

可乐听语

期刊 & 随想 · 总编

Colating Language Periodical & Capriccio Summary

R4-5

1

Total Issue 1

2021 · Irregular

(Open Version)

Ver: 0.1.0.6

Created: 2021/08/25 Update: 2021/10/12

Author: David Email: colating@yandex.com

[编号 No.]: Doc No.: Per 2021-1 / Cap 2021-1

内容说明: Open Version 只涉及 Colating 的实现原理。

[总目录 / ALL CONTENTS](#)

Note: When without Rx, It is mean that is the official release version.

总目录 (期刊 & 随想)

期刊 & 随想 · 总编.....	1
总目录 (期刊 & 随想)	2
期刊 & 随想 · 前言.....	4
阅读前置说明.....	4
Colating 阅读快速入门.....	5
可乐听语 · 期刊.....	25
■ 声滴算法原理 · 系列/ VOICE.....	26
世界语拼音原理分析.....	26
英语拼音原理分析.....	33
简明声滴算法原理.....	47
■ 概项联结框架 · 系列.....	65
通用概念表示项联结框架总论.....	65
■ 语文失效机理 · 系列.....	73
词法、句法、语法的失效机理分析 (待完成)	74
☆. 拉丁语概念字母构造长度的失效机理分析.....	76
随想录.....	78
■ 语文理论 · 系列.....	79
☆. 人类信息表达系统的 Colating 分类方法.....	79
■ 概念系统 · 系列.....	81
☆. 通用概念系统的设计步骤.....	81
■ 构词法 · 系列.....	84
☆. Colating 单词的分段构造模型图.....	84
☆. 同音字原理.....	86
☆. Colating 构词法思想指引.....	87
☆. Colating 单词的 3 种书写格式.....	90
■ 语法 · 系列.....	92
☆. “白马非马” 浅析.....	92
☆. 宙沏概念对象的种类与感知.....	98
☆. 限长多滴声节 1 维图言概念标记的本质.....	99

☆. 莫尔斯电码、手机号码与拉丁语单词的系统设计比较.....	102
APPENDIX 附录.....	109
APPENDIX A.....	110
语音类与声符类的概念.....	111
图形记号类.....	116
概念类.....	118
APPENDIX B.....	122
缩写。系列.....	122
APPENDIX C.....	124
☆. Colating 输入法.....	124
☆. Colating 的字母与音标.....	128
APPENDIX D.....	132
Colating 撰写规定的说明.....	132
APPENDIX X 名词索引.....	133
名词索引.....	133
APPENDIX X1.....	136
修改记录.....	136

表达思想的几种方式：哲理论证假定式，科学数据定义式，脑洞大开说教式，数学逻辑推理式。

[总目录 / ALL CONTENTS](#)

期刊 & 随想 ◦ 前言

在未来的一段时期，Colating 将以不定期的期刊和随想录的形式发布。

虽然 Colating 作品是为通用语文系统的创造而著作的，但是，如果能够理解语文创造的基本哲学原理，明白人类通用语言文字的创造原理，那么，对于语文的学习，是不是会有??帮助呢？

与绝大多数的语文作品不同，本作品将试图从无限哲学思想的高度，来揭示人类通用语文系统创造的原理和本质。

David

2021/08/25

(本文结束)

自动目录: [期刊_随想_总目录](#) [独立目录 / All Single CONTENTS](#)

阅读前置说明

本作品为拉丁语族著作，在文中的某些地方，可能会由于没有增加“拉丁语场景”之类的限制，从而导致在理解上的歧义。因此，此时需要按照拉丁语思维来理解这些内容。

本作品默认采用 10 进制计数系统作为标准计数参考系统，并以此来对其他进制的计数系统展开描述。

更多，请参阅：[Colating 撰写规定的说明](#)

(本文结束)

自动目录: [期刊_随想_总目录](#) [独立目录 / All Single CONTENTS](#)

Colating 阅读快速入门

Ver: 0.1.0.4

Created: 2021/09/24 Update: 2021/10/8

This article: A, Donor: Colating Group

1. Colating 是什么？

Colating Language 是一种经过重大改良、并具备可持续迭代改良能力、先进的通用拉丁语族新成员。

与英语、俄语、法语、德语之类的语言一样，都是属于同一类型的通用拉丁语言文字的智慧产物。

Colating 试图包容并蓄了其他通用语文的优点，并依照自创的语文理论理论来进行统筹与综合，这是一种面向未来的拉丁语族人类通用语言文字系统。

2. 与众不同的特点

Colating 的发明，既是一种通用语文系统的创造过程，也是一种通用语文哲学思想的诞生过程。Colating 总是试图建立在系统化语文哲学思想之上的产物。

人们可能并不需要去学习 Colating，但是，通过阅读 Colating 语文创造方面的哲学思想及理论，或许可以获得对其他语文学习的一些？？启示。

3. 世有数千语种，却难觅 1 本关于通用语文创造方面的系统理论

3.1. 世界上，人类通用语言的种类

网摘：世界上究竟有种多少语言？《大不列颠百科全书》称有“近千种”；世界最畅销书籍《圣经》的外语译本截止 1997 年就达 2197 种，语言种类多于文字种类，《圣经》至少证明世界语言种类应多于 2197 种。

迄今比较精确的统计数字来自人类学家，他们通过民族研究发现世界语言有 6809 种。

C 注：有些语种还处在比较原始的状态，是没有文字的。

3.2. 有创造，就会有新概念的诞生

由于通用语文创造方面的系统理论匮乏，因此，在 Colating 创造的过程中，必然就诞生了许多的新概念。在理解了这些基础的新概念之后，就能够顺畅的阅读本作品了。

请参阅：[阅读入门需要首先理解一些新概念](#)

4. 拉丁语单词的本质

4.1. 世界语

4.1.1. 世界语属于拉丁语族的一个分支

拉丁语音标声符单词，本质上就是一个声节。拉丁语字母图符单词，本质上就是一个优选数值

4.1.2. 世界语单词分析

世界语主要采用融合拼音算法（融合声滴拼读算法），声滴的理论压缩比为：143/28。

4.1.2.1. EO 音标声符单词

世界语声语声语单词的声节音标流发音（声语号/声语概念标记号），实质上是一个个在目标概念系统中独一无二的、限制最大可变长度的语音型音乐的声节，系统展开声滴总量的理论值约为 143 个。

需要注意的是：在音乐学上，音滴是非压缩的，1 个音滴对应于 1 个音符。在语文学上，构成声语号的声滴，是经过拼音算法压缩之后，采用压缩声滴来与声符（音标）相对应。

换言之，世界语要表示 143 个系统展开声滴，却只有 28 个声符，该怎样来处理呢？办法就是采用 1 个声符，以及 2 个声符的拼读组合来表示它们。理想状态是所有的辅音与所有的元音都能够有效地拼读，这样 28 个声符，在理论上就可以表示到的声滴总量为：23 辅音 x5 元音 +28 音标 =115 +28 =143 个。

4.1.2.2. 字母图符单词

世界语图言单词的构成字母图符流符号（图言号/图言概念标记号），实质上是一个个在目标概念系统级中独一无二的、限制最大可变长度的数学优选数值，该数学数制计数系统的基数为 28。

4.1.2.3. 声滴算法

将 28 个压缩声符（雷同 IPA 音标）映射到 28 个实际构词字母上，世界语的这种构词法，使得世界语单词能够做到“见词能读、听音能写”的效果。

4.1.2.4. 映射关系

EO 容易掌握的映射：

- 声滴与声符：利用融合拼音算法，将 143 个展开声滴压缩成 28 个压缩声滴，并与 28 个声符形成一一映射。
- 图粒与图符：直接转换，EO 将 28 个构词图粒直接转换成 28 个字母图符。

注 1：在单声滴声节概念系统中，图粒分为规则图粒和无规律图粒，规则图粒称为构词图符，简称图符；无规律图粒，称为构词图形。

声符与图符：一一映射，28 个音标声符映射于 28 个字母图符。

注 2：EO 为 Esperanto（世界语）的缩写。

4.2. EN 语

4.2.1. 英语属于拉丁语族的一个分支。

拉丁语音标声符单词，本质上就是一个声节。拉丁语字母图符单词，本质上就是一个优选数值

4.2.2. EN 语单词分析

4.2.2.1. EN 音标声符单词

英语声语单词的声节音标流发音（声语号/声语概念标记号）：实质上是一个个限制最大可变长度的语音型音乐的声节，或声符节，系统展开声滴总量的理论值约为 298 个。展开声滴相当于音乐学中的音符，不过数量达到了 298 个。

4.2.2.2. 字母图符单词

英语图言单词的构成字母图符流符号（图言号/图言概念标记号），实质上是一个个在目标概念系统中独一无二的、限制最大可变长度的数学优选数值，该数学数制计数系统的基数为 26。

4.2.2.3. 声滴算法

英语主要采用融合拼音算法，声滴的理论压缩比为：298/34。

4.2.2.4. 映射关系

- 展开声滴与标准压缩声滴的数量：EN 系统展开声滴总量的理论值约为 298 个，在经过压缩拼音算法之后，获得 34 个标准压缩声滴。
- 音标表的声符总量：标准压缩声滴总量 34 个 + 元音标压声滴连读组合 8 个 + 辅音标压声滴拼读组合声符总量 6 个 = EN 音标表全部混合声符总量 = 48 个。
- 各种声滴的混合，就构成了 EN 的音标表。EN 音标表的全部混合声滴总量 48 个，分别映射于 48 个声符。
- 图粒与图符：直接转换，EN 将 26 个构词图粒直接转换成 26 个字母图符。注：参见 4.2.2.1. EN 音标声符单词
- 难以掌握的映射，声符与图符：EN 主要表现为“1 对多”，“多对 1”的映射关系，实际 42 个声符压缩映射于 26 个字母图符，这点比世界语要复杂。
- 与世界语声符与构词字母的一一映射不同；EN 语将实际 34 + 8 个声符，采用不同的算法，一起混合映射到 26 个构词字母上，英语的这种构词法，使得大多数的英语单词，必须通过单纯的、简单的、额外的字母发音提示，才能完全做到“见词能读、听音能写”的效果。

4.3. EN 语与 EO 世界语的对比分析

4.3.1. 声语概念标记号

英语声语概念标记号采用 298 个声符构成的语音型音乐音符系统，也等同于 298 进制的计数系统，因此，声语号的构词效率比采用 143 进制的世界语要高。虽然 EN 更符合语文的实际应用场景，但由于它是在语文历史进程中不断累加创造的结果，因此，与世界语不同，其大多数的词汇，缺乏像世界语那样无需母语环境的、简单的、可学习的规律性。从这一点上来说，这是世界语简单性与英语合理性所发生的矛盾

4.3.2. 图言概念标记号

英语图言概念标记号采用 26 进制计数系统，因此，表面上，图言号的构词效率比采用 28 进制的世界语要低；但实际上，并不是什么样的数值都能作为图言号，图言号的总量与展开声滴的总量、拼读规则等，都有关系。受构词规则限制，实际上，世界语的构词效率应该比 EN 语要低？

4.3.3. 映射关系对比

4.3.3.1. EO 语声滴与字母映射系统的运算步骤:

- 第 1 步，第 1 次压缩，利用声滴压缩算法，将展开声滴转换成标准压缩声滴：EO 系统展开声滴总量的理论值约为 143 个，在经过压缩拼音算法之后，获得 28 个标准压缩声滴。
- 第 2 步，获得音标表中的声滴总量：等于标准压缩声滴，为 28 个。
- 第 3 步，EO 音标表中的声滴与音标表的声符，形成一一映射。
- 第 4 步，28 个声符，与 28 个构词字母形成一一映射。

4.3.3.2. EN 语声滴与字母映射系统的运算步骤:

- 第 1 步，第 1 次压缩，利用声滴压缩算法，将展开声滴转换成标准压缩声滴：EN 系统展开声滴总量的理论值约为 298 个，在经过压缩拼音算法之后，获得 34 个标准压缩声滴。
- 第 2 步，获得音标表中的声滴总量：标准压缩声滴总量 34 个 + 元音标压声滴连读组合 8 个 + 辅音标压声滴拼读组合声符总量 6 个 = EN 音标表全部混合声符总量 = 48 个。
- 第 3 步，EN 音标表中的声滴与音标表的声符，形成一一映射。
- 第 4 步，第 2 次压缩包含了 2 种压缩算法，针对 34 个标准压缩声符的声符总量压缩算法，以及针对 8 个元音标压声符连读组合的声符长度压缩算法；它们一起与 26 个构词字母，主要形成“1 对多”和“多对 1”的映射关系。另外，6 个的辅音标压声滴拼读组合，实际上只具有聚类学习的意义，它们与辅音构词字母为一一映射关系。

4.4. Colating 可乐语

与 EN 语相比，构词法的规律性与扩展能力，做了专门的大幅增强和改进；同时，和世界语一样，能够可以做到“见词能读、听音能写”的效果；而且，在构词能力、规律性、高效性等方面，Colaing 设计的目标，是为了超越当今所有的进取型语文系统的，理论上，已经超越了 EN。

请参阅：[错误: 引用源未找到](#)

4.5. 三种典型的拉丁语对比

4.5.1. 世界语

- 声语号：将 28 个压缩声滴在解压缩之后，形成了大约 143 个的展开声滴。这 143 个展开声滴，类似于音乐学中的音符发音。

- 与音乐学中的音节相比较，世界语的声语号，实质上，是一个采用了约 143 个展开声滴为单位声符的、限制最大可变长度的、语音型音乐的声节。
- 图言号：采用了 28 个构词字母组成的图言号系统，实质上是 1 个数学上的 28 进制计数系统。
- 映射规则：28 个压缩声滴与 28 个音标声符，28 个音标声符与 28 个构词字母，设计规定为一一映射关系。
- 映射分析：由于压缩声滴与音标声符，音标声符与构词字母形成一一映射。因此，1 个具有系统唯一性的世界语声语号，也就是 1 个声节，在宏观的理论规则上，它与唯一 1 个图言号形成一一映射关系。

从数学的角度来看，EO 语的 1 个声语号，本质上就是 1 个数学 143 进制计数系统的优选数值；1 个图言号，本质上就是 1 个数学 28 进制计数系统的优选数值。

注意：与音节类似，并不是什么样的数值，都能够成为声节。

4.5.2. EN 语

4.5.2.1. 声语号

将 34 个压缩声滴在解压缩之后，形成了大约 298 个的展开声滴，声语号类似于音乐学中的音符发音。

与音乐学中的音节相比较，EN 语的声语号，实质上，是一个采用了约 298 个展开声滴为单位声符的、限制最大可变长度的、语音型音乐的声节。

4.5.2.2. 图言号

采用了 26 个构词字母组成的图言号系统，图言号，实质上是 1 个数学上的 26 进制计数系统。

4.5.2.3. 映射规则

- 在当代英语中，将 34 个标准压缩声滴，8 个元音标压声滴连读组合，6 个辅音标压声滴拼读组合，一起共同组成音标表声滴；它们分别一一映射于 48 个声符。
- 实际上，需要关注与图符进行映射的声符总量，只有 42 个。另外的 6 个辅音标压声滴拼读组合，与辅音构词字母是一一映射关系，因为简单且单纯，所以无需花费额外的学习时间。

- 6个辅音标压声滴拼读组合：既没有增加系统展开声滴的总量，也没有像8个元音标压声滴连读组合那样，在压缩之后再与图符号进行映射；因此，它们属于高频聚类，主要的目的可能只是为了方便学习而已。

4.5.2.4. 映射关系分析

在当代英语中，音标表声符与构词字母图符主要同时存在“多对1”和“1对多”的映射关系，这点要比世界语复杂，但这是宇宙现实所造就的客观结果。

- EN语的映射过程，可划分为2个阶段，3种映射关系。
- 第1阶段，第1次压缩，产生音标表
 - 与世界语相同，即采用“声滴压缩算法”，将大约298个展开声滴压缩为34个标压声滴。
 - 在当代英语中，将34个标准压缩声滴，8个元音标压声滴连读组合，6个辅音标压声滴拼读组合，一起共同组成音标表声滴；它们分别一一映射于48个声符
- 第2阶段，第2次压缩，声符与图符产生映射。在此阶段，世界语的声符与图符为一一映射，
 - 音标表标压声符总量的压缩映射：英语标准压缩声滴，有：12个元音标压声滴，22个辅音标压声滴。需要将一个 $12+22=34$ 个的标压声符，映射到26个构词字母上，这属于声符数量压缩的映射。
 - 音标表元音声符连读组合长度的压缩映射：将“双元音”、“三元音”等，映射到1个构词字母上，这种做法，有利于缩短图言概念标记号的长度。
 - 混合压缩映射：在当代英语中，将34个标准压缩声滴，8个元音标压声滴连读组合，6个辅音标压声滴拼读组合，一起共同组成音标表声滴；它们分别一一映射于48个声符。然后，再与26个字母图符形成混合映射。
- EN语混合压缩映射算法，实际上将42个声符压缩成26个，再与26个构词字母发生一一映射。由于这是1种在单词中没有明确标志的聚类压缩算法，因此，这导致了大多数的EN语单词，并不能像世界语单词那样，可以简单地做到“见词能读、听音能写”的效果。
- EN语还存在着少量的声滴扩展算法，从而形成了“1对多”映射关系。如声语号“英[əɪ; æɪ]”同时映射于“an”。

4.5.3. Colating 语

4.5.3.1. 声语号

单词的构成采用的是 C52++ 系统。

Cola 大约有 1000 个左右的展开声滴，构成了类似音乐学上约有 1000 个近似单音的音乐音符系统，因此，声语号的表达效率比 EN 语高。Cola 最小声滴总量： $24 \times 14 = 336$ 个；最大声滴总量： $26 \times 16 \times 3 = 416 \times 3 = 1248$ 个。

4.5.3.2. 图言号

采用了 $52+x$ 个构词字母组成的图言号系统，实质上是 1 个数学上的 Colating 型 $52+x$ 进制计数系统，因此，图言号的表达效率比 EN 语高。

同时，Cola 具有标志位，因此，在发音和释义上，具有可预测性。

4.5.3.3. 映射规则

由于高效的声滴数量有限，因此，需要提高高效声滴的利用率。

由于人类个体的记忆能力有限，因此，需要压缩系统的图言号数量。

● 综合所述：

- 在声滴与图符的映射关系中，虽存在着“1 对多”、“多对 1”等映射模式，这是无法避免的。Colating 承认了这一客观的事实。

因此，在单词系统的设计中，由于所有的映射关系要么被显式标记，要么在语法中有明确的规定，因而，理论上完全可以做到“见词能读、听音能写”的效果。

- 如果抛开容易学习的规律性规则，那么，与 EN 语相比，Colating 稍微复杂了些。但是，由于 Colating 具有明显的、可学习的、简单的规律性规则，因而就显得比 EN 要简单得多。👁️ 详细内容，请参阅：Colating 其他部分的内容。

4.6. 进取型通用概念系统的设计，需要注意的一些问题

4.6.1. 不能过份简单化

宙沏概念对象的种类和数量，具有无限性；这就决定了在设计进取型通用语文概念系统的时候，不应该盲目地追求的简单化，无视宙沏概念对象的客观实际，而选取牺牲概念系统的总

体性能的做法；特别是缩减可用原子图言概念标记号的总量等的重要基础性能。简单化与合理化，有时往往会发生矛盾。

4.6.2. 具有合理的概念系统级更新换代能力

正如现代人无法替代未来人进行思想一样，古代人也无法替代现代人来思想。因此，概念系统的作用，就是不断的发明创造，合理的新陈代谢中进行标记、传播、延续，以及作为对比与借鉴的素材。

5. 克服不良的思维习惯，理性认识概念标记号的本质

5.1. 不要误以为只有数学的十进制。数学 N 进制计数系统：基数的本质

N 进制计数系统的构基数造模型：在数学数制计数系统中，基数的本质就是包含 0 在内的连续的 N 个自然数。最小的基数为 0，最大的基数为 N-1。

5.2. 电话号码，习以为常的 10 进制

采用数学 10 进制计数系统的电话号码，可以做到“见号码能读、听号码能写”的效果。

5.3. 世界语概念标记号（音标声节单词，单词构成字母队列）的声节与数值的本质

5.3.1. 世界语声语号的本质

将 28 个压缩声滴在解压缩之后，形成了大约 143 个的展开声滴。这 143 个展开声滴，类似于音乐学中的音符发音。因此，世界语的声语号，本质上，就相当于是音乐学中的音节，在 Colating 中，称之为“声节”。

5.3.2. 世界语图言号的本质

类似于 10 进制的电话号码计数系统，采用 28 个构词字母系统的世界语的图言号，实质上，就是在 28 进制基数系统中的优选数值。

5.3.3. 世界语的声语号与图言号的映射本质

在世界语中，声语号与图言号在设计上规定为一一映射关系，因此，以声语号（音标声节单词）为基准，站在语音型音乐家的角度来看，1个声语号就是1个以28个压缩声符为压缩型音符，实际展开声滴总量达到143个的音乐型声节。

以图言号（字母图符单词）为基准，站在28数制计数系统的数学家的角度来看，1个图言号就是1个以28（注：如无特别说明，描述采用10进制表示法，最大基数的数值为 $28-1=27$ ）为基数，以字母图符标记符号为数字的优选数值。

注意：与音节类似，并不是什么样的数值，都能够成为声节。

5.4. 计算机16进制计数法

参考计算机语言中的16进制系统的表示方法，即：0~9、A~F；世界语的28个构词字母，本质上，就是数学28数制中的28个基本数字。

由于思维习惯的缘故，人们常常误以为只有0~9才是表示数值，从而迷失了对概念对象的本质性认识。实际上，数值的符号表示，也是可以采用a, b, c,...；甲乙丙丁...；等等的。

5.5. 拼音的本质

拼音是一种声滴压缩与解压缩的算法，以及算法在解压缩过程中的拼读产物。

注意：拉丁语拼音和中文拼音，是有所区别的。共同点：它们都包含了声滴的压缩与解压缩算法。不同点：中文拼音还包含了拼音的声调扩展算法。

5.6. 只是改变了进制计数系统的基数，难道：熟识就会变成陌生？

5.6.1. 音符的数量，以及压缩表示法的变化

拉丁语声语概念标记号（音标声节单词）的本质：EO声语概念标记号系统，将语音型音乐的声符数量设置为约143个；EN声语概念标记号系统，将语音型音乐的声符数量设置为约298个。在音乐学上，与人们日常习惯使用的“哆来米发索拉西”音符系统相比，在经过了如此简单的声符（音符）数量，以及采用压缩声符（音符）表示法之后，还有多少人能够看得清楚，拉丁语声语号的构造本质呢？

5.6.1. 在数学数制计数系统中，基数数的变化

拉丁语图言概念标记号（字母图符单词）的本质：EO图言概念标记号系统，将数制计数系统的基数，设置为28；EN图言概念标记号系统，将数制计数系统的基数，设置为26。在数

学上，与人们日常习惯使用的 10 进制计数系统相比，在经过了如此简单的基数数量变换之后，还有多少人能够看得清楚，拉丁图言语号的构造本质呢？

6. 概念对象、概念标记号、概念本体属性、概念外部属性的关系

6.1. 实例的简要说明

概念对象：本体（存在对象张三）；

概念标记号：张三（名字）；

概念本体属性：张三（男性）；

概念外部属性：张三（关联手机号码）。

6.2. 在理解概念的时候，首先明确关注的目标

- 想关注“本体（存在对象张三）”，即张三这个人的本身？还是想关注“张三（名字）”这个概念标记号？
- 想关注“张三（名字）”，即张三的标记号？还是想关注指向“张三（本体，或名字）”所关联的概念本体属性张三（男性）？或是所关联的概念外部属性张三（关联手机号码）？
- 由于人类思维习惯的原因上，常常将“概念对象”与“概念标记号”不加区分地混淆起来使用，在一般场景之下，这种思维不会有问题。正是由于如此，使得在有些应用场景里可能会隐藏着概念指向的风险，甚至是巨大的风险。越是抽象的概念，隐藏的风险越大。
- 如果没有容易学习与创造原子图言语号的优秀通用语文概念标记号系统的支持，那么，母语群社会的概念数量将无法获得自觉的、有效的、可持续性的增长。

6.3. 概念标记号，也是非常重要的

如果概念释义的逻辑矛盾丛生，那么，将会造成人们难以正确地认知宇宙的事物和现象。

但是不论如何，概念标记号都只不过是概念对象的某种编号而已，人们真正应该关注的焦点是概念对象本身，而不是概念对象的编号。比如，财务人员给张三发工资，总不能看到帐号的名字是张三就往里面打钱吧？很可能此张三，非彼张三。

6.4. 综上所述，以此类推。

当在没有弄清楚所要关注的目标对象之前，本体、标记号、内部属性、外部属性这4类概念，常常被混淆，从而导致在理解上的本末倒置。

7. 怎样看待和理解《牛津英语词典》中的庞大词汇量？

7.1. 《牛津英语词典》的词汇量发展趋势

7.1.1. OED《Oxford English Dictionary》词汇量的递增足迹

- 1933年第1版，13卷，收词41万
- 1989年第2版：20卷，词汇达到61万余个；收录了301,100个主词汇，词汇数目达3.5亿个。
- 还在编纂中的第3版：OED编辑普罗菲特表示，OED3预计长达40卷。据报道，OED3自1994年起开始编纂，70名专家以每月50至60个新词为目标，至今完成约80万个，但进度仍然比估计落后20年。

7.1.2. 拉丁语构词字母的数量增加原因，以及拉丁语概念标记号系统设计的科学性问题简述

- 只要人类对宇宙概念对象的认识，在广度上和深度上的不断开拓与深入；只要人类的思想与行为在不断进步与锐意进取，其结果就是概念标记号的数量不断地增多。
- 随着概念标记号总量的不断增多，理论上，在未来会达到一个阶段性的相对新陈代谢平衡点。从无限的角度来看，绝对新陈代谢平衡点是不存在的；这是因为宇宙概念对象具有绝对的无限性，而人类智慧具有相对无限性的必然结果。
- 拉丁构词字母的增加，实质上，其目的是增加数学计数系统中的基数总量，从而达到扩大可用的概念标记号的总量。

拉丁文的字母约公元前7世纪~前6世纪时，由希腊字母间接发展而来，成为古罗马人的文字，古罗马灭亡前共包含23个字母，其中有16个辅音字母B、C、D、F、Z、H、K、L、M、N、P、Q、R、S、T、X，4个元音字母A、E、I、O，和一个音值不定的V，11世纪时增加了J、U、W，形成了今天的26个字母。

- 从OED来看，拉丁语的有效概念数量在不断地增长，意味着拉丁语群人们的思想在不断

进步。

- 哲学思考：原子图言号到底应该由谁来创造呢？概念标记号的本质，是概念对象的标记符号。概念标记号如果由创造家或发明家来创造，其目的一般是为了标记某种新的创造或发现。

7.2. 声节与数值型概念标记号系统，可以轻松应对知识大暴增下的庞大词汇量

拉丁语声语号与图言号的本质，就是限制最大可变长度的语音型音乐的声节，以及以构词字母总量为基数的优选数值。

声语号就如同音乐的音节那样，图言号就如同电话号码那样。

对于千变万化的音乐学音节，稍微学习过一点音乐乐谱基础知识者，在看到乐谱时，也许就能够唱出音节来了。

面对数以亿计的电话号码，稍微学习过一点 10 进制计数系统知识者，在看到数值符号时，也许就能够读出电话号码来了。

概念对象的标记号，本质上，只不过就是需要经过筛选过的声节型或数值型的标记号而已。

7.3. 假如概念标记号很简单，那么，是否人人都能成为发明家呢？

发明创造是一项开拓性的工作，并不见得人人都能够成为发明家的。就如音乐学中的音符，它们是那样的简单，可是，却并不见得人人都能成为作曲家，道理是一样的。

7.4. 拉丁语有效词汇量的不断递增，是语群在自觉不断进步的特征表现之一

倒退回到史前时代，那时候还不需要文字。回归到人类初生时的婴孩时期，不要说是文字就算是声语号，那也是仅有的几个“啊啊”、“咿咿”、“呀呀”。

这样的概念标记系统，是不是很简单呢？可是，作为通用语文概念标记系统，这样的系统又有什么实际的意义和价值呢？

7.5. 地球的生命是有限的，人类需要自觉不断地进步吗？

如果需要，那么，就意味着人类社会概念系统的词汇量，将不断地在增加，以及新陈代谢。因此，《牛津英语词典》中已经存在的庞大词汇量，而且词汇量还在继续性的递增和新陈代谢，这是不断在进步的信号。

7.6. 语文概念系统有效词汇量的自觉递增，需要开放型语文概念系统的支持

并不是什么样的通用语文概念系统，都能够支持有效词汇量的自觉可持续的递增。

7.6.1. 正如现代人无法替代未来人来进行思想一样，古代人同样也无法替代现代人来进行思想。

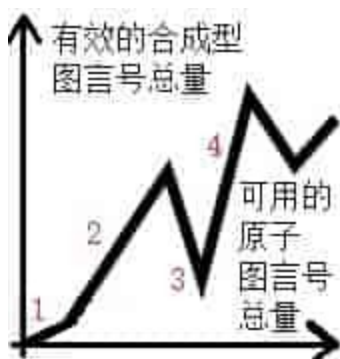
这将意味着除了与人类在历史进化中变化很小的那一部分数量不多概念之外，比如吃喝拉撒、衣食住行、喜怒哀乐、悲欢离合，其余的概念将在人类进步的历程中不断地诞生，OED 第 1、2 版，以及未来中的第 3 版的词汇量不断递增，这就是一个明显的标志性体现。

7.6.2. 如果没有原子图言号作为基础支持，那么，合成词在达到一定的总量之后，将难以再有效第标记宙泐概念对象。

当合成词的内部单元组合，再也无法避免概念释义的逻辑矛盾时。

单体合成概念（单词）的最大长度，必须控制在 1 个肺活量之内。

7.6.3. 合成词的有效性猜想图解



注：（猜想）

- 在图符定义的起始阶段，“可用的原子图言号总量”很少时，“有效的合成型图言号”的总量增长不明显。
- 当“可用的原子图言号总量”达到一定量级的拐点之前，“有效的合成型图言号”的总量增长明显，增长速率在上升。
- 当“可用的原子图言号总量”达到一定量级的拐点之后，“有效的合成型图言号”的总量增长速率在下降。
- 当语群思想进入到另外一个大境界时，比如进入第 5 维空间，那么，“有效的合成型图

言号”的情况，又将重复上述（1）~（3）的过程，并且是在原有的基础之上的递增。

7.7. 《牛津英语词典》与《康熙字典》的本质区别，是什么？

（略）。

Mb's 请参阅：[MEMBER 期刊目录 / MEMBER PERIODICAL CONTENTS]

8. 初学者怎样来看待和理解这么庞大的词汇量呢？

8.1. 首先需要端正心态，有效的词汇量越多，一般而言，代表当时的人类思想越进步

8.1.1. 不断递增与新陈代谢的有效词汇量

有效的词汇总量，代表了语种人们的思想水平。

面对有效的概念总量在不断递增中通用概念系统，特别是对于初学者们，健康的心态可能是感受到了欢欣和鼓舞，因为此时的人们生活在一个日新月异的不断进步社会里。脱离地球的束缚，进入宇宙社会，这可能是人类所希冀的下一个前行足印。

8.1.2. 公共词汇与专业词汇

8.1.2.1. 克服心理恐惧

对于初学者来说，在海量的词汇面前，会产生望而却步的想法，这其实只不过是心理上的恐惧导致，完全没有必要。

8.1.2.2. 理解手机号码

就像手机号码那样，人们总不能因为自己记不住数以亿计的号码，而产生恐惧吧？

通用型手机号码不是为了某个人或某些人所准备的，而是为了某种区域所有的人们而准备和预备的。面对数量如此庞大的手机号码，人们需要恐惧吗？其实，只要记住那些与自己有关的号码就可以了。在这些号码中，包含有一部分是公共的，比如号码前段 138，130 之类的还有一些非公共的，比如号码后段的数字，属于自己的亲朋好友的专门标记数值。

8.1.2.3. 理解概念

与通用型手机号码一样，在不断增长、且数量庞大的概念总量当中，除了吃喝拉撒、衣食住行、喜怒哀乐、悲欢离合等公共概念之外；那些被用来制造出手机的理论性的专业概念，对于绝大多数的人来说，可能一个都并不知道。但是，假如没有了那些可能与自己无关的、却在不断增长中的专业概念总量，那么，今日的人们也就没有手机可用了。

8.2. 动物性本能生活的日常词汇量

在有文字记载的数千年历史里，人类基本特征的进化和变化，都微不苛察。因此，日常生活中与人类本性有关的概念的种类和数量，虽有变化，但更多的是不变。

由此可知，对于初学者来说，常用只约有 4 级的词汇量；有人统计常用的高频词为 1368 个，[1368 个单词就够了]湖北教育出版 2017-10 作者:王乐平。

8.3. 人类不断进步的必然：日益庞大词汇量

可以将非日用高频单词，视为是专业单词，以及介于日常高频词与专业单词之间状态的追随单词，即追随伴随着时代不断进步而不断产生的那些与自己有关的单词。

那些进取型的通用概念系统，全人类都能够从中获益。不同行业的追随者都会努力地去掌握各自领域中的概念，而概念大多数都是由行业的开创者所创造的。

8.4. 将通用单词与专业单词混合在一起，使得对初学者造成心里困扰

人们假如想自己尝试一下创造一个新的概念，并为之创造一个新单词（图言号，以及对应的声语号），那么，就会明白，其实自己也是有可能为词典词汇量的递增做出贡献的。

宇宙内、表、外的一切，是无限的，专业学科可以有千千万万。因此，在通用语文概念系统里，拥有少变化的、有限的日常单词之外；同时，还拥有无限的日新月异的专业单词，这是常态。在学习英语的过程中，需要克服这个心理问题。

8.5. 应对庞大词汇量的 Colating 解决方案

- 进取型通用语文概念系统的设计，首先是为方便概念的创造者与发明者，其次才是为了概念使用者。 🗒️ 一般而言，能够满足概念的创造者与发明者所使用的通用概念系统，就一定能够满足普通的概念使用者所使用。这是因为概念的创造者与发明者所花费在语言文字上的学习时间，通常要少于普通人，像孔乙己那样懂得“回”字有好多种写法的人，它哪里还有太多的时间来进行实践与研究？
- 在《牛津英语词典》里，少变化的、有限的日常单词；与无限的日新月异的专业单词混杂在

一起，这样既不方便在非母语环境之下的概念学习，也不方便于新单词的创造，以及词典的维护。☑️为此，Colating 利用功能字母的方式，将概念系统分隔为 [基础追随概念系统] 与 [开拓进取概念系统] 来解决这个问题。这两个概念系统的构词法完全一致，只是有些标志字母的定义，以及 Colating 的专门拼读方法，可能会有不同，但都具有容易被学习的健硕规律性。

- [基础追随概念系统] 的编制受到严格的规则控制。☑️ [开拓进取概念系统] 的编制既受到与 [基础追随概念系统] 类似的规则控制，同时，通过标志字母，将追求自由个性发挥与创造的规则，也容纳于其中。

9. 莫尔斯电码、手机号码与拉丁语单词的系统设计比较

请参阅：☆. 莫尔斯电码、手机号码与拉丁语单词的系统设计比较

10. 阅读入门需要首先理解一些新概念

在现实生活中，关于语文学习方面的著作，随处可见。然而，关于怎样创造通用语文系统的作品，却难觅踪影。

为了避免在概念理解上可能产生的混乱，Colating 对可能产生释义混乱的概念，重新进行了概念创造，使得 Colating 的理论术语可以自成体系。

10.1. 新的声音名词

10.1.1. 声滴/语声滴

名词，在语音频率范围内，由单一频率基波与非基波（包含音色信息）共同构成的、近似于单位纯音的语音频谱，称为：声滴、或语声滴。

- 声滴是还没有映射任何其他标记的单纯语音。
- 对声滴进行概念释义，只是使用语音的方式。
- 声滴这个概念，是可以跨语种使用的。

10.1.2. 滴列

声滴的队列。

10.1.3. 声节

优选的适合作为声节的声滴队列。由优选的1个或多个声滴队列，就构成了1个声节。1个声节就是1个声语概念标记号。

- 声节：可分为“单滴声节”，“多滴声节”。
- 在限长多滴声节（1维）声语概念语文系统中，1个声节就代表了1个单词的发音；
- 在单滴声节2维图言概念系统中，1个声节就是1个字的发音。

10.1.4. 非声节

不适合用作声节的声滴队列，称为：非声节。

10.1.5. 声滴与声节的区别

用途：”声滴”用来描述“声节”和声语号。“声节”用来描述声语号和“句子”。

10.1.6. “单滴声节”与“单个声节”，“多滴声节”与“多个声节”，的区别

- 单滴声节：单个声滴声节。优选的适合作为概念标记用途的单个声滴，就构成了单滴声节
- 多滴声节：多个声滴声节。优选的适合作为概念标记用途的多个声滴队列，就构成了多滴声节。
- 单个声节：单个概号声节。每一个概念标记号，都是一个声节；可能是单滴声节，也可能是多滴声节。
- 多个声节：多个概号声节。用于表示短语、句子、段落，或其他。

注意：“多滴声节”不等于“多个声节”；“单滴声节”不等于“单个声节”

10.1.7. 声符/语声符

名词，在语文系统中，用于标记声滴的图形符号，即采用图形符号来映射声滴。如IPA（International Phonetic Alphabet）国际音标符号。

如IPA（International Phonetic Alphabet）国际音标符号。

10.2. 新的构词图形名词

10.2.1. 图粒/语图粒/纯图粒

名词，在语文系统中，图粒概念对象（图粒本体）是基础的构词图形单元，它存在有3种状态：“图粒”态，“符图粒”态，“纯图粒”态。

未被释义的构词图形单位，称为：图粒。

已被释义的“图粒”，称为：符图粒。

不需要、或无需考虑释义的图粒；这类图粒称为：纯图粒。

符图粒与纯图粒，共同组成图言号的基础构词图形单元。

石头（图粒）就是石头，当它被人用锤子敲了一个坑痕之后，就变成了一块与众不同的石头（符图粒），实际上，它原本就是石头。

图粒在被释义后，就变成了字母、偏旁部首，或者是其他任何的构词图形单元。

一块四不像的石头（图粒），它还就是石头。

图粒这个概念，是可以跨语种使用的。

10.2.2. 图符/图言符

名词，在语文系统中，图粒概念对象（图粒本体）主要存在有3种状态：未被释义的构词图形单元，称为“图粒”。已经被释义了的用于构造图言概念标记号的图粒，这类图粒称为“符图粒”。难以被规律化释义、或无需考虑释义的用于构造图言概念标记号的图粒，这类图粒称为“纯图粒”。

符图粒与纯图粒，共同组成图言号的基础构词图形单元。

10.3. 新的概念名词

10.3.1. 宙沕

宇宙内、表、外的一切。

10.3.2. 概念

心理学上认为，概念是人脑对客观事物本质的反映，这种反映是以词来标示和记载的。表达概念的语言形式是词或词组。概念是思维活动的结果和产物，同时又是思维活动借以进行的单元

10.3.3. 概念对象（宙沕概念对象）

宇宙内、表、外的一切，均可作为概念标示与描述的对象。

10.3.4. 声号/声语号

声语概念标记号，标示概念对象的声滴流符号。

10.3.5. 图号/图言号

图言概念标记号，标示概念对象的图粒阵列记号。

10.3.6. 原子图号（自由词素、根词）

概念释义由不包括后缀的、不可分割的图号整体构成。

10.3.7. 分子图号

概念释义由不包括后缀的、图号整体可以进行分割的合成词。

10.3.8. 单体声号

原子声号，或分子声号。

10.3.9. 单体图号

原子图号，或分子图号。

（本文结束）

自动目录: [期刊_随想_总目录](#) [独立目录 / All Single CONTENTS](#)

可乐听语 · 期刊

Colating Language Periodical & Capriccio

1

Total Issue 1

2021 · Irregular

(Open Version)

Ver: 0.1.0.11

Created: 2021/08/25 Update: 2021/10/8

Author: David Email: colating@yandex.com

[编号 No.]: Doc No.: Per 2021-1 / Cap 2021-1

内容说明: Open Version 只涉及 Colating 的实现原理。

[期刊目录 / Periodical Contents]

[总目录 / ALL CONTENTS](#)

⚠ 编写要求:

1. 特别注意: 只能运用中性语言, 理性客观态度, 以旁观者立场来撰写。切记!!!
2. **Open Version** 只涉及 **Colating** 的实现原理。过于深刻的内容, 以及与其他语种的深入比较, 应只在 Member Version 部分。

世界语拼音原理分析

The Principle of EsperantoPinyin System

Ver: 0.2.0.5

Created: 2021/08/26 Update: 2021/10/10

Author: David Email: colating@yandex.com

[摘要 Abstract]: 世界语属于拉丁语族的一个分支。

世界语声语声语单词的声节音标流发音（声语号/声语概念标记号），实质上是一个个在目标概念系统中独一无二的、限制最大可变长度的语音型音乐的声节，系统展开声滴总量的理论值约为 143 个。

需要注意的是：在音乐学上，音滴是非压缩的，1 个音滴对应于 1 个音符。在语文学上，构成声语号的声滴，是经过拼音算法压缩之后，采用压缩声滴来与声符（音标）相对应。

换言之，世界语要表示 143 个系统展开声滴，却只有 28 个声符，该怎样来处理呢？办法就是采用 1 个声符，以及 2 个声符的拼读组合来表示它们。理想状态是所有的辅音与所有的元音都能够有效地拼读，这样 28 个声符，在理论上就可以表示到的声滴总量为： $23 \text{ 辅音} \times 5 \text{ 元音} + 28 \text{ 音标} = 115 + 28 = 143$ 个。

世界语图言单词的构成字母图符流符号（图言号/图言概念标记号），实质上是一个个在目标概念系统级中独一无二的、限制最大可变长度的数学优选数值，该数学数制计数系统的基数为 28。

将 28 个压缩声符（雷同 IPA 音标）映射到 28 个实际构词字母上，世界语的这种构词法，使得世界语单词能够做到“见词能读、听音能写”的效果。

[关键词 Keywords]: 概念系统，声节，声滴压缩与解压缩，数制，双射

[编号 No.]: Doc Class: 语音/ voice; Doc No.: 20210826-2

[目录索引]

- | | |
|--------------------|------------------|
| 1. 引言 Introduction | 4. 声滴拼读算法分析 |
| 2. 声符与图符 | 5. 结论 Conclusion |
| 3. 声符与图符映射关系的指定 | |

[关联内容摘要]

- | | |
|----------------------|----------------------|
| ✧ APPENDIX A 名词对照表 1 | ✧ 通用概念标记关联糅合模型的对比与选择 |
|----------------------|----------------------|

[相关文章]

- | | |
|-------------|--------------|
| ✧ 简明声滴算法原理 | ✧ EN 语拼音原理分析 |
| ✧ 世界语拼音原理分析 | ✧ 声滴拼读算法原理 |

[名词解释]

- 声滴/语声滴/语音音素：名词，在语音频率范围内，由单一频率基波与非基波（包含音色信息）共同构成的、近似于单位纯音的语音频谱，称为：声滴、或语声滴。声滴是声语号的基本构建单位，它源于人类先天器官本能的声音处理能力。辅音 \subseteq 辅助标压声滴；元音 \subseteq 元母标压声滴。
- 滴列：声滴队列。
- 图粒/语图粒/纯图粒：名词，又称（语图粒），在语文系统中，图粒是声滴的图言概念标记号。声滴与声符、声符与图符发生映射关系是人为指定的结果。
- 标准压缩声滴：采用融合声滴拼读算法压缩后所获得的声滴。
- 专门压缩声滴：采用非融合声滴拼读算法压缩后所获得的声滴。
- 标准压缩声符（雷同 IPA 音标）：标准压缩声滴所对应的声符，即、音标。

- 专门压缩声符（雷同 IPA 音标）：专门压缩声滴所对应的声符。
- 实际使用图符：实际使用的图符。在拉丁语中为字母。

[缩写]

- 标压声滴：标准压缩声滴。
- 专压声符：专门压缩声符（雷同 IPA 音标）。
- 专压声滴：专门压缩声滴。
- 实用图符：实际使用图符。在拉丁语中为字母。
- 标压声符：标准压缩声符（雷同 IPA 音标）。

1. 引言 Introduction

世界语属于拉丁语族的一个分支。

世界语声语声语单词的声节音标流发音（声语号/声语概念标记号），实质上是一个个在目标概念系统中独一无二的、限制最大可变长度的语音型音乐的声节，系统展开声滴总量的理论值约为 143 个。

需要注意的是：在音乐学上，音滴是非压缩的，1 个音滴对应于 1 个音符。在语文学上，构成声语号的声滴，是经过拼音算法压缩之后，采用压缩声滴来与声符（音标）相对应。换言之，世界语要表示 143 个系统展开声滴，却只有 28 个声符，该怎样来处理呢？办法就是采用 1 个声符，以及 2 个声符的拼读组合来表示它们。理想状态是所有的辅音与所有的元音都能够有效地拼读，这样 28 个声符，在理论上就可以表示到的声滴总量为：23 辅音 x 5 元音 + 28 音标 = 115 + 28 = 143 个。

世界语图言单词的构成字母图符流符号（图言号/图言概念标记号），实质上是一个个在目标概念系统级中独一无二的、限制最大可变长度的数学优选数值，该数学数制计数系统的基数为 28。

采用数学 10 进制计数系统的电话号码，可以做到“见号码能读、听号码能写”。类似于 10 进制的电话号码计数系统，将 28 个压缩声符（雷同 IPA 音标）映射到 28 个实际构词字母上，世界语单词，声语号（音标声节单词）的构造为、采用了约 143 个展开声滴所构成的语音型音乐的声节，其展开声滴被采用压缩拼音算法处理后变成了 28 个压缩声滴；图言号的构造

为、采用数学 28 进制计数系统的优选数值。28 个压缩声滴与 28 个图符（构词字母）被设计规定为一一映射关系，因此，同样也可以做到“见词能读、听音能写”的效果。

参考计算机语言中的 16 进制系统的表示方法，即：0~9、A~F；世界语的 28 个构词字母，本质上，就是数学 28 数制中的 28 个基本数字。

2. 声符与图符

2.1. 声符

世界语一共有 28 个标准压缩声滴，分别对应于 28 个声符。参见：表 EO3-1。

2.2. 图符

在不讨论大小写的情况下，世界语一共有 28 个图粒，分别对应于 28 个构词图符（=字母）。参见：表 EO3-1。

3. 声符与图符映射关系的指定

3.1. 映射关系

世界语共有 28 个字母，书写形式采用拉丁字母，一个字母只发一个音，每个字母的音值始终不变，也没有不发音的字母，其语音和书写完全一致。见下表：

表 EO3-1：世界语字母和国际音标的对照表

字母	Aa	Bb	Cc	Ĉĉ	Dd
音标 0→4	/a/	/b/	/ts/	/tʃ/	/d/
字母	Ee	Ff	Gg	Ĝĝ	Hh
音标 5→9	/e/	/f/	/g/	/ĝ/	/h/
字母	Ĥĥ	Ii	Jj	Ĵĵ	Kk
音标 10→14	/x/	/i/	/j/	/ʒ/	/k/
字母	Ll	Mm	Nn	Oo	Pp
音标 15→19	/l/	/m/	/n/	/o,ɔ/	/p/

字母	Rr	Ss	Ŝŝ	Tt	Uu
音标 20→24	/r/	/s/	/ŝ/	/t/	/u/
字母	Ŭŭ	Vv	Zz	28 个音标映射 28 个字母	
音标 25→27	/w/	/v/	/z/		

3.2. 拼读规则

每个词的重音固定在倒数第 2 个音节上，学会了 28 个字母和掌握了拼音规则以后，就可以读出和写出任何一个单词。

4. 声滴拼读算法分析

4.1. 展开声滴的总量

假设每一个辅音与每一个元音之间的拼读都有效，那么，利用声滴拼读算法解压缩后所获得的“展开声滴”的总量（理论值） $=23 \times 5 + 23 + 5 = 143$ 个。

4.2. 声符与图符指定为一一映射关系

在世界语中：

- 标准压缩声符（雷同 IPA 音标）=世界语音标；实际使用图符=世界语字母。
- 音标与字母之间，被设计为一一映射。即：声滴 \rightarrow 声符（音标） \rightleftharpoons 图符（构词字母） \leftarrow 图粒。

这是世界语单词（单体图号：原子图号，分子图号）可以做到“见词能读、听音能写”的基础和原因。

4.3. 概记的编码结构

声语号，采用限长优选多个声滴流队列的声节结构形式。

图言号，采用限制最大可变长度的语音型音乐声节的映射图符结构形式。

4.4. 声节

具备标准拉丁语单词的共同特征：采用限制最大可变长度的、优选多声符结构。

声节的最短长度为 1 个声符，最大长度为限制最大可变声符长度。在世界语中，声符（世界语音标）与构词字母被设计为一一映射关系。

4.5. 数值

具备标准拉丁语单词的共同特征：世界语将 28 个压缩声符（雷同 IPA 音标）映射到 28 个实际构词字母上，世界语概念系统，等同于是一个以 28 为基数的标准数学数制计数系统；世界语单词，等同于是在该数学数制计数系统中的一个个独一无二的优选数值。

世界语单词优选数值的最短长度为 1 个图符，最大长度为限制最大可变图符长度。

4.6. 改良探索

增加系统全部声滴的数量：在相同声节长度下，可以倍增声语号的数量。

增加标准压缩图符的数量：在相同字母长度下，可以倍增图言号的数量。

5. 结论 Conclusion

5.1. 世界语展开声滴的总量

- 理论值约为：143 个。

5.2. 世界语声滴拼读算法

- 世界语的构成字母图符与声符之间，主要是“1 对 1”映射关系，总体拼读规则简单。

5.3. 概念标记号构成

- 概念标记号，采用限长优选多个声滴流队列的声节结构形式
- 概念标记号 = 声语号 | 图言号

5.4. 世界语单词的本质

- 实质上，是一个个限制最大可变长度的语音型音乐的声节，理论上，所使用的展开声滴的总量约等为 143 个。

- 实际上，是以 28 为基数的标准数学数制计数系统中的一个限制最大可变长度的优选数值。

EO 单词能够做到“见词能读、听音能写”的效果。

Appendix: 世界语网摘

世界语（Esperanto）是由波兰籍犹太人眼科医生拉扎鲁·路德维克·柴门霍夫（Ludwig Lazarus Zamenhof）博士在印欧语系的基础上于 1887 年 7 月 26 日发明创立的一种人造语言。

他期待这门语言能够为不同语言人群的交流提供便利，能够帮助人们跨越语言、肤色、种族、地域等界限，用同一个身份——世界公民来平等、友好地相处，而并不打算用取代任何民族语。世界语在文化交流、经贸往来和人类社会进步等方面作出了贡献。

Colating 简评：创造者的精神可贵，只是可自觉进步型的人类通用语文系统，没有将就这种说法。

（本文结束）

自动目录: [期刊_随想_总目录](#) [独立目录 / All Single CONTENTS](#)

英语拼音原理分析

The Principle of English Pinyin System

Ver: 0.2.0.5

Created: 2021/08/26 Update: 2021/10/12

Author: David Email: colating@yandex.com

[摘要 Abstract]: 英语属于拉丁语族的一个分支。

英语声语单词的声节音标流发音（声语号/声语概念标记号）：实质上是一个个限制最大可变长度的语音型音乐的声节，或声符节，系统展开声滴总量的理论值约为 298 个。展开声滴相当于音乐学中的音符，只是数量约可达到 298 个，而且采用压缩声符来表示。标准压缩声滴总量 34 个 + 元音标压声滴连读组合 8 个 + 辅音标压声滴拼读组合声符总量 6 个 = 音标表全部声符总量 = 48 个。

英语图言单词的构成字母图符流符号（图言号/图言概念标记号），实质上是一个个在目标概念系统中独一无二的、限制最大可变长度的数学优选数值，该数学数制计数系统的基数为 26。

与世界语不同；将实际 34 + 8 个声符，采用不同的算法，一起混合映射到 26 个构词字母上，英语的这种构词法，使得大多数的英语单词，必须通过单纯的、简单的、额外的字母发音提示，才能完全做到“见词能读、听音能写”的效果。

英语声语概念标记号采用 298 个声符构成的语音型音乐音符系统，也等同于 298 进制的计数系统，因此，声语号的构词效率比采用 143 进制的世界语要高。虽然 EN 更符合语文的实际应用场景，但由于它是在语文历史进程中不断累加创造的结

果，因此，与世界语不同，其大多数的词汇，缺乏像世界语那样无需母语环境的、简单的、可学习的规律性。

需要注意的是：在设计进取型通用语文概念系统的时候，不应该盲目地追求的简单化，无视宙沏概念对象的客观实际，而选取牺牲概念系统的总体性能的做法；特别是缩减可用原子图言概念标记号的总量等的重要基础性能。简单化与合理化，有时往往会发生矛盾。

[关键词 Keywords]: 概念系统，声节，声滴压缩与解压缩，数制，双射

[编号 No.]: Doc Class: 语音/ voice; Doc No.: 20210826-3

[目录索引]

- | | |
|--------------------|------------------|
| 1. 引言 Introduction | 4. 声滴拼读算法分析 |
| 2. 声符与图符 | 5. 结论 Conclusion |
| 3. 声符与图符映射关系的指定 | |

[关联内容摘要]

✧通用概念表示项联结框架 ◦系列

[相关文章]

- | | |
|------------|-------------|
| ✧简明声滴算法原理 | ✧EN 语拼音原理分析 |
| ✧世界语拼音原理分析 | ✧声滴拼读算法原理 |

[名词解释]

- 声滴/语声滴/语音音素：名词，在语音频率范围内，由单一频率基波与非基波（包含音色信息）共同构成的、近似于单位纯音的语音频谱，称为：声滴、或语声滴。声滴是声语号的基本构建单位，它源于人类先天器官本能的声音处理能力。辅音 \subseteq 辅助标压声滴；元音 \subseteq 元母标压声滴。

- 滴列：声滴队列。
- 图粒/语图粒/纯图粒：名词，又称（语图粒），在语文系统中，图粒是声滴的图言概念标记号。声滴与声符、声符与图符发生映射关系是人为指定的结果。
- 声号/声语号：声语概念标记号，标示概念对象的声滴流符号。
- 图号/图言号：图言概念标记号，标示概念对象的图粒阵列记号。
- 原子图号（自由词素、根词）：概念释义由不包括后缀的、不可分割的图号整体构成。
- 分子图号：概念释义由不包括后缀的、图号整体可以进行分割的合成词。
- 单体声号：原子声号，或分子声号。
- 单体图号：原子图号，或分子图号。
- 压缩声滴：采用融合声滴拼读算法压缩后所获得的声滴。
- 标准压缩声滴：采用融合声滴拼读算法压缩后所获得的声滴。
- 专门压缩声滴：采用非融合声滴拼读算法压缩后所获得的声滴。
- 标准压缩声符（雷同 IPA 音标）：标准压缩声滴所对应的声符，即、音标。
- 专门压缩声符（雷同 IPA 音标）：专门压缩声滴所对应的声符。
- 实际使用图符：实际使用的图符。在拉丁语中为字母。
- 变限声符长度/限制最大可变声符长度：限长多滴声节 1 维图言结构的概记，声号的构词声符的长度越长，表达效率就越低下；当声号的构词声符超出一定的长度时，声号将由于表达的效率太低，从而导致这个声号的失效。
- 变限图符长度/可变限制图符长度：限长多滴声节 1 维图言结构的概记，图号的构词图符的长度越长，表达效率就越低下；当图号的构词图符超出一定的长度时，图号将由于表达的效率太低，从而导致这个图号的失效。

[缩写]

- 标压声滴：标准压缩声滴。
- 变限声符长度：限制最大可变声符长度。

- 专压声滴：专门压缩声滴。
 - 变限图符长度：可变限制图符长度。
 - 标压声符：标准压缩声符（雷同 IPA 音标）。
 - 专压声符：专门压缩声符（雷同 IPA 音标）。
 - 实用图符：实际使用图符。在拉丁语中为字母。
-

1. 引言 Introduction

英语属于拉丁语族的一个分支。

英语声语单词的声节音标流发音（声语号/声语概念标记号）：实质上是一个个限制最大可变长度的语音型音乐的声节，或声符节，系统展开声滴总量的理论值约为 298 个。展开声滴相当于音乐学中的音符，只是数量约可达到 298 个，而且采用压缩声符来表示。标准压缩声滴总量 34 个 + 元音标压声滴连读组合 8 个 + 辅音标压声滴拼读组合声符总量 6 个 = 音标表全部声符总量 = 48 个。

英语图言单词的构成字母图符流符号（图言号/图言概念标记号），实质上是一个个在目标概念系统中独一无二的、限制最大可变长度的数学优选数值，该数学数制计数系统的基数为 26。

在英语中，由于采用 48 个混合型的声符映射到 26 个实际使用构词字母上，英语的这种构词法，使得大多数的英语单词，必须通过单纯的、简单的、额外的字母发音提示，才能完全做到“见词能读、听音能写”的效果。

参考计算机语言中的 16 进制系统的表示方法，即：0~9、A~F；英语的 26 个构词字母，本质上，就是数学 26 数制中的 26 个基本数字。

2. 声符与图符

2.1. 声符

在当代英语中，将 34 个标准压缩声滴，8 个元音标压声滴连读组合，6 个辅音标压声滴拼读组合，一起共同组成音标表声滴；它们分别一一映射于 48 个声符。

实际上，需要关注与图符进行映射的声符总量，只有 42 个。另外的 6 个辅音标压声滴拼读组合，与辅音构词字母是一一映射关系，因为简单且单纯，所以无需花费额外的学习时间。

6 个辅音标压声滴拼读组合：既没有增加系统展开声滴的总量，也没有像 8 个元音标压声滴连读组合那样，在压缩之后再与图符号进行映射；因此，它们属于高频聚类，主要的目的可能只是为了方便学习而已。

参见下表：

表 EN3-1: 英语 48 个声符 (音标) 发音表							
元音: 12 个元音标压声滴, 8 个元音标压声滴连读组合。							
单元音	前元音	0~3	/i:/	/ɪ/	/e/	/æ/	
	中元音	4~6		/ɜ:/	/ə/	/ʌ/	
	后元音	7~11	/u:/	/ʊ/	/ɔ:/	/ɒ/	/ɑ:/
	开合双元音	12~16	/eɪ/	/aɪ/	/ɔɪ/	/aʊ/	/əʊ/
	集中双元音	17~19	/ɪə/	/eə/	/ʊə/	声符压缩	
辅音: 22 个辅音标压声滴, 6 个辅音标压声滴拼读组合。							
爆破音	清辅音	20~22	/p/	/t/	/k/		
	浊辅音	23~25	/b/	/d/	/g/		
	清辅音	26~30	/f/	/s/	/ʃ/	/θ/	/h/
	浊辅音	31~35	/v/	/z/	/ʒ/	/ð/	/r/
破擦音	清辅音	36~38	/tʃ/	/tr/	/ts/	高频聚类 方便学习	
	浊辅音	39~41	/dʒ/	/dr/	/dz/		
鼻音	浊辅音	42~44	/m/	/n/	/ŋ/		
舌则音	浊辅音	45	/l/				
半元音	浊辅音	46~47	/j/	/w/			

2.2. 图符

在不讨论大小写的情况下，英语一共有 26 个图粒，分别对应于 26 个构词图符（英语构词字母）。如下：

Aa Bb Cc Dd Ee Ff Gg Hh Ii Jj Kk Ll Mm Nn Oo Pp Qq Rr Ss Tt Uu Vv Ww Xx Yy Zz。

3. 在当前英语系统中，映射关系总览

在英语中，主要存在有 3 种映射关系，理解清楚它们相互之间的算法与映射关系，对于学习 EN 语来说，或有所帮助。

3.1. 系统展开声滴与标准压缩声滴的映射

采用声滴压缩与解压缩算法，将约 298 个展开声滴压缩成 34 个标准压缩声滴。

系统展开声滴与标准压缩声滴之间的映射，包括融合拼音算法，扩展拼音算法，以及其他拼音算法。

3.2. 音标表声滴与声符的映射

“声符”指音标表声符，或系统声符。

为了提高构词效率，优化构词法，以及学习上的方便，将标准压缩声滴，高频的标压声滴连读组合、或者拼读组合等等，聚合在一起，形成了音标表声滴。

音标表的声滴与声符之间，通常为一一映射关系。

3.3. 声符与字母图符的映射

在当代英语中，声符与字母之间，主要有 2 种映射关系：标压声符总量的压缩映射，高频的元音声符连读组合长度的压缩映射。

3.3.1. 音标表标压声符总量的压缩映射

英语标准压缩声滴，有：12 个元音标压声滴，22 个辅音标压声滴。

需要将一个 $12 + 22 = 34$ 个的标压声符，映射到 26 个构词字母上，这属于声符数量压缩的映射。

3.3.2. 音标表元音声符连读组合长度的压缩映射

将“双元音”、“三元音”等，映射到 1 个构词字母上，这种做法，有利于缩短图言概念标记号的长度。

3.4. 在音标表中，6个辅音标压声滴拼读组合的意义

- 将高频的辅音拼读，聚合在一起，可以方便于学习。
- 这6个辅音标压声滴的拼读组合，可以视为是2个独立辅音的拼读组合。
- 它们并没有能够增加系统展开声滴的总量。
- 它们不像8个高频的元音标压声滴连读那样，与1个构词字母进行映射，从而具有压缩的用途。

4. 声符与图符映射关系的指定

在当代英语中，将34个标准压缩声滴，8个元音标压声滴连读组合，6个辅音标压声滴拼读组合，一起共同组成音标表声滴；它们分别一一映射于48个声符。

实际上，需要关注与图符进行映射的声符总量，只有42个。另外的6个辅音标压声滴拼读组合，与辅音构词字母是一一映射关系，因为简单且单纯，所以无需花费额外的学习时间。

6个辅音标压声滴拼读组合：既没有增加系统展开声滴的总量，也没有像8个元音标压声滴连读组合那样，在压缩之后再与图符号进行映射；因此，它们属于高频聚类，主要的目的可能只是为了方便学习而已。

但是，42个声符（音标）却只对应于26个图符（构词字母），因此，英语拼音与世界语拼音相比，英语的声符与图符之间的映射关系要复杂。

英语的声符与图符映射关系，主要有2种：1对多，多对1。

4.1. 英语声滴的分类

依照某个时代的英文单词使用频次，划分出21个辅音声滴和5个元音声滴，使得它们与26个构成字母形成一一映射关系，称为：虚构双射声滴。语文系统剩余的作为虚构双射声滴的补集声滴，称为：补构双射声滴。

虚拟的英语单词，英语虚构双射单词：全部采用虚构双射声滴来构造的英语单词。一般拉丁语诞生初期创造的高频单词，属于英语虚构双射单词。

4.2. 声符与字母图符映射关系

在当前英语系统中，需要与字母图符进行映射的声符总量，实际上有 42 个，却只有 26 个字母，因此，只有全部由虚构双射声滴所构成单词能够满足“见词能读，听音能写”的效果。

4.2.1. “1 对 1” 标准映射，1 声符对应 1 图符字母标准映射

英语声符（音标）有 48 个，却只有 26 个构成字母与之映射，因此，实际上，现代英语拼音不存在“1 对 1”标准映射关系。

可能“1 对 1”标准映射关系，只存在于概念数量级较小时代的某个古代时期。

4.2.2. “1 对多”，1 声符对应于多个图符（构词字母）的声符扩展映射

- 增加高效声节的利用率
- 声节的种类，根据发音容易辨别度等指标，将声节分为 3 个档次：高效、中效、低效
- 对于高效声节来说，1 个声节只对应于 1 个图言号，这样太浪费了。
- 1 对多，通过词性，或者与其他词搭配等办法来区分同音词，可以增加高效声节的利用率。由于高效声节的数量有限，因此，1 对多现象，似乎无法避免。
- 回收利用不符合融合声滴拼读算法的拼音
- 融合声滴拼读算法，使得某些辅音和元音的组合不成立。
- 在有限的构词字母总量的基础上，采用回收利用的方式，提高声符的利用效率。对于采用经典融合声滴拼读算法来说，有些元音可以与更多的辅音进行拼读，而有些元音只能与较少的辅音进行拼读。通过改变元音首次的原本发音，以达到利用不符合声滴拼读规则的哪些图符组合。
- 在相同的图符长度约束下，既要保证发音的合理，又要保证获得更多的图言号，由此，产生了“1 对多”现象？作用效果：相当于增加了可用声语号的数量。
- 句法或历史原因，调整单词在句子中的发音。还有可能，就是由于原始创造单词的发音，与在句子中左右两侧其他单词的发音，在句子中难以被识别，后来被修改了发音，因此，导致了后来者所看到“1 对多”现象？

词例：

[ʌ] ⇔ come [kʌm], cum [kʌm]; bus [bʌs]

注：

(1) 简短、发音响亮的优选声节的总量并不多。通过词性区隔，可以获得相同声语号、不同图言号构造的不同的概念标记号。

(2) 在 Colating 中规定，只会运用在高频动词、介词或其他语法词上。并且，数量必须控制在容易记忆的数量范围之内。

4.2.3. “多对 1”，多个声符对应 1 个图符字母的图符扩展映射

- 作用：相当于增加了图言号的数量。
- 由于英语元音有 20 个，却只有 5 个元音字母与之映射；辅音有 28 个，却只有 21 个辅音字母与之映射；因此，“多对 1”映射普遍存在。在相同的声滴长度约束下，“多对 1”这种做法，可以提高声语号的表达效率。

词例：

name [neɪm], bag [bæg], China [ˈtʃaɪnə] ⇨ a

[ən; æn] ⇨ an

4.2.4. 在“客观记录”之下，“多对 1”、“1 对多”现象是必然会出现的。

- 客观记录历史的代价，需要有足够多的、可用的词汇总量
 - 客观记录历史，是人类能够获得自觉进步的力量源泉
 - 1857 年，英国学者 R.C.特伦齐在英国语文学会提议编写客观记录英语词汇的新型词典。
 - 1928 年，《牛津英语词典》第 1 版以 12 卷本的形式面世，共收录 414825 个单词和短语。
 - 1989 年，《牛津英语词典》第 1 版以 20 卷本的形式面世，词典收录的词汇达到 61 万余个。
- 客观记录历史使得自觉不断进步成为了可能。概念数量的绝对无限性，高频概念长达数百年以上的有效寿命，以及人类声音器官先天本能表达能力的局限性；因此，这种现象是无法避免的。Colating 采用[基础追随概念系统]与[开拓进取概念系统]来解决这个问题。

- 即使在设计时，声符与图符被设计成一一映射关系；但是，只要在未来有了新增的图粒子，而构词字母的数量保持不变；又或者音标的发音发生了某种大幅度的改变，等等的这些原因，都将会导致一一映射关系的失效。Colating 采用[基础追随概念系统]与[开拓进取概念系统]来解决这个问题。

4.3. 拼读规则

4.3.1. 复杂的拼读规则

除了英语虚构双射单词之外，由于声符与构词字母的映射关系复杂，导致了英语的拼读规则，要比世界语复杂得多。

对于非母语学习者来说，由于部分构词规律化不明显、规则多且凌乱、或者难以记忆，从而造成英语的拼读规则难以掌握。

4.3.2. 拼读规则复杂的原因总结

- 客观记录历史的伟大思想，造成词汇总量不断增大，从而导致拼读规则不断走向复杂。
 - 1857 年，英国学者 R.C.特伦齐在英国语文学会提议编写客观记录英语词汇的新型词典。
- 只有客观的历史记录，才有可能做到明鉴。
- 宙泐概念对象、概号的总量，具有无限性。
- 大部分日用高频概号的平均有效历史寿命，可达数百、乃至数千年。
- 词汇总量绝对无限，而构词字母太少。
- 《牛津英语词典》从 1928 年的第 1 版，到 1989 年的第 2 版，词汇量从 41 万多增加到了 61 万多，并且，单词的数量仍然在不断地增加。
- 相比之下，英语构词字母却只有 26 个，这太少了。其结果就是现在新创造单词的构词字母长度，在不断加长。
- 在本质上，就是采用 26 进制优选数值来标记概念对象，结果就是基数 26 还是太少了。注意：并不是什么样的数值都可以作为概念标记，因为需要可考虑到数值的发音，以及长度限制最大可变长度问题。
- 声符的数量，还是少了一些，不过这个问题受限于人类发音器官的先天能力。这是产生

“1对多”映射的原因之一。

- 概号的构词字母长度受到表达效率的约束。概号的最小长度为1个，最大长度为变限图符长度。

5. 声滴拼读算法分析

5.1. 展开声滴的总量

标准压缩声滴：辅音 22 个，元音 12 个。

假设每一个辅音与每一个元音之间的拼读都有效，那么，采用融合声滴拼读算法压缩后所获得的声滴，即、标压声滴的总量（理论值）=22 x12 +22 +12 =298 个（k/3 级）。

5.2. 声符与字母图符的映射关系

5.2.1. 见号码能读、听号码能写

如同采用数学 10 进制计数系统的电话号码，可以做到“见号码能读、听号码能写”那样，世界语概念标记号同样具备这一特征。但是，英语却不能像世界语那样简单的拼读。

英语：标压声滴 = 标压声符 = 34 个；构词图粒 = 构词字母 = 26 个。

实际上，EN 语单词不存在一一映射关系。由于采用 48 个混合型的声符映射到 26 个实际使用构词字母上，英语的这种构词法，使得大多数的英语单词，必须通过单纯的、简单的、额外的字母发音提示，才能完全做到“见词能读、听音能写”的效果。

5.2.2. 比世界语的拼读复杂，但是，却是符合客观实际的需求的

英语声语概念标记号采用 298 个声符构成的语音型音乐音符系统，也等同于 298 进制的计数系统，因此，声语号的构词效率比采用 143 进制的世界语要高。虽然 EN 更符合语文的实际应用场景，但由于它是在语文历史进程中不断累加创造的结果，因此，与世界语不同，其大多数的词汇，缺乏像世界语那样无需母语环境的、简单的、可学习的规律性。

5.2.3. 明确可用原子图言概念标记号的总量，是概念系统设计的关键

需要注意的是：在设计进取型通用语文概念系统的时候，不应该盲目地追求的简单化，无视宙泐概念对象的客观实际，而选取牺牲概念系统的总体性能的做法；特别是缩减可用原子图言概念标记号的总量等的重要基础性能。简单化与合理化，有时往往会发生矛盾。因此，学习英语所需要的语言环境支持的力度，比世界语强，但是，却有可能远远低于其他语种。

5.2.4. 从声滴到字母图符，EN 经历了 2 次，3 种压缩算法

5.2.4.1. EN 语声滴与字母映射系统的运算步骤：

- 第 1 步，第 1 次压缩，利用声滴压缩算法，将展开声滴转换成标准压缩声滴：EN 系统展开声滴总量的理论值约为 298 个，在经过压缩拼音算法之后，获得 34 个标准压缩声滴。
- 第 2 步，获得音标表中的声滴总量：标准压缩声滴总量 34 个 + 元音标压声滴连读组合 8 个 + 辅音标压声滴拼读组合声符总量 6 个 = EN 音标表全部混合声符总量 = 48 个。
- 第 3 步，EN 音标表中的声滴与音标表的声符，形成一一映射。
- 第 4 步，第 2 次压缩包含了 2 种压缩算法，针对 34 个标准压缩声符的声符总量压缩算法，以及针对 8 个元音标压声符连读组合的声符长度压缩算法；它们一起与 26 个构词字母，主要形成“1 对多”和“多对 1”的映射关系。另外，6 个的辅音标压声滴拼读组合，实际上只具有聚类学习的意义，它们与辅音构词字母为一一映射关系。

5.3. 概记的编码结构

声语号，采用限长优选多个声滴流队列的声节结构形式。

图言号，采用限制最大可变长度的语音型音乐声节的映射图符结构形式。

5.4. 声节

具备标准拉丁语单词的共同特征：采用限制最大可变长度的、优选多声符结构。

声节的最短长度为 1 个声符，最大长度为限制最大可变声符长度。

与世界语不同；将实际 34 + 8 个声符，采用不同的算法，一起混合映射到 26 个构词字母上，英语的这种构词法，使得大多数的英语单词，必须通过单纯的、简单的、额外的字母发音提示，才能完全做到“见词能读、听音能写”的效果。

5.5. 数值

具备标准拉丁语单词的共同特征：英语单词是采用以基数为 26 的数学数制计数系统中的优选数值。

数值的最小长度为 1 个；最大长度为由变限图符长度。在英语中，图符 = 构词字母 = 26 个

5.6. 改良探索

增加系统全部声滴的数量：在相同声节长度下，可以倍增声语号的数量。

增加标准压缩图符的数量：在相同字母长度下，可以倍增图言号的数量。

简化复杂的构词法：拉丁单词的本质是限制最大可变长度的语音型音乐的声节，也是数学数制计数系统中的优选数值，因此，需要规律化，以便减少概念标记号的学习时间和精力。

6. 结论 Conclusion

6.1. 英语的展开声滴的总量

理论值约为：298 个。

6.2. 英语声滴拼读算法

采用融合声滴拼读算法。

英语的辅音与元音之间主要是“1 对多”和“多对 1”映射关系，对于非早期创造的高频单词，总体拼读规则比较复杂。

6.3. 概念标记号构成

概念标记号，采用限长优选多个声滴流队列的声节结构形式

概念标记号 = 声语号 | 图言号

6.4. 英语单词的本质

- 实质上，是一个个限制最大可变长度的语音型音乐的声节，实际使用声滴的总量约等为 298 个。
- 实际上，是以 26 为基数的标准数学数制计数系统中的一个个限制最大可变长度的优选数

值。

- 与世界语不同；将实际 34 +8 个声符，采用不同的算法，一起混合映射到 26 个构词字母上，英语的这种构词法，使得大多数的英语单词，必须通过单纯的、简单的、额外的字母发音提示，才能完全做到“见词能读、听音能写”的效果。
- 英语声语概念标记号采用 298 个声符构成的语音型音乐音符系统，也等同于 298 进制的计数系统，因此，声语号的构词效率比采用 143 进制的世界语要高。虽然 EN 更符合语文的实际应用场景，但由于它是在语文历史进程中不断累加创造的结果，因此，与世界语不同，其大多数的词汇，缺乏像世界语那样无需母语环境的、简单的、可学习的规律性。
- 需要注意的是：在设计进取型通用语文概念系统的时候，不应该盲目地追求的简单化，无视宙沕概念对象的客观实际，而选取牺牲概念系统的总体性能的做法；特别是缩减可用原子图言概念标记号的总量等的重要基础性能。简单化与合理化，有时往往会发生矛盾

(本文结束)

自动目录: [期刊_随想_总目录](#) [独立目录 / All Single CONTENTS](#)

简明声滴算法原理

The Simple Principle of Sound Drop Algorithm

Ver: 0.2.0.3

Created: 2021/08/26 Update: 2021/09/25

Author: David Email: colating@yandex.com

[摘要 Abstract]: 声滴算法，分为：声滴拼读算法和拼音扩展算法。

在本质上，声滴拼读算法属于一种声滴压缩与解压缩的算法。拼音扩展算法主要是对拼音声滴所采用的可具规律性的声滴扩增方法。

限长优选多个声滴流队列的声节结构形式的拉丁语概念标记的声语号（单词的音标），本质上是一个个在目标概念系统中独一无二的、限制最大可变长度的语音型音乐的声节，实际使用声符总量等于采用拼音算法解压缩后的系统压缩声滴总量。

与声符发生规律性映射、限长优选的字母图粒流队列结构形式的拉丁语概念标记的图言号（单词的图形标记），本质上是一个个在目标概念系统中独一无二的、限制最大可变长度的数学优选数值，数学数制计数系统的基数等于构词字母的总量。

单体概念标记号（原子图言号、分子图言号）能够做到“见词能读、听音能写”的关键，是将声符与图符设计成具有规律性的、容易被学习的映射关系。

[关键词 Keywords]: 声滴拼读算法，拼音扩展算法，声滴压缩与解压缩，声符与图符映射

[编号 No.]: Doc Class: 语音/ voice; Doc No.: 20210826-1

[目录索引]

- | | |
|-------------------------------|----------------------|
| 1. 引言 Introduction | 10. 拼音扩展算法原理 |
| 2. 声音 4 个主要特征属性 | 11. 声滴的分类 |
| 3. 人类语音器官及其频率范围 | 12. 滴、粒、符、号的关系 |
| 4. 语音的离散化与声滴总量 | 13. 标准压缩声滴系统设计的指标要求 |
| 5. 拉丁语单词（声语号/声语概念标记号）的本质
述 | 14. 拼音解压缩算法实例简 |
| 6. 系统展开声滴简单化的必要性 | 15. 声符与图符映射关系的指定 |
| 7. 拼音的生物听觉感受
说明 | 16. 拉丁语概念系统本质性问题的总结性 |
| 8. 声滴算法分类 | 17. 结论 Conclusion |
| 9. 声滴拼读算法原理 | |

[关联内容摘要]

APPENDIX A 名词对照表 1
声滴拼读算法原理

[相关文章]

声滴算法原理。系列

[名词解释]

- 滴列：声滴队列。
- 声滴/语声滴/语音音素：名词，在语音频率范围内，由单一频率基波与非基波（包含音色信息）共同构成的、近似于单位纯音的语音频谱，称为：声滴、或语声滴。声滴是声语号的基本构建单位，它源于人类先天器官本能的声音处理能力。辅音 \subseteq 辅助标压声滴；元音 \subseteq 元母标压声滴。
- 图粒（语图粒）：名词，又称（语图粒），在语文系统中，图粒是声滴的图言概念标记号。声滴与声符、声符与图符发生映射关系是人为指定的结果。

- 声号/声语号：声语概念标记号，标示概念对象的声滴流符号。
- 图号/图言号：图言概念标记号，标示概念对象的图粒阵列记号。
- 原子图号（自由词素、根词）：概念释义由不包括后缀的、不可分割的图号整体构成。
- 分子图号：概念释义由不包括后缀的、图号整体可以进行分割的合成词。
- 单体声号：原子声号，或分子声号。
- 单体图号：原子图号，或分子图号。

[缩写]

- 创发家：人类的创造家与发明家。
- 原子图号：原子图言概念标记号。类似：自由词素、根词。
- 概记：概念标记。
- 标准压缩声滴：采用融合声滴拼读算法压缩后所获得的声滴。
- 专门压缩声滴：采用非融合声滴拼读算法压缩后所获得的声滴。
- 标准压缩声符（雷同 IPA 音标）：标准压缩声滴所对应的声符，即、音标。
- 专门压缩声符（雷同 IPA 音标）：专门压缩声滴所对应的声符。
- 实际使用图符：实际使用的图符。在拉丁语中为字母。

1. 引言 Introduction

声滴算法，包括声滴拼读算法，拼音扩展算法。声滴算法的核心，是融合声滴拼读算法。

融合声滴拼读算法，又称经典拼音算法（经典声滴拼读算法），属于一种声滴拼读自然混音融合的衍生新声滴的方法，实际上，是一种声滴的压缩与解压缩的过程及结果。

Colating 声滴算法包含了声滴拼读算法、声调拼音扩展法、Colating 专门扩展法。

在本质上，声滴拼读算法属于一种声滴压缩与解压缩的算法。拼音扩展算法主要是对拼音声滴所采用的可具规律性的声滴扩增方法。

2. 声音 4 个主要特征属性

音调、响度、音色、长短，这是表述声音的 4 种不同类型的主要特征。在不同的学科中，这 4 种特征的名称有所不同；并且，同一种类型特征的具体参数也有所不同。但是，对它们进行研究的声理论，在本质上，却是几乎一样的，基本上可以直接相互引用或借鉴。

2.1. 音调：

人耳对声音高低的感受。音调主要由声音的频率决定；在音乐体系中称“音高”。

物体在 1 秒内振动的次数叫频率。物体振动得越快，频率越高。频率越高，音调越高；频率越低，音调越低。

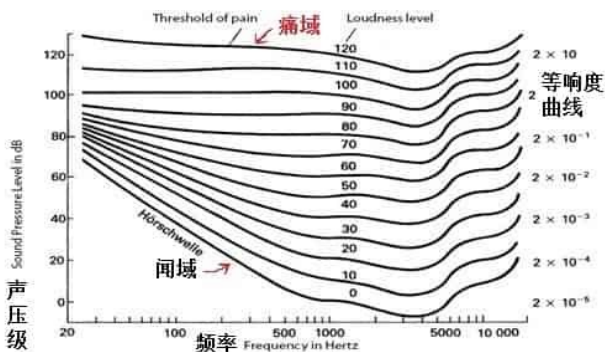
2.2. 响度

人耳对声音强弱的主观感觉称为响度。在音乐体系中称“音强”，又或“音量”。

2.2.1. 等响度曲线

在相同距离下，声波振动的幅度越大，产生的声压也越大。

人耳对声音的响度感觉，不仅和声波振动的幅度有关，还和频率有关。见下图：



图说：以图中任意一条等响度曲线为例，要想产生人耳感觉到同样大小的声音，不同的频率，所需要的声压级是不一样的；其中，500Hz~8KHz 这个频率范围所需要的声压最小。

2.2.2. 响度还跟距离发声体的远近有关系

声音是从发声体向四面八方传播的，越到远处声压越小，所以人们距发声体越远，听到的声音越小。如果能够想办法加大声音的声压，就可以使声音响度更大些。

2.3. 音色

对“声波频谱”的综合感受

人们能够区别具有同样响度、同样音调的两个声音之所以不同的特性；或者说是人耳对由相同基波，不同谐波所构成不同的“声波频谱”的综合反应。

人们能够分辨出各种不同乐器的声音，就是由于它们的音色不同。人的声音的音色也因人而异，所以人们闭着眼也能听出是哪位熟人在讲话。

2.4. 长短

声音延续时间的长短。

不同的声滴搭配不同的发音时间长度，可增强声滴之间的可辨别度。

3. 人类语音器官及其频率范围

3.1. 耳朵的听音能力

人耳的听觉频率范围在 20Hz~20KHz 之间；并且，人的听音能力会随着年龄的衰老而有所损失，见下表：

表 60 岁老人的听力损失值

纯音频率/kHz	1	2	3	4	6	8	12	15
听力损失/dB	0	-2.3	-7.7	-12.2	-15.7	-16.7	-32.6	-65.9

3.2. 声电话机标准语音频带的科学测定结果

在国际数字电话机通信标准中，规定电话的语音频带为：300~3400Hz/-3db（ITU-T P.310 06/2009）。舍弃一些不重要的细节不影响话质，这是能够满足电话机语音通话保真度要求的科学测定结果数据。

-3db 标准，就是说 300Hz 和 3400Hz 的传输电压幅度降低到正常的 0.707 倍，并不是一过这两个频率电压就完全消失了，现实中也做不到如此精确的滤波电路。

3.3. 人类自然语音频率范围

实际上，人讲话的频率范围主要集中在(1~3)KHz，低频段极少，高频也几乎没有。

在 ITU-T 通信标准中，为了满足最小保真度指标的要求，以及为了弥补电路处理过程中的频带损失，电话机标准语音频带的选取，要比实际的人类自然语音的频率范围要宽。

4. 语音的离散化与声滴总量

人类用来标记概念对象的语音，其频率不是连续的，而是以不同的声滴队列（声节）的形式出现。

4.1. 声滴

在语音频率范围内，由单一频率基波与非基波（包含音色信息）共同构成的、近似于单位纯音的语音频谱，称为：声滴、或语声滴。声滴是声语号的基本构建单位，它源于人类先天器官本能的声音处理能力。

4.2. 声滴流的产生

在语音频率范围内，将连续的语音频带离散化，从而产生出一个一个可被辨识的不同声滴。依照类似于音乐学中的音节原理，将多个声滴组合成声滴队列；当声滴队列依次发出声音时，就产生了声滴流。

4.3. 声滴流构成声语概念标记号

在符合目标语文系统概记声节发音规则要求的前提下，由 1 或多个声滴组成的队列，就构成一段可用声滴流（语声滴流），也就是 1 个声节。

将一段可用声滴流赋予一个特定的概念对象，这段声滴流就是该概念对象的声语号。声语号是采用不同声滴的有序排列来表示的，它不是连续频率的语音。从音乐学的角度来看，1 个声语号就是 1 个限制最大可变长度的声节。

4.4. 在理论上，声滴的数量范围

4.4.1. 从科学实验中，获取声滴的总量

数字电话机的标准语音频带为 300~3400Hz；而人类讲话的频率范围主要集中在(1~3)KHz 之内及附近；这意味着声滴的总量最多只有 K（千）量级别。不论如何，人类天然声滴最后的实际数量，一定是有限的。

4.4.2. 从社会实践中，获取声滴的总量

实际上，在世界现存的典型语文系统中，展开声滴的总量约在 143 ~400 ~1500 个左右。

其中，理论上，世界语约为 143 个，某典型多声符拉丁语文概念系统约为 298 个；某典型单声符语文概念系统约为 1500 个（实际统计值）。

5. 拉丁语单词（声语号/声语概念标记号）的本质

声语号采用限长优选多个声滴流队列的声节结构形式的拉丁语概念标记的声语号（单词的音标），本质上是一个个在目标概念系统中独一无二的、限制最大可变长度的语音型音乐的声节实际使用声符总量等于采用拼音算法解压缩后的系统压缩声滴总量。

与声符发生规律性映射、限长优选的字母图粒流队列结构形式的拉丁语概念标记的图言号（单词的图形标记），本质上是一个个在目标概念系统中独一无二的、限制最大可变长度的数学优选数值，数制计数系统的基数等于构词字母的总量。

在世界语中，实际使用声符总量 = 压缩声滴总量 = 构成字母总量。在英语中，实际使用声符总量 = 压缩声滴总量 > 构成字母总量。

5.1. 从声音的物理特征来看

乐音与语音，它们都具备音调、响度、音色、长短，这 4 种不同类型表述声音的主要特征。

5.2. 从音乐学的角度来看

声节与音节主要区别：多声滴声语号（拉丁语单词）的外形表现是一个个限制最大可变长度的语音型音乐的声节；音乐学上的音节长度，一般是固定的，或者只有少数几种变化的固定长度。

对于多声滴声语号系统，每一个声节就是一个声语号。声节的构成可以是1个声滴，也可以是系统规定的、有效的构词字母总长度，即、限制最大可变长度。

5.2.1. 不同学科的概念名称类比

如下：

名称	说明	名称	说明
声滴	语音频率范围在(1~3)KHz及附近	音符发 音	哆来米发索拉西, 27.5 Hz ~4186 Hz
声符	系统展开声滴总量, 数百、乃至上千个。压缩声滴总量一般为数十个。	音符	7的倍数, 目前不超过钢琴的108个琴键
声节	限制最大可变长度, 1个声节效率失效的可允许最大长度。	音节	多种固定长度

注1: 声滴, 类似与音乐音符所发出的纯音 + 表征音色的非纯音, 声符类似于音符。

注2: 语文学中的声滴, 类似于音乐学中音符的发音。实际上, 是一种基波 + 音色所构词的频谱。

注3: 声符 (语声符/语音符): 声滴的图形标记符号。1个声语号由1或多个声滴, 和或声符所构成。

注4: 声节 (语声节/语音节): 采用音乐的观点来看, 声节是概念级限制最大可变长度的优选声滴流, 是语文系统中的语音乐音节。类似于音乐学中音节。

5.2.2. 声节的特点和构成条件

声符的数量: 在单滴声节 2 维图言结构概念系统中 = 系统展开声滴总量; 在限长多滴声节 1 维图言结构概念系统中 = 压缩声滴总量;

声符的频率范围: 属于语音频带的子集;

声滴与图符构成双向可预测的映射。

5.2.3. 声节与音节的主要区别

在语文学上，声节的长度是可变的，声节的最短长度为 1 个声符，最大长度为限制最大可变声符长度。在音乐学上，音节的长度是一个固定长度，或者只有少数几种长度变化。

音乐的总体声音频率范围更宽，语言的声滴与声符的数量更多。

5.3. 从数学的角度来看

采用限长多滴声节 1 维图言结构形式的图言号（拉丁语单词），实质上，是一个个在目标概念系统中独一无二的、限制最大可变长度的数学优选数值，数学数制计数系统的基数等于构词字母的总量。

注 1：“优选”表示需要考虑声节或音节的发音规则，并不是什么样的发音组合，都能够组合成声节或音节。

注 2：“限制最大可变长度”，对于限长多滴声节 1 维图言结构形式概念标记系统，1 个声号或图号，当构成的声滴或图粒过长时，声号或图号将由于表达的效率太低，从而导致这个声号或图号的失效。

6. 系统展开声滴简单化的必要性

根据理论估算与实际统计，在不同的典型语种中，展开声滴的总量可达上百、乃至上千个量级。由于通用语文系统属于社会性思想交互的共用产品，而声号更是属于社会性共用的必需品，不仅需要自己掌握，还需要思想交互中的对方也要掌握。因此，需要对系统展开声滴进行简单化、规律化处理。

一种常用的解决方案就是对系统展开声滴进行压缩与解压缩处理，也就是所谓的融合声滴拼读算法，又称：经典拼音算法（经典声滴拼读算法）。

6.1. 系统展开声滴复杂性的体现

系统展开声滴的复杂体现在声滴总体的数量上，而不是指个别的声滴。

- 展开声滴的总量：约 400（单声符）~1500（多声符）个，每个个体都需要全部掌握。
- 群体的一致性：彼此的发音不应被误会。
- 声滴的主观性：在实际应用中，声滴一般只能采用主观评价，这也是造成跨地域语群

的发音难以保证一致性的一个原因。

6.2. 学习图号，离不开声号作为基础支撑

对于语音健全人类而言，声号，是学习与思考的根本基础支撑与构建单位。

6.3. 展开声滴简单化的常规方法：聚类与压缩

减少需要死记硬背的声滴总量，同时，使得声滴之间具有规律性，这可能是让声滴简单化至有效的方法。区分辅音与元音，为聚类算法；辅音与元音之间的拼读，是解压缩算法。

越具规律性的知识，就越适容易被学习。

在概念系统设计中，必须将总量达到上百、乃至上千量级的展开声滴加以规律化，使之简单到可以让跨地域的普通人群也能够容易掌握的程度，这就是融合声滴拼读算法存在的价值之一。

7. 拼音的生物听觉感受

7.1. 声滴的自然混音融合效应

对于2~3个不同的声滴，通过1个人快速连读、或者2个人分别同时读出时，有些声滴的组合连读会产生自然混音融合现象，从而自然衍生出一个新的声滴。这个过程就是“拼读”，在拼读时所衍生出来的这个新声滴，就是拼音。

7.2. 融合拼音与强制拼音

通过拼读自然混音融合现象衍生出来的拼读声滴，属于融合拼音。

对于不能产生自然混音融合现象的拼读，可以人为地设定一个声滴，让这个设定声滴充当混音融合声滴的角色。这个过程，以及在此过程中所设定的声滴，称之为“强制拼音”。强制拼音的目的，使得声符具有更高的构词效率。

8. 声滴算法分类

8.1. 声滴算法：声滴拼读算法，拼音扩展算法。

8.1.1. 声滴拼读算法：

- 按工作模式，分为：拼音压缩算法，拼音解压缩算法。
- 按拼读音感，分为：融合拼音算法（融合声滴拼读算法），强制拼音算法（强制声滴拼读算法）。
- 按使用阶段，分为：在声号设计阶段主要采用拼音压缩算法；在声号应用阶段主要采用拼音解压缩算法；
- 强制拼音算法：强附拼音算法，强定拼音算法。

8.1.2. 拼音扩展算法：声调扩展算法，Colating 专门扩展法，其他声滴扩展算法。

8.2. Colating 声滴算法：包含声滴拼读算法、声调拼音扩展法、Colating 专门扩展法。

8.3. 声滴算法的核心，是融合声滴拼读算法，又称“经典拼音算法（经典声滴拼读算法）”。

8.4. 说明：

- 声滴拼读算法 \subseteq 声滴压缩与解压缩算法。
- 拼音压缩算法 \subseteq 声滴压缩算法，拼音解压缩算法 \subseteq 声滴解压缩算法）：指对系统展开声滴所进行压缩与解压缩的算法。
- 强附拼音：通过连读产生的拼音音感效果，不如融合拼音。
- 强定拼音：连读时完全没有拼音的音感效果，人为强行指定的拼音。
- 拼音扩展算法，表示在拼读声滴的基础之上进行扩展，而不是指在系统展开声滴的基础之上进行扩展。如类似 ZH 拼音的音调法。

9. 声滴拼读算法原理

声滴拼读算法，在本质上，既是一种声滴压缩与解压缩的过程；同时，还包括了在这一过程中所产生的结果，就是通过拼读方法所产生出来的新声滴。

9.1. 不同声滴拼读算法的用途

拼音压缩算法主要用在声语概念标记号系统的创造过程。

拼音解压缩算法主要用在声语概念标记号系统的应用过程。

9.2. 拼音解压缩算法的种类

拼音解压缩算法，主要有 2 种：在说话时解压缩、在说话前解压缩。

9.3. 拼音解压缩算法选择的重要性

在通用语文系统设计中，拼音解压缩算法的选择，非常重要，对语文系统的整体性能会产生重要的深刻基础影响。

9.3.1. 在说话时解压缩

这种解压缩算法适用于多声符概念标记方法系统，理论上，“在说话时解压缩”所需要记忆的声滴和图粒的总量，一般在几十个左右。

将标压声滴的总量压缩到约 50 个左右，这有助于创造与之一一映射、至佳造型的基准图粒。

9.3.2. 在说话前解压缩

单声符概念标记方法系统通常适用这种解压缩算法。然而，受到人类先天声音器官的限制，与概念对象、概念、图粒等的数量相比，系统展开声滴的总量是极其有限的。

10. 拼音扩展算法原理

10.1. 声调扩展声滴算法

类似中文拼音的 4 个声调，但是，在限长多滴声节 1 维图言单词中的实现有所不同。

10.2. Colating 扩展拼滴算法

Colating 功能字母专门拼读方法。

10.3. 其他扩展拼滴算法

略。

参见：[Colating 声滴算法原理]。

11. 声滴的分类

11.1. 标准压缩声滴集合

采用混音融合现象的拼音压缩原理，将语文系统中所使用的所有的展开声滴进行压缩及分类处理，从而获得发音不响亮的辅音集合，发音响亮的元音集合；以及一些需要特别定义和处理的专门声滴集合。

由辅音集合、元音集合、专门声滴集合共同组成目标语文的标准压缩声滴集合。

11.2. 声滴的分类

标准压缩声滴（音标）=辅音标压声滴+元音标压声滴+其他声滴（如：Colating 独立功能声滴）。

辅音与元音分类的规则与条件是：当辅音与元音在进行拼读的时候，大多数（与标压声滴的设计有关）的拼读都可以通过自然混音融合衍生出不属于标压声滴集合中的其他展开声滴。

12. 滴、粒、符、号的关系

参见：Appendix A 名词对照表 1。

13. 标准压缩声滴系统设计的指标要求

13.1. 系统展开声滴的总量要求

在满足容易记忆和识别的前提下，以及在相同字母长度的约束条件下，声滴的总量越多，表明以声滴总量为基数的数学数制计数系统所能表达的不同数值的数量越多。

标压声滴的设计要求：在符合限长多滴声节 1 维图言概念标记方法的发音规则，以及融合声滴拼读算法的条件下，尽可能多量。

13.2. 标压声滴的总量要求

数以百、千计的系统展开声滴，在通过声滴拼读算法之后，最后一般会获得数十个标压声滴。在标压声滴与构词字母形成一一映射的场景里，在相同字母长度的约束条件下，标压声滴的总量越多，声语号的表示效率越高。

13.3. 标压声滴与构词字母的矛盾

标压声滴的总量越多，符合限长多滴声节 1 维图言“高中矮胖中瘦/如：hogmnl”至佳结构的构词字母的创造，也变得越困难。

14. 拼音解压缩算法实例简述

经典的拼音解压缩算法，主要有 2 种：在说话时解压缩、在说话前解压缩。

14.1. 在说话时解压缩

世界语采用多声滴声语号（拉丁语单词）概念系统，只需要使用 28 个声符，通过拼音解压缩算法，理论上，可以生成的展开声滴总量，约为 143 个。

世界语的声符与构词字母，设计为一一映射关系，这使得世界语单词的拼读简单容易。

采用在说话时解压缩方式的多声滴声语号，以及将声符与构词字母设计成一映射关系，这就如同全世界不同长度的电话号码那样，学习概号容易，创造原子图号简单。

14.2. 在说话前解压缩

对于通用语言概念系统来说，只有在图言号的总量不大于系统展开声滴总量的场景，才可在不需要额外标记的情况下，图言号就具备“见词能读、听音能写”的效果。

14.3. 目的与用途，会决定解压缩方式的最后选择？

有效的原子图号，应该由真正的、真实的创造家或发明家来创造吗？

15. 声符与图符映射关系的指定

声滴与图符之间的映射关系，典型有 4 种：数学一一映射，双向可预测的专用规则映射，弱规则的特别映射，无规则的零散映射。

15.1. 数学一一映射

在多声滴声语号（拉丁语单词）概念系统中，概念系统的主体结构，只有采用一一映射，单词的构造才能做到“见词能读、听音能写”的效果。

典型语种：世界语，23 个辅音，5 个元音。

15.2. 专用规则映射

定义可被容易学习与记忆的专用规则及其规则总量，可让概念系统具有更好的可用性。

15.3. 弱规则的特别映射

当概记号普遍使用这种弱规则映射时，将大大增加在非母语环境下的学习门槛与难度。

15.4. 无规则的零散映射

这种映射规则，应当控制在可简单地学习与记忆的数量，否则，概念系统将难以被学习和记忆。存在少量可容易被记忆的无规则的零散映射，或许可让语法系统的表示简洁明了。

15.5. 典型语种的声滴、声符，图粒、图符的关系

单体概念标记号（原子图言号、分子图言号）能够做到“见词能读、听音能写”的关键，是将声符与图符设计成具有规律性的、容易被学习的映射关系。见下图：

语种	系统全部声滴 (Nsa)	系统全部声符	标准压缩声滴 *注 1	专门压缩声滴 *注 1	标准压缩声符 (雷同 IPA 音标) (音标) *注 1	专门压缩声符 (雷同 IPA 音标) (音标) *注 1	实际使用图符 (拉丁语: 字母) *注 1	是否能够: 见词能读, 听音能写

世界语	≈143	--	28	--	28	--	28	是
EN 语	≈400	--	34	--	34	8	26	大部分高频词
Colating	? 1500	--	52 +5x2 +x	5x2 +x; x4	52	5x2 +x; x4	52 +5x2 +x	是（融合拼音与规则拼读）
某表音单音节	23 x21	--	44	--	44	--	44	是
某表义单音节	≈1500	n x Nsa	63	x4	63 x4	x4	n x Nsa（图符+图粒）*注4	否（特殊、或特别用途）

注1：参见：[缩写]，在文字开始部分。

注2：是否容易创造通用原子图号，是思想能够自觉获得不断进步的标志。

注3：是否能够“见词能读，听音能写”，是语群概念总量能够自觉获得不断增加的表现。

注4：拉丁语图言号的构成单元为：字母。典型单音节图言号的构成单元为：图符+图粒。

16. 拉丁语概念系统本质性问题的总结性说明

16.1. 概念对象的声语号：采用音标标记的单词

限长优选多个声滴流队列的声节结构形式的拉丁语概念标记的声语号（单词的音标），本质上是一个个在目标概念系统中独一无二的、限制最大可变长度的语音型音乐的声节，实际使用声符总量等于采用拼音算法解压缩后的系统压缩声滴总量。

16.2. 概念对象的图言号：采用图言标记的单词

长优选1维图粒流阵列的映射图符结构形式的拉丁语概念标记的图言号（单词的图形标记），本质上是一个个在目标概念系统中独一无二的、限制最大可变长度的数学优选数值，数学数制计数系统的基数等于构词字母的总量。

16.3. “见词能读、听音能写”的关键

单体概念标记号（原子图言号、分子图言号）能够做到“见词能读、听音能写”的关键，是将声符与图符设计成具有规律性的、容易被学习的映射关系。

在世界语中，实际使用声符总量 = 压缩声滴总量 = 构成字母总量。在英语中，实际使用声符总量 = 压缩声滴总量 > 构成字母总量。因此，世界语可以实现“见词能读、听音能写”的设计要求；但是，在庞大词汇量暴增的场景里，图言号的平均表现，要么变得比 EN 语长；要么将无法做到声符与字母保持一一映射。

与世界语不同；将实际 34 + 8 个声符，采用不同的算法，一起混合映射到 26 个构词字母上，英语的这种构词法，使得大多数的英语单词，必须通过单纯的、简单的、额外的字母发音提示，才能完全做到“见词能读、听音能写”的效果。

英语声语概念标记号采用 298 个声符构成的语音型音乐音符系统，也等同于 298 进制的计数系统，因此，声语号的构词效率比采用 143 进制的世界语要高。虽然 EN 更符合语文的实际应用场景，但由于它是在语文历史进程中不断累加创造的结果，因此，与世界语不同，其大多数的词汇，缺乏像世界语那样无需母语环境的、简单的、可学习的规律性。

16.4. 可用原子图言号的最大总量，是进取型概念系统设计的关键参数

需要注意的是：在设计进取型通用语文概念系统的时候，不应该盲目地追求的简单化，无视宙沕概念对象的客观实际，而选取牺牲概念系统的总体性能的做法；特别是缩减可用原子图言概念标记号的总量等的重要基础性能。简单化与合理化，有时往往会发生矛盾。

概念的声语号与图言号的映射关系设计，理想状态是“1 对 1”映射；但是，对于一种进取型的语文概念系统来说，实际上，“1 对多”和“1 对多”是无法避免的。承认问题并解决问题，请参阅：Colating 声滴算法原理。

在 Colating 语中，通过扩展声符的数量，以及采用功能字母的构词形式，创造性地解决了以上问题，以及其他的构词问题。请参阅：Colating 的有关文章。

16.5. 在语文概念概念系统设计中的其他博弈因素

在现实语言场景里，语文概念概念系统的设计，就像采用 0~9 这 10 个数字来设计莫斯电报码系统一样，可以设计得很简单，也可以设计得很复杂。这主要取决于通用语文系统设计者的 RX、YQ 与智慧。参考：EN 语拼音原理分析。

17. 结论 Conclusion

17.1. 通用语文系统的展开声滴总量

可达：1百多个，乃至上千。

多声符概念系统典型为：数百个。单声符概念系统典型为：上千个。

17.2. 声滴拼读算法的本质

一种声滴压缩与解压缩的算法。

17.3. 概念标记号构成

概念标记号 = 声语号 + 图言号

17.4. 拉丁语单词的本质

拉丁语单词的本质，是首先建立以构词字母总量为基数的数学数制计数系统，然后，利用该计数系统中的不同优选数值（符合声节要求），来对不同的概念对象进行数值编号。

世界语单词：是一种单词整体发音优选的标准 28 进制计数系统的数值。

EN 单词：是一种单词整体发音优选的非标准 298 进制计数系统的数值。

Colating 单词：采用专用 C52++系统进行编码的优选数值。

17.5. 概念的声语号与图言号的映射关系设计

理想状态是“1对1”映射；但是，对于一种进取型的语文概念系统来说，实际上，“1对多”和“1对多”是无法避免的。

(本文结束)

自动目录: [期刊_随想_总目录](#) [独立目录 / All Single CONTENTS](#)

通用概念表示项联结框架总论

General Concept Representation Item Connection Framework Pandect

Ver: 0.2.0.1

Created: 2021/08/25 Update: 2021/09/15

Author: David Email: colating@yandex.com

[摘要 Abstract]: 对于一个宙泐概念对象，人们应当怎样来认识和表示呢？

理论上，一个概念可以关联无数多个表示项。概念表示项联结框架，就是研究应该选取哪些表示项？以及如何将这些选定的表示项给恰当地组合起来？

在通用语文概念系统中，一个概念通常同时关联了 3 个主要部分的内容，即：图言、声语、释义；其中，图言号通常必含“人为分类信息”，或含[本能会意（可选）]。因此，概念表示项联结框架，实际上，就是研究如何将这 3 个表示项给恰当地组合起来。其中，概记联糅是其核心内容。

所谓的概记联糅，实质上，就是期望在图言概念标记中，能够包含到概念的声语信息、提示性释义信息；另外，最好还能包含到本能会意信息。

[关键词 Keywords]: 概项联结框架 概记联组 概记联糅 声图映射关系，融合声滴拼读算法，

[编号 No.]: Doc Class: 概念标记方法/ Concept Identifier; Doc No.: 20210825-5

[目录索引]

1. 引言 Introduction

4. 通用概记联糅模型

2. 概念对象的异类表示项的选择与联结 5. 结论 Conclusion
3. 通用概记联组模型

⚠ [注意事项]

- “通用”这个限定词，为了简化描述，常常被省略。
- Open Version 只涉及 Colating 的实现原理。与其他语种的比较内容，只在 Member Version 部分。

[关联内容摘要]

- 融合声滴拼读算法
- 声滴与图符的映射关系
- 可用原子图号总量的重要性：满足概念表达合理精确度要求的基础。

[相关文章]

通用概念标记关联糅合模型的对比与选择

人类信息表达系统的 Colating 分类方法

[名词解释]

- 宙沕：宇宙内、表、外的一切。
- 概念对象/宙沕概念对象：概念所描述与标记的对象。
- 可用原子图号总量：符合目标语文系统的构词法规则，容易被目标语群大众读、写、记忆的原子图号的总量。只与图言号本身的读写记有关，与概念释义无关。
- 概号/概念号：通用概念标记；包括优滴列/声语概念标记号、图粒阵列/图言概念标记号。
- 声语概念标记号：定义了图粒、或图符的概念标记号。
- 图言概念标记号：定义了声滴、或声符的概念标记号。

[缩写]

- 概记：概念标记
 - 概项联结框架：通用概念表示项联结框架。
 - 概记联组模型：通用概念标记关联组合模型。
 - 概记联糅模型：通用概念标记关联糅合模型。
 - 声号/声语号：声语概念标记号。关联了图符。
 - 声滴号：优滴列概念标记号。纯语音，没有关联图符。
 - 图号/图言号：图言概念标记号。
 - 图粒号：优粒列概念标记号。纯图形，没有关联图符，如图画等。采用纯图画（图粒阵列）的方式，同样可以表达简单的概念。
 - 概记法：概念标记方法
 - 可双预射：可被双向预测的映射关系。
 - 否双预射：不可被双向预测的映射关系。
-

1. 引言 Introduction

对于一个宙泐概念对象，人们应当怎样来认识和表示呢？

在设计通用概念标记系统的时候，对于一个具体概念对象的标记和表示，应当包含哪些必须的不同类型信息呢？这些不同类型信息之间的关系，应当怎样设计才合理呢？

以上的这 2 个问题，分别是概记联组模型与概记联糅模型的研究内容，统称为概项联结框架。

除了针对概念对象描述的不同类型信息之外，对于相同类型的信息，可以通过恰当的概念释义等方式，对概念对象在不同论域中的表现，展开更为具体论述；这部分内容在其他文章中阐述。

2. 概念对象的异类表示项的选择与联结

对于同一个概念对象，可以采用多种方法来对其进行标记和表示，以适应不同的人群，或者用途等。

对于一个具体概念对象的标记和表示，如果将所有可能关联的异类表示项，以及这些表示项之间的映射关系，全部关联结合在一起；那么，这样的概念标记系统要么无法设计，要么就会因为臃肿而导致效率低下，直到奔溃。

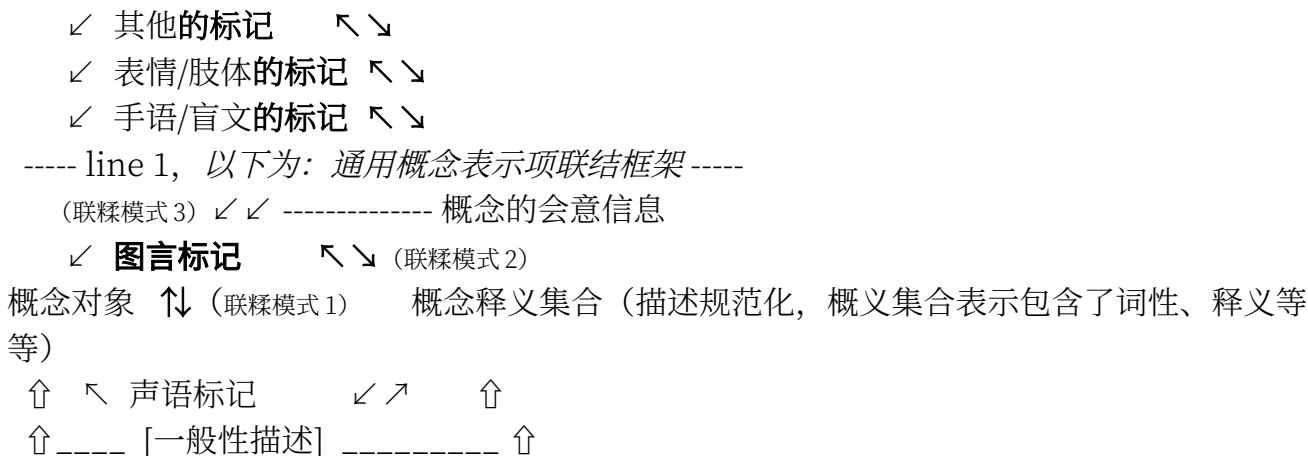
将概念对象的异类表示项联结在一起，这是一种有选择性的联结，并不是什么表示项都可以联结在一起。

就如“张三”这个概念，可以关联：名称发音、名称字符、性别、年龄、DNA、指纹、等等，但是，在具体的应用场景里，可能只需要选取其中的某些表示项。

2.1. 准全能概念表示项联结框架图

通过这个联结框架图，可以看出：一个概念对象，可以关联的异类表示项其实有很多；不过在一个具体的通用概念标记系统中，通常只选取了 line 1 以下的部分。

图 2-1 准全能概念表示项联结框架图



2.2. 联组模型与联糅模型的不同研究方向

- 概记联组模型，主要研究：图言、声语、释义，与概念对象之间的关系。
- 概记联糅模型，主要研究：图言、声语、释义，这3者之间的相互关系。

3. 通用概记联组模型

对于同一个概念对象的标记，将与概念对象有关的异类表示项关联组合在一起，称之为概记联组。这些异类表示项之间彼此相互独立，不存在、或不考虑这些属性之间的映射关系。

3.1. 参与通用概记联组的异类表示项

理论上，只要需要，一个概念就可以同时关联无穷多个表示项。

实际上，一个现代通用概念标记，通常只关联3个主要部分的内容，即：图言、声语、释义。

3.2. 概记联组的构成形式

标准概记联组 = 图言 | 声语 | 释义

声义概记联组 = 声语 | 释义 （注1）

未来概记联组 = 图言 | 声语 | 释义 | [*继续新增的表示项] （注2）

注1：利用声音来表达思想，可能这是地球上所有的语音动物都具有的共同特征之一。原始人类、成年文盲、或尚未学习图言号的幼童，以及在某些实际应用场景里，表达思想所采用的概念，一般只关联了2部分的内容，即：声语、释义。

注2：在未来时代，不排除一个概念会同时关联更多的表示项。

注3：概义与概号（声号、图号）本身的结构无关。在不同的语种里，同一个概念释义被映射到不同的概号上；在同一个语种里，在概念系统设计完成之前，理论上，可以将一个概义赋予任何一个概号联组。

3.3. 概念在释义典中的表现形式

一般的表示形式为：图言号，声语号，语法属性，论域及释义。

在论域中的释义，在有些释义典中也称为“条目”。

4. 通用概记联糅模型

在一个概念标记中，其不同类型的表示项之间，到底应当设计成什么样的关系呢？

也就是其“图言、声语、释义；其中，图言号通常必含“人为分类信息”，或含[本能会意（可选）]”之间，应该设计成什么样的映射关系呢？

在莫斯电报码系统设计中，同样是0~9这10个数字，可以设计成很简单的通用编码，也可以设计成异常难以被破解的非通用编码。在这一点上，其实，通用概念系统的设计，也与之相类似。

如果以成年文盲的立场，从学习和记忆的角度来看，那么，期望通过图言概念标记，能够获得哪些的信息呢？

如果从通用概念系统设计的角度来看，那么，期望在图言概念标记中，能够体现出哪些的信息呢？

4.1. 概记联糅的实质

从当今各个语种的概念标记系统来看，在通用语文概念系统中，一个概念通常同时关联了3个主要部分的内容，即：图言、声语、释义；其中，图言号通常必含“人为分类信息”，或含[本能会意（可选）]。其中，图言可以区分为整体形和根缀形；本能会意的方法有：拟声，象形，动画。

所谓的概记联糅，实质上，就是期望在图言概念标记中，能够包含到概念的声语信息、提示性释义信息；另外，最好还能包含到本能会意信息。

4.2. 联糅的意义

将声语信息糅合到图言标记中，以便通过图言标记就可以直接读取其优滴列发音。

将提示性释义信息糅合到图言标记中，通过已经记忆的概念内容映射表，这样就可以大致知晓目标图言标记的可能相关释义了。这就是词根和词缀的用途之一了。

将本能会意信息糅合到图言中，直接通过概记的声滴或图形的本能会意信息，这样就可以了解到概念释义可能会与什么事物有关了。

4.3. 联糅的实现方法

怎样将同一个概念，原本相互独立的、异类的表示项：声语信息、提示性释义信息、[本能会意信息] 糅合一个图言概念标记中呢？

注：[本能会意信息]不是必须的，但是，却可以有助于概念标记号的学习与记忆。一个有故事的概念标记号，如拟声、象形、动画等等，总比一个单纯的数值编号，要有趣得多。

4.4. 联糅的3种模式

声图联糅：将发声信息联糅到图言概念标记中。

义图联糅：将提示性释义信息联糅到图言概念标记中。

意图联糅：将[本能会意信息]联糅到图言概念标记中；只有相当少量的一些概念适用此法。

4.5. 联糅的优先级问题

在图言概念标记中，总是期望能够包含概念的全部信息，即：发声信息、提示性释义信息，[本能会意信息]。那么，当在不可兼得的情况下，应当如何取舍呢？其取舍的依据又是什么呢

4.5.1. 优先级

一个带有声语标记的图形，总是比一个不带声音的单纯图形，要容易学习、记忆、交流、传播。

发声信息的优先级最高，其次是提示性释义信息；[本能会意信息]的优先级最低。

- 实际上，在目标通用概念系统中，面对不断持续递增的庞大词汇总量，[本能会意信息]的作用和影响，变得并不重要了。又或者说，带有会意信息的概号外形被淹没在大量雷同概号外形的海洋里，从而显得渺小了。例如：Arrow 的 A 表示箭头，那么，arrow 的 a, Array 的 A，它们又该是什么东西的象形呢？
- 人类的语文学习从幼童时代就开始了，在那个时代的学习，只考虑声语概念标记，根本就不去关心图言概念的标记问题。
- 在一个概念标记中，除了发音与释义之外，还能带有某种可以想象的会意元素在里面

这无疑会更好；这样可以增加学习单词时的乐趣，不至于太枯燥，也有助于记忆等。

4.5.2. 象形文字系统

在[本能会意信息]中，象形标记由于能够跨语种，因此，常常被分割出来，构成独立的象形文字系统，非常适合在某些特定场合中应用。比如表示跌倒、男女之类的象形标记。

注：详细内容参见 [通用概念表示项联结框架。系列] 相关文章。

4.6. 联糅型概记在释义典中的表现形式

4.6.1. 联糅型限长多滴声节 1 维图言概念标记系统中的表现

部分、或全部的通过概念标记号的发音，可以直接被拼读出来。这取决于声图映射关系的设计构成概记的图粒、图符，与系统展开声滴、声符之间，存在某种映射关系。比如，这种关系可以是某种拼读规律、或语感等等。

有些概念标记号，隐含了[本能会意信息]。

4.6.2. 联糅型多声滴 2 维图言概念标记系统中的表现

这种概念构造模型，以及单声符概念构造模型，与 Colating 通用语文的创造理念不符；因此，在这里不做论述。

5. 结论 Conclusion

5.1. 概项联结框架的划分

概记联组模型：将不同类型表示项组合在一起。

概记联糅模型：不同类型表示项之间的映射关系。

5.2. 概念总量无限

5.2.1. 创造比继承重要，进步比保守健康。

5.2.2. 能够新增可用原子图号：是判断概念总量能够有效增加的主要标志。如果还能够新增可用原子图号，那么，说明目标概念系统还没有到达概念系统寿命的衰老拐点；此时，如非必

要，概念系统的架构可以保持原样。

5.2.3. 有效的合成型概念总量：参考某语种的概念系统，当原子图号总量在某个常用量级时，估算等于 4~7 倍的原子图号总量；当总量超过一定数量之后，其合成能力可能已经变得很弱了。

5.3. 概记联接的优先级

最高：发声信息。

次高：提示性释义信息。

最低：会意信息；在不可兼得情况下，可以被忽略。

5.4. 优选多声符概念标记的本质

声语号，是：以系统展开声滴总量为基数的进制计数系统的数值。

图言号，是：以构词字母总量为基数的进制计数系统的数值。

声语号与图言号之间的映射关系，有：可双预射，否双预射，其他类型的映射。

5.5. 拼音的本质

属于系统展开声滴的压缩与解压缩算法。

在概念系统设计阶段，采用声滴拼读算法。

在概念实际发声的运用阶段，采用“在说话时解压缩”、或“在说话时解压缩”的融合声滴拼读算法。

(本文结束)

自动目录：[期刊](#) [随想](#) [总目录](#) [独立目录 / All Single CONTENTS](#)

■ 语文失效机理 · 系列
A, Donor: Colating Group

词法、句法、语法的失效机理分析（待完成）

Ver: 0.1.0.0

Created: 2021/08/25 Update: 2021/08/31

Author: David Email: colating@yandex.com

[摘要 Abstract]: 。

[关键词 Keywords]: ,

[编号 No.]: Doc Class: 语法/ C; Doc No.: 20210825-8

[目录索引]

1. 引言 Introduction
2. 词汇的失效机理
3. 句法的失效机理
4. 语法的失效机理
5. 结论 Conclusion

[关联内容摘要]

[相关文章]

错误: 引用源未找到

[名词解释]

■ 概念: 心理学上认为, 概念是人脑对客观事物本质的反映, 这种反映是以词来标示和记载的。表达概念的语言形式是词或词组。概念是思维活动的结果和产物, 同时又是思维活动借以进行的单元。

[缩写]

1. 引言 Introduction

2. 词汇的失效机理

2.1. 失效类型

概念的释义失效

合成词的内部合成逻辑失效

概念标记号长度过长的失效

概念采用短语等表达方式，太长时，会导致句法失效。

2.2. 概念释义逻辑失效的实例

概念级失效，可能只有哪些具备自行创造概念的人，才有可能容易识别出来。

3. 句法的失效机理

当句子的平均长度接近 1 个肺活量的时候，句子的声语号流的表达，平均效率最高，舒适感最佳。

4. 语法的失效机理

5. 结论 Conclusion

6. 参考文献

参考文献格式：[序号]主要责任者. 文献题名 [J]. 刊名, 年, 卷 (期) : 起止页码。

[1]网络。

(本文结束)

自动目录: [期刊_随想_总目录](#) [独立目录 / All Single CONTENTS](#)

☆. 拉丁语概念字母构造长度的失效机理分析

Summary:

Version: 0.0.0.1 Created/Modified Date: 20211008 / 20211008

1. 肺活量生理因素

1.1. 拉丁语概念字母构造长度，不应超过 1 个肺活量

解决办法：需要限制单词的长度。

2. 合成词的合成逻辑失效

2.1. 处理表示词性的后缀之外，合成词的构成单元，不应超过 2 个

如果合成词的构成单元超过 2 个，当其中的一个构成单元存在不合理的逻辑矛盾时，错误可能将难以调整或被识别。

3. 不合理的合成词转义使用，产生的概念污染问题

3.1. 概念污染问题

采用不合理的转义法来创造的单词，对于普通的使用者而言，其脑袋未必也能够跟随着转义，于是，就形成了概念污染问题。

4. 过多的合成词转义使用，产生的系统混乱问题

4.1. 一个概念系统，只能有 1 种的主体构词法

当一个概念系统，同时存在 2 种或 2 种以上不同的主体构词法时，那么，将会导致概念释义系统的混乱。

(本文结束)

自动目录: [期刊_随想_总目录](#) [独立目录 / All Single CONTENTS](#)

随想录

Colating Language Capriccio

1

Total Issue 1

2021 · Irregular

(Open Version)

Ver: 0.1.0.2

Created: 2021/08/25 Update: 2021/08/25

Author: David Email: colating@yandex.com

[编号 No.]: Doc Class: 语音; Doc No.: 20210825-9

[随想目录 / Capriccio Contents]

[期刊_随想_总目录](#)

目录

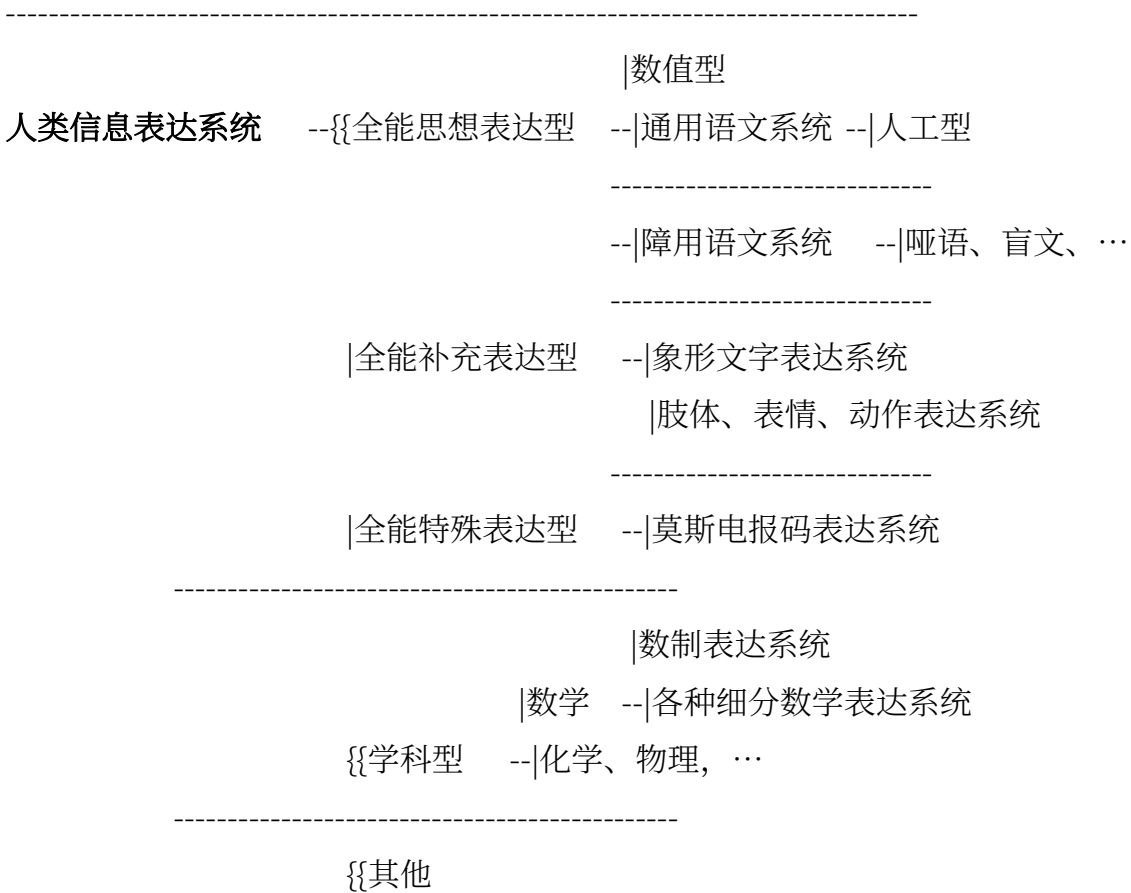
Contents

☆.人类信息表达系统的 Colating 分类方法

Summary:

Version: 0.0.0.1 Created/Modified Date: 20210912 / 20210912

按照用途来划分:



说明:

- ◇ 通用语文系统: 健用语文系统, 语文器官健全全能型语文系统。
- ◇ 障用语文系统: 语文器官障碍全能型语文系统。
- ◇ 全能思想表达型语文系统的局限: 由于采用了全能方法来表达思想, 因此, 在某些专门的场景里, 其表达的效率低下, 或者效果不佳。

(本文结束)

自动目录: [期刊_随想_总目录](#) [独立目录 / All Single CONTENTS](#)

☆.通用概念系统的设计步骤

Summary:

Version: 0.0.0.1 Created/Modified Date: 20211007 / 20211007

1. 声滴与声符子系统设计

1.1. 声滴子系统设计

展开声滴集合，包括压缩声滴集合、拼音扩展声滴集合、其他声滴集合。

其中，展开声滴，又称系统展开声滴，或概念系统展开声滴。

一般这几个集合之间的设计，是相互关联的，因此，需要统筹在一起进行联合设计。

1.1.1. 设计参数的至佳方向思考

1.1.1.1. 展开声滴设计参数的至佳方向

展开声滴的总量：越大越好。

- 优点：在相同声节长度的约束下，展开声滴的总量越大，可用的声语号总量也越多。
- 缺点：必须要考虑不同声滴之间的容易辨识度问题。
- 局限：主要受限于人类发音器官的天然本能发音能力。

1.1.1.2. 压缩声滴设计参数的至佳方向

在自然拼读的约束条件下，压缩声滴的总量：越小越好。

- 优点：压缩声滴的总量，要求适量小，这样声滴需要学习和记忆的代价就适量小。
- 压缩声滴的总量越小，与之映射的声符，以及于声符相映射的图符，就越容易创造。
- 缺点：在拉丁语中，当压缩声滴的总量小到某个拐点之后，如果在继续再小，将会导致与之对应的构词字母图符的总量，也会继续跟随着变小。
- 由此将导致采用构词字母组成的图言号数制计数系统的基数也变小。从而导致图言号

的构词效率变低。

- 局限：要想让所有的辅音与所有的元音，都能够通过自然拼读的方式来获得衍生的新声滴，并不是一件容易办到的事情。

1.2. 声符子系统设计

在单滴声节语文概念系统中，声滴的总量，等于声符的总量，一般它们之间是一一映射关系。

在多滴声节语文概念系统中，声滴的总量，是 5~20 倍的声符总量，一般它们之间可能是一一映射关系，也可能是 1 个声滴对应于多个声符的组合。

1.2.1. 在多滴声节语文概念系统中，设计参数的至佳方向思考

1.2.1.1. 压缩声符设计参数的至佳方向

在自然拼读的约束条件下，压缩声符的总量：越大越好。

- 优点：压缩声符的总量，越大，构造图言号的效率越高。
- 缺点：压缩声符的总量越大，与之对应的至佳构词字母的创造也会越困难。
- 局限：主要受限于人类发音器官的天然本能发音能力。

1.3. 声滴与声符映射子系统设计

一般声滴与声符的映射关系，主要表现为 2 种形式：规律性映射；无规律性映射。

2. 图粒与图符子系统设计

2.1. 图粒子系统设计

2.1.1. 在多滴声节语文概念系统中，设计参数的至佳方向思考

2.1.1.1. 图粒设计参数的至佳方向

至佳构词图粒的增长局限

虽然理论上，只要不超过最大展开声滴总量如 1500 个的图粒，都是可接受的。

实际上，假如想要创造数量超过 60 个左右，适合于多声符一维编码模式的”高中矮胖中瘦/如：hogmnl”特征的图符，可能会很困难。

2.2. 图符子系统设计

2.2.1. 在多滴声节语文概念系统中，设计参数的至佳方向思考

2.2.1.1. 图粒设计参数的至佳方向

至佳构词图粒的增长局限

虽然理论上，只要不超过最大展开声滴总量如 1500 个的图粒，都是可接受的。

实际上，假如想要创造数量超过 60 个左右，适合于多声符一维编码模式的”高中矮胖中瘦/如：hogmnl”特征的图符，可能会很困难。

2.3. 图粒与图符映射子系统设计

在单滴声节语文概念系统中，图言号由规则型图粒（图符）和非规则型图粒所构成的，图符一般指偏旁部首。单纯由图符只能构成很少的一部分图言号。

在多滴声节语文概念系统中，图言号一般由规则型图粒（构词字母图符）所构成的。图粒直接被用作图符，即图符就是构词字母。

3. 声符与图符映射子系统设计

3.1. 图粒子系统设计

3.1.1. 在多滴声节语文概念系统中，设计参数的至佳方向思考

3.1.1.1. 图粒设计参数的至佳方向

至佳构词图粒的增长局限

在单滴声节语文概念系统中，声符与图符的映射，由人为单独指定。

在拉丁语多滴声节语文概念系统中，声符与图符的映射，主体映射关系除了基准映射是人为指定之外，理论上，其余的都是希望采用规律型的映射关系。

(本文结束)

自动目录: [期刊](#) [随想](#) [总目录](#) [独立目录 / All Single CONTENTS](#)

☆. Colating 单词的分段构造模型图

Summary:

Version: 0.0.0.1 Created/Modified Date: 20210912 / 20210912

表 x-1: Colating 单词的分段构造模型图

	段 1		段 2		段 3			段 4	段 5			段 6		
位置	1	2	3	4	5	6	7	8	9	10	11	12	13	14
字母流	↑	前缀字母流	↑	↑	左扩字母流	↑	↑	主体词根 / 根词	↑	↑	右扩字母流	↑	↑	后缀字母流
功能字母标记	归档功能字母	...	前缀标记位 1	前缀功能字母	...	左扩标记位 2	左扩功能字母	...	右扩功能字母	右扩标记位 3	...	后缀功能字母	后缀标记位 4	...

注 1: 创造要求

- (1) A 级高频词: 一般不含功能字母, 力求简短。
- (2) 一般合成词: 段 3 和段 4, 不要同时使用。
- (3) 非高频的专业词汇: 可以包含全部的构词片段。
- (4) 标记位: 使用少数的几个普通构词字母、或功能字母组合来充当。

(本文结束)

自动目录: [期刊_随想_总目录](#) [独立目录 / All Single CONTENTS](#)

☆. 同音字原理

Summary:

Version: 0.0.0.1 Created/Modified Date: 20210917 / 20210917

1. 施氏食狮史原文

石室诗士施氏，嗜狮，誓食十狮。施氏时时适市视狮。十时，适十狮适市。是时，适施氏适市。施氏视是十狮，恃矢势，使是十狮逝世。氏拾是十狮尸，适石室。石室湿，氏使侍拭石室。石室拭，施氏始试食是十狮尸。食时，始识是十狮尸，实十石狮尸。试释是事。

2. 原因分析

- 2.1. 宙沕概念对象的绝对无限性
- 2.2. 人类智慧的相对无限性
- 2.3. 概念总体新陈代谢的历史漫长性
- 2.4. 人类语音器官的天然本能局限性
- 2.5. 地球寿命的有限性
- 2.6. 人类个体寿命的有限性，与人类整体寿命的漫长性
- 2.7. 人类渴望不断进步，渴望迈向宇宙社会的进取精神

(本文结束)

自动目录: [期刊](#) [随想](#) [总目录](#) [独立目录 / All Single CONTENTS](#)

☆. Colating 构词法思想指引

Summary:

Version: 0.0.0.1 Created/Modified Date: 20210914 / 20210914

1. 学习需要乐趣

对于限长多滴声节 1 维图言概念标记号，从数学上的理解，就是一组优选发音的数值，似乎不需要考虑[本能会意信息] 问题。但是，如果从学习的乐趣上来看，如果能够包含[本能会意信息]，比如象形，当然是更好。

怎样将同一个概念，原本相互独立的、异类的表示项：声语信息、提示性释义信息、[本能会意信息] 糅合一个图言概念标记中呢？

2. 不能单纯为了单词在[本能会意信息]造型上的那点乐趣，而忘乎所以。

事实上，随着学习到的概念数量的不断增加，越是高科技词汇的概念对象，概念将越来越抽象。

单词的[本能会意信息]，对于已经掌握了足够多日常高频概念的人们来说，这类人关注将不是所谓的如文字象形信息，而是词根的释义。在实际的语言使用中，甚至忘记了如象形之类信息的存在，因为习惯成自然了。

单词的[本能会意信息]，如所谓的象形构造的作用，对于母语文盲者来说，那只不过是一个传说。

手机号码 138xxx，从中可以看出它像中移动公司吗？邮政编码 0755，从中可以看出它像 SZ 吗？

3. 创造单词需要考虑不同词性的优先级

对于限长多滴声节 1 维图言概念标记号，太长的字母构造，将导致句子因长度太长而失效，所以需要考虑语文系统的句子平均长度问题。

- 动词、介词的发音：简短、响亮、有力；优先级最高。

4. 创造单词需要考虑[本能会意信息]的优先级

- 拟音：比象形的优先级高。
- 象形：前 3 个字母、或音节，最重要。
- 动画：前 3 个字母、或音节，最重要。love, olmove

5. 高频 0 级单词考虑，需要考虑[本能会意信息]的好处

- 隐含[本能会意信息]的单词，无法涵盖全部。
- 只考虑应用在高频 0 级单词上，这样入门学习最容易。
- 除了基础概念之外，概念越多，越抽象；越抽象的概念，要想隐含[本能会意信息]可能无法办到。因此，只有在入门学习期间，才需要考虑单词的[本能会意信息]问题。
- 名词尽可能将[本能会意信息]隐含进来。
- 介词、高频动词、小品词等的发音，尽可能与英文相同。

6. 怎样来认识[本能会意信息]在单词构造中的作用呢？

- 声语信息、提示性释义信息，在单词构造中是必须要考虑的要素。
- [本能会意信息]有 3 种：拟音、动画、象形。隐含[本能会意信息]的单词，只是对语文入门学习者有利。
- 思想越是进步的社会，抽象概念的数量越多；因此，不应该过分强调单词中的[本能会意信息]，以免迷失了对概念标记本意的理解方向。

- 过分强调非抽象概念的象形等类型的概念，说明思想还没有进入到抽象的世界。

(本文结束)

自动目录: [期刊_随想_总目录](#) [独立目录 / All Single CONTENTS](#)

☆. Colating 单词的 3 种书写格式

Summary:

Version: 0.0.0.1 Created/Modified Date: 20210917 / 20210917

1. 3 种书写格式

1 维护条形：常规书写模式。

2. 换行为 1 个单词的结束：

2 维 6 格（3 列 x2 行）：应用在、如在数值与普通单词的混合编辑场景。

2 维 9 格（3 列 x3 行）：应用在、如在数值与普通单词的混合编辑场景。

3. 换行为 1 个单词的继续：

2 维 8 格（4 列 x2 行）：应用在如广告牌等场景。

2 维 12 格（4 列 x3 行）：应用在如广告牌等场景。

4. 换行为 1 个单词的继续：

2 维 10 格（5 列 x2 行）：应用在如广告牌等场景。

2 格与 15 格（5 列 x2 行）：采用不同的空格符号来区分，长度为 3 的单词，采用不同的格行来区分。

5. 为了配合这 3 种书写格式，3 个字母长度的单词创造，被系统严格规定。

不能自由创造

优先级：基数词，单位（百、千、万。。），??

（本文结束）

自动目录：[期刊](#) [随想](#) [总目录](#) [独立目录 / All Single CONTENTS](#)

☆. “白马非马” 浅析

Summary: 语法错误造成了理解的困难。从表面上看是逻辑问题，实际上是语法缺陷问题。

Version: 0.0.0.1 Created/Modified Date: 20210909 / 20211003

1. “白马非马” 的故事原文

1.1. 原文与译文

原文:

“白马非马，可乎？”

曰：“可。”

曰：“何哉？”

曰：“马者，所以命形也。白者，所以命色也。命色者，非命形也，故曰白马非马。”

译文:

问：可以说白马与马不同吗？

答：可以。

问：为什么？

答：“马”是对物“形”方面的规定，“白马”则是对马“色”方面的规定，对“色”方面的规定与对“形”方面的规定性，自然是不同的。所以说，对不同的概念加以不同规定的结果，白马与马也是不同的。

2. 集合关系认识论

2.1. 集合与集合之间的关系

2.1.1. 子集与真子集

子集：如果集合 A 中的任何一个元素都是集合 B 中的元素，那么称集合 A 为集合 B 的子集。

真子集：如果集合 A 是集合 B 的子集，且集合 B 中存在不属于集合 A 中的元素，则称集合 A 是集合 B 的真子集。

2.1.2. 相等

如果两个集合的元素完全相同，则称两个集合相等

2.1.3. 其他集合的定义

空集，非空集合，全集，补集。

2.2. 概念推理

2.2.1. 全同关系： $A=B$ 。例：“马铃薯”：“土豆”。

2.2.2. 包含关系： $A>B$ 。例：男人：人

2.2.3. 其他关系

交叉关系： $A(A_1、A_2、B_1.....)$ ， $B(B_1、B_2、A_2.....)$ 。例：“学生”：“运动员”。

全异关系： $A-B$ 。例：奇数：偶数

3. 认知思维方法

在人类认识的现实世界中，“等同差异比较”和“隶属包含比较”，往往同时存在。

对于 1 个目标概念对象的认识，人们常常通过与参考概念对象进行某种模式的比较来获得的。

3.1. 等同关系：“等同差异比较”，比较与鉴别认知法

极为相似的双胞胎，大李和小李，一般是通过等同差异比较法来进行鉴别与辨识的。

对所有需要关注的指标参数分别进行“等同差异比较”，从而获得“判同”与“判异”的经典逻辑论结果，或者获得隶属度的模糊逻辑结果。当所有的比较结果都在判断相同的容许误差之内时，那么，这2个集合就是等同的。

3.2. 包含关系：“隶属包含比较”，概念蕴含推理认知法

“猪”是“动物”；“猪”集合包含在“动物”集合之内。一般是通过隶属包含比较法来进行鉴别与辨识的。

对所有需要关注的指标参数分别进行“包含比较”，当集合A包含了集合B所有的指标参数时，那么，集合 $B \subseteq$ 集合A。也称之为：集合B包含于集合A，或集合A包含了集合B。

3.3. “等同差异比较”与“隶属包含比较”的主要区别

3.3.1. 这2种不同的思维模式的比较

3.3.1.1. 等同差异比较

- 比较大小、程度，等等的概念对象属性或特征。
- 比较误差范围。
- 举例：在通常的场景下，买1斤菜，结果实际的重量是1.01斤；买家和卖家，彼此都不会觉得亏欠。

3.3.1.2. 隶属包含比较

- 比较大小、程度，等等的概念对象属性或特征。
- 不用考虑误差范围。
- 举例：在通常的场景下，买1斤菜，假如实际的重量是10斤，买家和卖家，彼此的感觉就是不一样，可能的结果是买家高兴，卖家郁闷。

3.3.2. 这2种不同的思维模式，不能采用相同的语法动词

3.3.2.1. 不能混用

假如这2种思维采用相同的语法动词，也就是混用，其结果就会像“白马非马”那样，在群体的理解上，产生难解的逻辑矛盾问题。参见：表X-1：“白马非马”的不同思维模式比较。

3.3.2.2. 不同问题，不同思维

从数学的角度上来说，一个是等价判断，一个是包含判断，这是不同性质的问题。就如在说明身高的时候，一般是不能够使用重量千克来解答的。

4. 看清楚“白马是马”问题

4.1. 怪圈现象图表

“是” / “非”	判断方法	白马非马	白马是马
比较鉴别思维	等价关系	正确	错误
隶属包含思维	包含关系	错误	正确

4.2. 怪圈现象图表的分析说明

4.2.1. 采用判同比较思维，得出的结论是：“白马非马”，是正确的。

在这个场景里，“白马是马”是错误的；

假如说“白马是马”是正确的，那么，“白马”和“马”是等价的吗？

“买白马”就是“买马”吗？“买马”就是“买白马”吗？这显然是谬论。

4.2.2. 采用包含比较思维，得出的结论是：“白马非马”，是错误的。

在这个场景里，“白马是马”是正确的；

假如说“白马是马”是错误的，那么，“白马”难道真的不是“马”吗？

4.2.3. 思维模式的对掐

4.2.3.1. 相互对掐的最后结果，就是都有道理，谁也不能真正说服对方。

- 一部分人习惯性地采用“等同差异比较”思维，得出的结论是：“白马非马”。
- 另一部分人习惯性地采用“隶属包含比较”思维，得出的结论是：“白马是马”。

4.2.3.2. 同一个人的不同思维习惯

- 在比较双胞胎时，习惯性使用了“等同差异比较”思维，用于判定“差异”。对极为相似的双胞胎，大李和小李的判定，不存在隶属关系。
- 在比较“蛋”时，习惯性使用了“隶属包含比较”思维，用于判定“隶属”。如鸡蛋、鸭蛋；当然，也可能同时使用了“等同差异比较”思维，比较不同“蛋”的大小。

4.2.3.3. 比较与鉴别认知法”与“概念蕴含推理认知法”，这是2种不同的思维模式。

- 采用“比较与鉴别认知法”提出来的问题，不能采用“概念蕴含推理认知法”来解决和处理；反之一样。
- “白马非马”的问题，就是采用1种思维模式来提问，却采用另一种思维模式来解答，这种相互交错的解释，导致了在思想上的陷阱和误区。

就如同提出一个身高的问题，通常是不能采用处理体重的方法来解答的，其道理是类似的。

4.2.3.4. 关键的问题是：

“比较与鉴别认知法”与“概念蕴含推理认知法”，这2种思维都是合理的，都是人类所必须的。并且，同时存在，纠缠在一起。

当这2中思维方式，在语法上使用相同的推理动词；同时，由于人们习惯思维混淆的惰性，从而使得“白马非马”问题陷入了难解的境地。

除非读者具有自创原子图言号的能力，否则，难解。

5. 简评理论分析

5.1. 公孙龙子的思维

5.1.1. 哪个正确？

“白马非马”正确？还是“白马是马”正确？

在这些论断中，没有理解清楚“概念对象”与其“内部属性”之间关系，并随意地割断它们之间的关系，从而导致了莫名其妙的鬼辩。

5.1.2. 一个非人造的宇宙概念对象，可能同时具有无数多种属性

“马”这个概念，是采用归纳法，在综合聚类之后，人为创造出来的一个非0级聚类集合概念。

“马（本体）”同时具有“形”、“色”，以及其他的属性。就如“张三（本体）”一样张三可以同时拥有肤色，身高，职位，不同的称谓等属性。具体分析见下面的小节。

5.2. 问题症结：是“语法”问题，采用不同的动词来区分，那就不是问题了

5.2.1. 解决方法举例，采用拉丁语语法，可以解决这个问题。

5.2.1.1. “is” : White horse is not horse. “白马不是马”。

5.2.1.2. “belong to” : White horse belong to horse. “白马属于马”。

5.2.2. **特别说明**：等价关系与包含关系，在语法上，需要分别采用不同的动词来区分。

否则，对“白马非马”的理解，将陷入到思维的陷阱里。

6. 理解“白马非马”的哲学意义

6.1. “比较鉴别思维”与“隶属包含思维”，是日常不可或缺的2种思维方式。

- 在句法上，需要采用不同的动词来区分它们。
- 假如采用相同的语法动词，那么，就会陷入在“白马非马”的思维陷阱里，相互对掐，难以解脱。

6.2. 理解“白马非马”问题的作用机理，是识别“白马非马”类问题概念的关键。

(本文结束)

自动目录: [期刊_随想_总目录](#) [独立目录 / All Single CONTENTS](#)

☆. 宙沕概念对象的种类与感知

Summary: 在通用语文概念系统中，图像与声音，是标记概念对象的最主要手段。

Version: 0.0.0.1 Created/Modified Date: 20210909 / 20210909

视觉：图像，可被眼睛察觉到的有形概念对象

听觉：声音，发出纵波的概念对象

嗅觉：气味，无形概念对象

味觉：味道，无形概念对象

触觉：触觉

意智觉：意识或智慧感知，属于人脑功能的一部分。

(本文结束)

自动目录: [期刊_随想_总目录](#) [独立目录 / All Single CONTENTS](#)

☆. 限长多滴声节 1 维图言概念标记的本质

Summary: 为什么 Colating 概念标记号采用的是拉丁语单词构造模型?

Version: 0.0.0.1 Created/Modified Date: 20210909 / 20210909

1. 通用概念标记的关联结合

通用概念标记联结模型 = 声语概念标记号 + 图言概念标记号 + 概念释义 + 语法属性。

其中，声语号不可或缺，这可能是整个地球语音动物的必需品。

图言号，可能有，也可能没有。即便是有，也会存在着各种各样的图形标记符号和标记方法。图言号不是天生本能就能够拥有的，它是人类智慧创造的产物。只有脱离了动物智慧层级的语文群体，图言号才会被人为创造产生。

2. 声语号的本质

2.1. 思想表达的声滴流形式

人类在利用语音来表达思想的时候，对于完全不懂当前外文的人来说，语音的对外展现，总是表现为声滴流形式。

2.2. 不同语种的声滴流之间的差异

可能存在差别的声滴基波频率

声滴总量的差异：展开声滴，或压缩声滴的总量差异。

声滴流编码与解码格式之间的差异

2.3. 声语号与声滴流之间的映射关系

多声滴声语号系统，1 个声语号对应于 1 或多个声滴流。

单声滴声语号系统，1 个声语号对应于 1 个声滴流。

2.4. 声语号的本质之一：声语号与声节的关系

1 个声语号对应于 1 个声节。

声节与音节的共同点：不是什么样的音符队列，都能成为音节；不是什么样的声符队列，都能成为声节。

声节必须要考虑其总体发音，以及声节的限制最大可变长度问题。

2.5. 声语号的本质之二：声语号与数学数制计数系统的关系

2.5.1. 在任何一种通用语文系统中，声滴的总量，都是有限的。

2.5.2. 在某主流典型优选多声符概念的通用语文系统中，理论上只有约 298 个标压声滴；在某主流典型单声符概念的通用语文系统中，实际统计只有约 1500 个标压声滴。

2.5.3. 限长多滴声节 1 维图言概念标记号，即拉丁语单词的本质，是一个个限制最大可变长度的语音型音乐的声节，如 EN 语实际使用声符总量的理论值约等为 298 个；同时，也是一个个在目标概念系统中独一无二的、限制最大可变长度的数学优选数值，该数学数制计数系统的基数、如 EN 语为 26。

3. 图言号的本质

3.1. 世界语

将 28 个压缩声符（雷同 IPA 音标）映射到 28 个实际构词字母上，世界语的这种构词法，使得世界语单词能够做到“见词能读、听音能写”的效果。

3.2. EN 语

与世界语不同；将实际 34 +8 个声符，采用不同的算法，一起混合映射到 26 个构词字母上，英语的这种构词法，使得大多数的英语单词，必须通过单纯的、简单的、额外的字母发音提示才能完全做到“见词能读、听音能写”的效果。

英语声语概念标记号采用 298 个声符构成的语音型音乐音符系统，也等同于 298 进制的计数系统，因此，声语号的构词效率比采用 143 进制的世界语要高。虽然 EN 更符合语文的实际应用场景，但由于它是在语文历史进程中不断累加创造的结果，因此，与世界语不同，其大多数的词汇，缺乏像世界语那样无需母语环境的、简单的、可学习的规律性。

需要注意的是：在设计进取型通用语文概念系统的时候，不应该盲目地追求的简单化，无视宙泐概念对象的客观实际，而选取牺牲概念系统的总体性能的做法；特别是缩减可用原子图言概念标记号的总量等的重要基础性能。简单化与合理化，有时往往会发生矛盾。

3.3. 可乐听语 Colating Language

Colating 语对声滴和图符都进行了扩展，不论是声语号，还是图言号，概念的标记能力都远远超过了当今所有的在用拉丁语。由于它具有容易被记忆和学习的显式规律性，因而，相比 EN 语还要简单明了。

(本文结束)

自动目录: [期刊_随想_总目录](#) [独立目录 / All Single CONTENTS](#)

☆. 莫尔斯电码、手机号码与拉丁语单词的系统设计比较

Summary: 为什么 Colating 概念标记号采用的是拉丁语单词构造模型?

Version: 0.0.0.1 Created/Modified Date: 20210925 / 20210929

1. 游戏场景

1.1. 人员与装备

1.1.1. 每个人的右手上都拿着 1 部可用的手机，每部手机对应着 1 个与众不同的唯一号码。

1.1.2. 这些人穿着五颜六色的衣服，以及在衣服上印刷着各自的手机号码。

1.1.3. 这些人在身上还带着各种各样的装饰品。

1.1.4. 每个人的左手都关联着 1 个与众不同的宇宙概念对象，或者说手里拿着 1 张标识该概念对象的唯一标记号。不同种类，不同属性，或者说是需要区分的、任何的与众不同。

1.2. 原理说明

1.2.1. 不同的手机号码与不同的概念标记号，它们都只不过是不同的编号而已。

- 不同手机号码的声语号，对应于不同的声语概念标记号。如拉丁语单词的音标。
- 不同手机号码的图言号，对应于不同的图言概念标记号。如拉丁语单词的字母组合。

1.2.2. 每个手机号码指向 1 个概念对象。

- 每个概念对象的声语概念标记号，就是在词典中 1 个单词的音标。
- 每个概念对象的图言概念标记号，就是在词典中 1 个单词所有构成字母的队列。

1.2.3. 不同的编号方法。

- 手机号码以图言号为基准，采用 10 进制固定长度的流水号编码。手机号码的声语号可以是各种各样的，只要求将不同的图言号映射到不同的声语号即可。

- 拉丁语的声语概念标记号，采用以“构成字母总量-1”为进制计数系统的限制最大可变长度的优选数值。拉丁语的声语概念标记号，采用以“系统音标总量-1”为声符的限制最大可变长度的语音型音乐的声节。

1.2.4. 通过不同的手机号码，就可以获得不同概念的释义。

- 当打电话到某一个手机号码时，就可以通过这个号码获知所对应概念对象的详细内容。

1.2.5. 五颜六色的衣服，以及在衣服上印刷着各自的手机号码。

本段内容，这里不展开讨论。

- 代表了不同字体形状构造的图言号。
- “人靠衣服，马靠鞍”，人们常常会因为“本体（存在对象张三）”的披红挂彩的外衣，而忽略了对张三本体的认知。

1.2.6. 概念对象，除了需要关联其内在属性之外，可以关联的概念外部属性，其实，还可以有很多。外部属性是概念对象与其他概念对象之间所发生的联系，关联的类型有：

- 概念对象与其他概念对象之间所发生的客观联系。
- 概念对象与其他概念对象之间所发生的联系，是人为指定的结果。

2. 三类系统设计比较表

表 SJ-1: 莫尔斯电码、手机号码与拉丁语单词的系统设计比较 (1) (后接续表 2, 3)

比较项目	标记号	摩尔斯电码 (类1)	手机号码 (类2)	世界语 (类3)	EN 语 (类3)	Colating (类3)
编码系统	声语号	*声语号系统设计基准: 2 进制 5 位限长计数系统	映射于图言号	*声语号系统设计基准: 约为 143 进制计数系统, 采用 28 个压缩声滴来表示	*声语号系统设计基准: 约为 400 进制计数系统, 采用 48 个压缩声滴来表示	*声语号系统设计基准: 约为 1000 进制计数系统, 采用 52+x 个压缩声滴来表示
	图言号	映射于声语号	*图言号系统设计基准: 10 进制计数系统	28 进制计数系统	26 进制计数系统	52++ 进制计数系统
声语号与图言号的映射关系	-	1 对 1 映射	1 对 1 映射	1 对 1 映射	主要为: 1 对 1, 1 对多, 多对 1	主要为: 1 对 1, 1 对多, 多对 1, Colating 规律声滴扩展
声语号输出的解压缩模式	--	不需要解压缩	在说话时解压缩	在说话时解压缩	在说话时解压缩	在说话时解压缩
读写效果	-	见电码能读、听电码能写	见号码能读、听号码能写	见词能读、听音能写	部分单词能做到: 见词能读、听音能写	几乎全部单词都能做到: 见词能读、听音能写。做不到的也会有提示标志
数值区别	-	优选数值	可以是流水号	类似音乐音节, 声滴排列要求有: 高低长短的起伏和节奏	类似音乐音节, 声滴排列要求有: 高低长短的起伏和节奏	类似音乐音节, 声滴排列要求有: 高低长短的起伏和节奏

表 SJ-1: 莫尔斯电码、手机号码与拉丁语单词的系统设计比较 (2) (后接续表 3)

比较项目	标记号	摩尔斯电码 (类1)	手机号码 (类2)	世界语 (类3)	EN 语 (类3)	Colating (类3)
图言编码长度 (数值长度)	-	2 个~限制最大可变长度为 5 个	固定	2 个~限制最大可变长度	1 个~限制最大可变长度	2 个~限制最大可变长度
独立使用	-	一般以排列组合队列形式使用	是	一般以排列组合队列形式使用	一般以排列组合队列形式使用	一般以排列组合队列形式使用
	声语号	常用缩写	可忽略	存在拟音的	存在拟音的	存在拟音的

隐含信息种类的数量注(3)	图言号	可包含的隐含信息多于手机号码	可包含数个隐含信息	可包含中低量隐含信息	可包含中大量隐含信息	可包含大量隐含信息
标记号的冗余设计	声语号	理论上不需要考虑	映射于图言号	受限于声节的限制最大可变长度	受限于声节的限制最大可变长度	受限于声节的限制最大可变长度
	图言号	理论上不需要考虑	号码的总量, 超过实际用户	映射于声语号。	映射于声语号	映射于声语号。
**原子概念标记的创造难度	声语号	简单	简单	简单	简单	简单
	图言号	简单	简单	简单	简单	简单
自创造标记号的释义权	-	编码者	运营商	创造者、发明者, 或其他	创造者、发明者, 或其他	创造者、发明者, 或其他

表 SJ-1: 莫尔斯电码、手机号码与拉丁语单词的系统设计比较 (3) (续表 3)

比较项目	标记号	摩尔斯电码 (类 1)	手机号码 (类 2)	世界语 (类 3)	EN 语 (类 3)	Colating (类 3)
概念学习难度	声滴流	专门训练人员可使用	容易	容易	容易	容易
	图符流	需要根据映射关系	容易	容易	容易	容易
	**声滴图符映射	容易	容易	容易	复杂, 部分单词没有明显的规律性。1 对多, 多对 1	容易, 具有明显的规律性。1 对多, 多对 1
	**图符流顺序映射	简单, 或异常复杂 (映射加密)	已是正常的队列顺序, 不需要。	已是正常的队列顺序, 不需要。	已是正常的队列顺序, 不需要。	已是正常的队列顺序, 不需要。
	**图号流的语法顺序映射					
	释义	与通用概念发生映射	简单	中等	中等	中等

注:

(1) 同音字总量: 按照牛津 OED 1989 年第 2 版收录了 301,100 个主词汇计算, $30000/1500=200$; 这意味着, 假如达到牛津第 2 版的主词汇量, 那么, 在单滴声节 Z 概念系统中, 每个声滴平均对应的同音字, 将达到 200 个。

(2) 多个声语号的排列组合构成队列，类似音乐音节，要求有：高低长短的起伏和节奏。

含“人为分类信息”，或含[本能会意（可选）]

(3) 隐含信息：包括“人为分类信息”，或含[本能会意]种类的数量。

网摘：摩尔斯电码也被称作摩斯密码，是一种时通时断的信号代码，通过不同的排列顺序来表达不同的英文字母、数字和标点符号。它发明于1837年，是一种早期的数字化通信形式。不同于现代化的数字通讯，摩尔斯电码只使用零和一两种状态的二进制代码，它的代码包括五种：短促的点信号“·”，保持一定时间的长信号“—”，表示点和划之间的停顿、每个词之间中等的停顿，以及句子之间长的停顿。

3. 概念与语法系统的简单化，或复杂化设计

出于不同的思想目的，不同的造字智慧，通用概念系统与通用语法系统，可以设计得很简单，也可以设计得很复杂，见下表：

表 SJ-2: 3. 概念与语法系统的简单化，或复杂化设计

关键参数	标记号	摩尔斯电码	手机号码	系统设计简单化	系统设计复杂化	说明
声滴	语音	简单	简单	**设计关键: 采用拼音压缩算法	**设计关键: 直接使用不压缩的系统展开声滴法	
	图言号	简单	简单	**设计关键: 按照数学数制计数系统，来进行图言号的设计	**设计关键: 采用非数学进制计数系统。或者说，采用数学无穷进制计数系统	
图粒	-	1对1映射	1对1映射	**设计关键: 1对1, 1对多, 多对1, 以及在单词中具有规律性的提示标志	**设计关键: 直接人工指定。声滴与图粒之间，不存在映射关系	
声滴与图符映射关系	--	简单，一般不需要考虑	简单，一般不需要考虑	**设计关键: 在说话时解压缩	**设计关键: 在说话前解压缩	
在图言号中隐含释义信息的种类		简单	简单	**设计关键: 尽可能多	**设计关键: 尽可能少	

图符流顺序映射		一般为正常顺序	一般为正常顺序	一般为正常顺序	一般为正常顺序	
原子图言号的创造难度		简单	简单	**设计关键: 简单。意味着概念标记号的精度可以不断提升	**设计关键: 高难度。意味着概念标记号的精度提升的代价昂贵	
图号流的语法顺序映射		**设计关键: 简单化采用正常顺序。复杂化采用高强度的加密顺序	一般为正常顺序	**设计关键: 按照数学逻辑推理原理来设计语法系统。	**设计关键: 按照违反数学逻辑推理原理来设计语法系统。	
释义		取决于目标通用语文系统	取决于运营商			

4. 结论

4.1. 拉丁语单词的特点

4.1.1. 声语号，即单词的音标：

- 音乐音符特性，声符的总量为展开声滴的数量，其最大数量受限于人类天然本能语言能力。
- 音乐音节特性，声语号应当容易识别，因此，1个单词的声语号构造，与音节一样，具有高低长短的音乐音节的特征，称之为“声节”。
- 声节的最大长度：受限于人类个体的最小肺活量。在1个声节的发音过程中，不可以进行换气，否则，会被识别成2或多个的单词。
- 与音乐音节不同：为了简单化系统展开声滴，通常采用拼音压缩与解压缩算法对展开声滴进行处理。

4.1.2. 图言号，即单词的图符：

- 数值特性：由字母构成的多声滴声节概念标记号系统，实质上，是以“构词字母总量-1”为数学进制计数系统中的优选数值。

4.1.3. 映射关系，在通用概念系统中，主要是：声滴与图符之间的映射

- 密码特性：映射关系可以设计得或很简单，或很客观，或很复杂。
- 单词设计成“见词能读、听音能写”的关键：声语号与图言号，设计成具有规律性的映射关系。

4.2. 概念标记号设计的依据

4.2.1. 概念对象的种类和数量

概念标记号是用来标记概念对象的，因此，确定概念对象的总量，是设计通用概念系统的第1个关键参数。

4.2.2. 概念标记号的，取决于宇宙概念对象的总量，以及概念标记号在新陈代谢过程中的平均有效寿命。

4.2.3. 声语号，受限于人类天然本能语言能力。

4.2.4. 图言号，受限于人类个体的记忆能力。

4.2.5. 是否需要客观：如是《牛津英语词典》那样，需要客观地记录英语词汇。客观就必然要求将原子图言号的创造权利，掌握在创造者或发明者的手上。

(本文结束)

自动目录: [期刊_随想_总目录](#) [独立目录 / All Single CONTENTS](#)

A General Human Language of Latin

This article: A, Donor: Colating Group

Appendix 附录

Colating Language Periodical & Capriccio

1

Total Issue 1

2021 · Irregular

(Open Version)

Ver: 0.1.0.2

Created: 2021/08/25 Update: 2021/09/2

Author: David Email: colating@yandex.com

[编号 No.]: Doc No.: Per 2021-1 / Cap 2021-1

自动目录: [期刊_随想_总目录](#) [独立目录 / All Single CONTENTS](#)

目录
CONTENTS

Appendix A

附录包含：名词解释 1：声音类、图形类、概念标记类。

引言

● Colating 概念重新创造的必要性

为了避免在概念理解上可能产生的混乱。Colating 对可能产生释义混乱的概念，重新进行了创造，这使得 Colating 的理论术语可以自成体系。

假如继续使用这些含糊不清的名词概念，那么，可能不论是读者，还是 Colating 的作者，最后都有可能将自己给搞混淆错乱了。👁️ 虽然这些名词是 Colating 重定义的结果，此时可以避免在目标语文概念系统不会出现逻辑矛盾，但是，如果这些名词又被其他人给再次定义，那么，就又出现矛盾了。因此，容易创造新原子图言号，是衡量一种通用语文概念系统健硕度的一个重要基础标志之一；同时，维护词典系统中的概念释义的准确性，也是必须日常工作的一部分。

● 概念释义混乱的举例

- 在不同的专业，同一个“名词”被定义为不同的概念对象，造成概念定义存在者逻辑矛盾。如，音节，在音乐体系，多个音符构词 1 个音节。汉语的音节是由声母和韵母组合发音，能单独发音的元音也是一个音节。
- 在不同的专业，由于概念对象之间的细微差别，需要定义为不同的“名词”。如，“音符”，在音乐体系，定义为用来记录不同长短的音的进行符号；全音符、二分音

符、…。在语文体系，语音的符号，虽然也具备了“音符”的意义；但由于存在声音频率范围和音符数量的差别，因此，被重新定义为“声符”，以便区分。

- 在不同语种之间，存在概念定义上的差异，直接使用概念映射翻译，这对于有些概念来说，并不恰当。

构词法。系列

Ver: 0.1.0.2

Created: 2021/08/25 Update: 2021/09/2

This article: A, Donor: Colating Group

语音类与声符类的概念

Summary:

Version: 0.0.0.1 Created/Modified Date: 20210418 / 20210424

语音类:

1. 声滴/语声滴

类似“语音音素”概念，名词，在语音频率范围内，由单一频率基波与非基波（包含音色信息）共同构成的、近似于单位纯音的语音频谱，称为：声滴、或语声滴。

声滴是声语号的基本构建单位，它源于人类先天器官本能的声音处理能力。

辅音 \subseteq 辅助标压声滴；元音 \subseteq 元母标压声滴。

1.1. 声滴与音素

声滴 \rightarrow (取代) \rightarrow 音素，声滴 \Leftrightarrow (区别) \Leftrightarrow 音素。

声滴与音素的定义，只有部分等价。声滴不只是包括元音与辅音，拼音也是一种声滴。

音素 (phone)：是根据语音的自然属性划分出来的最小语音单位。从声学性质来看，音素是从音质角度划分出来的最小语音单位。从生理性质来看，一个发音动作形成一个音素。音素分为元音与辅音两大类。(网摘)

在其他语文系统、或音乐学中，音素还可能与其他不同的、或与 Colating 相矛盾的定义。

站在还没有学会认字或字母的孩童、或文盲的角度，此时听到的声滴，它就只是单纯的声音而已，也就是声滴。

1.2. 展开声滴/系统展开声滴

在语文系统中，没有经过压缩处理的、所有的使用声滴。

✧ 在 Colating 中，系统展开声滴集合是系统全部音素的集合。

2. 声符类

声符，名词，在语文系统中，用于标记声滴的图形符号，即采用图形符号来映射声滴。如 IPA (International Phonetic Alphabet) 国际音标符号。

特点：由 1 个、或多的声符，构成 1 个声语号。

2.1. 声符与音符

声符 → (取代) → 音符。

音符：音乐名词，用来记录不同长短的音的进行符号。全音符、…。

音符 = 音 ∩ 符，这个概念的元集范围太广了，同样可用在语文系统中。

在音乐学中的音符，是非压缩的。而在语文学中的声符，一般是经过声滴拼读压缩算法处理之后的声滴。

Colating 采用了重新定义的方式，限定了其声音的频率范围与用途。

2.2. 声符 ⇔ (区别) ⇔ 音符。

共同点，都具备声音的4个关键属性。

不同点，频率范围、数量、用途。

2.3. 声符与声滴

声滴 \Leftrightarrow (区别) \Leftrightarrow 声符。

声滴，还没有被赋予其他标记号的语音概念。

“声滴”概念，可以在跨语文系统中使用，即、同一个声滴，在不同的语文系统中，其所对应的声符可以不一样；甚至在一种语文系统中存在的一个“声滴”，而在另一种语文系统中不存在，或者只是作为外来“声滴”来解释。

“声符”概念：同一个声滴，在不同的语文系统中的定义，可以不一样。

3. 声节

优选的声滴队列。由优选的1个或多个声滴队列，就构成了1个声节。1个声节就是1个声语概念标记号。

在限长多滴声节（1维）声语概念语文系统中，1个声节就代表了1个单词的发音；

在单滴声节2维图言概念系统中，1个声节就是1个字的发音。

3.1. 声节与音节

声节 \rightarrow (取代) \rightarrow 音节。

在语文学上，声节的长度是可变的，声节的最短长度为1个声符，最大长度为限制最大可变声符长度。在音乐学上，音节的长度是一个固定长度，或者只有少数几种长度变化。

音乐的总体声音频率范围更宽，语言的声滴与声符的数量更多。

声节 \leftarrow (雷同) \rightarrow 音节：音乐，音节--音 \cap 节。但是，它们的频率范围和用途，都不同。

3.2. 声节的表示方法

声节的 2 种表示方法：声滴流法，声符流法。

在纯语音表达场景，声号的声节，是由声滴流构成的。

在图形符号表达场景，声号的声节，是由声符流构成的。

3.3. 单滴声节

由单个声滴所构成的声节。

3.4. 多滴声节

由优选的包含 1 个声滴在内的多个声滴队列所构成的声节。

3.5. 单滴声节与多滴声节

声节可分为：优选的 1 个声滴的声节称为单滴声节。优选的 1 或多个声滴的声节称为多滴声节

3.6. 单滴声节系统

采用优选的单个声滴所构成的声节来作为单体图言概念标记号的通用概念系统。

3.7. 多滴声节系统

采用优选的包含 1 个声滴在内的多个声滴队列来作为单体图言概念标记号的通用概念系统。

3.8. 单滴声节与单个声节

单滴声节：由单个声滴所构成的声节

单个声节：1 个声节。

“多滴声节”不等于“多个声节”；“单滴声节”不等于“单个声节”。

3.9. 多滴声节与多个声节

多滴声节：由优选的包含 1 个声滴在内的多个声滴队列所构成的声节

多个声节：多个声节。

3.10. 限制最大可变长度

一个概念标记号，通常必含声语号，或含图言号。

一个概念声语号的有效长度，也就是声节的最大长度，并不是不限的。最佳的、高效的单位思想的表达，必须在 1 个句子之内完成；或者说在 1 个呼吸的肺活量之内表达完毕。

为了提高单位思想的表达效率，必须将句子声语号流的长度控制在一定的范围之内，为此：动词、介词等使用频率最高的单词，应当将声滴列的长度控制在尽可能短的范围。请参阅 Colating 构词法原理。

3.11. 限长多滴列

限制最大可变长度的多声滴声节。

例：世界语单词的字符为：限制最大可变长度的多声滴声节 1 维图言概念标记号。

世界语声语声语单词的声节音标流发音（声语概念标记号）为：限制最大可变长度的多声滴声节 1 维声语概念标记号；或限制最大可变长度的多声滴声节声语概念标记号，声节总是表现为 1 维的声滴流形式，所以在说明声语号时，“1 维”可以省略。

3.12. 微调定滴列

微调相同设定长度的多声滴声节 1 维图言结构形式：微调定滴列 1 维。暂无实例，虚构？一般为“限长多滴列”。

微调相同设定长度的多声滴声节 2 维图言结构形式：微调定滴列 2 维图言。如韩语。

图形记号类

4. 图粒

4.1. 图粒/语图粒/纯图粒

名词，在语文系统中，图粒概念对象（图粒本体）是基础的构词图形单元，它存在有 3 种状态：未被释义态，称为“图粒”；已被释义态，称为“符图粒”；难以被规律化释义、或无需释义态，称为“纯图粒”

未被释义的构词图形单位，称为：图粒。

已被释义的“图粒”，称为：符图粒。

不需要、或无需考虑释义的图粒；这类图粒称为：纯图粒。

符图粒与纯图粒，共同组成图言号的基础构词图形单元。

★ 石头（图粒）就是石头，当它被人用锤子敲了一个坑痕之后，就变成了一块与众不同的石头（符图粒），实际上，它原本就是石头。

图粒在被释义后，就变成了字母、偏旁部首，或者是其他任何的构词图形单元。

图粒这个概念，是可以跨语种使用的。

在限长多滴声节 1 维图言概念语文系统中，图粒的图形标记就是图符，图符为构词字母。 ■

☞ 在单滴声节 2 维图言概念语文系统中，图粒的图形标记，可以是表示偏旁部首的图符，也可以是其他的造字图形。声滴与声符、声符与图符发生映射关系是人为指定的结果。



4.2. 图粒与字元

4.2.1. 图粒→（取代）→字元。

在语文系统中，图粒是构成图形标识符的单位符号。如，字母、偏旁部首、或单字；以及其他

的造字图形。

站在还没有学会认字或字母的孩童、或文盲的角度，此时看到的图粒，是还没有关联声滴的原始图形。

“图粒”定义的必要性，可以更容易理解图形与图符的差异。  “图粒”是可以用来构成图言号的单纯图形，这种图形是可以跨语文系统的。  在其存在的语文系统中有意义，“图符”是“图粒”的图形标记符号。

4.2.1. 字元：

因汉字编码需要，将字形分解而成的实用部件。各种笔画、部首、偏旁、独体及其他部件，均可成为某种汉字编码的字元。不同体系的汉字编码，其所分解的字元不一定相同。（网摘）

4.2.1. 取代理由

字元的使用范围已被定义。

图粒，强调的是未被定义的某种构成图言号的图形。

5. 图符/图言符

名词，在语文系统中，图粒概念对象（图粒本体）主要存在有3种状态：未被释义的构词图形单位，称为“图粒”。已经被释义了的用于构造图言概念标记号的图粒，这类图粒称为“符图粒”。难以被规律化释义、或无需考虑释义的用于构造图言概念标记号的图粒，这类图粒称为“纯图粒”。

符图粒与纯图粒，共同组成图言号的基础构词图形单元。

5.1. 图符与字元

图符→（取代）→字元，图符 \rightleftharpoons （区别） \rightleftharpoons 字元。

图符与字元的区别

字元，似乎只是在方块字系统中的定义。

图粒，强调的是构成图言号的单位基本图形；可以适用于单声符或优选多声符概念系统中。

5.2. 图符与图粒

图粒 \Leftrightarrow (区别) \Leftrightarrow 图符。

“图粒”：“图粒”是还没有被定义（概念释义）的、可以用来构成图言号的单纯图形。这种单纯图形是可以跨语文系统使用的，即、同一个图粒，在不同的语文系统中，定义可以不一样；甚至在一种语文系统中存在的“图粒”，而在另一种语文系统中不存在这个“图粒”的定义，或者只是作为外来“图符”来解释。

“图符”只在它存在的语文系统中有意义。如，“a”在世界语和英语中，它们的发音不等价；而在其他语文系统中，甚至根本就不存在字母“a”。

5.3. 声符与图符

声符 \Leftrightarrow (区别) \Leftrightarrow 图符。

在世界语中，声符与图符为一一映射关系。

在某方块语中，声符 \neq 图符；拼音声符与图粒图符无关。

概念类

6. 思想概念类

6.1. 宙沕

宇宙内、表、外的一切。

6.2. 概念对象

宙沕概念对象：宇宙内、表、外的一切，均可作为概念标示与描述的对象。

7. 概念定义类

7.1. 概念

心理学上认为，概念是人脑对客观事物本质的反映，这种反映是以词来标示和记载的。表达概念的语言形式是词或词组。概念是思维活动的结果和产物，同时又是思维活动借以进行的单元

8. 概念标记方法类

8.1. 概念标记号

=声音图形标记符号 + 图言概念标记号。

8.2. 声音图形标记符号 \Leftrightarrow (区别) \Leftrightarrow 图言概念标记号

一个是概念的声滴流标记符号；另一个对应于概念的图粒阵列标记号。

在拉丁语优选多声符系统中，声音图形标记符号与图言概念标记号设计为一一映射关系。

在ZH单声符系统中，声音图形标记符号与图言概念标记号的映射关系，全部由人工单独指定。

8.3. 概号联组 (概念号联组)

通用概念标记关联组合，包含声号，图号。

8.4. 声号/声语号

声语概念标记号。属于优选的声滴流队列概念标记号。不研究盲文、哑语。

8.5. 图号/图言号

图言概念标记号。属于优粒列概念标记号。不研究盲文、哑语。

用图形来表达思想，概念的图形标记号。

Latin 单词 \leftarrow (等价) \rightarrow 图言概念标记号: Colating。

8.6. 可用原子图号总量

符合目标语文系统的构词法规则，容易被目标语群大众读、写、记忆的原子图号的总量。只与图言号本身的读写记有关，与概念释义无关。

例如：可以确定所以3个字母长度的优选发音词根构造，但还没有赋予、或不需要考虑其映射释义。这点是其与有效原子图号总量的区别。

8.7. 有效原子图号总量

符合目标语文系统的构词法规则，容易被目标语群大众读、写、记忆，且释义无逻辑矛盾的原
子图号的总量。

9. 概义/概念义

概念的释义。

10. 拼音声滴类

10.1. 标准压缩声滴

采用融合声滴拼读算法压缩后所获得的声滴。

10.2. 专门压缩声滴

采用非融合声滴拼读算法压缩后所获得的声滴。

10.3. 标准压缩声符（雷同 IPA 音标）

标准压缩声滴所对应的声符，即、音标。

10.4. 专门压缩声符（雷同 IPA 音标）

10.5. 实际使用图符

实际使用的图符。在拉丁语中为字母。

11. 杂类

11.1. 变限声符长度/限制最大可变声符长度

限长多滴声节 1 维图言结构的概记，声号的构词声符的长度越长，表达效率就越低下；当声
号的构词声符超出一定的长度时，声号将由于表达的效率太低，从而导致这个声号的失效。

11.2. 变限图符长度/可变限制图符长度

限长多滴声节 1 维图言结构的概记，图号的构词图符的长度越长，表达效率就越低下；当图号的构词图符超出一定的长度时，图号将由于表达的效率太低，从而导致这个图号的失效。

12. 名词释义 2

12.1. 创发家

人类的创造家与发明家。

(本文结束)

自动目录: [期刊](#) [随想](#) [总目录](#) [独立目录 / All Single CONTENTS](#)

Appendix B

缩写 ◦ 系列

Ver: 0.1.0.2

Created: 2021/08/25 Update: 2021/09/2

This article: A, Donor: Colating Group

1. 滴列类

1.1. **滴列**: 声滴流队列。

1.1.1. **优滴列**: 优选声滴流队列。

包括: 优单滴列, 优单滴列。

1.1.2. **优单滴列**: 优选单个声滴队列。

1.1.3. **优多滴列**: 优选 (1 维) 多个声滴队列。只有 1 维结构形式, 所以“1 维”可以省略。

2. 粒列类

2.1. **粒列**: 图粒流阵列。

2.2. **优粒列**: 优选图粒流阵列。

包括: 优选 1 维图粒流阵列, 优选 2 维图粒流阵列。

2.3. **优条粒列**: 优选 1 维图粒流阵列。

2.4. **优方粒列**: 优选 2 维图粒流阵列。

3. 概号类

3.1. **概号**：概念标记号。

3.1.1. **声号/声语号**：声语概念标记号。

3.1.2. **图号/图言号**：图言概念标记号。

(本文结束)

自动目录：[期刊_随想_总目录](#) [独立目录 / All Single CONTENTS](#)

Appendix C

引言

Colan 语言分为 Colating Language, Cofang Language. 因此, 输入法也分为对应的 2 种。

☆. Colating 输入法

Summary:

Version: 0.0.0.1 Created/Modified Date: 20210418 / 20210424

输入法: Colating 输入法用 Rime 输入法框架进行定制。

详细说明, 在此文档中: colating-ime-char.schema.yaml #采用 txt 文本编辑器可以打开, 如 notepad++。

下载:

Project Web1: <https://sourceforge.net/projects/colating>

Project Web2: <https://github.com/colating>

Colating 文档所使用的其他字库:

NotoSansCJKsc-Regular

NotoSerifCJKsc-Light

1. 说明

1.1. 字母代码与输入法相对应

((b: 表示在 IME 中, 前 26 个 Colating 字母, 可以直接输入。

(.b: 表示在 IME 中，后 26 个 Colating 字母，需要先输入前导符号 “.”，再输入电脑键盘上的普通构词字母。

(,b: / (;b: / ('b: / (-b: 与 “(.b:” 的用途类似，只是前导符号不同而已。

注1: 前位增加 “(”、或 “((” ，是为了后期整理时，直接替换成 Colating 字母使用。: **不要** 去掉。

2. 普通构词字母

2.1. 小写普通构词字母

字母与数制计数系统的基数相对应，序号从 0 开始。由于习惯问题：压缩声符（音标）与音乐音符相对应，序号从 1 开始。

2.1.1. 前段（上/下）9 个小写普通构词字母

前段键盘上面 9 个小写普通构词字母									
序号/前 9	0	1	2	3	4	5	6	7	8
小写字母	b	a	d	f	e	g	h	i	j
字母代码	((b	((a	((d	((f	((e	((g	((h	((i	((j
前段键盘下面 9 个小写普通构词字母									
序号/前 9	0	1	2	3	4	5	6	7	8
小写字母	ø	ɔ	ɒ	ɔ	ɔ	q	σ	ł	δ
字母代码	(.b	(.a	(.d	(.f	(.e	(.g	(.h	(.i	(.j

2.1.2. 中段（上/下）9 个小写普通构词字母

中段键盘上面 9 个小写普通构词字母									
序号/前 9	9	10	11	12	13	14	15	16	17
小写字母	k	o	l	m	u	t	y	p	r
字母代码	((k	((o	((l	((m	((u	((t	((y	((p	((r
中段键盘下面 9 个小写普通构词字母									

序号/前9	9	10	11	12	13	14	15	16	17
小写字母	Σ	g	ϣ	R	↓	a	ε	ϣ	ə
字母代码	(.k	(.o	(.l	(.m	(.u	(.t	(.y	(.p	(.r

2.1.3. 后段（上/下）8 个小写普通构词字母

后段键盘上面 8 个小写普通构词字母								
序号/后8	18	19	20	21	22	23	24	25
小写字母	s	q	z	x	v	c	w	n
字母代码	((s	((q	((z	((x	((v	((c	((w	((n
后段键盘下面 8 个小写普通构词字母								
序号/后8	18	19	20	21	22	23	24	25
小写字母	u	π	ο	ϕ	ε	ϑ	γ	Σ
字母代码	(.s	(.q	(.z	(.x	(.v	(.c	(.w	(.n

2.2. 大写普通构词字母

2.2.1. 前段（上/下）9 个大写普通构词字母

前段键盘上面 9 个大写普通构词字母									
序号/前9	0	1	2	3	4	5	6	7	8
大写字母	B	A	d	f	e	g	h	i	j
字母代码	(,b	(,a	(,d	(,f	(,e	(,g	(,h	(,i	(,j
前段键盘下面 9 个大写普通构词字母									
序号/前9	0	1	2	3	4	5	6	7	8
大写字母	ϣ	ϣ	ϕ	ϑ	κ	q	σ	τ	δ
字母代码	(;b	(;a	(;d	(;f	(;e	(;g	(;h	(;i	(;j

2.2.2. 中段（上/下）9 个大写普通构词字母

中段键盘上面 9 个大写普通构词字母									
序号/前9	9	10	11	12	13	14	15	16	17
大写字母	k	ο	l	m	u	t	y	p	r
字母代码	(,k	(,o	(,l	(,m	(,u	(,t	(,y	(,p	(,r

中段键盘下面 9 个大写普通构词字母									
序号/前 9	9	10	11	12	13	14	15	16	17
大写字母	Σ	g	ϝ	R	Ј	a	ε	ϙ	ə
字母代码	(;k	(;o	(;l	(;m	(;u	(;t	(;y	(;p	(;r

2.2.3. 后段（上/下）8 个大写普通构词字母

后段键盘上面 8 个大写普通构词字母								
序号/后 8	18	19	20	21	22	23	24	25
大写字母	s	q	z	x	v	c	w	n
字母代码	(,s	(,q	(,z	(,x	(,v	(,c	(,w	(,n
后段键盘下面 8 个小写普通构词字母								
序号/后 8	18	19	20	21	22	23	24	25
大写字母	u	π	o	ϝ	α	9	1	Σ
字母代码	(;s	(;q	(;z	(;x	(;v	(;c	(;w	(;n

3. 功能字母

3.1. 键盘（上/下）面普通功能字母

键盘上面的小写普通功能字母										
序号>5-5	0	1	2	3	4	5	6	7	8	9
小写字母	λ	Ω	π	N	χ	l	Ј	Ж	M	Д
字母代码	('m	('l	('k	('p	('o	('n	('j	('h	('i	('u
键盘下面的大写普通功能字母										
序号>5-5	0	1	2	3	4	5	6	7	8	9
小写字母	λ	Ω	π	N	χ	l	Ј	Ж	M	Д
字母代码	(-m	(-l	(-k	(-p	(-o	(-n	(-j	(-h	(-i	(-u

3.2. 键盘（上/下）面归档功能字母

序号>4	0	1	2	3
小写字母	‡	FC2	FC3	FC?
字母代码	('b	('g	('f	('y
.....				

序号>4	0	1	2	3
大写字母	𐀀	𐀁	𐀂	𐀃
字母代码	(-b	(-g	(-f	(-y

4. 临时字母

序号>4	0	1	2
大写字母	𐀄	𐀅	𐀆
字母代码	('w	('e	('r

序号>4	0	1	2
大写字母	𐀇	𐀈	𐀉
字母代码	(-w	(-e	(-r

(本文结束)

自动目录: [期刊_随想_总目录](#) [独立目录 / All Single CONTENTS](#)

☆. Colating 的字母与音标

Summary:

Version: 0.0.0.1 Created/Modified Date: 20211011 / 20211011

1. 普通字母与音标的映射

字母与数制计数系统的基数相对应，序号从 0 开始。由于习惯问题：压缩声符（音标）与音乐音符相对应，序号从 1 开始。

1.1. 前（上/下）9 个普通构词字母与音标

前段键盘上面 9 个小写普通构词字母 ⇨映射 ⇨音标									
序号/前 9	0	1	2	3	4	5	6	7	8
字母代码	((b	((a	((d	((f	((e	((g	((h	((i	((j

非 EN 音									
本音									
键上音标	[b]	[ʌ]	[d]	[f]	[e]	[g]	[h]	[i]	[j]
键上字母	b	A	d	f	e	g	h	i	j
键下字母	ɔ	ɔ	ɔ	ɔ	ɔ	ɔ	σ	ɫ	ɔ
键下音标	[uə/uə]	[ɜ]	[ei]	[ai]	[dɜ]	[əu/əu]	[au/aʊ]	[ʃ]	[ɔi]
本音									
非 EN 音									
字母代码	(.b	(.a	(.d	(.f	(.e	(.g	(.h	(.i	(.j
序号/前9	0	1	2	3	4	5	6	7	8
前段键盘下面9个小写普通构词字母 ⇨映射 ⇨音标									

1.2. 中（上/下）9个普通构词字母与音标

中段键盘上面9个小写普通构词字母									
序号/后9	9	10	11	12	13	14	15	16	17
字母代码	((k	((o	((l	((m	((u	((t	((y	((p	((r
非 EN 音									
本音									
键上音标	[k]	[ɔ]	[l]	[m]	[u]	[t]	[dr]	[p]	[r]
键上字母	k	o	l	m	u	t	y	p	r
键下字母	Σ	g	ɔ	R	ɔ	a	ε	ɔ	ə
键下音标	[ü]	[ŋ]	[eə]	[iu]	[tʃ]	[ɑ:]	[ɛ]	[ɪə]	[ə]
本音									
非 EN 音	✓		✓	✓			✓		
字母代码	(.k	(.o	(.l	(.m	(.u	(.t	(.y	(.p	(.r

音标										
小写字母	λ	Ω	π	N	κ	ι	∏	Ж	M	Д

3. 杂类字母与音标的映射

3.1. 归档功能字母与音标的映射

序号>4	0	1	2	3
字母代码	('b	('g	('f	('y
本音				
音标				
小写字母	#	FC2	FC3	FC?

3.2. 临时功能字母与音标的映射

序号>4	0	1	2
字母代码	('w	('e	('r
本音			
音标			
小写字母	тm	тp	FD3

4. Colating 专门发音方法及规则

(Appendix C 结束)

自动目录: [期刊_随想_总目录](#) [独立目录 / All Single CONTENTS](#)

Appendix D

Colating 撰写规定的说明

Summary:

Version: 0.0.0.1 Created/Modified Date: 20211007 / 20211007

如无特别说明，一般采用 10 进制来表述其他的进制计数系统。

(本文结束)

自动目录: [期刊_随想_总目录](#) [独立目录 / All Single CONTENTS](#)

Appendix X 名词索引

名词索引

Summary:

Version: 0.0.0.1 Created/Modified Date: 20210922 / 20210922

↕概念标记方法类, 132

标准压缩声滴, 132

↕概念类, 132

标准压缩声符（雷同 IPA 音标）, 132

↕拼音声滴类, 132

创发家, 132

↕声音类与声符类, 132

单滴声节, 132

↕图形记号类, 132

单滴声节系统, 132

变限声符长度, 132

单滴声节与单个声节, 132

- 滴列, 147-48
- 多滴声节系统, 132
- 多滴声节与多个声节, 132
- 概号, 147-48
- 概号联组 (概念号联组), 132
- 概念, 132
- 概念标记号, 132
- 概念对象, 132
- 概义/概念义, 132
- 可用原子图号总量, 132
- 粒列, 147-48
- 声滴/语声滴, 132
- 声滴与音素, 132
- 声符, 132
- 声符与声滴, 132
- 声符与图符, 132
- 声符与音符, 132
- 声号
- 声语号, 132
- 声号/声语号, 147-48
- 声节, 132
- 声节的表示方法, 132
- 声节与声节, 132
- 声节与音节, 132
- 图符
- 图言符, 132
- 图符与图粒, 132
- 图符与字元, 132

图号	优单滴, 147-48
图言号, 132	优滴列, 147-48
图号/图言号, 147-48	优粒列, 147-48
图粒, 132	优条粒列, 147-48
图粒与字元, 132	有效原子图号总量, 132
微调定滴列, 132	展开声滴/系统展开声滴, 132
限长多滴列, 132	宙沕, 132
限制最大可变长度, 132	专门压缩声符 (雷同 IPA 音标) , 132

(本文结束)

自动目录: [期刊](#) [随想](#) [总目录](#) [独立目录 / All Single CONTENTS](#)

Appendix X1

修改记录

Ver: 0.1.0.3

Created: 2021/08/25 Update: 202x/xx/xx

This article: A, Donor: Colating Group

[20210923] 废弃概念：“单-声-节”、“多-声-节”。

单-声-节，容易与“单-个-声-节”混淆。多-声-节，容易与“多-个-声-节”混淆。

[20211008] 修改错误内容；增加部分标题内容。

[20211010] 修改错误内容。

[20211012] 修改展开声滴的数制计数基数，Colating 音标，以及错误内容。

[20211013] 修改错误内容，以及废弃缺乏跨平台能力的自动项目符号功能。

(本文结束)

自动目录: [期刊_随想_总目录](#) [独立目录 / All Single CONTENTS](#)

Open Version

colating.win 本次有效期至: 2030-11-26



<http://www.colating.win>

<https://colating.github.io>

<https://sourceforge.net/projects/colating>



捐赠账号链接, 只会出现在以下网站中:



<https://colating.win/donate-account>

CSBN 123-1-123-12345-0

Version: 0.0.1.0
Date: 2021/9/23

意愿币: 0? / 1?

donate-account