

生成语言学理论模式下对语位化理论的再探究

摘要 由于书面语符号形式文字的概念一直以来是一个热点问题，与之产生的汉语本位之争也是国内学者争论的焦点问题，本文依托生成语言学相关理论，提出书面语“形+词/语素”的形式为语位，并提出语位是书面语的符号，接着我们通过对语言生成机制的再研究，提出了语言状态机制的运行方式，并提出认知态、形体态概念，以此提出语位是语言形体态与语言认知态在一定条件下进行指称功能的激活与使用后，来完成语言储藏传递的语言单位，并与之建立了语位化理论和汉语语位化理论及其相关标识。

关键词 语位 语言状态机制 认知态 形式态

汉语本位说法一直是国内学者争论的焦点问题，而最近国家大力发展“冷门绝学”——古文字，又给了我们研究语言文字的学人一个对文字这个书面语的符号形式系统的本质进行了讨论的机会，而近些年生成语言学的相关理论也在国内学界蓬勃发展，本文以期在生成语言学角度对文字以及汉字的本质予以解释和解答。

1 从名家理论中看文字及汉字的本质

裘锡圭（1988）认为，给拼音文字定性时，是看一个字中的字母表达了语言的什么成分；既然如此，给汉字定性也不能看一个字表示了什么成分，而应看一个字中的字符跟语言中什么成分发生关系。这样地地看法，是值得商榷的。

张玉金、夏中华（2001）总结诸家对汉字性质的总结，笔者选取关于裘锡圭先生的部分转录如下：

有的人是因为看到汉字里一个字通常代表一个语素，称汉字为语素文字的。像这样撇开字符的性质，仅仅根据文字书写的基本单位所代表的语言成分的性质，来给文字体系定名，也是不妥当的。

在谈论汉字性质时，不能把一个汉字跟西方拼音文字中的一个字等同。西方文字（如英文）中的一个字都是一个词的书写形式，比如kaiser是一个字。而现代汉语一个词的书写形式多数不是一个汉字，如“矛盾”中的“矛”不是词的书写形式。跟西方文字一个字相当的，是“人民”、“革命”、“布尔什维克”这类汉字组。

一个个汉字，其实只与西方文字中的一个个字母相当。理由如下：

第一，汉字体系中一个书写单位是一个汉字，而西方拼音文字中一个书写单元是一个字母。

第二，一个现代汉字来源于古代的一个表词字，而现代西方文字的一个字母也来源于古代的一个表词字。

第三，按铸造汉字所占用的面积来说，一个汉字也等同于西方文字的一个字母。

第四，西方文字以一个字母为基本单元跟语言要素发生联系；汉字体系以一个汉字为基本单元跟语言要素发生关系。

正因如此，以往考查文字性质时，都是把汉字、日本假名、英文字母这类符号放在同一层面上的。不这样就说不清楚。如“花”中的“艹”只指示语素的意义，“化”只表示语素的声音，两者都不表达语素；表达语素的是“花”字整体，总之，考察汉字表达语言中何种要素时，还应着眼于一个个汉字，而不是字符。

我们对裘锡圭先生的对于文字的相关理论进行了概括：

- ① 字是记录语素的符号，一个语素被一个字记录。如果两个符号记录的语素不同，那么这两个符号就是不同的字；如果两个符号记录的语素相同，那么这两个符号就是同一个字。
- ② 字是最小的书写符号，一个字就是一个方块字。如果两个方块字的字形相同，它们就是同一个字；如果两个方块字字形不同，它们就是不同的字。

我们发现裘先生在自己论述中实际上是只分析了汉字这个特殊的文字系统，而不是语言学中的文字。上述总结的第一点中提到说文字是记录语素的符号，实际上是说字携带了语素，这种视觉形式的意义是人为给予的，文字是书面语符号，笔者将其称为语位（将在下文中解释），笔者得到这个观点的原因是：语言符号的根本形式是它的语音，真正的符号是口语，用文字形式把口语从听觉变成视觉，形成书面语，书面语才是真正的符号，文字此时与语音相对应，是语言的形式之一。从书面语角度出发，裘锡圭（2020）所提到的“字”的第一种性质一样地它携带了意义，它是符号。

下面，我们简单讨论文字，来看看汉字学家如何认识汉字的。王宁（2014）的相关论述：

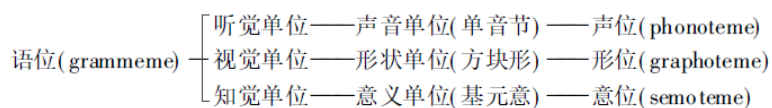
使用状态的汉字存在于记录汉语的文本里，带有语言环境，具有言语意义；其中记录当代文本的汉字处于动态当中，具有一定的流动性，字数、字频、覆盖率都会有少量的变动不居，也就是说，会有少部分汉字受社会语言运用的影响时进时出或存在频度变化。贮存状态的汉字存在于历来的词典、字书里，这些汉字虽然也都是从使用者或使用过的汉字中收集后编排起来的，但它们脱离了原有的文本，也就脱离了语言环境。它们依据辞书的体例聚合在一起，成为相互依靠的一群，并且一般都有形、音、义属性随之显现。中国的词典、字书不论如何编排，都是以汉字作为类聚的标志；也就是先贮存了汉字，才能贮存词汇及其音、义。很显然，汉字的使用状态与贮存状态有很大的区别。在贮存状态中处理汉字，比之在汉语文本中使用汉字，要复杂得多。

词是语言中最小的可以独立运用的意义单位，是音与义的结合体。字以它的字形将已具备语音、语义的词记录下来，因而成为形、音、义的统一体。由于古代汉语在相当长的历史时期内是以单音词为主的，训诂材料又是针对书面文献的，因此我们所见到的用汉字记录下来的训诂材料，绝大部分一字即一词，所以在古人心目中，对字词没有明确的区分。例如在重视名实之辨的先秦时代，人们思维中反映客观事物的概念、语言中表示概念的词、记录字的字被统称为“名”；在《说文解字》中，许慎又以“词”仅指虚词；前代训诂中则多以“字”的概念表示现代语言学意义上的“词”。

这里我们会发现此刻我们谈论的汉字的“字”它一直是以“形+词/语素”的形式出现的。王宁（2014）提到的所有的“字”是和裘锡圭（2020）所提到的“字”的第一种性质均是记录语言的一种单位，也就是以“形+词/语素”的形式，而词/语素是音、义的结合体，进而联系到形，会发生的是形与音、义的高度统一，笔者将王宁（2014）提到的所有的“字”和裘锡圭（2020）“字”的第一种性质称为语位（grammeme）。

2 语位（grammeme）及其性质

语位（grammeme）运用到汉字研究，是鲁川（2002）^[1]最先提出的，他解释的语位（grammeme）是：汉族人的听觉敏感单位是“音节”，所以汉语的基元单位应该定位在“单音节”上；汉语具有“以字形提示意义”的功能，所以汉语的基元单位要考虑“视觉形象”；汉字往往有多个“义项”，所以汉语的基元单位应该定位在“汉字的一个义项”上。并且给语位（grammeme）给出了自己的观点，见下图：



鲁川的相关论述实际上是向我们引出了语位，但是他的相关论述没有很明确的分析汉字的性质和汉字的基元单位——语位的本质，这和现在我们对于汉字本质的认识不足是有一定关系的。笔者在系统讨论语位（grammeme）前，我们先将符号、符号与语言关系进行总结。

赵毅衡（2012）给符号²下了一连串的定义：“符号是携带意义的感知：意义必须用符号才能表达，符号的用途是表达意义。反过来说，没有意义可以不用符号表达，也没有不表达意义的符号。”“意义就是一个符号能被另外的符号解释的潜力，解释就是意义的实现。”“意义必用符号才能解释，符号用来解释意义。反过来，没有意义可以不用符号来解释，也没有不解释意义的符号。”“符号就是意义，无符号即无意义，符号学即意义学。”符号学的讨论实际上也是回答文字学一直以来的“文字是记录语言的符号，符号具有意义，所以文字具有意义。”这个理论逻辑悖论，下面我们通过符号与语言关系再次讨论。

在符号与语言关系上，我们要清晰的认识到³：（1）符号和符号以外的事物。它们都是实体：形式和内容结合的实体。但是符号是把符号以外的实体当做自己的内容。（2）语言符号的两种形式。语言符号的根本形式是语音，但是语音不是符号，真正的符号是口语。用文字形式把口语从听觉变成视觉，是书面语符号，其中的文字不是符号，只是符号的形式。（3）书面语对口语的“强化”。应该是书面语在把口语变成视觉形式以后，可以修改得更加简洁流畅，可以反复和跳跃阅读。

依据符号与语言关系，笔者认为语位有如下特性和定义：

首先，我们明确将王宁老师提到的所有的“字”和裘先生“字”的第一种性质称为语位

（grammeme），依据鲁川（2002）常康杰（2021）的相关分析，笔者提出语位

（grammeme）是形体态与认知态高度统一的最小又自由的语言单位，是书面语的符号系

1 鲁川. 汉语的“语位”[J]. 语言教学与研究, 2002(04):9-18.

2 赵毅衡. 符号学[M]. 南京: 南京大学出版社, 2012.

3 本部分是笔者私下请教湖南师范大学彭泽润教授而得到的彭老师观点，不能代表彭老师观点的全部，如有错漏是笔者的责任与彭老师无关。

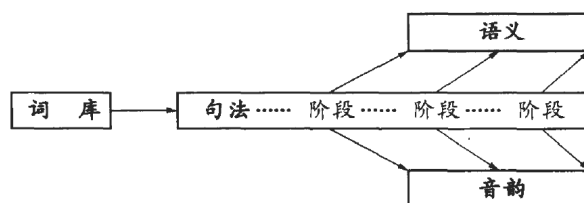
统，是语言生成机制与语言储藏传递机制平行运用的产物。

3 生成语言学语言机制下的语位化理论

下面我们来谈一下生成语言学中给语言机制下的定义，实际上邓思颖（2019）^[4]提出：根据最简方案的设想，语言机制主要由“认知系统”（cognitive system）组成。认知系统包括了人类语言的“运算系统”（computational system）和“词库”（lexicon）两个部分。词库是储存词的地方，标明词的特性，也应该是负责构词的部门——词法。运算系统又称为“创法部门”（syntactic component）或“狭义句法”（narrow syntax）。

我们说词是音、义的结合体，那么音、义属于语言机制吗？实际上，邓老师在引形式句法的几位名家的观点^[5]中也提到这个问题：Chomsky（2000）以及Hauser, Chomsky and Firth（2002）等倾向于认为音和义的本质不属于语言机制的一部分。所以，实际上由音、义经过语言机制各部门加工后的产物才属于语言机制，这个产物实际上最终储藏在语言机制的词库（lexicon）中，我们称这个产物为认知态（Cognitive state）。

具体语言机制如下图^[6]：



在最新的分布形态学(Distributed Morphology,DM)理论中，取消了词库，并认为句法运算是在运算计算原子后进行排列的，而是直接在句法运算时生成的，胡旭辉(2021)充分的解释DM模式下构词相关的理论：

如果词库中没有词，并且词库不负责生成词供句法运算使用，那么词库中究竟包括什么呢？分布形态学理论（Distributed Morphology，DM）认为词库的内容是：

列表1：抽象语素⁷（abstract morpheme，梅广（2019）翻译为“成品词素”）与词根（root）（句法终端节点成分，句法运算的对象）

列表2：词项（vocabulary item）（句法推导的语音实现）

列表3：百科知识（encyclopaedia）（句法推导的语义实现）

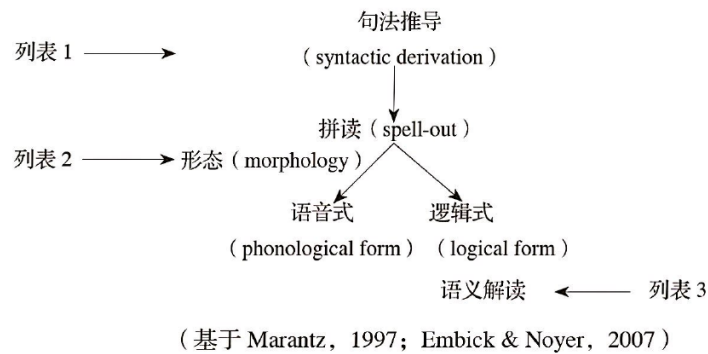
以上词库列表内容在语言结构生成的不同阶段被提取，依据此我们可以得到分布形态学理论（Distributed Morphology，DM）模式下T模式图示：

4 邓思颖著.形式汉语句法学(第2版).上海：上海教育出版社,2019.04:10.

5 邓思颖著.形式汉语句法学(第2版).上海：上海教育出版社,2019.04:12.注释[2]

6 邓思颖著.形式汉语句法学(第2版).上海：上海教育出版社,2019.04:12.

7 这里的语素并不传统结构主义语素



根据这个图式，笔者认为：实际上，DM模式下，语言生成机制生成的是语言序列根据一定条件进行的排列，我们将其称为语言认知态（Cognitive state），而书面语系统中的形体特征则是依据一种特殊规则进行的形状组合生成的产物，进而我们也可以将其称为语言形体态（Structure state）。据此，我们给予语位一个新的定义——**语位是语言形体态与语言认知态在一定条件下进行指称功能的激活与使用后，来完成语言储藏传递的语言单位。**

进而，我们在常康杰（2021）的基础上提出修正语位化理论（The theory of Revise Universal Language Grammeme）取消了音位、义位和形位的说法，取消词库；认为句法运算是在运算计算原子后进行排列的，认知态不再进入词库，而是直接由句法运算时生成的，在需要记录时，再进行语言储藏传递机制（Language Storage-transfer faculty）映射激活并使用的，而且语言生成机制与语言储藏传递机制在需要记录和传递语言编码序列时是并行运算的。

修正语位化理论中认为语言机制包括语言生成机制（Language generation faculty）和语言储藏传递机制（Language Storage-transfer faculty），为了不与形式句法学中语言机制概念相混，我们称修正语位化理论中的语言机制为语言状态机制（Language generation state）。

下面我们将语言状态机制（Language generation state）中的语言生成机制（Language generation faculty）和语言储藏传递机制（Language Storage-transfer faculty）进行解释：

语言生成机制（Language generation faculty）就是形式句法学中的语言机制，这里语言生成机制实际上产生的是语言的口语。

语言储藏传递机制（Language Storage-transfer faculty）是指语言认知编码序列（下文称认知态序列）总需要通过一种视觉平面的，可以跨距离传递的储藏形式进行自然传递。

这是语言是交际工具的本质导致的，原因是语言生成机制运行后生成的认知态序列（也就是词组，句子）在历史上因相关原因（技术，通用语的弹性过大等原因）无法储藏传递，进而影响语言的交际，所以就必须使用一种可储藏传递的形式，来进行认知态与认知编码序列的记录与储藏，所以实际上语言储藏传递机制机制产生的是语言的书面语。

语言储藏传递机制由**编码储藏，传递，储藏解码**三个部分组成。

编码储藏部分分为生成储藏阶段，编码记录阶段，其中视觉平面生成阶段与语言生成机制中的词库（词法）平行运行，储藏记录阶段与语言生成机制的运算系统（句法）平行运行。

视觉平面生成阶段时，句法推导出的认知态与视觉平面的形体特征在一定条件下进行指称功能的映射激活，进而产生形体态，注意此时认知态是由自然存在的句法推导生成的，是已经存在的，形体特征是依据认知态的某些特征进行的视觉平面编码，认知态和形体特征都是已经存在的，所以这种视觉平面编码只不过是依据一种规则进行形状的生成罢了，所以我们把这种视觉平面编码叫做形体态，进而我们给形体态给出了一个定义——**形体态是形体特征依据一种规则进行的形状生成**，其实我们在这个视觉平面生成阶段中产生的形体态就像一个不透明瓶子，在生成这个不透明瓶子同时我们将认知态装进这个瓶子里进行储藏，此时的这个装满的不透明瓶子就是一种形体态与认知态叠加状态，这种叠加的状态就是语位，所以我们说此时的这种语言状态形，音，义是一个统一体。

编码记录阶段，实际上是将生成储藏阶段产生的那个装满的不透明瓶子，依据运算系统（句法）进行排序的阶段，我们可以叫语位序列。

传递部分，其实是把装满的不透明瓶子排序（语位序列）进行的传递。

储藏解码过程时，我们把装满的不透明瓶子排序（语位序列）进行解码还原出不透明瓶子（形体态）和里面装的内容（认知态）。

所以，实际上我们研究古文字时，是看到的那个不透明的瓶子，我们考释古文字、同源词研究时实际上也是储藏解码的过程，进而笔者也提出了传统语文学的研究本质是语位。

4 语位的性质与标识

我们提出**语位是语言形体态与语言认知态在一定条件下进行指称功能的激活使用后来完成语言储藏传递的语言单位。**

下面我们将语位的性质予以介绍：

(1) 由于百科知识进行语义解读在逻辑式中是唯一的，所以语位（grammeme）具有唯一性；由于抽象语素与词根的结合具有统一性，且语位化理论认为形体态与之映射激活的认知态之间是一一对应的，所以只要认知态改变，无论形体态是不是一致，均为两个语位。

(2) 语位（grammeme）的生成性是指人的大脑中的不同语言状态机制映射生成语位特征（grammeme feature），然后语位特征（grammeme feature）经过分别通过语言生成机制和语言储藏传递机制生成形体态（Structure state）和认知态（Cognitive state），认知态与形式态经过语言状态机制运行后激活形成语位（grammeme）。

生成性的发现，说明语位统一体是会重组的，不是固化的。

根据性质，我们以第一人称的语位范畴中的“予、余、舍”为例提成语位的标识体系：

{_{第一人称}予、余、舍}

语位（grammeme）：{予}、{余}、{舍}表示；

形体态（Structure state）：[<予>],[<余>],[<舍>]表示；

认知态（Cognitive state）：[_{第一人称}<予>：/y⁵⁵/]表示；

由于认知态 (Cognitive state) 是由形体态记录的所以采用“[_{第一人称}< >:/ /]”表示, 而不使用{ }, 如果形式态的只是由一个形符组成的可以直接使用[]表示。

在不进行太深入的语位研究时可以只使用“{ }”语位符号表示认知态, “[]”表示形式态, 若形位超过两个时可以用“|”分隔形位, 但是必须提前声明望大家注意。

5 汉语语位化理论 (The theory of Chinese Grammeme) 建立

对话位定义我们在上文已经给出了, 下面我们具体的谈一下汉语语位 (Chinese Grammeme) 是汉语认知态 (Cognitive state) 和形体态 (Structure state) 经过语言状态机制形成的产物, 是语言书面语的符号。

汉语语位 (Chinese Grammeme) 具有特殊的可成范畴性, 可成范畴性是指汉语语位的形体态之间的关联性可以组成形体-语位范畴 (Structure-Grmmeme category); 由汉语语位的认知态 (Cognitive state) 之间的关联性组成认知-语位范畴 (Cognitive-Grmmeme category); 形体-语位范畴 (Structure-Grmmeme category) 和认知-语位范畴 (Cognitive-Grmmeme category) 可分别简称为形体范畴 (Structure category) 和认知范畴 (Cognitive category)。

语位范畴 (Grmmeme category) 是语位与语位之间具有形体态 (Structure state) 之间的关联性或者认知态 (Cognitive state) 之间的关联性所组成的最大的映射范畴。

下面给大家一组例子来更好的对汉语语位化进行理解:

{数数}语位的认知态_[查点数目]<数><数>: /ʃu²¹⁴ʃu⁵¹/, 形体态是[数|数], 在{数数}中{数₁}语位的认知态是_[成词后表一种计算数字的动作]<数>: /ʃu²¹⁴/, 形体态是[数]; {数₂}语位的认知态_[成词后表数字的总称]<数>: /ʃu⁵¹/, 形体态是[数]; {数₁}和{数₂}实际上是传统意义上记录语素的语位, 我们可以将{数₁}和{数₂}两个语位叫做{数数}语位的次语位 (subgrmmeme), 简称次语位, 此时在不研究形式态的学术讨论中可以使用字形、语素和词等术语, 研究形体态的学术讨论中则建议必须使用汉语语位化理论术语, 可以加注等价术语。

6 汉语语位 (grammeme) 与汉语传统概念的对应

6.1 汉语语位 (grammeme) 与传统语文学概念的对应

(1) 传统语文学中研究的“字”的本质是语位 (grammeme);

(2) 传统语文学中的文字学实际上研究的是语位 (grammeme) 中形体态 (Structure state) 的组成生成的变化和形体态 (Structure state)、认知态 (Cognitive state) 之间的关系;

(3) 传统语文学中的训诂学实际上研究的是分析语位与语位之间的语言范畴问题, 共分为形体范畴 (Structure category) 和认知范畴 (Cognitive category), 前者是传统学术上, 字形源流的同源, 后者指的词源上的同源。

(4) 传统语文学中的音韵学实际上研究的是语位 (grammeme) 与语位范畴 (Grmmeme category) 的音系问题, 构拟的古音实际上就是构拟语位 (grammeme) 的音韵特征。

(5) 造字是形体特征最后形成形体态的过程，实际上是语位的可成范畴性与唯一性的结合。因此对于一个汉语语位的构形分析，不只是为了根据形体态，还必须结合其认知态，这是语言储藏传递机制的解码过程。

6.2 汉语语位 (grammeme) 与现有带字、词概念的对应

(1) 语位的认知态就是词 (词项)；形体态就是字形。

(2) 同音字是语音式相同，逻辑式不同且形体态不同的语位集合，可以叫同音语位组；

(3) 同义词是逻辑式相同，语音式不同且形体态不同的语位集合，可以叫同义语位组；

(4) 同源词是语位范畴 (Grammeme category)，其中可分为形体-语位范畴 (Structure-Grammeme category) 和认知-语位范畴 (Cognitive-Grammeme category)。

7 汉语语位化理论下对汉语词的分析

下面，我们以“人民”为例来介绍语位化理论的实际应用。

{人民}是一个语位，讨论{人民}中的[人]和[民]，此时的[人]和[民]好像只是形式态和两个音节，但是其实不然依据生成性，它们会映射出不同的义位出来，重新构成两个新的认知态，在组成两个次语位{人}与{民}，这就是语位是语言状态机制中最小又自由的语言单位的概念之所在。

下面我们讨论{人间}和{人民}两个语位，这里的人、间，人、民四个语素可以由认知态也就词的理论进行推解，但是的人、间，人、民四个语素是由三个形式态组成的，但是要注意的是最终组成的{人间}的次语位{人}和{人民}次语位{人}并不是一个次语位，因为他们的此时的认知态是不同的。

根据唯一性，实际上我们可以对语位进行鉴定，看形体态和认知态是不是一致的，但是我们要注意的是汉语认知态的词汇意义具有显性，而抽象语素结合后的功能意义是隐性的，此时我们汉语词的词汇与功能意义是叠加。在制图理论中实际上也将功能词的显性与隐性作为普遍性中的参数进行处理，这个现象可以称为汉语词的词汇、功能意义叠加，这里要注意一个点，我们所说的词的词汇与功能意义叠加是在语言生成后，通过与形体态激活使用后，在语言储藏传递机制中储藏环节产生的，其实我们所说的词所具有的词性就是语言储藏传递机制中储藏环节的产物，其实就是“词典词”，这是人为创造的，并不能说“词典词”就是词库，而与之形体态激活后产生的语位则起到储藏编码的作用，在语言表达中语位也起到记录作用，这就是王宁 (2014) 提出的“汉字的使用状态和储藏状态”的本质。

对于汉语词的词汇、功能意义叠加观点，我们认为每个词的功能意义实际上在该词中处于隐性作用，而句子中所显现的功能词作用其实是在句子有语言生成机制加工后生成的，经过语言储藏传递机制后由接收信息的人在认知解码过程中，将语位序列之间这类处于隐性特征体现出来的，这种情况现在叫作语法化。

当然由于语位化理论是建立生成语言学上的，具体一定的普遍适用性，所以同样的分析，也适用于世界任何语言的书面语系统。

8 综 述

陆俭明先生曾经对笔者说提出新的思路，寻求新的解释，这很好；但必须注意，提出新的思路与解释，必须是实实在在的，不能仅仅是提出个新名词。笔者在汉语史与文字学研究中，已经部分使用了语位化的思路，陆老师的一句箴言，也促进着笔者在语位化思路上的探索，关于“字本位”与“词本位”的讨论由来已久，笔者希望通过本文提供一个新思路，供学界讨论，同时语位化思路也适用于任何语言的文字，进而这种思路的提出也是汉语文字学向普遍文字学研究的一次探索。

注 释

- 1.本文先后交予北京大学陆俭明先生、台湾师范大学赖贵三先生、上海师范大学潘悟云先生，郑州大学李运富先生，江苏师范大学金立鑫先生，浙江大学程工先生，北京外国语大学尹洪波先生，寻求诸位先生的意见，其中北京大学陆俭明教授建议将本文的语位改为文位，在此表示感谢，其余诸位先生也为本文的成文给予笔者一定的帮助。
- 2.这里我们说语位是形、音、义的统一体，统一体指各种矛盾或矛盾的各方在一定条件下形成的事物整体。

参考文献

- [1] 常康杰.形式句法学基本理论（未刊稿）,2022.
- [2] 常康杰.语位化及汉语语位化的研究[C].浙江省语言学会第20届年会书面报告,2021.
- [3] 邓思颖.形式汉语句法学(第2版)[M].上海：上海教育出版社,2019.04:10.
- [4] 胡旭辉.生成构式语法理论：构词、事件结构与名词结构[J].语言学研究,2021(31):52-64.
- [5] 刘馨茜.分布形态学理论与最简方案[J].海外英语,2019(21):249-251.
- [6] 裘锡圭《文字学概要（修订版）》[M].商务印书馆,2020.
- [7] 王宁.《通用规范汉字表》与辞书编纂[J].辞书研究,2014(03):1-8+93.
- [8] 张玉金,夏中华著.汉字学概论[M].南宁：广西教育出版社,2001.01.