汉语言文学研究文库

面向计算机的现代汉语“得”字研究

骆 琳 著

Computer-Oriented Analysis on the Chinese Character “得” in Modern Chinese

華中科技大學出版社
http://www.hustp.com

汉语言文学研究文库

- 教育部人文社会科学研究规划基金项目“面向计算机的现代汉语‘得’字识别研究”（项目编号：11YJA740061）研究成果
- 华中科技大学自主创新研究基金项目“汉语国际推广背景下的汉语学科创新”（项目编号：2012WZX012）研究成果
- “华中科技大学文科学术著作出版基金”资助项目
- 华中科技大学“985”项目资助

面向计算机的现代汉语“得”字研究

骆 琳 著

Computer-Oriented Analysis on the Chinese Character “得” in Modern Chinese

華中科技大学出版社
http://www.hustp.com
中国·武汉

内 容 简 介

目前我国计算语言学理论和方法的研究还不能为开发汉语信息处理应用系统提供足够的支持。“得”在现代汉语中是一个使用频率高、意义用法复杂的汉字，“得”的研究在语言学和汉语信息处理方面都是重要的课题。该书从汉语信息处理的角度对语言学的这个重要课题重新展开研究，既具有重要的学术价值，又具有充分的应用价值。

本书首次以为计算机识别服务为目的，立足于面向计算机的自然语言信息处理，使用大规模真实语料作为研究材料，对现代汉语的“得”字进行包括语体分布特征、左右邻接特征、语法结构及语义关系等在内的全方位的观察与研究，借助形式标记的发掘，实现对不同类型“得”字结构的鉴别，以适应计算机对不同“得”字“理解”的要求。其研究成果显现出鲜明的文理有机结合的特色。本书的出版不仅为现代汉语“得”字结构的研究和计算机处理方面研究提供重要的参考价值，而且其研究模式对相关研究也具有重要的借鉴作用。

图书在版编目(CIP)数据

面向计算机的现代汉语“得”字研究/骆琳著. —武汉：华中科技大学出版社，2015.10
(中国语言文学研究文库)
ISBN 978-7-5680-1354-3

Ⅰ.①面… Ⅱ.①骆… Ⅲ.①“得”字-研究 Ⅳ.①H146.2

中国版本图书馆 CIP 数据核字(2015)第 263276 号

面向计算机的现代汉语“得”字研究

Mianxiang Jisuanji De Xiandai Hanyu “De” Zi Yanjiu

骆琳 著

策划编辑：周小方 杨 玲
责任编辑：刘 烨
封面设计：原色设计
责任校对：张会军
责任监印：周治超
出版发行：华中科技大学出版社(中国·武汉)
武昌喻家山 邮编：430074 电话：(027)81321913
录 排：武汉正风天下文化发展有限公司
印 刷：武汉鑫昶文化有限公司
开 本：710mm×1000mm 1/16
印 张：13.5 插页：2
字 数：266 千字
版 次：2016 年 1 月第 1 版第 1 次印刷
定 价：39.80 元

汉语言文学研究文库

Computer-Oriented Analysis on the Chinese Character "得" in Modern Chinese

序

《面向计算机的现代汉语“得”字研究》是骆琳博士利用计算机辅助现代汉语语法研究的学术专著。

在汉语研究中，语法研究起步最晚，现代汉语语法研究的第一部专著应该是黎锦熙先生在1924年发表的《新著国语文法》。现代汉语语法研究至今虽不足百年，但其研究成果丰硕，名家辈出，足以傲视学林。但到了20世纪80年代，现代汉语语法研究遇到了来自国际汉语教育和计算语言学两个方面的挑战。这是因为中国学者研究现代汉语语法，不仅需要能科学分析汉语的语言学的学养——专家知识，还依赖于熟练运用汉语的能力——言语能力，以及正确理解汉语语义所需的一切知识——文化背景。国际汉语教育的学习者，本身的母语和作为第二语言的汉语二者包孕的言语能力和文化背景或同或异，于是原先现代汉语语法研究不认为是问题的，可能成了问题，小问题可能成了大问题，简单问题可能成了复杂问题。作为非生命体，计算机不具备任何自然语言的言语能力和文化背景，在理解和处理汉语时，原先现代汉语语法本体研究的一切问题和不是问题的问题都成了问题，要求学者面向计算机，对现代汉语语法重新进行审视，从零开始展开研究。显而易见，这种面向计算机的现代汉语语法研究，不仅对计算语言学具有重要的学术价值，对于面向国际汉语教育的和语言学本体的现代汉语语法研究也具有重要的参考价值。

本书就是在这样一种学术背景下应运而生的。骆琳博士长期从事国际汉语教学和科研，又具有计算语言学的工作经验，她选取“面向计算机的现代汉语‘得’字研究”作为研究课题，正是从自己的现代汉语教学和科学研究中提炼而得，同时这些丰富的教学和科研实践经验也保证了她能很好地完成这个课题并撰成这部专著。

将“得”字作为研究对象是一个很好的选择。我们曾随机抽取大规模的现代汉语文本进行字频统计，结果显示五百万汉字的文本共使用了8364个字种，“得”字排在32位，共使用21223字次，覆盖率为0.0042446，使用频率极高。下面是一些可供参考的数据。

序　号	汉　字	字　次	字　频	累计覆盖率
32	得	21223	0.0042446	0.302879
100	长	8254	0.0016508	0.4729788
1000	球	749	0.0001498	0.8589312
1252	封	549	0.0001098	0.8911852
3778	祳	49	0.0000098	0.9900084

可以看出，计算机在处理现代汉语文本时，每一千个汉字就会遇到 4 个“得”字；或者说，如果能厘清“得”以前 32 个最常用高频汉字，计算机就可能能够处理约三分之一的现代汉语文本。“长”以前 100 字覆盖率已经达到 47%，现代汉语常用单音虚词基本分布在这 100 字内。从 1000 号“球”到 1252 号“封”，字频已经降至在一万字文本仅仅出现一次，下至 3778 号“祳”，累计覆盖率达到 99%，这个区段的字主要是实词。可见，虽然以 100 字为界，前后覆盖率各约为 50%，但如果考虑到虚词的分布，那么在前 100 字中一个字一个字地厘清这些虚词所构成的句法结构及其句法功能，就可以在极大程度上满足计算机处理现代汉语大规模真实文本的需求。骆琳博士就是基于这种逐字解决高频字的处理策略，选取“得”字作为研究对象的。

汉语大规模真实文本的计算机处理是一个复杂而困难的工作，举步维艰。甄选汉语中出现频率高的字，识别这些字组成的各种句法结构，分析其句法功能，形成一个组块，从而一块一块地覆盖整个文本，计算机处理时先识别各个组块，再根据上下文依附关系组合相关组块为完整的句法树。这种语言处理策略就是 20 世纪 90 年代从国外计算语言学界传入的组块分析(chunk parsing)，或称部分句法分析(partial parsing)。我第一次接触组块分析，马上就想到汉藏历史比较语言学中的拼“七巧板”的构想。正如汉语大规模真实文本的计算机处理，汉藏历史比较语言学也是一个复杂而困难的工作，西方印欧历史比较语言学的理论、方法和经验往往不能照搬到汉藏语言的历史比较中，特别是在同源词的择词、匹配、确认上遇到了巨大的困扰，引起了激烈的争论。同源词是建立语言亲属关系的基石，是语言谱系树赖以构建的不可或缺的枝干，正如同构成完整句法树的句法组块。为了解决这个棘手的难题，1969 年，张琨先生在那篇著名的论文《汉藏语系的“针”字》中提出：“我想着重研究一下‘针’这个词，看看是否能在汉藏语系诸语言中建立起同源关系，而不是以不足的证据去建立汉藏语系诸语言的语音对应关系。本文所论都是尝试性的。我仅仅试图给散放着的七巧板的第一块划板找到它应有的位置。我希望最终能建立起足够的同源词，为汉藏语系语言的比较研究和历史研究提供一个坚实的基础。”(《汉藏语系语言学论文选译》72 页，张莲生译)二十年后，计算语言学家提出的句法处理的解决方案与此惊人地相似。古人云“集腋成裘”，佛家说“聚沙成塔”，人同此心，学同此理，化整为零，化繁为简，这实际上就是科学研究中最基本的研究方法——分析。本书

就是分析法在现代汉语大规模真实文本计算机处理方面的一次很好的实验，给完整句法树添加了一个坚实的组块。

计算机作为“电脑”，一切工作都是对人脑的模拟。人类至今对自己大脑是如何工作的还不甚了解，我们重视思维的结果，但不一定了解思维的过程，而要想让电脑获得如同人脑的思维结果，必须回到思维的起点，重现思维的过程。部分句法分析的工作主要是组块的识别与分析，而识别是分析的前提。对于“得”字相关结构的分析，语言学本体研究主要着眼于“得”字作为助词的述补结构。学者搜集这些准备用于分析的实例的过程，是浏览语料，遇“得”字都须判断，一边识别述补结构，一边筛汰非述补结构。人们往往以为自己只是在进行识别，而意识不到同时还进行着筛汰。实际上识别和筛汰都是资料搜集缺一不可相辅相成的两个侧面，没有筛汰就不能识别。所以，如果要利用计算机在现代汉语大规模真实文本中对“得”字进行处理，首先必须对所有不同“得”字结构进行分类，然后根据研究的目的，提取需要的类型，排除不需要的类型，以备下一步的分析工作，这就是计算机识别。

从为计算机识别服务的目的出发，本书不论来源、不论读音、不论词性，凡字形为“得”的字均纳入讨论范围，据此分为六种类型：“得 1”为普通动词，“得 2”为能愿动词，“得 3”为构成述补结构的结构助词，“得 4”为动态助词，“得 5”为构词语素，“得 6”为专名、借词用字等其他用法。另外还有一些误为“得”的错别字，如果着眼于计算机识别，不妨称作“得 7”。“得 7”与其他类型性质根本不同，前六类根据需要或提取或排除，而“得 7”是在文本预处理阶段即应予以校正的对象，永不会被提取。

我们根据“得”字结构在各类文本中的频次及频率列表中的数据加以统计整理成表 1-1，用以观察六类“得”字的分布情况。统计文本总字数为 4353836 字，“得”共出现 12103 次，字频 0.003，与我们五百万字文本“得”字字频 0.004 略有差异。表中数字斜杠前为字次，斜杠后为各类“得”字在 12103 个“得”字中所占的百分比。

表 1-1 “得”字结构在各类文本中的频次及频率列表

得 12103/100.00			
得 3		3539/29.24	
非得 3	得 1	490/4.05	8564/70.76
	得 2	2712/22.41	
	得 4	21/0.17	
	得 5	5226/43.18	
	得 6	115/0.95	

从表中数据可以看出，语言学本体研究重点关注的“得 3”（述补结构）只占全部“得”字的 29.24%，非“得 3”竟占到 70.76%，而不大被关注的“得 2”（能愿

动词)也占22.41%,大致与"得3"持平,至于极少被关注的"得5"(构词语素)所占比例更达到43.18%。可见在六类"得"字结构中,虽然"得3"的结构和功能复杂,但"得5"覆盖面最大。语言学本体研究将研究的重心投注于"得3",显然是出于专家的学术兴趣,而计算语言学如果解决了"得5"的识别问题,就几近完成了"得"字计算机处理的一半任务,这是基于大规模计算机文本处理的策略,这充分反映了两个学科的不同性质。本书将所有类型的"得"字结构全部纳入研究范围,正突显出面向计算机的学术取向。

近年来,计算机自动研究形成了计算语言学的新的发展方向,但目前还没有比较成熟的研究成果。科学研究是人脑最复杂最精密的思维活动,研究的过程包括现象的观察,问题的发现,数据的搜集、校正、整理、加工,信息的提取、定性、分类、排序、统计、比较、综合,进行判断、推理,最终形成符合逻辑的结论。这种能力,即使是人类自己,没有经过长期的学习和严格的训练,也绝不能自动获取。要想让电脑模拟人脑进行科学研究的全部功能,实现完全的计算机自动研究,就目前的计算机技术还很困难,但是让电脑模拟人脑的某些功能,实现部分的计算机自动研究,却是完全可以做到的。科学研究的两大支柱:一是材料,一是逻辑。电脑长于材料的形式化处理,人脑善于逻辑判断和推理,可以取长补短,相辅相成,共同进行工作。计算机根据专家给定的条件对材料进行形式化处理,可能必然和偶然、规律和例外、通例和特例混而不分。这时专家必须进行人工干预,通过逻辑判断推理对计算机处理结果进行分析,再进一步给定条件,计算机再一次对材料进行形式化处理,剔除特例和例外。专家不断地发现问题,计算机不断地按照人的要求解决问题,如此人机互动,反复运作,直至最终得出人所预期的结果。这就是我们长期以来积极提倡和努力实践的计算机辅助研究。(computer assisted research,CAR。)

计算机辅助研究必须依照材料的不同性质,采用不同的计算机技术。根据我们的观察,基于自然语言的真实文本数据大致可以分成四种类型。

第一种是语段式信息,一般采用全文检索(fulltext retrieval)技术进行处理。例如,搜索"得"即可汇聚所有含有"得"字的语段,搜索"得不得了"即可汇聚所有如同"好得不得了"的语段。第二种是格式模信息,指具有固定格式的语段。例如,"好得不能再好"一类的语段,即具有"Z+得+不能再+Z"的固定格式,"说得好不如干得好"一类的语段,即可抽象出"X得Z不如Y得Z"的格式模,其中Z为长度为1或2个字符的字符串,X和Y是两个长度相同但字形不同的字符串。一般可以利用正则表达式(regular expression)构建一个表示这个固定格式的表达式,用以搜索汇聚所有可以纳入这个模板中的语段。第三种是结构化信息,指可以结构化为二维逻辑数据表构建关系数据库(relational database)的数据。例如,颗粒最粗的"得3"数据表的列可以设置"述语"、"补语"、"原句"、

“出处”等几个字段，行收录所有“得 3”结构事例，组合成二维逻辑结构。关系数据库一般采用 SQL(structured query language，结构化查询语言)查询和汇聚信息。第四种是语义型信息，所谓“语义”指研究所需的研究对象的各种属性。例如，有关“得”字的结构类型、词性、句法成分、语义、语气、标点符号、出处等等，计算语言学将这些“语义”设计为各种不同的标记标注到数据上，计算机根据这些标记来辨识访问所需要的数据。传统的现代汉语语法研究所用的标记系统是一套统一预订的字母符号，功能有限。本书采用 XML(extensible markup language，可扩展标记语言)规范进行标记。XML 并没有规定任何具体的标记，只是提供设计和使用标记的规则，学者可以使用自然语言制定种类不限数量无限的标记，所以才叫做“可扩展标记语言”，而不是“某某标记系统”。XML 文档可以采用 XQuery(XML 查询)查询和汇聚信息。本书设计了 36 个 XML 标记，分别表示含“得”结构、形容词、名词、动词、副词、代词、数词、量词、得 1、得 2、得 3、得 4、得 5、得 6、句子出处、结构助词、连词、人名、状态词、助词、副动词、副形词、名动词、形名词、处所词、地名、方位词、机构团体名、介词、区别词、时间词、专有名词、成语、语气词、习惯用语、标点符号等信息，完全满足了研究的需求。本书率先在现代汉语大规模真实文本的计算机处理中使用了 XML 技术，是计算语言学和计算机辅助研究的一次值得重视的实验。

第一、第二两种信息源于文本，形式同为字符串，二者有的可以互通。例如，“好得不得了”一类的语段也可以抽象出“Z 得不得了”的格式模，而“Z 得不能再 Z”的格式模信息，也可以通过检索“得不能再”来提取。但也有的不能相通，例如，“X 得 Z 不如 Y 得 Z”的格式模信息就不能通过字符串的全文检索来提取。第三、第四两种信息也可以通过计算机技术互相转换，XML 就被认为将来可以完全替代关系数据库，但 XML 显然更为灵活，可以处理更为复杂的数据，即使仅有一条实例也可以处理。对这两种信息，研究者注重的不是源于文本的字符串，而是通过字段或标记附加的各种信息，这些信息源于专家知识。本书综合使用这些计算机技术处理六类“得”字结构，先设定条件搜索汇聚特定格式的“得”字结构，再标注各种所需的学术信息，再搜索，再标注，循环往复，直至穷尽全部数据。对于大规模真实文本，这显然是一种劳动密集型工作，需要耗费大量时间和精力，而且更是一种知识密集型工作，需要学科专业学养和计算机科学训练，绝非一个不是专家的操作员或一个不懂计算机技术的专家所能胜任。这就是计算机辅助研究的特点和魅力所在。

早期的计算语言学研究，大多是在规模有限的经过训练的文本中进行的。1990 年 8 月，在赫尔辛基召开的第 13 届国际计算语言学大会上，大会组织者首次提出了处理大规模真实文本的战略目标。大规模真实文本的意义，不仅在于数据量巨大，更在于文本没有经过训练，数据没有经过提纯，类型繁多，可用价值

密度低，语言素材芜杂，处理难度大，对于传统的计算语言学无疑是一个巨大的挑战。骆琳博士勇敢地接受了这个挑战，本书就是她交出的一份答卷，对现代汉语大规模真实文本的计算机辅助研究做出了可供借鉴的探索。

以上所说是我觉得本书可以注意的几点。至于本书的长短得失，相信能阅读这本书的应该都是专家学者，仁者见仁，智者见智，不用我赘言。

2014 年 2 月 2 日

前言

随着信息时代的到来，中文信息的自动化处理越来越显示出其重要价值。然而缺少细致的致力于规则的句法描写已成为严重制约中文信息自动化处理的瓶颈。句法分析作为自然语言信息自动化处理中的重点和难点，虽然走过了几十年的研究与发展历程，但是当面对海量数据，需要对真实文本进行分析和处理时，由于汉语句法结构的复杂性和灵活性，使对汉语句子结构的整体分析无论是在时间上还是空间上都面临着巨大的挑战。部分句法分析(partial parsing)作为近年来新出现的一个语言处理策略，主要着眼于组块(chunk)的识别与分析。已有的部分句法分析研究成果证明，尽管部分句法分析得到的结果并不是一棵完整的句法树，但分析得到的每一个组块都是完整句法树的一个子图(subgraph)，只要在组块间加上彼此的依附关系(attachment)，就可以构成一棵完整的句法树。这样就能够在某种程度上简化句法分析的任务，同时也使句法分析技术在大规模真实文本处理系统中得以利用成为可能。

本书以“面向计算机的现代汉语‘得’字研究”为题正基于此，我们希望通过对“得”字结构的识别研究，使之成为完整句法树的一个子图，从而最终实现计算机的自动识别。由于纯粹从为计算机识别服务的目的出发，立足于面向计算机的自然语言信息处理，我们将研究范围限定在无论来源、无论读音、无论词性，凡字形相同的“得”字均在我们的讨论之列。

研究思路可以概括为：①完成对封闭性训练语料的核对与标注。②使用开发工具 Visual Studio . Net 编写 Visual Basic . Net 应用程序。③自建数据库，完成对数据的分析和统计。④对数据库中的数据进行穷尽性研究，并在借鉴前贤们已获得的研究成果的基础上，归纳出“得”字在不同语用环境中的句法特征和语义表现，分析共现成分特征。⑤分析“得”字述补结构的语法及语义关系。

研究重点主要集中在三个方面。

第一，“得”字结构的分布特征研究。在对各类“得”字的句法功能及语义特征进行明确界定的基础上，对“得”字结构的语体分布特征进行了详细的描述，并对表现出来的明显倾向性进行了适当的分析。着重观察“得”字述补结构中“得”

前成分与不同语体的对应关系，以及“得”后不同补语类型在各类语体中的分布情况，并分析其分布状况及产生对应关系的原因。

第二，“得”字结构的组合特征研究。在对各类“得”字左右邻接特征分布进行统计的基础上，结合对“得”字左右邻接限制特征的调查，对“得 1”、“得 2”、“得 3”、“得 4”的左邻接和右邻接特征及其限制性特征进行了包括隐性邻接在内的详细描述，发现其邻接规律，并就“得”字的左右显性邻接共现情况进行观察和描述；引入“熵”的计算，通过数据的演算进一步说明各类“得”字对左右邻接词语所具有的选择性。

第三，“得”字述补结构的语法及语义分析研究。在借鉴前人研究成果的基础上，从利于计算机识别与处理的观点出发，对“得”字述补结构的结构类型，即可能式述补结构和非可能式述补结构，从句法模式到句法成分间的语义选择进行了明确的界定；并就非可能式述补结构中补语的结构类型进行分类，确立了非可能式述补结构的结构形式与语法意义的对应关系。

我们希望研究得出的结论及建构的框架能为类似字词结构的计算机处理研究提供借鉴，并为今后计算机相关中文信息处理的应用系统的开发提供语言学上的支持。

目录

第一章 绪 论

第一节 问题的提出

自然语言的计算机处理技术(natural language processing by computers,简称 NLPC)从 20 世纪 50 年代起步,至今已经成为现代科学技术研究的一个热点。现已建成的计算机自然语言处理系统虽然很多,但是尚没有一个专门的真正意义上的汉语词汇—语法计算机自动检索系统,尤其是在汉语语法的自动检索方面更显不足。分析其原因关键还在于我们的语言研究者并没能给出全面、细致的句法规则描述。计算机由数据处理、信息处理发展到知识处理,对语言文字处理在深度和广度上提出的要求都越来越高。语言研究成果直接推动了计算机技术的发展,语言学研究的进程也直接制约着计算机技术发展的进程。计算机的使用,不仅更新了语言研究的手段,同时也向语言研究者提出了新的要求。传统的语言研究用于人与人的交际,人理解语言可以凭借背景知识和语感来对语言现象进行判断,而计算机却不行。计算机强调规则化和可操作性。它要求用机械的方法推演和计算,从而对语言进行理性的分析。缺少全面、细致的致力于规则的句法描述已成为严重制约中文信息自动化处理的瓶颈。而这些细致、具体的规则描述的缺失,使得即便编程人员的水平再高也无能为力。因此,致力于规则化的汉语句法描写应该成为当前汉语语法研究的总体方向和目标。

汉语句法体系的建构吸收和借鉴了大量英语句法描写的相关知识和内容。20 世纪以来,虽然经过许多汉学家的艰苦努力,提出了不少行之有效的汉语句

法分析方法，但究竟哪种方法更适合汉语，仍有许多争论。近年来，针对在大规模真实文本中运用完整句法分析法所遇到的困难，一些学者开始尝试把一个完整的句法结构分解为几个易于处理的子结构，以降低完整句法结构分析的难度，提高分析效率，从而提出了部分句法分析思想。部分句法分析不以得到完整的句法分析树为目标，而只要求识别其中某些结构相对简单的成分，即组块，因此，其分析结果并不是一棵完整的句法树。然而，各个组块却是完整句法树的一个个子图，只要对其间的依附关系加以考量，就能得到完整的句法树。这一思想使句法结构分析任务在某种程度上得以简化，并有利于句法分析技术在大规模真实文本处理系统中得以推广和使用。本研究课题的提出正基于此，我们希望通过对“得”字结构的识别研究，使之成为完整句法树的一个子图，从而最终实现计算机的自动识别。

现代汉语中的“得”字是一个使用频率极高，意义用法相当复杂的汉字。在不同的语境和上下文组合中，它代表了几种不同层次、不同类属的语言单位，具有不同的功能，表达不同的意义。

《汉语大词典》[1]在义项的选择上古今兼顾，在释义上注重溯源，例证大部分为历史文献；《现代汉语词典》[2]则是为推广普通话、促进汉语规范化和汉语教学服务，在义项、词形、读音的选取和释义上，以现在通行为标准的现代汉语为标准，不列古形、古音、古义；而《现代汉语八百词》[3]选词以虚词为主，每一个词按意义和用法分项详加说明，主要供非汉族人学习汉语、一般语文工作者和方言区的人学习普通话时使用。

在《汉语大词典》[1]中，“得”字被分别标示为：

“得 1 dé”：(1)获得、得到；(2)捕获；(3)成功、完成；(4)演算产生结果；(5)贪得；(6)得利、得益；(7)得生；(8)有；(9)适宜、得当；(10)中；(11)知晓、明白；(12)满意、得意；(13)亲悦、融洽；(14)值、遇；(15)与、给；(16)到、抵达；(17)待、等到；(18)犹言行、可以；(19)犹言算了；(20)用在动词前，表示能够；(21)岂、怎；(22)通“德”。(例子略)

“得 2 děi”：(1)需要；(2)必须；(3)将要。(例子略)

“得 3 de”：助词 (1)用在动词后，表示可能、能够；(2)用在动词后连接表示程度或结果的补语；(3)用在动词后表示动作已经完成；(4)用在动词后表示动作持续进行；(5)犹的。

在《现代汉语词典》[2]中，“得”被解释为：

“dé 得 1”：(1)得到；(2)演算产生结果；(3)适合(得体)；(4)得意；(5)完成；(6)用于结束谈话时，表示同意或禁止；(7)用于情况不尽如人意时，无可奈何。

“dé 得 2”：(1)用在别的动词前，表示许可；(2)用在别的动词前，表示可能这样。

“de 得 3”：助词 (1)用在动词后面，表示可能；(2)用在动词和补语中间，表示可能；(3)用在动词或形容词后面，连接表示结果或程度的补语；(4)用在动词后面，表示动作已经完成。

“děi 得 4”：(1)需要；(2)表示意志上或事实上的必要；(3)表示揣测的必然；(4)舒服、满意。

而在《现代汉语八百词》[3]中，对“得”的解释为：

“得 1 de”：[助] 连接表示程度或结果的补语。基本形式是“动/形＋得＋补”，动词不能重叠，不能带“了、着、过”。(1)动/形＋得＋形。如“说得快/雨下得急/收拾得干净极了/颜色绿得可爱/茶沏得酽酽的”，表示否定在“得”后加“不”字。(2)动/形＋得＋动。“得”后不能是单个动词。如“跑得一个劲儿地喘/大厅里亮得如同白昼/高兴得大声笑着/团结得像一个人一样/乱得理也理不清”。(3)动/形＋得＋小句。如“累得气都喘不过来/跑得满身都是汗/伤心得眼泪围着眼圈儿转/气得手直发抖”。(4)动＋得＋名＋动。“得”后不能是单个动词。名词是前面动词(使动意义)的宾语，这个名词都可以用“把”字提到动词的前边去。如“忙得他团团转/逗得我们哈哈大笑/乐得他跳了起来”。(5)一般的动宾短语加“得”时，要重复动词。如“他唱歌唱得好听极了/我说话说得忘了时间/孩子们听故事听得不想回家”。(6)动/形＋得＋四字语。如“说得一清二楚/说得头头是道/搞得乱七八糟/忙得不亦乐乎”。(7)形＋得＋很。如“好得很/糟得很/清楚得很”。(8)动/形＋得。“得”后的话不说出来，有“无法形容”的意味。如“看把你美得！/瞧你说得！/这番话把他气得！”。(9)以上格式的动词或形容词前如意思上容许加否定词，一般限于“别、不要”。如“别说得太过分/不要弄得太响”。

“得 2 de”：[助] 用于表示可能、可以、允许。(1)动＋得。动词限于单音节。否定式是在“得”前加“不”，动词不限于单音节。如“用得/吃得/这东西晒得晒不得？/这双鞋穿得/这件事放松不得/”，这种格式里的动词一般都是被动意义，不能带宾语。但是“顾得、顾不得、舍得、舍不得、怨不得”等是主动意义，可以带名词、动词做宾语。如“顾得这个，顾不得那个/舍不得吃/怨不得你”；(2)在动结式和动趋式复合动词的中间插入“得”或“不”，表示可能或不可能。如“看得清楚，看不清楚/扯得断，扯不断/吃得了，吃不了/爬得上去，爬不上去”，与上面(1)不同，这里的动词只要是及物的都可以带宾语。如“看得清楚那几个字/我们拿得下这块大油田/这个东西，我叫不出名字”，这一类有的已经凝固为熟语，没有相应的不带“得、不”的格式；或者虽有，但是意思不同。如“对得起，对不起/来得及，来不及/这儿坐得下，那儿坐不下”，这类熟语，否定式比肯定式用得多。

“记得、认得、晓得、觉得、显得、值得、省得、免得”里边的“得”是构词成分，不是动词后面的助词。

“得 děi”[助动](1)表示情理上、事实上或意志上的需要；应该；必须。不能