Food Nomenclature, Terminology: Standards and Harmonization for Food Composition Databases and Food Trade

B.A. BURLINGAME
INFOODS Coordinator and Nutrition Program Leader, New Zealand Institute for Crop and Food Research, Palmerston North, New Zealand

Keywords food composition, nomenclature, standards, harmonization

Introduction

Food composition activities are being undertaken by a variety of agencies, programs and people, for an ever-increasing number of reasons, sometimes without a cohesive framework or proper management guidelines. The result is that poor information is produced, rather than valuable information that could be widely utilized. This problem has national, regional and international implications, and it is being dealt with by INFOODS, the International Network of Food Data Systems. It is a joint UNU/FAO project. Among the requirements for solving the problems is the need to establish, quantify and resolve the technical issues. Some of the technical issues are identification of components; sampling, sample handling and sample documentation procedures; methodological issues; data quality issues; and data interchange issues. Fundamental among the technical issues is food identification.

With food identification, the technical harmonization issues have proved difficult. A number of solutions have been recommended over the years. These include systems developed in the 1970's and 1980's, such as the International Feed Information Center system (Harris et al. 1980), Eurocode 2 (Arab et al. 1987), the Factored Food Vocabulary system (McCann et al. 1988), and the CoData system for nutritional epidemiology (Butrum 1985). These proposed solutions relied on words, alphanumeric codes, position-specific facets, etc., and they did offer an approach to standardizing the way foods are identified.

Only two of the proposed food identification systems presented to the international food composition community have gained acceptance: Langual (Pennington and Butrum 1991) and The INFOODS Nomenclature and Terminology System (The INFOODS System) (Truswell et al. 1991). Langual's system was developed to meet the regulatory responsibilities of the US Food and Drug Administration (Pennington and Hendricks 1992). It was a rigid hierarchical food description language, relying on alphanumeric codes, not text, within the facets, and thesauri to provide the translations (Hendricks 1992). This approach appealed to the EUROFOODS regional group, because it was suited for the multiplicity of languages on the European continent. A European Langual Working Group was established in the early 1990's, to communicate needs to the US Langual Committee where the computer programs were housed. Langual has been incorporated into the country databases in France, Denmark, Sweden, and Slovakia.

The INFOODS System is text based and relies on a multifaceted descriptor approach. It is open-ended without standardized terms, allowing national data generators to describe foods in their own words. The system distinguishes between single and mixed (multiple ingredient) foods with different facets. The INFOODS System is used in New Zealand, the South Pacific, several ASEAN countries, at least two African countries and ten Latin American countries. It is also being incorporated into the working systems in Middle Asia and South Asia.

Critical analysis of Langual and The INFOODS system

New Zealand's Food Composition Data System, managed in Advanced Revelation® relational database management systems software (ARev), on a Pentium® file server with multi-user access, was adapted to allow incorporation of Langual and The INFOODS Nomenclature Systems. Langual was evaluated by obtaining coding for the entire finfish section of the New Zealand Food Composition Database. The data file was submitted by electronic mail to the United States Food and Drug Administration (FDA). Langual coding was undertaken by the Technical Information Specialist of the FDA's Center for Food Safety and Applied Nutrition. The coding was prepared, printed and mailed back to New Zealand. A field was created in ARev to allow entry of this information.

The INFOODS System was applied to the entire food composition database by modifying the multifaceted naming system that had been in use since 1988. The criteria for evaluation of the two food identification systems were constructed from informal and formal discussions and focus group sessions at regional and international meetings and subsequently confirmed and expanded during meetings of the New Zealand Food Composition Program team. These criteria included the technical usage of the systems by New Zealand's database systems' analysts and data compilers, the usage by food composition users in the health sector and food industry within New Zealand, and through international data interchange. Assessment of the systems with several different national databases, and with regional databases, provided the opportunity for differentiation of both the national and regional applicability of the two systems.

Survey of regional data center coordinators

Based on these experiences, a questionnaire was constructed that would test the applicability of the New Zealand results, or determine if a broader set of criteria was important. Regional data center coordinators worldwide were selected as the most appropriate test group because of their experiences managing food composition programs at both the national and regional levels. For ranking national importance, regional data center coordinators were asked to base their responses on their experience managing their countries' food composition programs, including their knowledge of the personnel involved in compiling data, and their understanding of the needs of their national data users.

For ranking regional importance, they were asked to base their responses on either their actual experience of managing the compilation of a regional database, or the future likelihood that they would receive data files from countries in their region for the purpose of compiling a regional database.

Critical analysis of Langual and The INFOODS system

In the assessment of the suitability of the systems for endorsement as the international standard, eight criteria emerged as relevant. These are shown in Table 1. They cover language and culture (items 1 and 2), the labor and skill required for learning and maintaining the system (items 3 and 4), usefulness for local users of food composition tables and databases (items 5 and 6), adequacy for visual documentation, such as colors (item 7), and their benefits and/or requirements for international trade of foods (item 8).

Table 1: Ratings of two international food identification systems, based on assessment in Crop and Food Research with local data and overseas databases.

Criteria	Langual	INFOODS System
1. language independence	good	poor
2. culture independence	good	poor
3. compiler-friendly	very poor	very good
4. ease of maintenance	very poor	very good
5. local usefulness in food composition tables	very poor	very good
6. usefulness in dietary assessment software packages	very poor	very good
7. ability to visually document foods (e.g., colour, packaging, bar codes, etc.)	poor	poor
8. usefulness in international food trade/regulation	potentially good	neutral

Results show that Langual and The INFOODS System were different when viewed on the basis of local development and use, and on the basis of international interchange of food composition data. Langual scored better in relation to addressing issues of barriers of language and culture. The INFOODS System scored better in relation to 'friendliness' to data compilers and local usefulness. However, some fundamental limitations are common to both systems. Both systems of food identification failed to adequately address requirements for international applicability, such as adequately identifying colors of foods, different cultivars of the same species, label details on manufactured foods (e.g., bar codes) and other features that cannot be adequately differentiated with text or code descriptors. Neither system has played a role in international food trade.

Language independence

Langual scored better in addressing barriers of language when assessed with data sets in English, Spanish, and Thai. That is, for common foods, there are Langual codes that can be computer-read and converted using a faceted thesaurus (part of the 1997 Langual package), allowing the 'foreign' national language to be ignored. However, in assigning Langual codes to data files, it was apparent that the system was indirectly highly language-dependent. Difficulties became evident when coding for foods where descriptors that had not been assigned Langual codes, and where there seemed to be subtle differences between Langual facets. Raw, for example, was not among the facets that could be chosen, and many difficulties arose when trying to differentiate between 'product type' and 'food source', 'extent of heat treatment' and 'preservation method'. A second and potentially more serious problem was the human element of using the local knowledge and language in coding. The ambiguities inherent with the English language resulted in mistakes relating to inconsistencies in terminology applied to foods. A Langual coder must understand a certain form of American English in order to properly assign codes.

European countries have adopted Langual because it serves the fundamental function of language independence, notwithstanding the other difficulties it presents. The lack of rigidity with The INFOODS System has been viewed as a liability of the system, rather than an asset, by most in Europe (Schlotke 1996). The INFOODS System was viewed as highly language-dependent in the files examined. The Thai descriptor files had both English language and Thai character set, the Chilean descriptor files had only Spanish language, and New Zealand had English with some Maori food descriptors as alternative names. In making international comparisons, translations and interpretations were required in order to match equivalent foods and processes. The matching required people familiar with the data sets and food supply for each country to explain the foods, the processing and preparation, describe the edible portion, and then determine if there were equivalent foods.

An interesting dilemma emerged in this evaluation of language independence: the rating of 'poor' at the national level was actually a desirable attribute. For example, New Zealand is (mostly) a monolingual country, therefore a system that is independent of language, as the Langual system is, is not especially useful for local data users. English is spoken by all, so a system relying on alphanumeric codes that require translation is less appropriate than a system that is text-based, like The INFOODS System. Nevertheless, when interchange takes place with other countries, China for example, the English language descriptors are not useful and alphanumeric codes have more value.

Cultural independence

All the food composition data systems evaluated to date rely on grouping, or classification of food. Food source or type (e.g., fruits, vegetables, meat) generally groups foods. Most food composition databases have between ten and 25 food groups. Even though the concept of food grouping is internationally accepted, classification of has been shown to be highly culturally dependent and most national databases have unique examples. The Pacific Islands food composition tables have coconut products as a group because of the economic and cultural importance of this food. Other countries divide coconut products into several food categories such as fats and oils for coconut oil; nuts and seeds for coconut flesh; beverages for coconut water. Also unique to the Pacific Island tables

is the category of Wild Animal Foods. The Central America and Panama (INCAP) database has three groups that are unique: Bananas, Maize, and Cornbreads. The Thai food composition database has Edible Insects as a group. Langual requires standardization on food groupings as the first facet, 'product type', while The INFOODS System does not. Hence, the defined Langual groupings effectively dismiss many food-culture relationships in the attempt to tightly manage food identification.

Compiler-friendly and ease of maintenance

The criteria of 'compiler-friendly' and 'ease of maintenance' relate to the labor and skill involved in learning and maintaining the system. This is of great concern to managers of food composition programs where staff expertise and turnover rate can compromise the functioning of the program when the computer systems are difficult to learn and operate.

Langual was rated as compiler-unfriendly. Difficulties experienced included long lead time in developing familiarity with the coding system and codes, delays in coding when local terms were not identical to Langual terms, and the inability to code when descriptors had not been assigned Langual codes. Further difficulties related to the base language of Langual being American English.

The INFOODS System was considered highly 'compiler-friendly'. The INFOODS System allows food descriptors to be used with the nomenclature and terminology of the compiler; it is only the faceted arrangement of the descriptors that imposes a structure for the terms. Because New Zealand's original system was a multifaceted, field specific structure, the integration of The INFOODS System was a simple exercise.

Maintenance of Langual codes in a database is time consuming. Often, assigning codes requires developing a consensus within the small group of experienced people participating on the international Langual Steering Committee. Exchange of information presents a number of problems; there were many foods and several processes that can not be matched with Langual codes. However, for international standards, this is a necessary procedure and one that is used in The INFOODS food component nomenclature (Tagname) system (Klensin 1992).

The results of the first round of a European trial with Langual (Deary 1993) showed that many food and nutrition professionals found the system difficult to use. Coding correctly was termed the 'hit rate'. The worst hit rate for all foods was less than 40% for the facet 'treatment applied' and the best was just over 80%. Deary reported improvements with learning, but the hit rate for 'treatment applied' was fewer than 50% in a second round. Several recommendations from the coders were reported: the need for further clarification of some facets; the need to improve completeness with regard to ability to further describe, or discriminate between food characteristics of interest; the need for the system to evolve through a central committee to manage the model; the need to review the entire vocabulary and eliminate ambiguity.

Local usefulness in food composition tables and dietary assessment software packages

Langual was assessed as having little usefulness in food composition tables and dietary assessment software products for conventional users. The INFOODS System was assessed as being useful and appropriate for food tables and software packages. Nevertheless, a problem was identified in both systems: a facet that incorporated the most significant information for identifying a food with a fixed character length for use in printed tables and computer products where long text fields would be unsuitable. The 'short name' facet was created with a 32 character limit and it is the single most useful facet in the New Zealand system used in the body of the Concise New Zealand Food Composition Tables (Burlingame et al. 1997), in Diet1/NZ (Xyris Software 1990-1997), and most commonly used in FOODfiles (Hapanyengwi and Burlingame 1995). When a short name facet is not created in a database system, dietary assessment software packages often truncate the name (Xyris 1990-1997), and extensive manual typesetting is required for printed food composition tables (as opposed to preparing a camera-ready report straight from the databases) (Burlingame 1996).

In 1993 when Deary conducted his study, the Langual coding procedure was 'paper-based'. All candidates reported that this made Langual unacceptable, and that 'computer aided tools need to be developed to support Langual coders.' Some of those tools have recently been developed (Schlotke 1996), and these could improve the local usefulness of Langual.

Usefulness in food trade and food regulation

The name of a food carries significant information related to national, regional and international regulations. Often, what a food is and therefore what it can be named, is specified in food standards of individual countries, and regions with trading bloc agreements. When, where, and how much can be traded is often dependent on what the food is. Improperly named foods available for sale in the domestic market, or food names that do not comply with food standards when traded internationally, can result in legal action, loss of markets or market share, and other tangible and quantifiable problems (The Press On-Line 1997).

Although Langual was developed for regulatory use by the United States Food and Drug Administration, to date it has never been used in the US for regulatory purposes (Chatfield 1995). As of early 1997, the USA has declined to participate further in its management, updating, and maintenance, leaving that job entirely to the European Langual Committee. And, in spite of being widely used in Europe by the food composition professional community, Langual has no regulatory role in the European Union.

Adequate for visual documentation

Neither the INFOODS System nor Langual was adequate for color identification, cultivar differentiation when the cultivar was not named, or several other criteria requiring visual documentation because these characteristics are difficult to explain with precise words. Dissatisfaction among data users and data base compilers in this area led to successful trialing, and subsequent

inclusion of images in food composition databases as a routine documentation procedure (Burlingame et al. 1995).

Survey of regional data center coordinators

Regional data center coordinators from ten countries were surveyed to determine what criteria are most important in designing or selecting a system for identifying foods in food composition data programs. AFROFOODS (Zimbabwe), ASEANFOODS (Thailand), CEECFOODS (Slovakia), LATINFOODS (Chile), NORAMFOODS (United States), SAARCFOODS (Pakistan), and OCEANIAFOODS (Fiji) replied.

The most important criterion for national food composition programs was usefulness in international food trade with six of the seven respondents ranking this criterion as absolutely essential. Compiler-friendliness and ease of maintenance both scored highly, as did usefulness in food composition tables and databases and usefulness in diet assessment packages. Least important was language independence. When compared to the assessment criteria in Table 2, they demonstrate that The INFOODS System satisfies the national requirements better than Langual.

The most important criteria for regional purposes were compiler-friendliness and ease of maintenance. Three criteria tied for last place: language independence, culture independence and visual information. The range of scores was tighter for regional than national. The results for regional food composition programs do not demonstrate clearly that Langual or The INFOODS System would be the preferred system.

Conclusions and recommendations

In spite of the recognized need for food identification harmonization, and the years of effort from many people and agencies, no system is adequate for adoption as the international standard. Images can not be used independently of other descriptors or coding systems. Furthermore, it is not possible to create an acceptable, international system that solves all the current problems that exist in food identification. Nevertheless, there is no inherent incompatibility between the systems and text-based, code-based, and image-based systems can all be used in the same database.

For national usage, the INFOODS System seems effective and useful, and meets most requirements. The principles and philosophy behind Langual are sound but its implementation is difficult. To solve the remaining problems, an expert committee on food nomenclature, terminology and descriptors, should be convened. The tasks for this committee should be the following:

- examine food descriptor files from databases around the world, and identify common and unique features;
- prepare an update, as a continuation of the development of The INFOODS system, previously published in the Journal of Food Composition and Analysis (Truswell et al. 1991);
- link the system(s) to food standards, such as the Codex Alimentarius, and to widely used systems such as E-numbers for additive identification;
- assume responsibility for the compilation of an electronic international food description dictionary-thesaurus-concordance, utilizing the existing Langual thesauri, and possibly including food images.

A demand for a system to use for regulatory purposes in food trade, internationally or in some economically strategic regions (like Europe and the USA), linked to an efficient, practical system for food identification in food composition databases, would make a strong case for the creation of a minimum set of standards or a harmonized approach for describing and identifying foods worldwide.

References

Arab, L., Wittler, M. & Schetiler, G. (1987) Eurocode 2. In European Food Tables in Translation. Springer, Heidelberg.

Burlingame, B.A., Cook, F.M., Duxfield, G.M. & Milligan, G.C. (1995) Food data: Numbers, words and images. In: Quality and Accessibility of Food-Related Data, 1st ed (Greenfield, H., ed.), vol. 1, pp. 175-182. AOAC International, Arlington, Virginia.

Burlingame, B.A., Milligan, Spriggs, T.W. & Athar, N. (1997) The Concise New Zealand Food Composition Tables 3rd edition. Institute for Crop and Food Research Limited, Palmerston North: New Zealand.

Burlingame, B.A. (1996) Development of food composition data base management systems: the New Zealand experience. Food Chemistry 57: 127-131.

Butrum, R.R. (1985) Considerations in designing a system for description of foods and flexible retrieval of relevant data. Proceedings of a CoDATA Meeting on the Role of Data in Scientific Progress. Elsevier Science, Amsterdam.

Chatfield, M. (1995) Personal communication.

Deary, J. (1993) Report of Langual Coding Experiment. FLAIR Concerted Action Program N°12: Eurofoods-Enfant Project. Improvement of the Quality and Compatibility of Food Consumption and Food Composition Data in Europe. MAFF, London.

Hapanyengwi, G. & Burlingame, B.A. (1995) Survey of FOODfiles users. Unpublished data.

Harris, L.E., Haendler, H., Riviere, R. & Rechaussat, L. (1980) International Feed Databank System:: An Introduction into the System with Instructions for Describing Feeds and Recording Data. INFIC Publication No. 2. International Feedstuffs Institute, Utah State University, Logan, UT.

Hendricks, T.C. (1992) Langual: an automated method for describing, capturing and retrieving data about foods. In: International Food Data Bases and Information Exchange, World Review of Nutrition and Dietetics (Simopoulos, A.P. & Butrum R.R. eds.), vol. 68, pp. 94-103. Karger, Basel.

Klensin, J.C. (1992) INFOODS Food Composition Data Interchange Handbook. The United Nations University, Tokyo.

McCann, A., Pennington, J.A., Smith, E.C., Holden, J.M., Soergel, D. & Wiley, R.C. (1988) FDA's Factored Food Vocabulary for Food Product Description, J Am. Diet. Assoc. 3: 336-341.

Pennington, J.A.T. & Butrum, R.R (1991) Food descriptions using taxonomy and the Langual system. Trends in Food Science and Technology 2: 285-288.

Pennington, J.A.T. & Hendricks, T.C. (1992) Proposal for an international interface standard for food databases. Fd. Ad. Cont. 3: 265-275.

Schlotke, F. (1996) Langual - harmonization of different national versions. Report of the European Langual Technical Committee, Institute of Scientific Computing, CH-8092, Zurich.

The Press On-line (1997, April 24) Butter on trial.

Truswell, A.S., Bateson, D.J., Madafiglio, K.C., Pennington, J.A.T., Rand, W.M. & Klensin, J.C. (1991) INFOODS Guidelines for Describing Foods: a systematic approach to describing foods to facilitate international exchange of food composition data. J. Fd. Comp. Anal. 4: 18-38.

Xyris Software (1990-1997) DietI/NZ, versions 1-5. Brisbane, Australia.