

数学与文学

■文 / 胡伟文



法国数学家哈达玛

法国数学家哈达玛

(Hadamard, 1865—1963)

曾说：“语言学是数学和人文科学之间的桥梁。”语言学作为人文科学之一比较直接地与数学发生联系，数学通过语言学而与更多的人文学科建立联系。语言学是所有人日常生活最必需的，而数学则是所有科学生活最必需的。正式利用数学研究语言现象却并不是很早的事情，甚至还不如数学

之应用于音乐、雕刻那样早。

语言学作为数学与人文科学之间的这座桥梁是语言学家和数学家们共同修建的。数学家们不仅利用数学研究了语言，而且在研究语言的过程中发展了数学。俄国数学家马尔可夫的例子是很著名的。马尔可夫在研究俄语字母序列的数学研究中，提出了随机过程论，后来成为一个独立的数学分支，对现代数学的发展产生了深远的影响。

电子计算机的出现使语言与数学更加靠近，使数学渗透到了语言学的更多领域。例如，自动句法研究带有浓厚的数学色彩，主要应用于机器翻译研究以及基于模式匹配的自然语言理解系统研制，并在应用的牵引和推动下不断发展。在计算机科学领域中，众多计算机专家和人工智能学者，也都学用数学方法来研究句法。这样，便从不同的领域涌现出了大批兼通语言学、数学和计算机科学的人才。例如，语音的自动合成与分析是语言信息处理的一个重要方面，也是语言学研究的重要方面。由于语音频谱提供

出来的信息实在太多了，所以研制语音合成器，使之能把语音频谱转化为语音变得十分困难。即使如此，这一研究仍取得了令人瞩目的进展。

图像识别与文字学研究也结合起来，并涉及许多数学问题。如何研究书面文字结构，是一个极有意义的课题。电子计算机的出现和广泛使用大大促进了数学与语言学的结合。数学渗透到了形态学、句法学、词汇学、语言学、文字学和语义学等语言学的各个分支，既推动了语言学的发展，又促进了数学自身的发展。

一、数学与文学

文学以语言为基础，但文学不仅仅是语言问题。文学与数学的联系是经过了语言学这座桥梁的，并且这种联系的性质已大不一样。如果我们以托尔斯泰 (A. N. Tolstoy, 1828—1910) 也喜欢数学来说明文学与数学联系的必然性，那可能是没有说服力的。但是，我们从许多的例子也许还是可以看到一些现象。

(一) 托尔斯泰欣赏的“割草人”

故事

俄罗斯的大文学家托尔斯泰确实也十分喜爱数学。他非常欣赏并经常向人们推荐下面的“割草人”故事：“一组割草人要割完两块地上的草，大的一块草地的面积比小的一块大一倍。全组人员上午都到较大的那块草地上去割草；下午，全组人员平均分成两半，一半人仍留在较大的草地上，另一半人到较小的那块草地上去割草。到傍晚时，大草地上的草刚好割完，小草地上的草还剩下一小块，第二天正好可由一个人一天割完。问这组割草人共有多少人？”

这个问题当然可以利用代数方法来解决，而且可以算出每个人每天能割整个草地的 $\frac{1}{9}$ ， $\frac{8}{9}$ 的草地在一天内就被割完了，故全组人员为 8 个。但是，托尔斯泰却特别赞赏用几何方法来回答这一问题。

全组的人在大块草地上工作了半天，相当于半组的人工作了两个半天；半组人又工作了半天才完成大块草地的割草，因此，相当于半个组的人割

草 3 个半天（或一天半）割草地 1，也就是说，半个组的人半天能割草地 $\frac{1}{3}$ 。半个组的人在小块草地上又割草半天，因此也割了 $\frac{1}{3}$ ，小块地面积为 $\frac{1}{2}$ ，于是还剩下 $\frac{1}{6}$ 。这意味着每个人每天能割草地 $\frac{1}{6}$ 。而全组人员一天的割草地面积是 $\frac{4}{3}$ ，它是多少人完成的呢？ $\frac{4}{3} \div \frac{1}{6} = 8$ 。

前面说用代数方法可算出每个人每天割草地 $\frac{1}{9}$ ，那是将整个草地面积记为 1。用几何方法算时，将整个面积记 $\frac{3}{2}$ ，而 $\frac{3}{2}$ 的 $\frac{1}{9}$ 正好是 $\frac{1}{6}$ 。

托尔斯泰对算术教学还很感兴趣，他甚至编写过一本算术课本。他特别喜欢那些貌似复杂、但能找到简便解法的算术题。对那个“割草人”问题，他一直十分欣赏，到晚年时还见人就称赞不已。实际上，这道题也确实还为数学工作者所喜爱，直到今天，这类题还经常出现在一些青少年的数学读物之中。

托尔斯泰还善于用数学概念做比喻来说明某些道理。例如，他曾用分数来表示人的真实价值。他把别人对一个人的评价比作分子，这往往比较符合客观实际；把一个人自己对自己的评价比作分母，一部分人往往容易夸大这个分母。在分母固定时，分子越大分数值越大；在分子固定的时候，分母越大分数值越小。这个比喻确实发人深省。

托尔斯泰还曾经利用数学知识写过一篇小说，讽刺那些贪婪成性、要财不要命的人。题目为“一个人需要很多土地吗？”小说大意是：

有一个叫巴河姆的人到草原上去买地，卖主卖地的方法很特别。任何一个来买地的人，只要交 1000 卢布，他便可以在一天之内，从太阳出山开始行走，

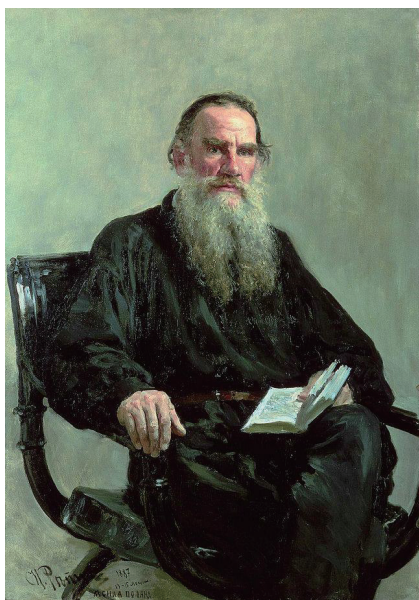
由草原上任一点出发；在草原上走到太阳落山，如果在日落之前，他回到了出发点，那么，他这一天所走的路线所围住的土地，就算他买到的土地了；如果他在日落之前没有回到出发地点，那么，他就一寸土地也得不到，白白丢掉 1000 卢布。

巴河姆认为，这样的规定真是有利可图，便爽快地交了 1000 卢布。第二天太阳刚刚升起，巴河姆就在草原上迈开了大步。他先沿一条直线一口气走了 10 俄里，然后向左拐弯 90°，又沿直线走了很远很远，才又向左拐弯 90°，继续前进了 2 俄里。这时他发现天色已经不早，至此他总共已走了 24.7 俄里的路程。于是他不得不改变前进方向，直向出发点跑去。巴河姆终于在日落前又跑了 15 俄里赶回到了出发点。但是，当巴河姆停下来时，脚跟尚未站稳，便两腿一软，扑倒在地，口吐鲜血，一命呜呼了。

巴河姆付出了生命的代价，究竟换来了多少土地呢？1 俄里等于 1.0668 公里，巴河姆这一天共跑了 $(24.7+15) \times 1.0668 = 42.35$ 公里的路程，所围的图形是一个直角梯形，面积约为 86.72 平方公里，亦即 8672 公顷、约 13 万亩地。

一下子弄到了 13 万亩土地，不可谓不多，但是人都死了，再多的土地还有什么意义呢？

这个故事对于那些贪婪的人是一个讽刺，作品的讽刺意义还不止于此。喜爱数学的托尔斯泰在作品中还有更深刻的寓意。巴河姆贪心，并且也很愚蠢。他走的那个直角梯形的两底分别是 2 俄里和 10 俄里，而两腰则是 12.7 俄里和 15 俄里。如果按一个正方形的路线行走，围同样大的面积只需行走 37 公里，少走 5 公里多；



托尔斯泰

如果按一个圆形的路线行走，围同样大的面积则只需要行走 33 公里，可以少走 9 公里多。亦即，这里包含着另一个数学问题：在同等面积的情况下求最小周界。如果巴河姆懂得这个道理，也许他既能得到 13 万亩土地，又不至于累死；贪婪加无知葬送了他的性命。关键仍在他的贪婪，同时，他还不会去计算另一个数字：他应走多少公里才不致送命？

（二）文学作品中的计算

北宋著名的文学家苏轼，不仅诗词写得精彩，而且还是绘画的高手，如诗的画，如画的诗。有一次，他画了一幅《百鸟归巢图》，明代一位名叫伦文叙的状元，在他的画上题了一首诗：归来一只复一只，三四五六七八只，凤凰何少鸟何多？啄尽人间千石食。

画名既是“百鸟”，而题画诗中却不见“百”字的踪影。诗人开始好像是在漫不经心地数数：一只，又一只，三只，四只……七只，八只；数到第八只，好像诗人已不耐烦了，突然感慨横生，笔锋一转，发了一通议论。

那议论之中，诗人感叹官场之中廉洁奉公、洁身自好的“凤凰”太少；而那贪污腐化的“害鸟”则太多，他们巧取豪夺，把老百姓赖以活命的千石（dàn）、万石“啄尽”了。那么，诗人仅仅在发议论吗？诗人完全没有顾及百鸟吗？

我们先把诗中的数字顺序写下来：1，1，3，4，5，6，7，8；然后，再用诗中实际暗示的运算关系把它们连起来，便可得到一个算式 $1+1+3\times 4+5\times 6+7\times 8$ ，算起来恰好等于 100！原来，诗人巧妙地将 100 分成了两个 1，3 个 4，5 个 6，7 个 8，



苏轼

用诗又“画”了一橱“百鸟图”，“百”字含而不露地藏藏在诗中。

歌剧《刘三姐》中有一个精彩的片段，说的是刘三姐与 3 位秀才的对歌。双方用唱山歌的方式相互问难。3 位秀才自然自恃有“学问”，对歌中给刘三姐出了一道难题：

罗秀才：小小麻雀莫逞能，三百条狗四下分，一少三多要单数，看你怎样分得清？

刘三姐：九十九条打猎去，九十九条看羊来，九十九条守门口，还有三条狗奴才。

刘三姐一下子把 300 分成了 4 个奇数之和： $300=99+99+99+3$

歌中“三条狗奴才”系指陶、李、罗三个助纣为虐的秀才。对歌既表现了刘三姐的机智与幽默，又表现了刘三姐对这些秀才的鄙视和嘲弄。

文学作品中使用不定数词表达概数也是常见的现象，而且也有不同的

使用风格。

二、中国古代经典数字诗

中国古代的数字诗就是将数字巧妙地嵌入诗中，形成一种独特的表现形式，取得别样的效果。数字诗大致分为以下五类。

（一）“十”字令诗

“十”字令诗是数字诗中最为常见的一种，就是将“一”到“十”十个数字嵌入诗中。以下，给出 5 首“十”字令诗作为示例。

1. 宋代邵雍的村景诗

山村咏怀

一去二三里，烟村四五家。

亭台六七座，八九十枝花。

这是宋代邵雍描写一路景物的诗，共 20 个字，把 10 个数字全用上了。这首诗用数字反映远近、村落、亭台和花，通俗自然，脍炙人口。

2. 明代林和靖的雪梅诗

雪梅

一片二片三四片，五片六片七八片。

九片十片无数片，飞入梅中都不见。

这是明代林和靖写的一首雪梅诗，全诗用表示雪花片数的数量词写成。读后就好像身临雪境，飞下的雪片由少到多，飞入梅林，就难分是雪花还是梅花。

3. 苏轼给同窗的劝勉诗

相传苏轼与同窗赴京赶考，学友因为天气耽搁担心迟到，感叹成诗曰：

一叶孤舟，坐二三个骚客，启用四桨五帆，经由六滩七湾，历尽八颠九簸，可叹十分来迟。

苏轼劝勉道：

十年寒窗，进九八家书院，抛却七情六欲，苦读五经四书，考了三四五次，今天一定要中。

4. 吴承恩的暮景诗



吴承恩

吴承恩在《西游记》中的一段“红轮西坠”后天色已晚的景色描写：

十里长亭无人走，九重天上现星辰。
八河船只皆收港，七千州县尽关门。
六宫五府回官宰，四海三江罢钓纶。
两座楼头钟声响，一轮明月满乾坤。

5. 《归田琐记》中的贬官诗

《归田琐记》中有一首诗，刻画的是过去当官时作威作福，搜刮民脂民膏，而后“荣归故里”“赋闲在家”的丑恶嘴脸，淋漓尽致，值得玩味。

一命之荣称得，两片竹板拖得，
三十俸银领得，四乡地保传得，
五十嘴巴打得，六角文书发得，
七品堂堂考得，八字衙门开得，
九品补服借得，十分高兴不得。

(二) “一”字诗

这是一种只用“一”字嵌入诗中的诗。如下：

1. 唐代王建的一首《古谣》

古谣

一东一西垄头水，一聚一散天边路。
一去一来道上客，一颠一倒池中树。

这里，短短的28个字中有着矛盾的8个“一”，但都统一在一幅风

景画里，可见“一”字有多么强的表现力。

2. 元代无名氏有一首散曲

一年老一年，一日没一日，一秋又一秋，一辈催一辈。一聚一离别，一喜一伤悲。一榻一身卧，一生一梦里。寻一伙相识，他一会咱一会，那一般相知，吹一会唱一会。

全曲嵌“一”字竟多达二十二个。作者在年华易逝、光阴催老、人生如梦的叹息中，寻求着自己的人生位置，情绪看似旷达，实则透露着辛酸，具有强烈的感染力。

3. 清人王士禛的《题秋江独钓图》

清人王士禛作的这首诗总共用了九个“一”，也是有名的数字诗：

题秋江独钓图

一蓑一笠一扁舟，一丈丝纶一寸钩。

一曲高歌一樽酒，一人独钓一江秋。

清人陈沆（也有人认为是北宋苏轼）有一首与此相似的诗：

一帆一桨一渔舟，一个渔翁一钓钩。

一俯一仰一顿笑，一江明月一江秋。

这首诗中出现了十个“一”，比上一首还多一个。两首诗都创造了一个渔翁怡然自乐的生活情景。

清代女诗人何佩玉也写过一首同样的诗：

一花一枝一砚石，一抹斜阳一鸟飞，

一山一水一寺中，一林黄叶一僧归。

连用十个“一”字却不使人感到重复，而是意境悠远，妙趣横生。

4. 清代侯善渊的《一剪梅》

清代侯善渊的《一剪梅》是这么写的：

一个尘劳一个忙。一自离别，一得真常。一天精秀一天凉。一点清光，一带凝阳。一气相交一气张。一结神丹，一命延长。一灵透人一云房。一对金童，一引仙乡。

全词60个字里竟有16个“一”，比起上面的两首诗，有过之而无不及，又是别有一番情趣和风味。

(三) 数字回环诗

数字回环诗就是诗歌中一到万（或十）几个数字回环使用。如下：

1. 卓文君写的“从一到万，从万到一”

一别之后，两地相思，说的是三四月，却谁知是五六年。七弦琴无心弹，八行书无可传，九连环从中折断。十里长亭望眼欲穿。百般想，千般念，万般无奈把郎怨。

万语千言道不尽，百无聊赖十凭栏。重九登高看孤雁，八月中秋月圆人不圆。七月半烧香秉烛问苍天，六月伏天人人摇扇我心寒，五月石榴如火偏遇阵阵冷雨浇花端，四月枇杷黄，我欲对镜心意乱，三月桃花随流水，二月风筝线儿断。噫！郎呀郎，巴不得下一世你为女来我为男。

2. 无名氏写的从万到一共十三个数字

柳絮飘飞，万朵入云间，一去无归期，千般情不舍。望断天涯，魂随魄也依，百转盼回归，寝无言，食厌味，倚门伫立，怎奈十月雪花漫天舞，不见伊人归，空叹息，重阳九，赏菊桂，登高远眺，形影孤单独自泣。眼看八月中秋，叹，月圆人残缺，遥对蟾宫祷佳期。七巧之期时，牛郎织女亦相会，使我独饮相思泪，谁怜惜？炎炎夏日六月心，心寒不觉身子温，只字片影无踪迹，一片忧伤没法医。五五端阳划龙舟，江上游客皆成对，思念化作江水流，欲诉衷肠无知已。转眼四月采桑节，忙里偷闲问鱼雁，可带佳音归？雁无语，振翅飞，鱼我声，潜水底。看，三月桃红柳绿时，春色好，我心更劳累，枉费了一片心机，莫不是，二人相携漫步路，一切随风吹？犹自瞎



李调元

猜疑。忽听喜鹊啼不住，望君，笑挥手臂，喜孜孜，欲笑泪先滴！

3. 清代李调元的从一至十、从十至一

清代李调元写过这样一首数字诗，从一至十，再从十至一。诗韵味浓厚，即情随性，没有一点矫作。

一大乔二小乔，三寸金莲四寸腰，
搽得五六七次粉，妆得八九十分娇。
十九月亮八分明，七个才子六个癫，五
更四时鸡三唱，怀抱二乔一枕眠。

(四) 杂数诗

杂数诗就是诗歌中运用数字没有定数和规律，只根据诗歌的需要来写，古代的数字诗大多属于这一种。如下：

1. 宋代王安石的《麻雀》

一窝二窝三四窝，五窝六窝七八窝，
食尽皇家千钟粟，凤凰何少尔何多。

这是宋代政治家、文学家和思想家王安石写的一道《麻雀》诗。他眼看北宋王朝很多官员，饱食终日，贪污腐败，反对变法，故把他们比作麻雀而讽刺之。

2. 清人黄焕中的《闺怨》

百尺楼台万丈溪，云书八九寄辽西。

忽闻二月双飞雁，最恨三更一唱鸡。
五六归期空望断，七千离恨竟未齐。
半生四顾孤鸿影，十载悲随杜鹃啼。

(五) 数字隐秘诗

数字隐秘诗就是将数字用谜语的形式表现出来，十分有情趣。例如：

1. 宋代才女朱淑贞的《断肠谜》

宋代才女朱淑贞，临终前写的《断肠谜》算是这一类数字诗。

下楼来，金钱卜落；问苍天，人在何方？恨王孙，一直去了；骂冤家，言去难留。悔当初，吾错失口；有上交，无下交；皂白何须问，分开不用刀。从今莫把仇人靠，千里相思一撇消。

整首诗暗含“一”至“十”十个数字，句句充满对负心人的谴责，倾诉自己无限的哀怨！诗中隐藏的数字，你能找出来吗？

2. 《玉房怨》

后人填了一首《玉房怨》，与朱淑贞的《断肠谜》大同小异。

元宵夜，兀坐灯窗下。问苍天，人在谁家？恨玉郎，全无一点真心话。叫奴欲罢不能罢，吾今舍口不言他。论交情，曾不差。染成皂，难说青白话，恨不能一刀两断分两家。可怜奴，手中无力难抛下，我今设一计，教他无言可答。

三、“红学”官司与数理语言学

《红楼梦》前八十回与后四十回的作者是否相同？是一场著名的“红学”官司。用数学方法对作品进行写作风格分析、词汇相关程度分析和句型频谱分析，对一些文学界的争论问题进行研究，是数理语言学的一个重要应用。

1980年6月，在美国威斯康星大学召开的国际首届《红楼梦》研讨会上，来自威斯康星大学的华裔学者陈炳藻先生宣读了一篇名为“从词汇上的统计论〈红楼梦〉的作者问题”的博士论文，引起了国际红学界的关注和兴趣。

1986年，陈炳藻教授公开发表了专著《电脑在文学上的应用：〈红楼梦〉与〈儿女英雄传〉两书作者》。利用计算机对《红楼梦》前八十回和后四十回的用字进行了测定，并从数理统计的观点出发，探讨《红楼梦》前后用字的相关程度。他将《红楼梦》一百二十回分为三组，每组四十回，并将《儿女英雄传》作为第四组进行比较，从每组中任意取出八万字，分别挑出名词、动词、形容词、副词、虚词这五种词汇，运用数理语言学，通过计算机程序对这些词进行编排、统计、比较和处理，进而找出各组相关程度。结果发现《红楼梦》前八十回与后四十回的词汇相关程度达到78.57%，而《红楼梦》与《儿女英雄传》的词汇相关程度是32.14%。由此他推断出《红楼梦》的作者为一个人的结论。

这个结论是否被红学界所接受，还存在一定的争论，但是这种方法却给很多人留下了深刻的印象。

苏联的著名长篇小说《静静的顿河》，也曾经有过关于作者的争论。有人认为，该书是肖洛霍夫(M. A. Sholokhov, 1905—1984)剽窃了一个无名作者的作品后加工而成的。后来，用与上述方法类似的数学方法，还了肖洛霍夫的清白。[科技](#)