

标题：叶典第一次会议——有关新两分的初步讨论

日期/时间：2021年3月18日 21:30~23:00

与会人员：白易、李佳宸、沈天珩、林成世、姜兆勤、葛炳仑
(排名不分先后，按姓氏笔画排序)

撰稿人：沈天珩

页数：4

目录

- 1.会前讨论
- 2.迭代原则
- 3.“莽”、“辩”相关字形的特殊拆分
- 4.删除上部左右相同取中的特殊规则
- 5.镜像字、倒转字的处理方式
- 6.最终结论

1.会前讨论

按原“首部正向取大”原则而不考虑其他，沈认为“藁”字应被拆作“草，人”。然而此拆法有失直观性，沈本人并不推荐此拆法作为主码，同时李亦持相同观点。关于此问题，经讨论后一致决定，在对汉字进行拆分时不能破坏其中包含的独体字结构，如“藁”字中的“木”即为独体，此时不应被分为“十”和“人”而分别归至首尾部。

而对于独体字的定义，则参照国家教育部、国家语委组织研制的《现代常用独体字规范》，以此为基础进行繁体与异体的适当增补。

新两分的基本原则确定为不考虑字义，只考虑字形，在此基础上进行“首部正向取大”与“尾部逆向取大”。如“鸿”应当取“汙，鸟”，而“汙，鸡”则可作为容错码出现。“屨”等字同理，字形拆分“屎，女”作为主码，而字义拆分“尸，娄”作为容错码。此举意在将新两分的拆字原则规范化，同时兼顾以遵循字义为主的常规拆分避免查不到、打不出字的情况出现。

此外，由“首部正向取大”原则，林提出将原两分中“戠”与“幾”的首部“戈”分别改为“从”、“么”，获得一致同意。林原先还提出将“戠”的首部定为“丛”，因其较“从”更大，但因不能破坏独体字“戈”的结构的原则被否决。

2.迭代原则

由于两分字元限定为常用字或常用部件(具体哪些字将由日后会议再讨论决定)，因此经常会存在首尾两部无法将整个字的字形取完的情况，如“籛”作“台，林”，非“辶，林”或“臬，彡”，因其均存在非常用字。会议决定，将“首部正向取大”与“尾部逆向取大”的描述规范为一条“迭代原则”，具体操作即先将一个字一分为二，如果分开之后的“首尾部”中存在非常用字，则首部将取其首部，尾部将取其尾部，即“首部的首部”与“尾部的尾部”，并以此类推。如“籛”字，若先分为“辶，林”，由于首部“辶”非常用字，故再取其首部“台”得到“台，林”；若先分为“臬，彡”，由于首尾皆非常用字，因此取首部的首部与尾部的尾部结合得到“台，木”。两者相比，由“尾部逆向取大”原则，定义“台，林”为最终拆分。

3.“莽”、“辯”相关字形的特殊拆分

由于首尾部相同而中间不同的字过多，若“辯”系列的字统一拆作“辛，辛”则会导致重码暴增，对于该系列的字，经讨论一致同意保留原两分的规则，即首部取两边而尾部取中间。对于该特殊拆分的描述，林首先提出“首尾部相同且字形未用完，则尾部取中间”的规则。

对于“噐”字，与会人员一致同意拆分为“口，页”。而对于“口，页”的来源却说法不一。林认为按该描述先得到“口，口”，首尾相同，则尾部变为中间由“口、页、口”三部分组成的部件。显然该尾部非常用字，于是按迭代原则再取其尾部。由于中间的“口、页、口”亦符合该特殊拆分的原则，因此尾部应取中间的“页”，最后得到“口，

页”。

而对于“啮”、“罍”此类结构的字，若按该描述则将变为“口，氏”、“口，一”，受此影响而有失直观性的字存在不少。虽林认为按其“激进派”的观点将“罍”拆作“口，一”亦能接受，但其余与会人员反对该拆分。经讨论决定该类字仍需保留“口，口”的拆分，于是林的描述便存在漏洞。

此外，葛提出是否应将“器”拆作“口，犬”，沈表示反对，因按首部正向取大原则此处存在一个更大的常用字首部“哭”可以使用。其余与会人员同意此观点，保留“哭，口”。因此当且仅当无法按首部正向取大原则取出常用字首尾时才适用此特殊规则。

最后经讨论得出的描述应当为：当一个字为左中右ABA结构或上中下ABA结构时，取“A，B”作为其初步拆分，如若存在非常用字，则继续套用迭代原则。因此对于“器”字的拆分，其应先拆作“皿，页”，由于首部“皿”非常用字，再取其首部“口”最终得到“口，页”。

4.删除上部左右相同取中的特殊规则

原两分中存在一条规则，当一个上下结构的字上半部分为左中右结构且左右相同时，首部取中间部分。例如“樂”字应取“白，木”。而原两分在这一点上的实际拆分上较为混乱，输入法码表中可以清晰看到“樂”的实际拆分为“么，木”，但“樂”的拆分却又为“白，木”，规则不统一。虽按笔顺“樂”字也应是先写“白”，但林提出对于生僻字、尤其是结构复杂的字而言根本不可能提前预知笔顺，故如同“不知道生僻字的字义而无法对其进行字义拆分，只能从字形入手”一般，一致认可不得将笔顺作为拆分时考虑的因素。

由于此类字并不多，本着尽量将拆分规则简单化的原则，最终决定将该特殊规则删除，“樂”仍保留与输入法码表中相同的“么，木”，而“樂”则需改为“糸，木”。“鬱”一类字同理。对于“學”字，则将首部定为“白”，因其可以作为“白”的形近变体从而认同至“白”。与之相关的简体的“学”字，原先是将“灬”作为常用部件单列，但考虑到简繁的一致性，会议决定取消该部件作为常用部件，将“学”字拆为“小，子”，其中“小”为“灬”的可认同变体。

5.镜像字、倒转字的处理方式

对于扩展区的少数存在镜像或倒转的部件的情况，例如“𨾏”、“𨾏”等字，与会人员一致认为应将其视为未镜像、倒转的常规字进行处理，如上两例则分别应拆作“邑，邑”与“見，鬼”。而林提出该描述不够严谨，否则“后”与“司”等镜像字亦将会被包含其中。因此对该规则的描述最终添加了一个“存在非常规笔画”的前提条件。

6.最终结论

对于本次会议的最终结论，由姜整理 并得到一致通过。现将最终结论总结如下：

- 1、规范独体字：不拆为主码
- 2、非规范独体字：拆为主码，拆时不破形
- 3、非独体字：
 - 3.1、迭代原则，迭代于3.2~3.4
 - 3.2、首尾两分，其中首部取认识的最大部件为主码
 - 3.3、半包围，先外后内
 - 3.4、首尾一样取中间(ABA型取AB)
 - 3.5、取尽有冲突，首部取大为主码

(End of document)