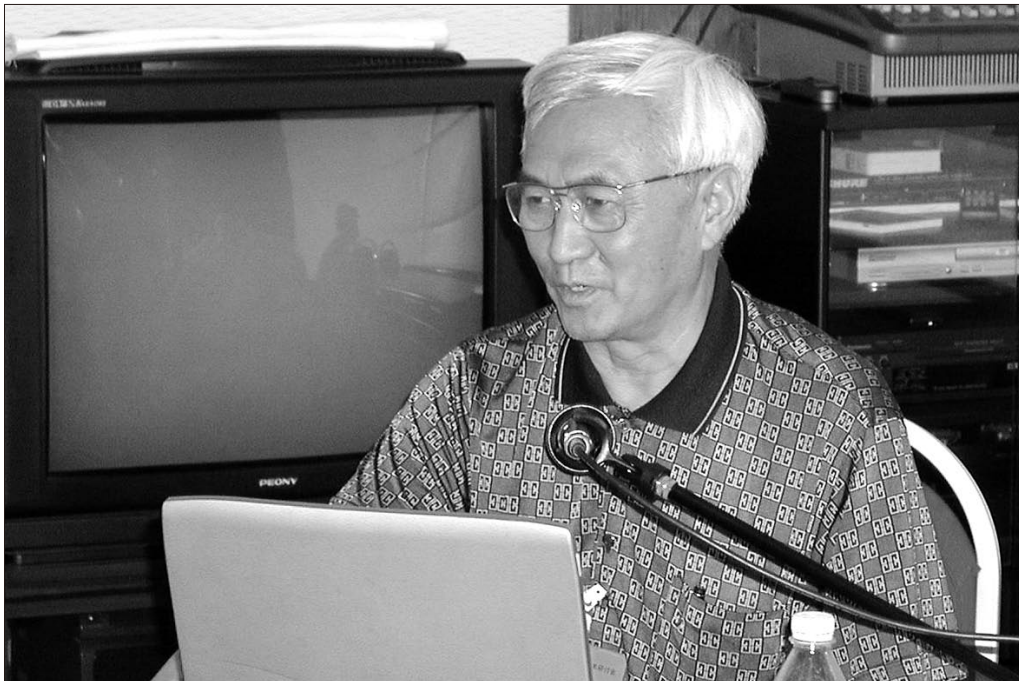


□赵晶

冯志伟，1939年4月生，云南省昆明市人，计算语言学家，专门从事语言学和计算机科学的跨学科研究。他先后在北京大学和中国科学技术大学研究生院研究生毕业，获双硕士学位。先后到法国格勒诺布尔理科大学、德国夫琅禾费研究院新信息技术与通信系统研究所(FhG)、德国特里尔大学语言文学院、德国康斯坦茨高等技术学院国际术语学和应用语言学中心(CiTaL)、桑夏自然语言处理研究院、韩国科学技术院(KAIST)电子工程与计算机科学系(E ECS)学习、工作。现为国家教育部语言文字应用研究所研究员、博士生导师、学术委员会委员。



# 冯志伟 与数理语言学共舞

## 跨学科奇才

语言学家冯志伟教授是极为罕见的跨学科奇才。过去的著名语言学家有的只懂社会科学，不懂自然科学。许多人只懂古代汉语、现代汉语或普通语言学，一般只着重研究汉语的语音、词汇、语法或文字等某一个方面的问题，研究的问题和领域比较单一。他们中的一些佼佼者，至多也懂得两三门外语，视野不开阔，语言的纵横对比研究不够，有一定的局限性。而冯志伟先生却懂得理科中的数学、物理、化学和计算机科学，又懂得语言学中的古代汉语、现代汉语、文字学、音韵学和普通语言学，深研过汉、英、法、德、俄、日等语言的语音、词汇和语法的自动处理，并能把各方面知识紧密地结合起来综合应用，在计算机上加以实现，成为文理兼通的语言学专家。冯教授不但在国外著名大学的电子工程与计算机科学系讲授理科的“机器翻译的方法和技术研究”、“自然语言处理的算法研究”、“计算语言学专题研究”等艰深的博士课程，而且还在国内外大学的中文系讲授“汉魏六朝散文”、“唐诗”、“宋词”、“古代汉语”、“现代汉语”、“汉字的历史与结构”等饶有风趣的课程，他还能给学生们辅导英、德、法、俄、日等外语课程，提高学生们的英语口语表达能力和书面写作能力。他还是一位翻译专家，出版过翻译著作。这样的人才确实是很罕见的。

冯志伟自幼聪慧，18岁就以优异成绩考入北京大学地球化学专业。他在北大图书馆偶然看到了美国语言学家乔姆斯基(N. Chomsky)的论文《语言描写的三个模型》(Three models for the description of language)，被乔姆斯基在语言研究中的新思想深深地吸引，继而从理科转到中文系语言学专业从事语言学的学习。在中文系读书期间，冯志伟一面学好传统语言学的各门课程和外语，一面利用课余时间继续研究数理语言学的问题，他充分地利用北京大学图书馆丰富的藏书和最新的杂志，跟踪国际上数理语言学发展的足迹，成为了班上名列前茅的学生。1964年，冯志伟考上北京大学语言学理论专业的研究生，经导师岑麒祥教授同意，他的研究生毕业论文的题目定为《数学方法在语言学中的应用》，在我国语言学研究中，首次系统地全面地来研究数理语言学这个新兴学科。这样，我国的数理

语言学研究便首先在北京大学正式地开展起来。

北京大学中文系的著名语言学家王力先生和朱德熙先生都支持冯志伟的数理语言学研究，王力先生曾对冯志伟说：“语言学不是简单的学问，我们应该像赵元任先生那样，首先做一个数学家、物理学家、文学家、音乐家，然后再做一个合格的语言学家。”北京大学的这些第一流的学者，总是站在科学的最前沿来看待学术的发展，他们的鼓励给了冯志伟以巨大的力量。

1978年，之前毕业于北京大学中文系研究生的文科学生冯志伟，以优异的理科成绩考上了中国科学技术大学研究生院信息科学系，又开始了理科的学习。1979年，入学不到一年，他就在《计算机科学》杂志创刊号上发表了《形式语言理论》的长篇文章，用严格的数学表达方式向计算机科学界说明数理语言学中的形式化方法如何推动了当代计算机科学的发展。后来，当计算机界的一些专家了解到，这篇论文的作者是“文革”前北京大学中文系的一个文科研究生的时候，感到非常惊讶。1990年~1993年，冯志伟应邀在德国特里尔大学用德语讲授一系列的汉语语言学课程，并且用德语出版了语言学的专著，显示了他的外语才能。

## 梅花香自苦寒来

在多学科间游走是需要深厚的知识积淀的，知识积淀则来自于勤奋。冯志伟在上世纪60年代初就已经学会了4门外语，而且能够使用这4种外语阅读数理语言学的外文文献。他取得这样的学习成绩，不仅是天赋，更多的是因为勤奋。比如，“文革”期间，冯志伟在朋友们的帮助下，用了将近10年的时间，对数百万字的现代汉语文本(占70%)和古代汉语文本(占30%)进行手工查频，从小到大地逐步扩大统计的规模，建立了6个不同容量的汉字频度表，最后根据这些不同的汉字频度表，逐步地扩大汉字的容量，终于计算出了汉字的熵。在当时的政治环境和技术环境下，能实现如此繁重的工作，难度可想而知。

冯志伟不仅自己学习的时候勤奋有加，对待教学也是一丝不苟。在特里尔大学文学院任教期间，冯志伟用德语给德国学生讲授“汉魏六朝散文选”、“唐诗宋词

选”、“中国现代散文选”、“汉字的发展与结构”、“汉语拼音正词法”、“汉语词汇史”、“机器翻译的理论和方法”等课程。为了讲好课，他苦练德语口语，认真用德语备课，在上每一节课之前，他都要先用德语把讲课的内容自己对自己叙述一遍或多遍，直到能够熟练地背诵为止，他把“备课”当作特别认真，课堂教学效果很好，他的讲课受到德国学生们的一致好评。

冯志伟在韩国科学技术院任教期间，利用全部的业余时间来翻译《语音和语言处理 自然语言处理》(原文为英语)一书，晚上加班到深夜，连续工作了11个月，当翻译完14章(全书的三分之二)的时候，他不幸患了黄斑前膜的眼病，视力出现障碍，难以继续翻译工作，剩下的7章由中国科学院软件研究所的一位年轻的副研究员协助他完成。这位副研究员把剩下的7章逐一翻译成中文，通过计算机网络一章一章地传给在韩国的冯志伟，冯志伟使用语音合成装置，让计算机把书面的文本读出来，通过读出来的语音进行译文的校正。2004年，全书的翻译大功告成，由电子工业出版社以《自然语言处理综论》的书名出版。这本书的出版受到广大读者的欢迎，而冯志伟为此却损害了自己的视力，他不得不借助于语音合成装置来阅读了。

## 硕果累累著作丰

天赋和勤勉成就了跨学科奇才，冯志伟在数理语言学领域不断前行，解决一个又一个难度很高的问题，硕果累累。

在法国留学期间，经过在计算机上编写程序，进行潜心的钻研和反复的试验，冯志伟提出了“多义多标记树模型”(Multiple-labeled and Multiple-branched Tree Model, 简称MMT模型)。这个模型提出后，立即引起了国际计算语言学界的高度重视。冯志伟根据他提出的MMT模型，于1981年完成了汉-法/英/日/俄/德多语言机器翻译试验，建立了FAJRA系统(FAJRA是“法语-英语-日语-俄语-德语”的法语首字母缩写)，在IBM-4341大型计算机上，把20多篇汉语文章自动地翻译成英文、法文、日文、俄文、德文，当时在实验室工作的外国朋友亲自目睹他们视为天书的一篇篇汉语文章被计算机翻译成他们

看得懂的5种外国语之后，无不拍手叫绝。这是世界上第一个汉语到多种外语的机器翻译系统，开创了多语言机器翻译系统之先河。从法国回国之后，冯志伟在中国科技信息研究所计算中心担任机器翻译研究组的组长，在王力先生的鼓励之下，他利用当时北京遥感技术研究所的IBM-4361计算机，于1985年进行了德-汉机器翻译试验和法-汉机器翻译试验，建立了GCAT德-汉机器翻译系统和FCAT法-汉机器翻译系统，进一步检验了MMT模型分析汉语和生成汉语的能力，试验结果良好。

冯志伟上世纪80年代在德国夫琅禾费研究院新信息技术与通信系统研究所(FhG)担任客座研究员期间，克服了重重困难，使用UNIX操作系统和INGRES软件，建立了数据处理领域的中文术语数据库GLOT-C，并且把这个数据库与FhG的其他语言的术语数据库相连接，可以快速地进行搜索和检索，并且能够处理汉语的查询和检索，而且能够处理汉语(当时计算机上还没有成熟的中文操作系统)。这是世界上第一个使用汉字的中文术语数据库，具有开创意义。2006年，联合国教科文组织奥地利委员会、维也纳和国际术语信息中心给冯志伟教授颁发了维斯特奖，以表彰他在术语学理论和术语学方法研究方面作出的突出贡献。

今年71岁的冯志伟著述颇丰，其内容涉及不少领域，得到国内外不少同行专家的广泛关注和高度评价。他现在已在多家著名出版社出版了20余部专著，翻译国外重要论著数十篇(部)，并用汉、英、法、德、俄文等撰写了有代表性的重要学术论文百余篇，在国内外许多著名刊物上发表。

冯志伟对于他的这些成就却看得很平淡，他很少对别人谈起他的成就。平时除了喜欢喝白开水和游泳之外，他几乎没有什么特殊的嗜好，每日粗茶淡饭，过着非常清贫的生活。他从不认为自己是“专家”，没有任何的架子，总是谦和地对待他的学生和周围的同志。在北京大学中文系59级同学纪念册上，冯志伟写下了这样的人生感言：“先天不足，后天失调；岁月蹉跎，艰辛备尝；老当益壮，穷且越坚；平生无悔，褒贬由之。”他对于生活的信条是：“在科学探索的过程中，我所知道的东西终究是有限的，而我不知道的东西始终是无限的，只要平生无悔就很好了！”

## 冯志伟的学术专著



《语言与数学》/冯志伟著  
/世界图书出版公司/定价：  
38.00元

本书共分7章论述了语言与数学的关系：语言符号的随机性与统计数学，语言符号的冗余性与随机过程，语言符号的离散性与集合论，语言符号的递归性与公理化方法，语言符号的层级性与图论，语言符号的非单元性与复杂特征运算，语言符号的模糊性与模糊数学。本书对于语言符号的各种特性的论述是作者多年来在机器翻译和语言信息处理研究的实践过程中潜心探求得出的新见解，关于这样的论述的著作至今在国内外还是第一本。

## 冯志伟印象

### 著名数学家 丁石孙教授

1982年，北京大学中文系开设了“语言学中的数学问题”，冯志伟的这门课程是给汉语专业学生开的选修课程，许多同学对这门学科产生了很大的兴趣，经过一个学期的学习，同学们初步认识了现代数学的发展给语言学注入了生机，觉得获益匪浅，对语言学这门古老的学科分支的发展充满了信心，而且这一举动冲击了相当多的人的旧概念，使闭塞的中国学术界认识到，即使在人文科学教育中，数学也在逐渐起作用。

### 浙江大学外语学院院长 何莲珍教授

冯志伟教授的学术活动中有很多个第一：第一位测定汉字熵值，第一位把从属关系语法和配价语法应用于机器翻译中，第一位开发汉外多语言机器翻译系统，第一位建成中文术语数据库，第一位研究生成语法的公理化方法，第一位提出多义多标记树模型，第一位在国内开设数理语言学课程，第一位出版《数理语言学》专著，第一位出版《现代术语学引论》专著，第一位出版《现代语言学流派》专著。他那朴实的治学风范为学界所称道，在同行中有很好的口碑。

### 云南大学成人教育学院 张在云教授

冯志伟是教育部语言文字应用研究所的学术委员会委员，又是中国科学院自动化研究所国家模式识别重点实验室的学术委员会委员，横跨了语言学科和自动化学科；冯志伟是中国语文现代化学会副会长，又是中国人工智能学会理事，横跨了语文研究和人工智能研究；在国外，他是德国特里尔大学文学院的教授，又是韩国科学技术院电子工程与计算机科学系的教授，同样也横跨了文科和理科的不同学科。