such as a British Approved Name (BAN) in the UK, or a *Denominazione Comune Italiana* (DCIT) in Italy.

1 International Nonproprietary Names

This chapter presents a qualitative analysis of the International Nonproprietary Name (INN) nomenclature, focusing in particular on the underlying conceptual taxonomy and semantic transparency. INNs will be the focus of this study as they are the most commonly used system of generic names, and their form is used by default in both the UK and the EU with only a few notable exceptions (Aronson 2000). There are over 8,000 INNs currently in use. INNs are designated by the WHO, and formally placed in the public domain to promote consistency of global communications between manufacturers, clinicians, prescribers and patients. The nomenclature is published in six languages (WHO 1997). Given their international status, the name designation process in place must encompass a broad conceptual system and naming guidelines must be robust and stringently applied.

INNs are designated according to a set of guidelines (WHO 1997), which aim to achieve usability (pronounceable, legible, audibly perceptible, comprehensible and memorable), clarity (free from confusion) and taxonomy (showing relationship within the conceptual system). The WHO dictates that pharmacological relationship be shown by using a common ‘stem’, which may be a prefix, infix, suffix, or a ‘freefix’, which

can appear anywhere in the name. A 'stem' in this context is a word part to which a particular pharmacological meaning has been assigned, and which is used to signify the relationship between substances. By using a common stem, substances are placed into pharmacological groups, related by anatomical target, therapeutic action, or chemical composition. The use of stems creates a taxonomic conceptual system for INN, and allows users to exploit this systematicity to increase retention, pronunciation and recognition of the names.

The INNs programme began in 1952, and between 120 and 150 new names are designated each year. They are first created in Latin, and this form is translated into the six official languages of the United Nations: English, French, Spanish, Russian, Chinese and Arabic. The Latin form of the name is used for translation of the nomenclature into other European languages, such as Italian and Portuguese (Mareckova *et al*. 2002).

Morphosemantic analysis of INNs is possible because their meaning is highly compositional, i.e. meaning is derived from the meanings of constituent parts (Deléger *et al*. 2009). In contrast to medical terminology in anatomy and general medicine, INNs are not full neoclassical compounds in that they cannot be parsed into elements directly derived from classical languages. INNs are composed of a random element, normally a prefix, and at least one stem. Stems are formed from three types of component. These types are listed in (1).

(1)
a. abbreviations, such as the sub-stem *-tu-* in *situximab* denoting targeting tumorous tissue, or the stem *-kin* in *ilodecakin* denoting interleukin type substances;
b. acronyms, such as the stem *-mab* in *urtoxazumab* denoting monoclonal antibodies; and
c. elements of chemical nomenclature. These can be seen as adapted neoclassical forms, such as the stem *-fos* (from Latin *phosphorous*) in *clofenvinfos*, denoting phosphorous derivatives.

2 Why is this important?

The World Health Organisation (WHO) cites globalisation, consumerism, growth in free markets, increased cross-border communication and the ubiquity of the Internet as agents of change in medicine and pharmaceuticals, giving rise to new safety concerns. Furthermore, the increasingly global trade in pharmaceuticals and higher levels of regulatory complexity have impelled many intergovernmental

organisations to make efforts towards harmonisation of regulatory activities to ensure consistent efficacy of pharmacovigilance efforts (WHO 2002).

Medication errors make up a high proportion of all patient safety events (Jordan & Kyriacos 2014; Ostini *et al*. 2012), and some result in overdose or adverse drug reactions, and can cause serious harm to patients (Aronson 2009; Runciman *et al*. 2003). Medication incidents in the UK resulted in 50 deaths between October 2011 and September 2012 (Jordan & Kyriacos 2014). It is estimated that medication errors cost the USA between $15 and 28bn each year, and that the USA spent an additional $213bn (8% of total healthcare spend) in 2012 on costs arising from medicines' mismanagement, including medication errors (Aitken & Valkova 2013).

Medication errors may be a result of medicines having names that look alike or sound alike, and are referred to as *LASA errors*. Examples of confused LASA pairs are given in (2).

(2) a. mercaptopurine-mercaptamine. A 9-month-old infant presented with nephropathic cystinosis, and was prescribed mercaptopurine by the GP instead of mercaptamine. After a month on the wrong medication, she developed pancytopenia but ultimately made a full recovery (MHPRA 2010).

b. *hydromorphone-morphine*. An elderly patient was discharged after being administered hydromorphone instead of the prescribed morphine, by a nurse in the Emergency Department. He suffered a fatal respiratory arrest on his way home.

LASA errors are estimated to account for around 25% of all medication errors in the US (Emmerton & Rizk 2012), and occur in all aspects of medications management – during prescribing, dispensing and administration of the medication. LASA errors thus represent a significant threat to patient safety.

The bulk of extant literature on LASA errors focuses on mitigating their occurrence (Emmerton and Rizk 2012; Ghaleb et al 2010, Aronson 2009; Kovacic and Chambers 2011), and very little research has been conducted into linguistic properties of the nomenclature to elucidate properties that may prime the risk of the errors occurring. Profiling of such properties could inform the name formation process and thus prophylactically reduce the risk to patient safety. It is also possible that elucidating external factors contributing to the likelihood of confusion

error (such as high syllabic similarity) will encourage reporting of adverse drug events (ADEs) and near misses, since these may be under-reported due in part to fear of reprisal, blame and reputation damage (Aronson 2009).

More needs to be known about the formal and semantic properties of the main global medication nomenclature of International Nonproprietary Names. This study examines semantic transparency in the nomenclature, and the underlying conceptual taxonomy of pharmacological relationship. In the context of this study, semantic transparency is defined as the correspondence between form and meaning within a lexical unit, and the extent to which meaning motivates form and meaning is derived from form.

3 Medical taxonomies and ontologies

There are many systems of classification in medicine, such as the HUGO (HUman Genome Organisation) gene nomenclature, Medical Subject Headings (MeSH) used to index published research on Medline, and the University of Washington Digital Anatomist (UWDA) (Shapiro *et al*. 2005; Segura-Bedmar *et al*. 2008). Due to the exponential growth of published research in medicine, it is now impossible for specialists to keep abreast of developments in their field, and the need has arisen to automate recognition of key concepts in the literature (Coletti & Bleich 2001, Segura-Bedmar *et al*. 2008). The Unified Medical Language System (UMLS) is an example of an ontology by which automated software can read and assimilate information in published research (Segura-Bedmar *et al*. 2008), and encompasses various nodes such as the UWDA for anatomy. Some systems determine nomenclature, such as the HUGO gene nomenclature, and others are used to assign conceptual relations, such as the UWDA (Shapiro *et al*. 2005). The UWDA uses various semantic links, e.g. the oesophagus is *part-of* the foregut, *continuous-with* the pharynx and stomach, and *adjacent-to* the trachea and thoracic aorta and thoracic vertebral column.

The terms *classification*, *taxonomy* and *ontology* are often used interchangeably to refer to any system of categorisation, but for the purposes of this study, *ontology* is taken to mean any system that categorises concepts (Stevens *et al*. 2000) and a taxonomy should be seen as a methodology for categorisation. There are several key distinctions to be made. An ontology is "the concrete form of a conceptualisation of a community's knowledge of a domain" (2000: 1), whereas a taxonomy does not necessarily include added knowledge beyond the necessary and

sufficient criteria for categorisation. Ontologies may be multidirectional and include multiple types of semantic relation, such as meronymy and metonymy, whereas a taxonomy is an upside down tree structure (Shapiro *et al*. 2005) and is based upon intrinsic properties of its members. Taxonomies are typically 'tree-like' hierarchies, employing hyponymy (*is-a*, class membership) to express semantic relationship. Under classical Jackendovian theory, the organisation of systems will inevitably depend upon our conceptualisation of the world (Jackendoff 1983), but further consideration of that is beyond the scope of this chapter. The prototypical taxonomy is the plant or animal kingdom used in biology (Shapiro *et al*. 2005, Coletti & Bleich 2001).

According to the WHO, the INN system is a 'classification', but can be more specifically defined as a taxonomy since it only employs *is-a*, hyponymic semantic relations. Although there is a global taxonomic system for pharmaceutical substances, the Anatomical Therapeutic Chemical (ATC) index, INNs use a different taxonomy that does not align with the ATC (Segura-Bedmar *et al*. 2008) and is not used by any other organisation. For example, the medication name *selegiline* in the ATC system would be found by drilling down into the taxonomy: Nervous system > Anti-parkinson drugs > Dopaminergic agents > Monoamine oxidase B inhibitors, but in the INN system by Psychopharmacologics > Antidepressants > Monoamine oxidase inhibitors.

The INN system employs at most a four-level taxonomy, and assigns alphanumeric codes to each level. Although there is room for four levels, currently names fill only two levels, so the INN system can be seen as a flat taxonomy, or a collection of individual taxa under an undefined hyperonym. There is sparse information on the taxonomy beyond the statutory guidance of the WHO, and neither definitions nor necessary and sufficient criteria for inclusion in the taxonomy are provided. The INN system is unique in the world of medical ontologies and taxonomies in that the nomenclature it motivates is used by people at all levels of society with all levels of knowledge.

4 A typology of taxa in the INN nomenclature

Pharmacological relationships between substances are demonstrated by the use of a common stem (WHO 1997: 1), which may be a prefix, infix, suffix, or a 'freefix'. By using a common stem, the INN indicates that its denoted substance belongs to a group of substances with similar pharmacological activity (WHO 1997: 1). The common stem or sub-stem is combined with a "random, fantasy prefix", normally chosen by the

submitter of the new substance (WHO 1997: 6), and "the only requirement is to contribute to a euphonious and distinctive name" (WHO 2004: 128). Displaying taxonomy from right to left, starting at the end of the name, is a predictable approach for the user as they can first categorise the name under its stem, and further sub-categorise under sub-stems by reading to the left. The reverse would be impossible due to the meaningless prefix. The INN taxonomy is based upon hyponymy, and in this chapter, *stem* will be used to denote hyperonym, and *sub-stem* to denote hyponym.

This chapter presents a qualitative typology of taxa found in the INN nomenclature, and reviews the implications of these types in the usability of INNs. WHO guidance stipulates that names must not be liable to confusion and that relationship must be shown by the use of a common stem. Therefore, there must be a robust and structured underlying conceptual taxonomy in place to facilitate correct usage of the medication names. The typology that follows is a qualitative analysis of the author's database of monolexic INNs (n=7,111) and the *WHO Stem Book* 2011, which provides information on the INN taxonomy and lists of INNs containing each stem and sub-stem (WHO 2011).

4.1 Single-level taxa

There are many INNs that are regularly formed, some with only a single-level taxon represented by a single stem. This taxon has no hyponyms. Examples are given in Table 1. These stems occur as all four types of affix: infix, freefix, prefix and suffix.

Stem	**Affix type**	**Pharmacology**	**Examples of INNs**
arte-	prefix	antimalarial agents, artemisinin related compounds	arteflene, arterolane
-coxib	suffix	selective cyclo-oxygenase inhibitors	etoricoxib, tilmacoxib
-formin	suffix	antihyperglycaemics, phenformin derivatives	benfosformin, metformin
nab	freefix	cannabinoid receptors agonists	menabitan, nonabine

-pris-	infix	steroidal compounds acting on progesterone receptors	ulipristal, asoprisnil

Table 1: Examples of single-level taxa

These single-level taxa illustrate the longevity of the INN nomenclature: from its inception in 1952, the taxonomy has allowed for developments in pharmacology by creating empty pharmacological taxa. Stems are created, but may not appear in names immediately – the system is proactive rather than reactive. This future-proofing is similar to Dmitri Mendeleev's periodic table, in which gaps were left for elements not yet discovered. It is possible that in future a sub-category of cannabinoid receptors agonists may be discovered, and in that case a sub-stem of *nab* can be created.

4.2 Regular taxa

Many stem taxa clearly display their taxonomy in names that can be interpreted from right to left. The stem is the suffix, and sub-stems are distinguished from their co-hyponyms as infixes directly before the suffix stem. The taxon for "antiasthmatic, antiallergic substances not acting primarily as antihistaminics" has the stem *-ast*, and sub-stems *-lukast*, *-milast*, *-trodast* and *-zolast*. Montelukast is a substance in this group, and its meaning can be easily derived from the order of word parts: the suffix stem *-ast* can be used to categorise the substance as part of the antiasthmatic taxon, and the infix *-luk-* can be used to further sub-categorise it as a leukotriene receptor antagonist.

In regular taxa such as these, morphemic concatenation is ordered as in Table 2.

Random prefix	**Sub-stem (distinguishing part)**	**Stem**	**INN**
andol		ast	andolast
monte	luk	ast	montelukast
teto	mil	ast	tetomilast

sero	trod	ast	serotrodast
qua	zol	ast	quazolast

Table 2: Morphemic concatenation in the *-ast* stem taxon

4.3 Monoclonal antibodies: a complex and regular taxon

Monoclonal antibodies are a relatively new but rapidly growing branch of biochemistry, and their INN taxon is complex but well defined. There is a regular correspondence between names in this group and their meaning, and meaning is predictable for the user. INNs for monoclonal antibodies comprise a random prefix, followed by two infixes and a stem referring in a specified order to:

- the target class or disease class;
- the source class on which the immunoglobulin sequence is based; and
- the hyperonym *-mab*.

Each of the three meaning-bearing elements must be defined in the name and concatenated in a certain order, but they combine freely with each other.

In this complex and regular taxon, morphemic concatenation is ordered as in Table 3.

Ran-dom prefix	**Sub-stem 1 (target of medication)**	**Sub-stem 2 (source of antibody)**	**Stem**	**INN**	**Description**
icru	c	u	mab	icrucumab	targeting cardiovascular system, of human origin
siru	k	u	mab	sirukumab	targeting interleukin, of human origin
ce	tu	xi	mab	cetuximab	targeting tumours, of

					chimeric origin
ur	toxa	zu	mab	urtoxa-zumab	targeting toxin, of humanized origin

Table 3 Morphemic concatenation in the *-mab* stem taxon

The complete *-mab* taxon is shown in Table 4.

sub-stem 1 - target	**sub-stem 2 - source**	**stem**	
		-mab	monoclonal antibody
	-a-		rat
	-axo-		rat/mouse
	-e-		hamster
	-i-		primate
	-o-		mouse
	-u-		human
	-xi-		chimeric
	-xizu-		chimeric/humanised
	-zu-		humanised
-b(a)-			bacterial
-c(i)-			cardiovascular
-f(u)-			fungal
-k(i)-			interleukin
-l(i)-			immunomodulating
-n(e)-			neural

-s(o)-			bone
-tox(a)-			toxin

Table 4 -*mab* stem taxon

There are eight infixes to denote the target class and nine infixes to denote the source class, although not all of these are currently used in INNs. These may be combined freely with each other but the order in which they appear in the word is fixed. As an example, *urtoxazumab*, which refers to a humanised monoclonal antibody directed against a type of toxin-producing Escherichia Coli (E. Coli), can be decomposed as *ur-toxa-zu-mab*, in which *-toxa-* indicates that it targets a toxin, and *-zu-* indicates that it is humanized (derived from a non-human antibody, which has been engineered to be more homologous with human antibodies). This is not an example of a three-tiered taxonomy, but rather two mutually independent parameters of classification under a single hyperonym.

Infixes have two forms depending on the following letter; for example, *-tox(a)-* is realised as *-toxa-* in the above example because it is followed by a consonant. In *actoxumab*, it is realised as *-tox-* because it is followed by the vocalic infix *-u-*. Without optional linking elements, INNs may contain phonemes outside the phonotactics of English (such as **urtoxzumab*) and thus run the risk of being mispronounced and misunderstood. In names for monoclonal antibodies, meaning is conveyed in almost every component of the word, and even in a single letter, thus optimizing the space available and minimizing character redundancy.

4.4 Irregular display of taxonomy

For names within some taxa, there are very clear concatenation rules for stems and sub-stems. For example, in names for monoclonal antibodies the distinguishing part of the hyponym is prefixed to the stem of its hyperonym in the word. However, often the formation of sub-stems for hyponyms of a main stem is not consistent in INNs, given the variation in types of affix used. The taxon for antivirals, with the stem *vir*, will be used as an example. As a freefix, the stem *vir* can appear anywhere in the name, meaning that the user cannot rely on a right-to-left display of taxonomy to derive the meaning. The stem has sub-stems *-amivir*, *-cavir*, *-ciclovir*, *-fovir*, *-gosivir*, *-navir*, *-previr*, *-virine* and *-viroc*.

In irregular taxa such as these, there are two orders of morphemic concatenation, exemplified in (3).

(3) a. alamifovir (random prefix: *alami*; distinguishing part of sub-stem: *-fo-*; stem: *vir*)
b. vicriviroc (random prefix: *vicri*; stem: *vir*; distinguishing part of sub-stem: *-oc*)

The meaning of *alamifovir* in (3a) can easily be derived from the order of morphemic concatenation in the name. However, in *vicriviroc* in (3b), the distinguishing part of the hyponymic sub-stem *-viroc* follows its hyperonym in the name, and thus it is not immediately obvious to which taxon the name belongs. A user may mistakenly categorize the name under the stem **-oc*, which does not exist. This can also be seen in other stem taxa, such as the hyperonym *prost* with hyponym *-prostil*, and in any taxon for which the stem is not a suffix. This problem arises from the phenomenon of freefixes in INNs. As stems such as *vir* and *prost* do not have a set position in the word, semantic transparency may be low since the user has no predictable way of recognizing the main stem.

4.5 Morphosemantics

As is clear from regular stem taxa and the monoclonal antibodies taxon, morphemic order is important for the easy recognition of INNs. When stems are suffixes and sub-stems are infixes, taxonomy is displayed and semantic transparency is high.

In some stem taxa, meaning is motivated solely by morphemic order: sub-stems take the same form as the stem and are distinguishable only by their affix type. The *-fos* stem taxon denoting "insecticides, anthelminthics, pesticides etc, phosphorous derivatives" will be used as an example. When *-fos* is used as a suffix, it is the hyperonym of the taxon. When it appears as an infix or a prefix, it is the hyponym denoting "various pharmacological categories belonging to fos, other than those above" (an insufficient differentiation, but nonetheless confirming its hyponym status). This messy taxon may be mistaken for a single category with a freefix stem, but in fact the position of the stem in the name motivates meaning.

Freefixes disturb semantic transparency in INNs by creating unpredictability in morphemic order, and they cause inconsistency. Users must learn that in some cases meaning depends on the position of the stem in the name, and that in other cases position does not matter.

4.6 Allomorphy

A stem may have several orthographic allomorphs, illustrated by the examples in Table 5. These are variant forms of the stem that do not indicate a change in meaning. Allomorphy in this sense can be dangerous: when pharmacological relationship is differentiated in units as small as a single letter, allomorphs such as *-profen* and *-profene* may lead to confusion in other areas. For example, *-fenin* and *-fenine* are separate stem families denoting "diagnostic aids; (phenylcarbamoyl)methyl iminodiacetic acid derivatives" and "analgesics, glafenine derivatives", respectively, but are only distinguished by a final letter *-e*. Variation in the orthographic form of stems under a common hyperonym can obfuscate meaning, and may also falsely suggest a relationship where there is none.

Official stem	**Examples**	**Allo-morphs**	**Examples**
-azepam (diazepam derivatives)	diazepam, lorazepam	-azam;	arfendazam, clobazam
-cillin (antibiotics, 6-aminopenicillanic acid derivatives)	penicillin, amoxicillin	-cillide; -cillinam	libecillide; bacmecillinam, pivmecillinam
-eridine (analgesics, pethidine derivatives)	morpheridine, properidine	-ethidine	pethidine[2], hydroxy-pethidine
-izine (diphenylmethyl piperazine derivatives)	cetirizine, cyclizine	-yzine	hydroxyzine
-mantadine (adamantane derivatives)	amantadine, somantadine	-mantine; -mantone	memantine, dopamantine; idramantone
-profen (anti-	ibuprofen,	-profene	aprofene,

[2] The example of pethidine highlights the difference between the accepted linguistic use of 'stem' and the use of 'stem' in the context of INNs. In pethidine, -ethidine is the stem, and p- is the 'random prefix' assigned to distinguish it from others in its 'stem family'.

inflammatory agents, ibuprofen derivatives)	tetriprofen		diprofene
-tril (endopeptidase inhibitors)	dexecadotril, candoxatril	-trilat	omepatrilat, sampatrilat
-triptyline (antidepressants, dibenzo[*a*,*d*]cycloheptane or cyclopheptene derivatives)	nortriptyline, butriptyline	-tiline	levoprotiline, oxaprotiline

Table 5: Allomorphy

5 Discussion

As with Saussure's two sides of a piece of paper (Sanders 2004), the formal and semantic aspects of language are inextricably linked: formal aspects of INNs are motivated by their underlying semantics, and although the formal realisation of INNs may be conducive to conveying meaning, it can equally misrepresent meaning and increase the risk of confusion. Although meaning is primarily derived through the taxonomic system of stems and sub-stems, it is also motivated by the class of affix and the ordering within the name of stems and sub-stems.

The pharmaceutical nomenclature and its peripheral systems of nomenclature are large and complex. A taxonomic conceptual system developed over a number of decades will inevitably contain some broken links and general inconsistencies, but these inconsistencies should not work to the detriment of the overall aim of the system. The structure of the INN taxonomy does not conform to the archetypal tree structure, and is at times messy and fractured. Properly structured taxonomies help to bring substantial order to a model, whereas improperly structured taxonomies make models confusing and difficult to use (Guarino and Welty 2000). A robust taxonomy is also important for automated recognition systems, and this will increase the life-span and usage of the INN nomenclature (Segura-Bedmar *et al*. 2008).

Semantic motivation in International Nonproprietary Names is multifaceted, encompassing multiple methods of conveying meaning. The INN nomenclature does not sufficiently exploit formal aspects of language

by using a systematic and linear ordering of stems and sub-stems in the word, thus resulting in pharmacologically related sub-stems that appear to be unrelated, such as *-ciclovir* and *-viroc*. This type of hyponymy defeats the object of a nomenclature to reflect a deep classification of concepts, as the formal realization gives no indication of its paradigmatic relations. Of course, there are many instances of nested concepts that are mirrored in the formal realisation, such as the naming of monoclonal antibodies with the stem *-mab* and two systems of infix sub-stems, and the designation of single stems that do belong to a taxon of related stems does not present too many problems.

This means that there is no single way for a user to predict meaning, and the burden of learning on clinical users (pharmacists, nurses, physicians, medical students) is high. They must understand not only the meanings of stems and the layout of the taxonomy, but also the inconsistencies peculiar to each stem taxon. They need to know that meaning may be motivated at the morphemic level by the class of affix and the morphemic concatenation of stems and sub-stems, and to know when to ignore spelling variation and when to take note of a single letter bearing meaning. It is little wonder that in practice, clinicians rely instead on the gradual learning of whole names and memorize them based upon their spelling and phonology (personal communication with Dr. Sue Jordan, Swansea University). Many stem taxa, such as those for monoclonal antibodies, are complex but regular, and there is a predictable display of taxonomy and correspondence between form and meaning. Semantic transparency in many of the resulting names is low, and marred by inconsistency.

The exact interplay between transparency and similarity is currently unclear, since preliminary observations I made suggest that in some cases, transparency will increase orthographic similarity between names and thus increase the risk of confusion, yet without transparency, users would need to learn thousands of names by rote. Given the risk to patient safety in this rapidly expanding field, the system of international pharmaceutical nomenclature certainly warrants further linguistic investigation.

6 Conclusion

INNs carry out different functions depending on the user, and thus they must be understandable at multiple levels. For patients and non-professionals they must be recognisable and pronounceable, but at the same time, they must ‘whisper’ in the ear of health professionals by communicating meaning through neoclassical compounds and the general

norms of pharmaceutical nomenclature. To facilitate the correct realisation of each INN in all four modalities of language (listening, reading, speaking and writing) and thus to prevent confusion between INNs, it is imperative that formal aspects are controlled and fully optimised.

The WHO has been described as the "locus of efforts to improve global health … which is the foundation for peace and prosperity" (Council for Foreign Relations 2012). Global stewardship is an essential role of the WHO, in how it identifies needs to be met and takes a leadership role in setting global norms (Clark *et al*. 2010). Given the primary objective of the WHO, "the attainment by all peoples of the highest possible level of health", the INN programme must always be viewed from the perspective of patient safety. By creating an international nomenclature that is publicly available worldwide, it acts as a fulcrum between various institutions operating within separate strata of the pharmaceutical industry and inevitably paves the way for more global consistency and communication, ultimately enhancing the safety of patients.

References

Aitken, Murray & Valkova, Silvia (2013), *Avoidable Costs in U.S. Healthcare*, Parsippany (NJ): IMS Institute for Healthcare Informatics.

Aronson, Jeffrey K. (2000), 'Where name and image meet – the argument for adrenaline', *BMJ* 320(7233): 506-509.

Aronson, Jeffrey K. (2009), 'Medication errors: what they are, how they happen, and how to avoid them', *QJM: Monthly Journal of the Association of Physicians* 102(8): 513-521.

Clark, William C.; Szlezak, Nicole A.; Moon, Suerie; Bloom, Barry R.; Keusch, Gerald T.; Michaud, Catherine M.; Jamison, Dean T.; Frenk, Julio & Kilama Wen L. (2010), 'The Global Health System: Institutions in a Time of Transition. Harvard University, Centre for International Development: Working Paper 193', http://www.hks.harvard.edu/var/ezp_site/storage/fckeditor/image/center-programs/cid/publications/193.pdf (viewed 19 Nov. 2014).

Council on Foreign Relations (CFR) (2012), 'Global Governance Monitor: Public Health', http://www.cfr.org/global-governance/global-governance-monitor/p18985?breadcrumb=%2Fthinktank%2Fiigg%2Fpublications#/Public%20Health/Overview%20Video (viewed 1 Oct. 2014).

Coletti, Margaret H. & Bleich, Howard L. (2001), 'Medical Subject Headings Used to Search the Biomedical Literature', *Journal of the American Medical Informatics Association* 8(4): 317-323.

Deléger, Louise; Namer, Fiammetta & Zweigenbaum, Pierre (2009), 'Morphosemantic parsing of medical compound words: transferring a French analyzer to English', *International Journal of Medical Informatics* 78(1): 48-55.

Emmerton, Lynne M. & Rizk, Mariam F. S. (2012), 'Look-alike and sound-alike medicines: risks and "solutions"', *International Journal of Clinical Pharmacy* 34(1): 4-8.

Ghaleb, Maisoon A.; Barber, Nick; Franklin, Bryony D. & Wong, Ian C. K. (2010), 'The incidence and nature of prescribing and medication administration errors in paediatric inpatients', *Archives of Disease in Childhood* 95(2): 113-118.

Guarino, Nicola & Welty, Christopher (2000), 'Ontological Analysis of Taxonomic Relationships', in Juris Hartmanis & van Leeuwen, Jan (eds.), *Lecture notes in computer science*, Germany: Max-Planck Institute of Computer Science, 210-224.

Institute for Safe Medication Practices (ISMP) (2011), 'List of confused drug names', http://www.ismp.org/tools/confuseddrugnames.pdf (viewed 3 Oct. 2014).

Jackendoff, Ray S. (1983), *Semantics and Cognition*, London: MIT.

Jordan, Sue & Kyriacos, Una (2014), 'Medicines' management: a public health problem on nursing's agenda', *Journal of Nursing Management* 22(3): 271-275.

Kovacic, Laurel & Chambers, Carole (2011), 'Look-alike, sound-alike drugs in oncology', *Journal of Oncology Pharmacy Practice* 17(2): 104-118.

Marecková, Elena; Simon, Frantisek & Cerveny, Ladislav (2002), 'Latin as the language of medical terminology: some remarks on its role and prospects', *Swiss Medical Weekly* 23: 581-587.

Medicines and Healthcare Products Regulatory Agency (MHPRA) (2010), 'Mercaptamine and mercaptopurine: confusion between drug names', https://www.gov.uk/drug-safety-update/mercaptamine-and-mercaptopurine-confusion-between-drug-names (viewed 18 Feb. 2015).

Ostini, Remo; Roughead, Elizabeth E.; Kirkpatrick, Carl M. J.; Monteith, Greg R. & Tett, Susan E. (2012), 'Quality Use of Medicines-medication safety issues in naming; look-alike, sound-alike medicine names', *International Journal of Pharmacy Practice*, 20(6): 349-357.

Runciman, William B.; Roughead, Elizabeth E.; Semple, Susan J. & Adams, Robert J. (2003), 'Adverse drug events and medication errors in Australia', *International Journal for Quality in Health Care* 15(90001): 49-59.

Sanders, Carole (2004), *The Cambridge Companion to Saussure*, Cambridge: Cambridge University Press.

Segura-Bedmar, Isabel; Martínez, Paloma & Segura-Bedmar, María (2008), 'Drug name recognition and classification in biomedical texts. A case study outlining approaches underpinning automated systems', *Drug Discovery Today* 13(17-18): 816-823.

Shapiro, Linda G.; Chung, Emily; Detwiler, Landon T.; Mejino, José L. V.; Agoncillo, Augusto V; Brinkley, James F. & Rosse, Cornelius (2005), 'Processes and problems in the formative evaluation of an interface to the Foundational Model of Anatomy knowledge base', *Journal of the American Medical Informatics Association (JAMIA)* 12(1): 35-46.

Stevens, Robert; Goble, Carole A. & Bechhofer, Sean (2000), 'Ontology-based knowledge representation for bioinformatics', *Briefings in bioinformatics* 1(4): 398-414.

Tuohy, Nancy & Paparella, Susan (2005), 'Look-alike and sound-alike drugs: errors just waiting to happen', *Journal of Emergency Nursing* 31(6): 569-571.

World Health Organization (WHO) (1997), 'Guidance on the Use of International Nonproprietary Names (INNs) for Pharmaceutical Substances', http://whqlibdoc.who.int/hq/1997/who_pharm_s_nom_1570.pdf (viewed 21 Oct. 2014).

WHO (2002), 'The importance of pharmacovigilance', http://whqlibdoc.who.int/hq/2002/a75646.pdf (viewed 21 Oct. 2014).

WHO (2004), 'The use of stems in the selection of International Nonproprietary Names (INN) for pharmaceutical substances', http://whqlibdoc.who.int/hq/2004/WHO_EDM_QSM_2004.5.pdf (viewed 27 Oct. 11).

WHO (2011), 'WHO StemBook 2011: The use of stems in the selection of International Nonproprietary Names (INN) for pharmaceutical substances', http://www.who.int/medicines/services/inn/StemBook_2011_Final.pdf (viewed 3 Feb. 2014).