

特别声明：本文为新华网客户端新媒体平台“新华号”账号作者上传并发布，仅代表作者观点，不代表新华号的立场及观点。新华号仅提供信息发布平台。

我们为什么喜欢网上聊天、发朋友圈和微博？至少作者我认为，很大一部分原因是因为有各种表情符号、颜文字（Emoji），乃至人民群众喜闻乐见的表情包。在一个用0和1构成的数码世界，没有表情包根本不能聊天好么.....

实际上，很多互联网特色的表情，反而传递了更多更生动的情感，进可战斗，退可卖萌，实在是太好用了。（表情包有2、3个G、现实生活中的面瘫，你们躺枪了么？）

在表情包和表情符号诞生之前，我们

只能用字符来表示情绪

，于是有了很多虽然线条简单、但却萌力爆表的字符表情（Emoticon，和Emoji不太一样），由各种各样的或正常、或稀奇古怪的符号构成。这些表情也一直保留到了现在，甚至还常常被使用——但你知道这些符号都来自哪里吗？

你的表情也是别人的语言

(⋮;⋮⋮), (ノ°;°)ノ

——这么惊讶是在干什么？这个惊讶的“嘴” ⋮，是西里尔字母（Kirillica）的一部分，读音是De。俄语、塞尔维亚语、哈萨克斯坦语，都用的是西里尔字母。

(=^ω^=)

□□ 这个圆圆的兔唇ω，如果把它单拿出来，可能很多人也认识。它是希腊语里面的“欧米茄”（Omega），希腊字母表里面的最后一个字母，物理里面用大写的它Ω表示电阻。而惊慌脸

(°;Δ°)

里面的三角形Δ，念“德尔塔”（Delta，小写是δ），初中数学里面是方程根的判定。

希腊字母表 | Wikipedia

很多你觉得稀奇古怪的字符，实际上是另外的书写体系的一部分。我们所熟知的现代西方大部分语言都是用拉丁字母拼写的（有的带有个别的特殊字符和变音符号，例如德语字母上面的“眼睛”）。

要说的话，古老的希腊语可是拉丁字母和西里尔字母的“妈”。灿烂的古希腊文明孕育出了最早的数学和哲学，而在文艺复兴以后的现代科学发展中，简洁又意蕴深刻的希腊字母也被频频用作数学和物理中的符号，算是西方学术界的一种传统。

西里尔字母用希腊语字母改写，得名于向斯拉夫蛮族们传教的圣人西里尔。这套字符通行于斯拉夫语族大部分语言，曾经苏联用它进行“民族语言改造”，更扩大了它的势力范围（例如哈萨克语属于突厥语系，也使用了西里尔字母）。

斯拉夫民族使用的西里尔字母表 | Wikipedia

以上这几个字母，仅仅是我们颜文字考据的入门第一步。而之后.....你根本想不到，我们都用了多么遥远的语言来表达情绪！！

更稀奇古怪的语言符号

如果说西里尔字母和希腊字母还能算混个脸熟，下面的这些符号，就有点

“是不是专门为表情包设计的”

嫌疑了。

比如

(□~□~□)

这真的不是桃心么？！

□(□□□□)

这真的不是手指么？！其实，这是格鲁吉亚的文字“骑士体”，分别念做ghani和Iasi，但却圆滚滚的萌萌的呢。格鲁吉亚文由他们的民族先祖创制，有人推断是学希腊文字而来，但正本之源却一直没有定论。最早的格鲁吉亚文出现在公元400多年，这套独特的文字和高加索山里的民族一样，拥有古老而独特的历史。

（作者随手截图的）格鲁吉亚文 | Wikipedia

至于像：

(□•_•)□

这个捏拳的姿势，还有

(□•. •□)

的“爪子”，来自弯弯扭扭的泰文。这个

(□_□)

流泪的怨念之眼，和

(□_□)

这个皱眉的大眼睛，来源是卡纳达文。这两个稀奇古怪的文字系统，都来自一个大的文字系统之母——

婆罗米文字

。

印度孔雀王朝的阿育王流传下的石刻诏书就是用婆罗米文写就的 | Wikipedia

在公元前3-4世纪，盘踞在印度半岛的强大的“孔雀王朝”使用的就是婆罗米文字，而这种文字派生出了多种文字系统，例如泰文、天城文（印地语使用的文字）、古吉拉特文、僧伽罗文、泰米尔文，甚至还有藏文，多流行于南亚和东南亚。

这些风格各异的符号，大家可以感受一下.....

一些婆罗米系文字的辅音 | Wikipedia

不过，要往上追溯的话，婆罗米文字是由一种古老的闪族文字

“阿拉米文”

演变而来，阿拉米文跟腓尼基文血缘甚近，都来自原始的迦南字母；前面提到的拉丁、西里尔字母的妈希腊字母，也来自腓尼基-

迦南书写体系。晕了吧？一句话总结，

上面出现的所有稀奇古怪的字符，都有一个同样的祖宗

。

上文中提到的婆罗米文字，拉丁、西里尔字母，还有它们的妈希腊字母，都来自腓尼基-迦南书写体系 | Wikipedia

还有更稀奇的。像：

☐(☐ ☐ v☐ ☐)☐

这里面两个“举拳”的符号，以及奔跑的小人：

☐(☐)☐

实际上来自北美的一个原住民部落纳斯喀皮人（Naskapi），他们使用文字的方向表示不同的元音读音。神奇吧！

纳斯喀皮文字的读音。你找到☐了吗？ | Wikipedia

天道好轮回

既然我们的表情用了别人的语言，那别人的表情岂不是要用我们的？

当然啦！比如，在英国的西汉姆球迷中，许多的推特名都带着

“父”

字。

作者我百思不得其解，直到我看到了他们的队徽.....

(☐—☐☐☐—)`我服了，你们随使用吧。

不仅仅是歪果仁借用我们的符号，

中国网友们也在重定义着自己的文字

。比如人民群众喜闻乐见的“囧”，现在看到之后都会条件反射出一张“囧脸”，恰好读音又是“窘迫”的“窘”字，简直就是完美的文字表情。

实际上，“囧”原意为光明，《说文解字》里解为

“窻牖麗廔闓𠄎”

——别看一个字你都不认识，但都是与采光的窗户有关的，囧本身是个象形字。

“囧”字的甲骨文形式。是不是少了一些囧感？ | Wikipedia

若一定要说表情丰富的象形字，“观”字的甲骨文和金文要更生动一些，瞪着大眼睛警觉的即视感。

左边第二个字就是金文的“观”，整幅字是“听鸟观鱼” | 网易网友“风之行”博客

其实这也没有什么值得奇怪的，

不就是强行象形嘛

。汉字天生就有很多象形字，被老外拿去那是再正常不过。比如，

(^_^) o 自自 o (^_^) 代表碰杯，

(𠄎 𠄎益 𠄎) 代表极度厌恶，

、(o`皿'o)/代表愤怒，

凸(`0´)凸代表竖中指，

川´;·ω·`川 代表熊(的毛)，

(个_个) 代表眼睛，

((┌^┐井)代表不爽，

□(((#^_^)八(^_&^*)))□ 代表high five，

(" ' ▽ ')_中☆{{{Δ}}} 代表拿锤子敲，

(oT-T)尸 代表举白旗，如此等等.....

等等，这怎么看起来这么别扭啊.....

你拿汉语做表情？

炎黄子孙好出戏！

别扭就对了！因为你已经知道这些字的“含义”，这些含义和它的外形存在了冲突

。

实验心理学有一个经典成果叫

“斯特鲁普效应”

，描述的就正是类似这样的现象。心理学家约翰·斯特鲁普在纸上涂上几种颜色的墨水，让被试把颜色念出来。区别是，第一组里的墨水组成了表示颜色的单词，而且还和墨水的颜色不一样；第二组里墨水就是方框框。

就像这样：

红 黄 蓝 绿

■ ■ ■ ■

是不是很困扰？是不是看到第一个字总是想念“红”而不能正常地念出墨水的颜色“绿”？斯特鲁普发现被试要花很久才能成功念出第一组的墨水颜色，而第二组没有困难。反过来，如果让被试念出字本身的内容，那么用什么墨水就无关紧要了。看起来，当我们阅读一个字的时候，它的含义是“优先”的；我们会先自动读出它的意义，然后才想起来实验员的要求，努力把它纠回去。

我们作为汉语母语使用者，阅读汉语表情文字遇到的就是这样的问题。当然，如果你反复地盯着这些表情文字，慢慢地它们好像又变得正常了。

“语义饱和”

在其中发挥了作用——反复想一个单词会让它变成无意义的笔画组合，仔细盯着一

个字看就越来越觉得不认识它。而在表情文字里这样正好能让我们识别出它的象形含义。

从ASCII到Unicode

的信息进化史

好了，说完了文字，我们再来说表情符号的另一个重要来源：

特殊符号

。

前一阵子非常流行的

(' ') ^ _ _ _

掀桌表情（对，如果你用的是搜狗输入法，输入“掀桌”二字即可出现）。那个“桌”，实际上就是特殊符号的一种——表格符号。在稍微有些时日的DOS系统里面，人们只能使用纯文本，那么碰到图和表格怎么办呢？于是，人们加入了一些特殊符号来满足这个需求。

但是往哪儿加这些符号呢？

一开始，人们在电脑上用的是一套叫做ASCII（读作“阿斯克”，American Standard Code for Information Interchange，美国信息交换标准代码）的编码系统，主要用于显示英语和常用的标点，于1967年发布、1986年更新，一共只有128个字符（其中有95个是可以显示的，剩下的都是“控制符”），算上大小写、数字和标点，基本也就是你的键盘上有什么就是什么了。

即使这样，用ASCII也能搞艺术.....都是大触啊！ |

所以呢，早期大家聊天的表情，笑脸都是“:)”这样的。上古网络小说《第一次的亲密接触》里面，宅男主角可是能够对着这样一个微笑符号yy好久的。

但是ASCII的128个字符肯定是不够用的，别说表情了，你让人家打个naïve都没法点上“很天真”的两个点，很难受是吧，欧洲的各国工程师更是不能忍，于是他们在ASCII的基础上，制定了

扩展版的ESCII

，将原来的7位扩展为8位，空间一下子从128升级到了256。但是由于做扩展ASCII的人太多，标准并不兼容，关键时刻还是乱码一片。

大名鼎鼎的IBM代码页437 (Code page 437) ，是比较流行的扩展ASCII编码系统之一，也是MS-DOS使用的字符编码。可以看到，这时候就已经有前面提到的表格符号了，

此外还有不少数学符号、货币符号

——对表情包事业也是巨大的贡献啊！

代码页437从128位到255位的字符编码 |

其实不仅仅是这些符号，前面空出来的“控制符”区域，其实被一些系统暗戳戳地规定成了

“特别图像字元”

。里面藏了黑白两个笑脸呢！这大概就是

“官方表情包”

了？

但.....其实并非每个系统都采用了这些字符，并不能算是标准。

看到左上角的8比特笑脸了吗？ | Wikipedia

Anyway，随着信息事业扩展到世界各地，一个更大、更包容的编码系统势在必行，那就是

Unicode (万国码)

，1991年正式发布的编码标准，从8位一下提升到了16位，一口气把前面说到那些奇奇怪怪的语言字符都包含进来了。最关键的是，全部的简体和繁体中文也在这个系统里！（中日韩的汉字是1992年加进来的）

至2020年3月，Unicode已经发布到了13.0.0版，收录超过13万个字符。表情包的花样也跟着指数级上升，玩法被脑洞大开的世界各地网友不断翻新。

比如流行的所谓“Lenny face”：

这个和doge有点异曲同工的邪恶表情，“眉毛”和“嘴”其实是“双弱音符”（Double Breve），和其它字符结合出现，并不是一个独立的字符。这种玩法也常见于一个泰语的上标字符不断叠加成“搭楼”的效果，是一个意思。这些都是Unicode编码系统内实现的。

其它的什么太阳☀️ 花草🌿 桃心💕 萌萌大眼睛👁️👁️（其实是四分之三）之类，也都是Unicode编码系统里面的字符啦。

总之，表情符号从很大程度上，是我们对于符号的重新赋意，也就是一种使用既有工具和内容的再创造了。很多也是从小圈子里面传出来的，比如著名的2chan和Reddit。虽然现在各个软件系统里面都用图标定义了真正的“颜文字”，但我们依然热衷于用自己的脑补功力，在简单的线条和符号里制造生动的表情。

你们还闲着干啥？还不赶紧晒一下你最得意的表情.....

作者：李子

编辑：望江湖

一个已经乱码的AI

【ㄱ】字其实是韩文里【雪】的意思啊！

然而Google的图片结果.....我整个AI都ㄱ_ㄱ了！！

你弃疗了是不是？？

本文来自果壳，未经授权不得转载.

如有需要请联系sns@guokr.com

来源：新华号 果壳