

Abstract LanguaL: Controlled Vocabulary for Indexing Dietary Supplements in US Databases

#28

Johanna Dwyer¹, Leila Saldanha¹, Joanne Holden², Janet Roseland², Jayne Ireland³, Anders Møller³, Karen Andrews², Regan Bailey¹, Joseph Betz¹, Jaime Gahche⁴, Constance Hardy⁵, and Susan Pilch⁶

¹Office of Dietary Supplements, NIH; ²Nutrient Data Laboratory, Agricultural Research Service, USDA; ³Danish Food Information; ⁴National Health and Nutrition Examination Survey/Planning Branch, NCHS/CDC; ⁵Center for Food Safety and Applied Nutrition, FDA; ⁶MEDLINE/PubMed Database, NIH Library



INTRODUCTION

Several dietary supplement databases are available in the U.S. and other countries. It is necessary that conventional food¹ and dietary supplement databases use consistent descriptive and classification systems if data are to be used efficiently by multiple stakeholders with different goals. This can be achieved by use of a faceted thesaurus consisting of mutually exclusive terms to describe features of a product.

The concept of using a faceted thesaurus to index foods originated at the U.S. Food and Drug Administration (FDA) in the mid-1970s, to overcome barriers in accessing and exchanging information about food. The LanguaL thesaurus (<http://www.langual.org>) was created to answer the need for a consistent cataloging system.



LanguaL is derived from the Latin words “*langua*” and “*alimentaria*” or “*language of food*”. It is based on the principle that items within a database can be described by a combination of uniform terms chosen from “*facets*” or features characterizing various mutually exclusive attributes of these products. Descriptors are arranged in a hierarchical order, within each facet in the thesaurus from broader to narrower terms, to facilitate indexing and retrieval.

“Scope notes” accompany descriptors to explain when a particular descriptor should be used and ensure uniform use of terms by indexers and searchers. The thesaurus also provides additional information for many descriptors, such as the specific definitions in the U.S. Code of Federal Regulations (CFR).



The LanguaL Dietary Supplement Structured Vocabulary or LanguaL DS Thesaurus is a proposed modification of the food product thesaurus, designed to capture the unique features of dietary supplement products sold in the U.S.

Adaptations of the LanguaL Thesaurus for Use in Dietary Supplement Databases.

(See Table 1) Presently foods (other than dietary supplements) can be described using descriptors chosen from 14 facets. Of these, 12 are applicable to dietary supplements.

The scope and descriptors in each facet were chosen to be consistent with:

- The scope of dietary supplements outlined in the Dietary Supplement Health Education Act of 1994 (DSHEA).
- Relevant regulations governing dietary supplements as published in the CFR, and U.S. FDA guidance and practice in regulating dietary supplements were also considered (21CFR101; primarily 21 CFR 101.36, and 21 CFR 101.4 and 21 CFR 101.9).

For example, only ingredients that meet the definition of dietary supplements according to DSHEA are described as “dietary (supplement) ingredients”. All other ingredients are treated as “non-dietary (supplement) ingredients”.



The table also indicates whether the corresponding information can be obtained from the product label.

The two food facets that were dropped are **F**. Extent of Heat Treatment; and **G**. Cooking Method. The titles of some food facets did not seem fitting for dietary supplements, so modifications to the descriptors were proposed. For example, Facet **H**, Treatment Applied, was provisionally changed to “Ingredients” for the LanguaL DS Thesaurus, as under this facet all ingredients, other than major source, are indexed.



¹Under the Federal Food, Drug and Cosmetic Act and its amendments, food labeling is required for most prepared foods, such as breads, cereals, canned and frozen foods, snacks, desserts, drinks, etc. Nutrition labeling for raw produce (fruits and vegetables) and fish is voluntary. Such products are referred to as “conventional” foods. Dietary supplements are a special category of products that comes under the general umbrella of foods, but which has separate labeling requirements.

Table 1. Proposed LanguaL facets for dietary supplements.

Characteristic	Facet Letter	Food Product Descriptor	Dietary Supplement Descriptor	Information Available From DS Labels Yes/No/Partial
Food Group	A	Product Type	Product Type (as defined by DSHEA)	Yes
Food Origin	B	Food Source	Dietary Ingredient Source (major or primary ingredient that characterizes product).	Yes
	C	Part of Plant or Animal	Part of Source (major or primary ingredient)	Yes
	E	Physical State, Shape or Form	Physical State, Shape or Form	Yes
Processing	F	Extent of Heat Treatment	Not relevant to dietary supplements	
	G	Cooking Method	Not relevant to dietary supplements	
	H	Treatment Applied	Suggest using “Ingredients” as alternate facet name for dietary supplement databases.	Yes
	J	Preservation Method	Preservation Method	Partial
Packaging	K	Packing Medium	Packing Medium	
	M	Container or Wrapping	Outside package container or wrapping (physical container or package, wrapping of dietary supplements).	No
Dietary Uses	N	Food Contact Surface	Dietary Supplement Contact Surface: (Material(s) in direct contact with product.)	No
	P	Consumer Group/ Dietary Use/Label Claim	Label Claims/Consumer Group/Dietary Use	No
Geographic Origin	R	Geographic Places and Regions	Geographic Places & Regions (Place of Manufacture/Origin)	Partial
Miscellaneous Characteristics	Z	Adjunct Characteristics of Food	Adjunct Characteristics of dietary supplements - (Distribution Channels)	Partial

Indexing dietary supplements using the LanguaL DS Thesaurus

Table 2 details the proposed categorization scheme, the narrower (more descriptive) terms, and scope for each facet in the proposed LanguaL DS Thesaurus. The Narrow Terms (NT) are hierarchical. When the LanguaL DS Thesaurus is operational, it is envisioned that there will be pop-up boxes to guide the indexer with the coding. Based on the level of detail available, the indexer would check the appropriate boxes and code the information accordingly.

Every term in LanguaL is assigned an alphanumeric code, which is the indexing system. The letter refers to the facet. Specific terms/attributes that describe a product are assigned a numeric code. For example, the code H0311 refers to Facet **H** (an ingredient other than the primary ingredient) and 0311 is the code for niacin within that facet.

Under Facet **P**, the exact language authorized by U.S. FDA for health claims is included in the Scope Notes so that the indexer will not confuse a health claim with a structure/function claim.

Table 2 also shows how a typical adult multivitamin/mineral (MVM) product would be described using the LanguaL DS Thesaurus. Based on information provided in Supplement Facts panel, the product is described as a MVM (combination) product under Facet **A**. Calcium carbonate is identified as the “primary ingredient”, as it is the major ingredient by weight. Information described under Facets **B** and **C** is the “source” and “part of the source” of calcium carbonate. The “source” in this case is not specified and therefore assumed to be synthetic. All other ingredients captured in the Supplement Facts panel are dietary ingredients and are listed under Facet **H** “Ingredients”. All remaining ingredients listed outside the Supplement Facts panel, e.g., cornstarch, silicon dioxide, hydrogenated palm oil, are listed as non-dietary ingredients under Facet **H**.

CONCLUSION

The LanguaL DS Thesaurus will assist with classifying products and retrieving information about dietary supplements in databases in the U.S. This tool should be valuable to researchers and stakeholders, as it will improve the precision of matching dietary supplement products recorded in food consumption surveys with the composition of these products as indexed in dietary supplement databases. This more precise matching should result in more accurate estimates of nutrient intakes.

The LanguaL DS Thesaurus enables database developers to catalogue and link data and facilitates the sharing of databases within and between countries.

The LanguaL DS Thesaurus will aid in better estimates of nutrient intakes from dietary supplements, because there will be a uniform approach to how dietary supplement product data will be described, recorded, and retrieved.

The LanguaL DS Thesaurus is not intended to and will not resolve all the challenges faced in the development and maintenance of dietary supplement databases; however, it will help address many challenges faced by database developers.

The authors welcome comments and suggestions for improving the indexing system.

Reference

Saldanha LG, Dwyer JT, Holden JM, Ireland JD, Andrews KW, Bailey RL, Gahche JJ, Hardy CJ, Møller A, Pilch SM, and Roseland JM. A structured vocabulary for indexing dietary supplements in databases in the United States. *J Food Composition and Analysis* 2012; 25:226-233.

Table 2. LanguaL Dietary Supplement Structured Vocabulary and with indexing of an adult multivitamin/mineral (MVM) product.

Facet	Broad Term (BT)	Narrow Term (NT)	Narrower Term (NT)	Narrower Term (NT)	MVM Example
A: Product Type	Dietary Supplements (USA ²)	Vitamin			Generic MVM
		Mineral			
		Herbal/botanical	Yeast, algae, fungus		
		Amino acid/protein			
		Other dietary substance to supplement the diet	Electrolytes		
			Fiber		
			Metabolite, constituent, extract, isolate, or combination of any of these.	Hormone precursors; steroid precursors; 7-dehydrocholesterol, lutein, omega-3s, CoQ10	
			Combination of any of the above in “ingredients listed above”	MVM	
				Mineral(s) & botanical	
				MVM & amino acids	
B: Product Source (major or primary ingredient that characterizes product)		Algae, fungus, yeast (lower plant forms).			
		Animal (fish oil, bone meal, thyroid or glandulars, desiccated liver)			
		Chemical (synthetic source)			
		Higher plant (according to botanical genera/species)			
		Bacteria (probiotics).			
		Not identified			
		Extract, concentrate, metabolite, isolate or isolated part of plant or animal			
		Anatomical part of plant or animal	Lower plants: algae part or fungus part	Flower	
			Animal body part (such as gland, etc.)	Leaf	
			Plant part (seed, root, stem, leaf, bark, or whole plant)	Root	
C: Part of Source (major or primary ingredient)			Bark		
			Seed		
			Whole plants		
			Stem & parts		
			Bulb		
			NA		
			Tablet/caplet		
			Chewable		
			Non-chewable		
			NA		
E: Physical state shape or form		Capsule			
		Soft gel			
		Liquid			
		Powder			
		Other forms.	Wafer		
			Lozenge (includes strips)		
			Gel		
			Teabags		
			Gummies (balls, etc.).		
			Encapsulated (coated/covered)		
H: Ingredients (all other ingredients that are not included in Facet B)	Also contains dietary ingredient (as defined by DSHEA).	Vitamin		From LanguaL list of terms	X
		Herbal/botanical		From LanguaL list of terms	X
		Amino acid/protein		Insert list of amino acids	
		Other dietary substance to supplement the diet (fiber, enzymes).		Electrolytes	
				Enzyme	
				Fiber	
		Also contains non-dietary ingredient (everything that is not a dietary ingredient, and listed below Supplement Facts panel)	Antioxidants		
			Preservatives		
			Binding Agent		
			Bulking agents (fillers, excipients)		
	Colors				
	Emulsifiers				
	Flavors				
	Sweeteners				
	Other				
	Coating materials	Vegetarian			
		Non-vegetarian (e.g. gelatin)			
		Hala/Kosher			
J: Preservation Method		Irradiation			
		By chemicals			
		By gases, e.g., steam			
K: Packing Medium		Alcohol			
		Gas, other than air			
		Other			
		No packing medium used			

Facet	Broad Term (BT)	Narrow Term (NT)	Narrower Term (NT)	Narrower Term (NT)	MVM Example		
M: Outside package container or wrapping (physical container or package, wrapping of DS).	Container or wrapping by form	Bottle			X		
		Blister packet					
		Packet/pouch					
		Dropper					
		Envelope					
		Multicontainer package					
		Can					
		Tetrapac (box)					
		Wrapper, material unspecified					
				Glass			
N: Dietary Supplement Contact Surface: Material(s) in contact with product.		Metal					
		Paper or paperboard					
		Plastic					
		Cotton					
		Metal foil					
		P: Label Claims/Consumer Group/Dietary Use	Claims	Nutrient content claims (FDA approved list for nutrient content claims appropriate for supplements)	More of nutrient X		
					Percent of nutrient X (% of DV)		
					Comparative percent: Nutrient X compared to Y		
					Good source of nutrient X (10-19% or more of DV)		
					High/Excellent/Rich in nutrient X (20% or more of DV)		
Health Claims (will use approved FDA language for health claims)	Calcium and osteoporosis						
	Folate and neural tube defects						
	Soluble fiber from certain foods and risk of coronary heart disease						
Qualified Health Claims (will use approved FDA language for QHC)	0.8 mg folic acid and neural tube birth defects						
	Vitamins and vascular disease						
	Selenium and cancer						
	Antioxidant vitamins and cancer						
	Phosphatidylserine and cognitive dysfunction and dementia						
	Omega 3 fatty acids and coronary heart disease						
	Monounsaturated fatty acids form olive oil and coronary heart disease						
	Green tea and cancer						
	Chromium picolinate and diabetes						
	Calcium and colon/rectal cancer						
	Calcium and recurrent colon/rectal polyps						
	Calcium and hypertension.						
	Calcium and pregnancy induced hypertension and preeclampsia						
	Unsaturated fatty acids from canola oil and reduced risk of coronary heart disease						
	Corn oil and corn oil containing products and a reduced risk of heart disease						
	Structure function claims. (Note: unlike health claims, there is no approved FDA language for these claims.)	“General health”					
		“Immune function”					
		“Sports/performance enhancing”					
		“Energy”					
		“Weight loss”					
		“Mental health/brain health”					
		“Occasional insomnia”					
		“Supports mood”					
		“Menopause”					
		“Cardiovascular heart health”					
		“Joint health”					
		“Sexual health”					
		“Bone health”					
		“Gastrointestinal health”					
		“Modulates blood glucose levels”					
		“Hair, skin and nails”					
		“Vision/eye health”					
		Other					
		Structure function claim disclaimer	Included				
			Not Included				
R: Geographic Places & Regions (Place of Manufacture/Origin)	Intended User Group	Statement of material facts					
		One or more ingredients is claimed to be standardized.					
		Four years and above		All-ages			
				Weight loss			
				Body builders/ Performance enhancing			
				Age-specific men			
				Age-specific women			
				Adult	X		
		Infants					
		Children less than 4 years					
Pregnant and lactating women							
Z: Adjunct Characteristics of DS - (Distribution Channels).							

²As defined in Dietary Supplement Health and Education Act of 1994 (DSHEA). Pub L 103-417, 108 STAT. 4235, (Oct. 25, 1994)