

字书字料库汉字本借关系类型字段研究

冯宝成

(渤海大学 文学院, 辽宁 锦州 121013)

摘要: 汉字本借关系类型字段是字书字料库本借关系界面的核心字段, 字段设计是本借关系界面建设的重中之重。汉字本借关系分类应以古籍文献例证为主要依据, 并应综合考查文献例证的年代、构成本借关系字组的形义关系、相关社会用字习惯等因素。字段设计应以科学性、全面性、有限性及标注的便利性为原则。文章基于前期渤海大学字书字料库汉字本借关系类型字段的标注情况, 针对汉字本借关系分类存在的外延交叉、内涵模糊等不足, 综合考查相关因素, 对汉字本借关系内部情况进行梳理与分类, 并依照设计原则对字段设计进行修改和完善, 以进一步提高汉字本借关系界面的应用价值。

关键词: 字书字料库; 汉字; 本借关系类型; 字段设计

中图分类号: H12 **文献标志码:** A **文章编号:** 1672-349X(2022)01-0091-07

DOI: 10.16160/j.cnki.tsxyxb.2022.01.011

Study on the Field of the Relationship Between Orthograph and Borrowed Graph in Chinese Character Form Database

FENG Bao-cheng

(School of Literature, Bohai University, Jinzhou 121013, China)

Abstract: The field of the relationship between orthograph and borrowed graph is the key field for the relationship interface in the Chinese Character Form Database (CCFD), and the field design enjoys a top priority in the interface construction. The classification of the relationships should be mainly based on the classical literature. In the process of classifying the types, the following factors should be considered comprehensively, including the age of the relationships, the morphological and semantic relationships of the word groups, and the social usage habits, with the principles of being scientific, overall, finite, and convenient in labeling. Based on the labeling situation for the above field of CCFD in Bohai University and aiming at the deficiencies of crossing denotation and vague connotation in the classification, this paper comprehensively studies the relevant factors, categorizes the inner situation and modifies the field design according to the above principles, so as to improve the application value of this interface.

Key Words: CCFD; Chinese characters; the type of the relationship between orthograph and borrowed graph field; field design

基金项目: 国家社会科学基金重点项目(20AYY018); 辽宁省教育厅 2020 年度人文社会科学基础研究项目(WJ2020015)

作者简介: 冯宝成(1996—), 男, 辽宁朝阳人, 硕士研究生, 研究方向为汉字学。

本借关系界面是字书字料库字际关系属性界面的重要组成部分,其核心字段是汉字本借关系类型字段。目前渤海大学字书字料库(CCFD)本借关系界面汉字本借关系类型字段下设五种关系类型,即古本字与后造本字、本字与通假字、通假字与通假字、假借字与假借字、假借字与后造本字^[1]。在前期标注实践中发现,此字段对汉字本借关系的分类尚存在一些不足,并在一定程度上影响了汉字本借关系界面字料信息的标注工作,也将对标注完成后此界面的应用产生不利影响,因此亟待对现有汉字本借关系类型字段进行调整与完善,并制定出一套符合汉字本借关系实际的本借关系类型字段标注标准。

一、本借关系内部具体情况梳理

受汉字形义关系发展演变复杂性的影响,汉字字形承载的词义差别较大,包括本义、引申义、假借义等,构成本借关系的各字形所承载固定义的复杂多样是本借关系类型多样化的主要原因。因此在对本借关系类型字段进行修改、完善前,有必要对构成本借关系的一组字结合文献例证时代是否共存、字形承载的固定义类别、形义关系的历时演变情况以及社会用字习惯等因素进行梳理,以理清汉字本借关系内部的复杂关系,为本借关系类型字段的修改与完善做好前期准备。构成本借关系的字组内部主要存在两种情形,即同源通用字与同音借用字。

(一)同源通用字

同源通用字这一概念是由王宁先生首次提出的^[2]。“同源通用是词语分化推动文字孳乳的过渡阶段界限模糊的表现”^[3],是在汉字大量孳乳阶段必然产生的文字现象。按构成本借关系的字组中各字在同源关系中的角色,可以分为原字与分化字间的通用和由同一原字分化而来的分化字间的通用。

1. 原字与分化字间的通用关系

从汉字孳乳分化角度来看,原字即孳乳分化出新字的字形,而分化字是指用来记录原字形所承载的部分词义的新字形,这些字形一般是由原字通过构形方式派生出来的。按原字

所承载的常用义类别,可将原字与分化字间的通用关系分为专造本字与分化字间的通用和假借本字与分化字间的通用;专造本字所承载的固定义包括本义和引申义,假借本字所承载的固定义为假借义。

对原字与分化字间的关系依字形词义历时演变情况进行梳理,按文献例证所处时代分化字是否已经产生进行区分,有以下三种情况:其一,后出分化字形产生之时,文献中使用原字,此时原字当为本用,但部分训诂学专著或语文教材中的古诗文注解将原字注为后出分化字的同音借用字,即通假字。其二,孳乳分化阶段,原字与分化字词义界限模糊不清,两字形共同承担部分词义,形成一种“词义共享”现象^[4],两字此时当为通用关系,即同源通用。其三,孳乳分化阶段结束后,部分文献中弃分化字形不用,仍使用原字,此阶段两字间关系学界存在争议,争议的焦点在于此种情况是否属于本有其字的通假现象。有学者认为:“当初文与后出本字从所记词义到文字形体上已产生分化以后,再用初文来记录它的原始本义,就是通假。”^[5]张严匀则认为此时原字的使用当为古字的沿用,其与后出分化字构成古今字^[6]。因汉字孳乳分化、词义模糊的时间段很难确定,因此将此种情况下的原字与分化字的关系统一标注为原字与分化字的通用关系。

2. 分化字与分化字间的通用关系

由同一原字分化而来的若干字形在孳乳分化阶段同样存在词义不固定和各字形间相互通用的现象,构成同源通用关系。如“返”与“叛”,两字皆由“反”字分化而来,各自承担“反”字部分引申义。“返”本义为“回”,“叛”本义为“背叛”,但在有些文献中常用“返”表“背叛”义,如《淮南子·奇俗训》中有“谈话而不称师,是返也”,此处“返”当与“叛”为分化字与分化字间的通用关系。

(二)同音借用

同音借用现象分为本无其字的假借和本有其字的通假,将假借和通假两个术语进行明确界定对于汉字本借关系研究来说有着重要意

义。下面分别对其内部情况进行梳理。

1. 本无其字的假借

本无其字的假借自汉字使用之初就不可避免地存在着,这种现象的产生主要是由早期汉字数量少与汉语词汇量不断增多的矛盾决定的。王宁先生提到:“早期文字不够用的时候,一个字兼记两个词,这其实就是一般所说的‘本无其字的假借’。”^[7]随着人类思维的发展、认知能力的提高,汉语词汇量大大增加,但已有汉字的数量却无法记录所有新词汇的需求,因此就产生了借用已有字形表义的现象,即本无其字的假借。在汉字本借关系中本无其字的假借字多与它所承载假借义的其他假借字形构成本借关系。

本无其字的假借按照各字形在文献中的使用情况又可分为三种关系类型:其一,假借之初,多个字形同时被假借用作表同一本无其字的词义,这些字形的使用频率差别不大,无所谓“正”与“不正”。如“胡”“奚”“曷”“何”四字皆被假借用作疑问代词,皆只能用作宾语和状语,用作宾语时一般皆表“什么”义,用作状语时一般皆表“为什么”义,且四字在共时层面上的使用频率差距不大。这种情况属于假借字与假借字的关系。其二,两个或多个假借字形中的一个假借字形在后代成了固定表该假借义的正字,其他假借字已很少被使用。如在后代文献中较多地使用“何”字作疑问代词,“胡”“奚”“曷”三字使用的频率大大降低,“何”成为疑问代词的正字,若此时仍使用“胡”“奚”“曷”三字作疑问代词,那么依据“正”与“不正”的通假字判定原则,三字当属“何”的通假字,与“何”构成本字与通假字之关系。其三,同一词义在不同时代的文献中所假借的字形不同,各字形在各自时代皆为某词义的正字,如“女”与“汝”,两字分别在不同时代被假借为第二人称代词,皆为各自时代的假借正字,但部分训诂学家不了解第二人称代词所依附字形的历时变化,误将“女”注作“汝”的通假字,其实二字当为假借字与假借字之关系。

2. 本有其字的通假

本有其字的通假是指某义的表达在共时文献中有其固定使用的字形,此字形即“正字”,但在共时文献中又临时借用了音同或音近的其他字形,且被借字形与“正字”从汉字孳乳分化角度看不存在分化关系,此时“正字”与被借字就构成本字与通假字的关系。“正字”又可称作“本字”,与“通假字”组成一组相对的概念,“正字”与“通假字”构成通假关系的本义不一定是其造字本义,曹先擢先生提出:“字的本义包括造字本义和字的常用义。”^[8]字的本义既包括造字本义,又包括常用的引申义和假借义,明确字形所承载的词义性质对于通假关系的确立和判定有着重要意义。

本有其字的通假关系按通假字在后代文献中的地位又分为两种情况:其一,通假字在当时文献中临时代替本字,在后代文献中未发展成为正字。若某字形所携带的同一词义在不同古籍文献中借用了两个不同的字形来表达,被借用的两个字间构成通假字与通假字的关系。如“棗”“蚤”二字在古籍中皆见被借表“早晨”义,是“早”的通假字,前者可见湖北云梦秦简甲种《日书》“利棗不利莫”,后者可见《史记·秦始皇本纪》“公何不蚤告我”?“棗”“蚤”二字即为通假字与通假字的关系。其二,通假字在当时仅为正字的临时借用字,但在后代逐渐发展成为表此通假义的正字,此为“代替型通假”,情况较为复杂。如“谊”与“义”。“谊”字《说文·言部》释为“人所宜也”,本义为合宜的道德、行为或道理,后借“义”表此义,彼时二字为本字与通假字的关系,“谊”为本字,“义”为通假字;后“义”演变为表通假义的正字,“义”“谊”二字又构成本字与通假字的关系,“义”为正字,“谊”为通假字。

综上,汉字本借关系内部情况复杂,其中既有同源通用关系,又有同音借用关系,因此,在设计字书字料库汉字本借关系类型字段时,不仅要综合考虑以上情况,还应综合考查构成本借关系的字组在文献例证时代是否共存、字形承载的固定义类别、形义历时演变情况及文献

例证时代社会用字习惯等,并作为本借关系分类的主要依据。

二、字书字料库汉字本借关系类型字段设计原则与本借关系分类依据

字书字料库汉字本借关系界面中的所有字段皆与汉字所携带的本借关系信息相关,因此,汉字本借关系类型字段设计应立足于汉字本借关系内部情况及字书字料库标注的具体要求。

(一)字书字料库汉字本借关系类型字段设计原则

应用于汉字学及相关学科的研究是字书字料库汉字本借关系界面建设的主要目的,字段设计是此界面建设的重要过程,而汉字本借关系类型字段设计则是此界面建设的重中之重。因此,综合考虑相关因素提出如下设计原则。

1. 科学性是首要原则

科学性是指对汉字本借关系的分类要准确,符合本借关系的实际情况;名称设置要符合汉字学相关规范,分类标准与分类角度要统一。主要体现在三个方面:其一,分类结果的内涵与外延要清晰,避免交叉;其二,本借关系类型的名称要清晰,要让使用者通过名称就可清晰了解此类型本借关系的具体情况;其三,本借关系类型名称设置要符合汉字学相关规范,并采用学界通用的汉字学术语对汉字本借关系类型进行命名。

2. 全面性是重点原则

全面性是指分类结果要全面,避免出现某些构成本借关系的字组无类型可选择的情况。要保证分类结果的全面性,字段设计者首先需要明确汉字本借关系的内涵与外延,同时综合考察、梳理汉字本借关系内部情况,这样才能使本借关系类型字段设置的本借关系类型更加全面。

3. 有限性是补充性原则

有限性是指在本借关系类型字段的设计过程中,设计者对于汉字本借关系的分类要有一定限度,不能进行赘余式分类,否则就会既占用本借关系界面的空间,又给标注工作带来不必

要的干扰。

4. 便利性是内在要求

便利性是指本借关系类型字段的设计要为标注工作提供便利。本借关系类型字段所含信息较为复杂,目前仍需进行人工标注。标注者首先要进行本借关系类型判定,而分类的科学与否是决定标注者能否快速标注的重要因素。

在汉字本借关系类型字段设计中要兼顾以上原则。其中全面性以科学性为前提,有限性与便利性则互为补充,四条原则兼顾,字段设计才能更加完善,进而更好地为汉字学及相关学科研究提供帮助。

(二)字书字料库汉字本借关系类型字段本借关系分类的依据

判定一对汉字是否构成本借关系的主要依据是参证文献及两字的声韵关系,而判定其具体关系类型则主要依据文献例证,同时应综合考查其形义关系历时演变情况及文献例证时代的社会用字习惯等因素。

1. 文献例证及其时代

文献例证是汉字本借关系判定和分类的基础。汉字本借关系是文献角度的字际关系,其关系仅存在于文献例证之中,一旦脱离文献,这一关系就不复存在。同时,文献例证的时代也是汉字本借关系判定和分类需要重点考查的因素。汉字的使用具有历史性,其形音义并不是一成不变的,因此在对汉字本借关系进行判定和分类时,必须要对文献例证的时代进行考查;不仅如此,还要对构成本借关系的字组中各字形产生的时代及其各时代所承载的词义变化进行历时梳理。

2. 本借关系字组的形义关系

形义关系即某字形所承担词义的历时演变情况及词义的性质,是汉字本借关系分类的核心依据。词义的历时演变情况指因词义的引申及汉字的同源通用和同音借用等情况导致的字词义对应关系的变化,亦即字形所承载词义的变化;词义的性质是指本借关系中的某字形所承担的词义在共时层面是固定义还是临时通假义。

3. 社会用字习惯

社会用字习惯是指共时层面上汉字使用者多用哪个字形记录某个词义,使用频率最高的字形即为“正字”,“正字”这一概念是相对于通假字而言的。“正”是指某词义在共时层面的文献中多由某字形承担,且已经形成社会共识,此词义即为该字形的固定义。与“正”相对的“不正”是指某词义在共时层面的文献中极少用某字形承担,此字形在文献中表达此词义仅属临时借用。“随着‘字’与‘词’概念的引入和发展,人们从语用角度研究通假,根据某时代某词义所用字形的频率确定本字,这一原则的提出和确立,有助于处理通假研究中的一些通假字的判定与确立。”^[9]因此,在进行本借关系类型字段设计时需综合考查文献例证时代的社会用字习惯。

综上,在对汉字本借关系类型进行研究时,要以文献例证为基础,以历史的眼光对构成本借关系的字组从例证时代、形义关系、社会用字习惯三方面进行综合考查,切忌以当代情况为根据判定文献例证时代的字际关系,以免造成判断失误。

三、字书字料库汉字本借关系类型字段设计存在问题及修改建议

“字书字料库能够提供具有客观性、真实性、规模化、多功能的字书汉字数据信息”^[10],而建设一个使用价值高的字书字料库,字段设计则是其中重要的工作环节,也是字料信息标注与应用的基础。以下在总结目前字书字料库汉字本借关系类型字段存在问题的基础上,提出针对性的修改、完善建议。

(一)汉字本借关系类型字段存在的问题

目前字书字料库汉字本借关系类型字段下设古本字与后造本字、本字与通假字、通假字与通假字、假借字与假借字、假借字与后造本字五种类型,这五种类型的划分是对学界已有本借关系研究进行总结与创新的结果。但在前期的标注实践中发现,本借关系类型字段在设计上还存在一些问题,主要表现在三个方面。

1. 本借关系分类结果不具备全面性

上述五种关系类型并不能涵盖所有在文献例证中构成本借关系字组的具体本借关系小类,出现了部分文献例证中构成本借关系的字组在标注过程中无对应类型的情况。如由同一原字分化而来的两个分化字间的通用关系无对应类型,因此无法标注,分化字与分化字间的通用当属同源通用的一种,因此无法将其标注为本字与通假字的关系,需要增加标注选项。

2. 部分本借关系类型字段的名称内涵模糊、外延交叉

类型名称内涵模糊、外延交叉容易导致标注工作陷入两难境地,即文献例证中的某字组既可标注为A关系,又可标注为B关系。如“古本字与后造本字”“假借字与后造本字”,其中古本字内涵模糊且与后造本字外延交叉,假借字与后造本字外延交叉。

3. 本借关系类型字段无法显示同源通用或同音借用信息

汉字本借关系实际上包含同源通用与同音借用两种不同的汉字字际关系。同源通用包括原字与分化字、分化字与分化字、假借本字与分化字间的通用,同音借用主要包括本字与通假字的借用,但现阶段的本借关系类型字段无法对此进行区分与标注,亟待完善、优化。

(二)汉字本借关系类型字段设计建议

1. 将本字与通假字的关系细分为专造本字与通假字、假借本字与通假字两种类型

构成本字与通假字关系的字组中的本字在形义关系上存在多样性,本字所承担的词义不仅包括本义,同时还包括引申义、假借义。因此,遵循本借关系分类原则中的科学性原则,可将本字与通假字关系细分为专造本字与通假字、假借本字与通假字两种类型。这是由于专造本字与假借本字在词义来源上存在较大差别。在专造本字与通假字的关系中,专造本字承担的词义包括本义和引申义;在假借本字与通假字的关系中,本字承担的词义为假借义,主要包括某些虚词、联绵词等。

2. 将古本字与后造本字类型名称改为专造本字与后出分化字

相对于后造本字,假借字是由后出分化字充当的假借本字,亦可称为古本字,并与后造本字构成古本字与后造本字的关系。为消除古本字这一名称内涵模糊、外延交叉带来的困扰,本着类型字段设计的科学性原则,将古本字与后造本字的类型名称改为专造本字与后出分化字,用来专指同源通用关系中的原字与分化字间的关系。

3. 增加后出分化字与后出分化字类型

现阶段本借关系类型字段设计忽略了由同一原字分化而来的分化字间的通用关系,前期标注过程中一般将这种关系当作本有其字的通假关系,标注为本字与通假字。这一标注方式忽略了分化字与分化字之间在汉字孳乳分化、词义过渡转移阶段也存在着因词义界限模糊而相互通用的情况,显然不符合汉字发展的实际情况,因此可在本借关系类型字段添加分化字与分化字间的通用关系选项。

4. 增设“本借关系大类”新字段,并与修改后的本借关系类型字段相关联

同源通用与同音借用虽同属汉字本借关系大类,但二者存在着本质区别。在共时层面上,构成通用关系的一组字词义界限模糊,而构成借用关系的一组字词义界限清晰。因此可将同源通用与同音借用区别开来,在本借关系界面增加本借关系中最重要标注信息,使本借关系类型字段的设计更具科学性。

(三) 汉字本借关系类型例析

经调整后的本借关系类型字段分为专造本字与后出分化字、分化字与分化字、专造本字与通假字、假借本字与通假字、通假字与通假字、假借字与假借字、假借字与后出分化字七种类型,下面分别进行举例说明。

1. 专造本字与后出分化字

此类型指同源分化字中原字与后出分化字的通用关系,其中不包含假借字与后出分化字间的通用关系。如“取”字,《周礼》解为“大兽公之,小禽私之,获者取左耳”,《说文·又部》解为

“取,捕取也”。其甲骨文字形为用手割耳之状,本义为捕获到战俘或野兽时割下耳朵,后引申为娶妻之义,如《易·蒙》有“勿用取女,见金,夫不有躬,无攸利”。后因“取”字形承担词义过多,故为娶妻义专造“娶”字形,两字从汉字孳乳角度构成原字与分化字的关系;又因两字在汉字孳乳阶段常常通用,故从文献角度又构成专造本字与分化字的通用关系。

2. 分化字与分化字

此类型主要指由同一原字分化而来的两个或多个分化字间的通用关系。例如“避”与“僻”,两字皆由“辟”字分化而来。《说文·辟部》有“辟,法也”,本义为施加刑罚,后假借表回避、躲避、荒僻、邪僻等义,后为回避、躲避义专造“避”,为荒僻、邪僻义专造“僻”,但在分化之初,“避”“僻”二字的词义范围模糊且不固定,常常通用。如在汉桓宽《盐铁论·地广》“今推胡越数千里,道路回避,士卒劳罢(疲)”中,“避”表荒僻义;在湖北云梦睡虎地秦简《语书》“是以圣王作为法度,以矫端民心,去其邪避,除其恶俗”中,“避”表邪僻义;在《史记·樗里子甘茂列传》“韩,公之雎也,今公言善韩以备楚,是外举之僻雎也”中,“僻”表躲避义。在以上文献例证中,“避”“僻”二字当为分化字与分化字的通用关系。

3. 专造本字与通假字

此类型主要指本有其字的同音借用中本字与通假字的关系,二字从汉字孳乳角度看具备分化关系。如“早”在《说文·日部》“早,晨也”中的本义为早晨,“蚤”在《说文·虫部》“蚤,啮人跳虫”中的本义为跳蚤,但在部分文献中却借用同音字“蚤”表早晨、月初义。如《淮南子·天文》“日至于曾泉,是谓蚤食”中,借用“蚤”表早晨义。在以上文献中,“早”“蚤”二字构成专造本字与通假字的关系。

4. 假借本字与通假字

此类型中,假借本字承担的词义是假借义。如“才”在《说文·才部》“才,草木之初也”中,本义为草木初生,后假借表方、始、仅仅义,成为以上副词的假借本字;在李时珍《濒湖脉学》“伏脉

推筋著骨寻,指间裁动隐然深”中,又借用“裁”表始义。在以上文献中,“才”“裁”二字构成假借本字与通假字的关系。

5. 通假字与通假字

此关系类型指在文献中借用了两个或多个不同的字形表某字形所记录的某一词义,被借用的两字构成通假字与通假字的关系。如“故”“辜”二字在文献中都被借作表固定义,前者见《庄子·秋水》“夫物,量无穷,时无止,分无常,终始无故”,后者见《大戴礼记·夏小正》“鵙则鸣。鵙者,百鹑也;鸣者,相命也。其不辜之时也”。两字表固定义时同为“固”之借字,当为通假字与通假字的关系。

6. 假借字与假借字

此关系类型是指本无其字的假借关系,两个或多个假借字形在使用频率上无太大差别,无法判断哪个是正字,因而构成假借字与假借字的关系。如“余”与“予”二字皆被借作表第一人称,并且属于本无其字的假借,因而二字构成假借字与假借字的关系。前者如《论语·述而》“天生德於予”,后者如唐柳宗元《小石潭记》“余弟玄宗”。部分通假字字典中将两字判定为本字与通假字的关系,但两字从社会用字习惯上考查无法区分“正”与“不正”,因此当为假借字与假借字的关系。

7. 假借字与后出分化字

此关系类型指本无其字的某一词义假借了一个音同或音近的字形,后又为该词义专造了新字形,被借用的字形与后造字形就构成了假借字与后出分化字的关系。如“舍”在《说文·亼部》“舍,市居曰舍。从亼中口,中象屋也,口象築也”中的本义为屋子,后假借作表舍弃义。后为舍弃义专造“捨”字,“舍”“捨”二字就构成假借字与后出分化字的关系。

综上,基于前期渤海大学字书字料库

(CCFD)本借关系界面中本借关系类型字段的标注情况,对汉字本借关系内部情况进行梳理与分类,并对类型字段设计进行修改和完善,但不排除仍有其他类型的汉字本借关系未被发现,一旦发现新的本借关系类型仍需对以上分类进行修改、补充,以便更好地为汉字学及相关学科的研究提供便利。

参考文献:

- [1] 柳建钰. 字书字料库的理论、实践与应用[M]. 北京:中华书局,2021:268.
- [2] 陆宗达,王宁. 训诂方法论[M]. 北京:中华书局,2018:54.
- [3] 李国英. 试论“同源通用字”与“同音借用字”[J]. 北京师范大学学报(社会科学版),1984(4):53-57.
- [4] 鲁普平. 原字与分化字“词义共享”现象及相关问题探索[J]. 语文学刊,2019(4):62-67.
- [5] 芮东莉. “同源通用”的类型及其性质[J]. 阜阳师范学院学报(社会科学版),2003(2):33-36.
- [6] 张严匀. 论“本有其字”的假借[J]. 佳木斯大学社会科学学报,2019(4):119-121.
- [7] 王宁. 汉字构形学导论[M]. 北京:商务印书馆,2015:170.
- [8] 曹先擢. 通假字的字义移位问题[G]//吕叔湘. 语言文字学术论文集. 北京:知识出版社,1989:522.
- [9] 冯靓芸. 《汉书》通假研究[D]. 上海:复旦大学,2005.
- [10] 柳建钰. 字书字料库中字料标注若干问题刍议[J]. 语言文字应用,2015(3):133-140.

(责任编辑:白丽娟)