

7
东北师范大学图书馆学系讲义

图书文献主题法

(上)

《汉语主题词表》的理论、结构和应用

文甲龙 编

共
三
本

东北师范大学图书馆学系
长春市业余图书馆学院

1980年11月

《汉语主题词表》出版的时候，我正在长春市业余图书馆学院讲授《分类与主题》一课。我怀着异常兴奋的心情迎接词表，并且恭敬地、认真地学习它。学院的同学们要求我介绍它。东北师范大学图书馆学系也有同样的意见。正好当时我在吉林工业大学图书馆讲授《科技文献情报工作方法》一课，也有检索词表的章节。这几方面的需要，促使我加紧学习，赶写了这个东西，作为东北师范大学图书馆学系《图书文献分类法—分类与主题》讲义的补充，也是在上述三单位授课的讲稿。

《汉语主题词表》是新生事物，它自有其理论、结构和应用方法。对于这些，我都是小学生。这里只是记录下了自己的学习心得，而且又是赶写的“急就章”，由讲稿原封搬来而成，错误和缺点定然难免。恳请同志们批评指正。

文田龙 于吉林工业大学图书馆
1980年11月

勘 误 表

页款	行数	错误	正
5	6	语族	正族
14	倒19	最本	最本
15	倒12	主题 右下	主题 左上
18	12	高山滑雪 ↙ 非正式	高山滑雪 ↘ 正式
21	图8	C*	(括号位置应改为) C*

我们经过一个半学期学习了图书文献分类法——图书文献的分类处理方法。

图书文献分类，就是从科学分类为基础，主要依据图书文献内容的科学属性，用逻辑的方法，对图书文献分门别类地、系统化地加以组织。

图书文献经分类后组织起来的系统，基本上是用逻辑的演绎法建立起来的。在这样的系统中，由于把握住了文献的本质属性，再根据属性的共同点和差异点决定系统最基本的构成单元——类的建立。这些类，以科学（学科）分类为基础，按着一定的逻辑规则，纵向深入横向展开地发展，构成了一个系统。这样的—个系统当然反映了文献内容的联系，着重于揭示和组织文献的内容。因此，这个系统从文献内容的学科族性方面集中文献，使图书情报机构大量繁杂的文献资料条理化、系统化，组成了一种便于查考、便于利用、便于研究、便于以恒文献资料的存取系统；为人们检索某一文献提供一种认识上的向导，在建立人—系统—书的最佳联系上不失为一种方法。这便是分类法的主要功绩，或者说是它的主要方面，千百年来为文献的管理、检索、编目创造了良好的条件，并继续发挥着作用。

但是，~~这个系统并不是绝对的~~何是人为的，加上其检索语言的固有特点，~~因此这个系统有其局限性。~~

首先，~~分类法具有滞后性。~~因为分类系统集中文献是以内容的~~族性为基础的~~不能超出系统按文献内容反映的对象集~~合的界限~~之及时性。系统是事先确立的，这就很~~难适应文献中~~出现的新事物、新内容。

再次，缺乏预见性。无论是分类表的编辑者还是使用者，都很难预见到在系统的什么地方会出现或编辑者会在什么地方补充新类目。第四缺乏适应性。系统受着标记的制约，有时类目的扩充便会破坏整个标记体制。第五缺乏直接性。因为检索语言是一种人为的标记，所以不能直接提问。第六，不同系统间不易统一。第七，缺乏机械性。当然，分类法是一种从学科特性检索角度满足用户要求的组织方法，是一种和管理工作密切结合在一起的组织方法。因其优点掩盖其局限性，或因其局限性而否定其优点，都不是唯物主义的态度、辩证的态度。正视其局限性，研究其局限性，一方面可以促进分类法的改进，一方面为建立另外的检索、存取系统提供条件。这便是我们的目的。

第一节 主题法原理

图书情报机构收藏着大量的文献，目的是为了利用人们利用文献，是要利用它的内容，或者说利用它记载的知识和成果。文献是一种知识载体。任何文献都有其研究对象，反映着一定的主题内容。而任何对象都有着一定的相应的概念；概念是反映对象的本质属性的思维形式。科学认识的成果，都是通过形成各种概念来加以总结和概括的。“自然科学的成果是概念。”（列宁）知识的表述必须通过概念来进行。所以一切文献都有概念性。

一切文献，必须在作者—标引提供者—用户间运动，才能起作用。在这个产生—加工—利用的运动全过程中，关于一篇具体的“红宝石激光器”、“园艺拖拉机”或“苏联军”的文献，作者必须使用一定的概念，才能使文献得以产生；标引者必须得运用一定的概念，才能使文献得以报导、传递；用户要求得到它的提问和检索，也必须得使用一定的概念来进行；用户之所以理解文献的情报信息，给他从一定的知识，使他做出某种判断，文献之所以被利用，也要

通过概念来进行。所以说，文献运动的整个过程中，无论处于“产生—加工—利用”的各个阶段，或是对于“作者—标引者—用户”的各人，无一不是概念在运动、在传递、在作出成果，或者说终究是概念在运动、在传递、在作出成果。

当然，概念是通过它的物质外壳语言来表达的。文献表达概念的语言形式是词或词组。也就是说，文献内容的质是知识，而它的外壳是表达概念的词（词组）。文献的内容必须通过词来表达。如果我们能用最能表示内容的词（一个或几个）来组织、检索文献，这样既不失却揭示内容这个中心，又能克服分类法上述的局限性，这就是主题法。这样我们就必须研究一下词的性质，影响着检索的性质。

表达文献内容的词，和标引者、用户使用的词有时是不统一的。因为同一个对象、同一概念涵义，有着不同的词来表达，如关于“地瓜”就有很多表达方法：地瓜、蕃薯、红苕、红薯、山芋、甘薯等。自行车、单车、脚踏车；激光器，光度计、聚光、集光等。也是同样的情况。而有些词则是同形异义词。如：“黑旗军”，就有1861年宋景诗在山南领导起义的黑旗军和刘永福在广西领导的黑旗军之分。“六书”一词的涵义，可从指造字法的象形、假借、会意、形声、指事和转注等六种造字方法，也可指书法上的大篆、小篆、隶书、行书、楷书和草书等六种书法。这些由于历史、地域原因及不同译名，新称旧称，简称繁称，学名俗名等形成的同义词，以及同形异义词在自然语言中大量存在着。而文献产生在先，标引加工在后；加工在先，利用检索在后，所以，在作者、标引者、用户在使用词时，很难做到一致，也是可能和允许对同一对象使用不同的自然语言。但是如果在检索过程中出现这种情况，将会影响到检索质量，势必严重影响检索效果。因此，如果同词在标引、组织、

存储, 检索文献, 就必须对词进行规范化, 该词具有单义性、持指性。一种对象, 一个事物, 一个概念用单一的词(词组)表达。特定的词(词组)表达单一的涵义。就是说, 作为检索标志的词——检索词要求具有单义性。与它同义的自然语言词(组组)只是一种非正式检索词或称引导词(异义词), 要建立它们之间的规范与转换关系。

概念不是永恒的, 词也是有动态性的。词会随着社会历史和人类的发展而变化。新词在生长, 旧词在消亡, 词的涵义在改变。如: 过去只有固态、液态、气态三个词表达物质的状态。随着社会实践、科学技术的不断发展, 供人们对事物的认识日益深化, 发现了物质的第四态, 出现了“等离子态”这个词。又如电子计算机这个事物本身的发展, 随之就出现固态计算机等这样的词。这些, 都势必及时到文献中来。所以检索词是具有动态性的。词的书面形式是文字, 文字有它固有的顺序(音顺或形顺), 而这种自然顺序, 不会因词的增加和消亡而受影响。

概念是有内涵和外延的, 词就有它的涵义和范围(这里指的是实词中的名词, 含动名词)。因此, 词和词之间存在着一定的语义关系, 形成一定的语义层次。当两个概念的内涵和外延完全重合时, 表达它们的词之间的关系便是同一关系。表示为同义词。如: 自行车、脚踏车和单车。如果概念间内涵和外延的关系, 为一个概念的外延被另一个概念的外延所包含, 前者为种(下位)概念, 后者为属(上位)概念。表达它们的词之间的关系便是属种关系, 表示为广义词和狭义词。当然, 概念不但可以划分, 也可以分解, 分成许多部分的形成整体与部分的分解。表达它们的词之间的关系便是一种整体与部分的关系。这种属种关系、整体与部分的关系, 各学科、各产品及各工艺手段之间的包含关系统称为属分关系。如激光器与固体激光器、晶体激光器、红宝石激光器; 生物学与动物学; 中国与江西; 原子与电子; 电子对抗与电子干扰、电子侦察、电子

反干扰、电子反侦察等。此外，概念之间还有一种除同一关系、属分关系以外的某种密切关系，如形式与内容，原因与结果等。表明它们的同便是相关关系。如：资本聚集与资本垄断。同总是从这些同一关系、属分关系和相关关系表明着它们之间的语义层次关系。由这些关系，便会形成一个语族，和一个语义网络。词是表述某一概念的语言形式，而概念是表述某种对象的，某一事物的。一事物与另一事物的不同，最根本的是它们之间表现在概念上的固有属性的差异。所以，表现在表述概念的词的层次关系也就是对象（事物）间的关系。词族，也就反映了对对象属性之间共同属性差异形成的族性系统。这样的系统，决定于对象间相互关系的层次，而不受人为主的系统的影响，而且，系统位置的、关系的标志，不是人为的标记，而是经过规范的、与概念—对象—对应的检索词，直接提问一索即得。这就是说检索词不但有规范性、动态性，还有同族性（语义层次性）。

概念是可以限定和概括的，即能够缩小和扩大。增加某一概念的内涵依该概念变为外延较小的概念的逻辑方法，便是概念的限定；减少其内涵，便变为外延较大，称概念的概括。这就使得表述概念的词具有组配性质。如“喷气式飞机”与“垂直起落飞机”组配，便得出“喷气式垂直起落飞机”。同理：“飞机”+“起落架”=“飞机起落架”；“肾疾病”+“结石”=“肾结石”；“水”+“蒸发”=“水蒸发”；“晶体管”+“可靠性”=“晶体管可靠性”；“球墨铸铁”+“齿轮”=“球墨铸铁齿轮”。“初等教育”+“理化”=“初等教育理化”；“心理学”+“会议录”=“心理学会议录”；“农民起义”+“中国”=“中国农民起义”等等。由于词的这个组配性质，使得能控制检索词的轻重，提高检索词的概念深度，扩大了检索途径。如“西藏”、“石油”、“东北地区”、“中国”，四个词就能组配表述“中国东北地区石油西藏”，控制一司重，如法，提高了“西藏”这一概念的深度。可从八个词的

途径检索，扩大了检索途径，这就是检索词的组配性。

正是由于经过规范化了的、表达对象的词具有上述这些性质，所以我们能得以用文献内容论述的对象来集中文献，再根据与对象一一对应的检索词的层次关系去组织文献，用检索词直接作为检索标志。文献主题处理方法就是，主题法是根据文献内容，从文献论述的对象集中文献，从检索词的自然顺序（音顺或形顺）组织成一个隐含层次、结构关系的、以检索词作为检索标志的文献网络的方法。

主题法和分类法都是从文献内容出发组织文献的方法，在这一点上，它们是有共性的。它们根本的区别在于，主题法建立的从对象集中、从表达了对象的属性的检索词的层次、结构关系的词顺组织起来的网络，这种网络，归根结底是由对象的性质决定的，而分类法建立起来的系统，则是内容的学科属性决定的人为分类系统；主题法使用的检索标志是直接表达对象的检索词一标识，而分类法则使用的是人为的符号一标识，也就是说，一个是所建立的网络与系统不同，一个是使用的检索标志不同。或者说，从组织文献的方法看：主题法按对象集中文献，按检索词的自然顺序排列，充分体现独特性，分类法从表达客观事物学科体系的确切类目集中，按属性进行系统排列，充分体现系统性。从功能范围看：主题法更适应广泛检索的要求，分类法系统满足族性检索的要求；分类法便于图书文献的组织排列，与管理、统计、分析、通报工作结合较密切。从检索手段看：主题法使用的检索标志是经过规范化的自然语汇，标识表征直接反映文献的主题内容，容易掌握、比较直观，是一种直接的检索语言；分类法的标识是人为的标识，标识是类目的表达符，表征不能直接（必须经过“翻译”）反映文献的主题内容，比较间接、不易掌握，是一种间接的检索语言。从适应程度看：主题法不受体系及其标识约束，能比较及时地反映新学科、新技术的新发展，增删灵活、修订方便，检索途径广，利用于机检；分类法受体系及其标识的约束，类目安排比较固定，

需落后于新学科、新技术的发展，增删修订较困难。

总之，主题法最大的特点是专指性，即能集中与研究课题有关而分散在各学科中的文献资料，但它把按事物属性建立起来的学科体系折散了。主题法和分类法都是从检索方法矛盾运动中发生发展起来的。它们都是侧重于揭示文献内容的一种组织方法。由此产生的分类系统和主题网络，都是检索文献的一种工具。从检索角度看，它们是一个事物的两个对立统一的方面。由于客观需要的多样性，必将产生组织方法的多样性，也使主题法和分类法的侧重点相异，带来各自的特殊性。它们都是一分为二的，各有长短；不宜互相排斥，而应互相补充，彼此借鉴，弥补相互不足。在图书馆文献情报工作的实践中正在产生一种有趣的现象，在主题法中包含着分类法的因素，在分类法中蕴含着主题法的因素，两者相互渗透。在分类表建立要旨的时候，尽量列出专指性的类目，在分类法的实践中，越来越多地利用内外组配的方法，参见、交替、分析方法，使图书馆文献更能趋于专指主题集中。而主题词表的建立，本身就是一种特殊的划分，包含着分类的思想。而且主题表中的参照系统以及范畴表、词族表的建立，体现着分类法中的属种、并列关系，引进着分类系统的长处，创造了专指性检索的条件。因此我们认为：主题法和分类法都是组织文献，建立“人—系统—书”联系的方法，都不可偏废。

主题法跟一切事物无不在运动发展一样，它同样也在不断演变中。最先出现的主题法，是一种标题法。那些表示事物概念的主题词——“主标题”和表示事物各个方面的词——“子标题”均予预先固定排列给出。这时，当然组配固定，排列固定，也就查检方便，依循有据，不易混乱。但也因其固定，没有列出的新组配形式（专指概念）得不到及时反映，不能适应新发展。为了克服这一缺点，出现了无词法。这种单元词（无词）是从图书文献内容中抽取出来的最基本、内、不可分割的概念，进行组配而成主题词，也同时是标题。如

“激光”、“电视”、“雷达”，这个词再分割便没有了独立的和明确的检索意义或意义会有很大变化，这个词便称为单元词。单元词可以组配为较为复杂的特指概念，如“激光电视”、“激光雷达”等。这时，当然灵活，但如组配原则掌握不好，不能统一，新产生的特指概念便会含糊不清，从而降低检索效率。此外在主题法中还有一种“关键词法”。主题标引时直接取自图书文献的题目和内容中足以表达其主题内容的词语（关键词）进行标引。这时当然标引速度快，灵活异常，但因没有统一的依据，关键词很不统一。加上文献用语漫无规范，标引的结果会造成漏检和误检的。

在上述发展基础上，并吸收了分类法的长处，产生了一种叙词法。我们说的主题法，目前多指叙词法。

关于能否用主题法集中、组织文献，以及如何组织、排列文献，还要参考《图书文献分类法——分类与主题》讲义本身，也是补讲义，恕不系统阐述。

第二节 检索词表 (Thesaurus)

用主题法组织图书文献，必得有一种颇便进行工作的工具，像分类法必定有分类表一样；这种工具称为检索词表。

The saurus 一词，源出希腊语的 *Thesaurus*，意为“宝库”。可广义地理解为“词典”或“百科全书”；也有人用作“词库”注。到了十九世纪五十年代，由英国目录学家伯尼埃 (Bernier, C.L.) 首次用于情报检索领域。国内外称呼很不一致。国际标准化组织 (ISO) 称为检索词表，联合国教科文组织 (UNESCO) 称为叙词表，苏联科学院科学技术情报研究

注：如 Roget, P.M.: 《*Thesaurus of English Words and Phrases*》。这里用作“词典”。第一次使用“词库”这个含义的是 Lotini, F.B. 他编了一种《百科词库》。

所所长米哈依洛夫(Михайлов, А. И.)称为检索词库,日本的渡边诚称为检索词表,而中国科技情报研究所及,某图书馆称为主题词表,各家对它下的定义也不尽相同(注)

我们认为, *Thesaurus* 一词自从被用于情报检索以后,它的本义是指为了情报检索的需要,而把检索词(名 *Descriptors*)按不同的检索范畴以及这些词的语义层次、係从及结构关系编成的词表。因此,应根据它的特定目的——情报检索的需要,根据 *Descriptors* 的性质,定名为“检索词表”。这里涉及 *Descriptors* 的语义问题,在采用 *Descriptors* 时,本意是指标引人员用来标引描述文献主题的词,也可以说是标示一类文献的词或符号,它並沒有叙述什么,而只是一类文献的标示符,它的最终作用是用于检索,它是检索语言中的基本单位,相当于自然语言中的基本词汇。因此,应称为“检索词”,由它组成的词表,应称为“检索词表”。

检索词表可以定义为: 检索词表(*Thesaurus*)是一种为情报标引和检索的需要,将规范化的、动态性的检索词(*Descriptors*)通过参照系统表达其层次、结构关系,并按字母顺序和等级体系编列的情报检索语言词汇表。

词表首先是检索词的表,检索词包括标引和检索时用作检索词的标准检索词(有时供同检索词作指标准检索词),而那些在标引和检索中不使用的、只用来转换成标准检索词的词叫引导词或弃入词(*non-descriptor*、*forbidden term*, or *lead-in term*)。所谓规范化,就是检索词具有单义性,规定一个词只代表一种概念。也就是说,要求在同义词中选取一个检索词来唯一地代表某一概念;在多义词则各抒意义中概

注: ISO: “既可以从它的功能方面,又可以从它的结构方面下定义。从功能方面: 检索词表是一种控制术语的工具,以便将文献作者、标引者和检索者的自然语言转换成更加强制性的‘系统语言’(文献工作语言、情报语言),又能将这种‘系统语言’转回成自然语言。从结构方面: 检索词表是某一种

代表其某一主要意义。所谓动态性,是指检索词在经常不断地更新,或前增加表示某学科的新概念或检索词,或者是根据标引和检索实践中的经验和资料改造现有的词义及结构。词表是为标引和检索文献情报而编制的,所以所选的词必定是在文献中出现(使用)频率高的、代表一定对象的、有检索意义的、有组织性能的词。词表是按检索词的字母(音顺或形顺)和等级体系,又要通过等级系统来表述检索词的层次、结构关系而编制而成的,因此它本身便是一种系统。关于词表的选词、编制原则及方法,这些都是一项很重大的问题,要深入探讨,必将涉及逻辑学、布尔代数、语言学以及电子计算机等方面的知识。本书又不做全面的阐述,只将关于检索词表的结构及使用方面的内容做阐明。

检索词表结构

检索词表自有其结构原理,在前一节主题法中已约略提到一些。本节只就在应用中涉及的一些问题在有关部分阐述。检索词表的结构组成可用下图表示:

的学不。中在语义种属种关系上相关的受控制的与动态性的词汇表。” UNESCO:“叙词表是某一特定学科领域内,语义相关和类属相关的词汇的标准和动态的词表”。Muxadrov, A.H.:“所谓情报检索词库就是指帮助情报使用者用叙词语言来表达他们的情报需求,以及使用这些术语来编制精细的文献索引和进行情报检索的一种专门词典。在这个词库中必须包括该系统检索语言所使用的所有叙词,並十分明确地表现他们在情报上的相互联系;同时还必须包括系统内部认为与叙词有关的同义关键词”。日本的渡边武则认为:“所谓检索词典,就是展示检索词及其相互关系的词典。”中国科技情报研究所和北京图书馆对主题词表的定义请参考《汉语主题词表(Chinese Thesaurus)》。

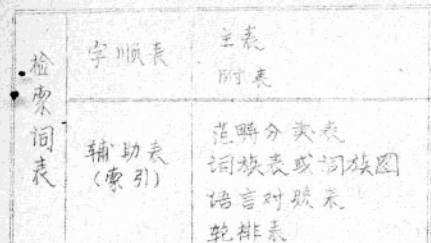


图1. 检索词表组成图

目前全世界的检索词表种类很多。据有人统计，美国一国就有三百多种，名称也是各种各样。由于各种词表的编制条社编制原则以及系统设计的不同，各表在结构成份上也不尽相同。

一般说来，词表在体系上分两大部分：字顺表和辅助表（也称辅助索引）。字顺表又分主表和附表两部分，但大多数词表将这两部分合二而一统一编排。辅助表分图1所示四种，实则是对字顺表从各个角度的一种索引或全显示检索词的层次等级关系的表。辅助表和字顺表里的参照系统一起构成的参照网络的设计，各种词表也各有千秋。现在结合《汉语主题词表》进行详述，各种名词术语也尽量与《汉语主题词表》求得一致。

《汉语主题词表(Chinese Thesaurus)》(简称CT或词表，下同)的体系结构如图2：

汉语主题	字顺表	主 表	社会科学	第一卷第一分册
			自然科学	(A-T) 第二卷第一分册
		附 表	社会科学	(G-L) 第二卷第二分册
			自然科学	(M-T) 第二卷第三分册
	辅 表	范畴分类表	社会科学部分	(U-Z) 第二卷第四分册
			自然科学部分	社会科学合编 第三卷
		词 族 表	社会科学部分	第一卷第二分册
			自然科学部分	第二卷第六分册
		词 族 表	社会科学部分	第一卷第二分册
			自然科学部分	第二卷第五分册

词 表	助 表	語 言 對 照 表	英 漢 對 照 索 引	社會科學部分 第一卷第二分冊		附 表 合 編	第 三 卷
				自然科學部分 第二卷第七分冊			
				世界各國政區名稱部分			
				自然地理區劃名稱部分			
				組織機構名稱部分			
				人物部分			

图2:《汉语主题词表》体系结构图

从这图可以看出《汉语主题词表》在体系结构上除附表外,分成社会科学和自然科学两大部分,分别编出其主表及范畴索引和词族索引,附表则为社会科学和自然科学合编。而英汉对照索引,除分社会科学与自然科学两部分,附表部分也独立编出,成六大块。这样显得结构支离破碎,头绪繁多,不见完整,查检不便。全书共分十大册,过于分散零碎。但就结构组成而言,还是比较完备全面的。

《汉语主题词表》在结构体例上,有不一致的地方。如社会科学的词族索引,只属主表的款目词(《汉语主题词表》称主题词)均列出范畴分类号,而附表款目词就没有范畴分类号;自然科学部分的词族索引,无论是族首词还是其下位词,均无范畴分类号。这样的一些缺点,使《汉语主题词表》严格地说,还没有构成完整的,各部通畅异常的完善的网络关联。

当然,所有这些并不影响词表的大局,提出来只是为了在使用时引起注意了。

下面分词表结构各部分进行阐述。

(一) 字顺表

《汉语主题词表(Chinese Thesaurus)》的字顺表,从体系结构上分为主表和附表两部分,从编制上分社会科学(共一册)和自然科学(共四册)主表和附表(共一册)。

主题词表是将全部主题词款目按汉语拼音字母顺序排列而成的，它是标引情报文献的主要工具。附表是从主表中析出的几种专用词汇表，附表是主表的一种特殊，是主表的补充，因此，在标引和检索情报文献时，应特别注意对附表检索词的使用，不要因一查主表没有某词，马上就判定为漏过。

主题词表最基本的组成单元是主题词款目。

《汉语主题词表》的主题词款目构成如下：

主题词款目	款目词 汉语拼音		决定款目体系位置		
	款目词	正式主题词 (检索词)	作标引和检索标识		
		非正式主题词 (导入词)	不作标识		
	款目词 英译名		与对照索引联系		
	款目词 范畴分类号		与范畴索引联系		
	款目词 注释项	范畴限定词		赋予检索词单义性	
		含义、观点注释			
	参照系统	参照项	用(Y)项	图一关系	指引检索词
			代(D)项		表明导入词
			属(Y)项	部分关系	指出上位概念
分(F)项			指出下位概念		
族(Z)项			指出族属词		
考(C)项			表明相关检索词		
		相关关系			

图3:《汉语主题词表》主题词款目组成图

《汉语主题词表》的主题词款目由两大部分构成：一是款目词部分，包括汉语主题词本身，及其汉语拼音、英译名、范畴分类号及款目词注释项；二是参照系统。

主题词款目是款目的基础，是每条款目的必备部分。款目词有两种：正式主题词(Descriptor)和非正式主题词(Non-descriptor, forbidden term, or lead-term)。

正式主题词是经过规范化的、可用于标引和检索的主题词。当正式主题词有与它对应的同义词时，同时这些同

义词使用频率较高，则被选为非正式主题词。非正式主题词不得用作标引和检索的标识，只当转换成正式主题词之用。

主题词的选词必须遵循下列原则：

(一) 概念性。主题词必须是表示某事物、某对象的、能表达文献的主题和使用者查询课题的特定概念、科学上的通用名称。它包括下列范围：

- 1, 表示具体事物的名称术语，如拖拉机，商品；
- 2, 表示事物的状态或现象的名词术语，如强度，经济危机；
- 3, 表示学科门类的名称术语，如电子学，心理学；
- 4, 表示工艺方法和工艺过程的名词术语，如焊接，热处理；
- 5, 表示研究方法的名词术语，如分析，总结；
- 6, 表示地域、区域、人物、组织机构等的专有名词，如中国、北京、地中海、鲁迅、中华全国总工会；
- 7, 文献资料类型名称，如年鑑、手册；
- 8, 表示化学元素、化合物的名词术语，如氧原子、硅化物、吡啶、聚烯烃等
- 9, 金属材料、合金，如金属板、耐候合金。

(二) 频率性。选定的主题词，主要是各学科领域内经常出现的、在实际检索工作中有一定的使用频率的、并能汇集一定的文献量的或具有组配意义的最基本的名词和词组。但使用频率、出现频率的高低不是取舍的绝对标准，新事物、新学科领域出现的主题词，虽然频率较低，也应选入。

(三) 兼容性。除本专业学科和基础学科，注意选取那能在各学科内具有检索意义的基本名词术语。这些词汇，大多比较典型、涵义明确、变化缓慢、具有长久生命力和强大的组配力。利用组配方式能形成许多适合现代科学技术发展的特有性概念。一般不选用那以在各学科中太特指的名词术语。如产品型号、著作名等可作自然语言标引。

(四) 思想性。主题词的选录，主要应从三大革命运动需要出发，贯彻为四化服务的原则，体现四个坚持的精神，对不同的政治观点用词及其语义失保，应注意其思想性和政治内容。

《汉语主题词表》是一部大型综合性检索工具。它的收词范围包括马列主义、毛泽东思想、哲学、社会科学和自然科学等学科领域的名词术语，它宣称，既适合于社会科学的书刊资料的标引和检索的需要，又适合于自然科学的各种书刊、研究报告、学术论文、会议录、专利说明书、标准和样本等资料的标引和检索的要求；它适用于机检，又考虑了手检的要求。当然检验结果最终要由实践判决，但它的确是收录了主题词108568条，其中正式主题词91158条，非正式主题词17410条。这以主题词中，自然科学占78113条（其中正式主题词65200条，非正式主题词12913条，自然科学正式主题词中包括族首词2821条，成族词50039条，无关联词12123条），社会科学占30455条。（其中正式主题词25958条，非正式主题词4497条）。

每一条主题词上方，均注明汉语拼音，用以组织主题词顺序，按其排列。

在汉语主题词旁，注有英文译名，供与英汉对照索引联系，以便在标引和查寻英文文献时参考。同一汉语主题词具有几个英文译名时，则同时列出。有的汉语主题词和尚未见诸英文文献，或一时未查到者，英译名暂缺，在这个基础上另编有英汉对照索引，字顺表英译名与英汉对照索引一起，实则起了“汉英、英汉对照词汇”的作用，这是《汉语主题词表》的特色之一。

在主题的汉语主题后，注有该词在范畴索引的范畴分类号，用以通过范畴分类号的联号，在范畴索引中查寻同一类目中的有关主题词。

注释项在《汉语主题词表》中没有归入分类系统。在汉语主题词下，有注释项注了一些简要注释，目的是明确主题