

# 汉语日常会话的对话行为分析标注研究\*

周 强

清华大学信息技术研究院语音语言技术中心

清华信息科学与技术国家实验室, 北京 100084

**摘要:** 对话行为分析是进行更深入的对话理解模型探索的合适切入点。本文综合前人研究成果, 设计了一套针对汉语日常会话的对话行为标注体系。引入主客观阐述和正反向反应子类, 加强对话行为依存对和连贯修辞对的结构描述, 同时引入话题线索分析机制, 有效组织会话中的话题变化趋势。在 500 个日常会话片段上进行的对话行为标注实验显示出了 90%左右的两人独立标注宏一致率。表明目前的对话行为标记集设计具有很好的可操作性, 可以适应汉语日常会话的行为功能模式描述需求。

**关键词:** 会话分析; 对话行为标注; 话题线索

**中图分类号:** TP391

**文献标识码:** A

## Research On Dialog Act Annotation In Chinese Daily Conversation

Qiang Zhou

Tsinghua National Laboratory for Information Science and Technology

CSLT, Research Institute of Information Technology, Tsinghua University, Beijing 100084

**Abstract:** Dialog act (DA) analysis is a suitable breakthrough point for further exploration of dialog understanding model. By absorbing the essence of previous studies in dialog act annotation schemes, the paper designs a new dialog act annotation scheme for Chinese daily conversation. The core function subclasses for the subjective and objective statement and the positive and negative response are introduced to enhance the DA descriptive ability. Two descriptive mechanisms of coherent rhetorical pairs and DA dependency pairs are combined to form a more complete description system for session structure. The topic thread analysis mechanism is introduced to effectively organize the topic-changing trend in conversation. DA tagging experiments on 500 daily conversation fragments show about 90% tagging macro consistency on two independent annotators. It indicates that the designation of the current DA tag set has good operability, and it can meet the needs to describe the functional behavior patterns of Chinese daily conversation.

**Keywords:** Conversation Analysis; Dialog Act Annotation; Topic Thread

### 1. 引言

对话(Dialog)是人类最基本的信息交互方式。对话理解模型需要重点关注信息交互过程中对话参与者的认知机制变化及其实现机理, 这与独白(monolog)理解中关注的内容连贯性结构有很大不同。近年来, 认知心理学家和心理语言学家在这方面进行了大量研究, 提出了许多有效的对话理解模型。其中的典型工作包括: Clark(1996)提出的交际作为转换(Communication-as-transfer)模型<sup>[9]</sup>、Pickering & Garrod(2004)提出的交互对齐(Interactive Alignment)模型<sup>[10]</sup>、Fusaroli et. al. (2014)提出的人际协同(Interpersonal Synergy)模型<sup>[11]</sup>。

\* **收稿日期:**

**定稿日期:**

**基金项目:** 互联网话语理解的心理机制与计算建模 (61433018); 汉语语篇中连贯关系和隐含角色的分析标注研究 (61373075)

**作者简介:** 周强 (1967—), 男, 博士, 研究员, 主要研究方向: 自然语言理解, 语料库语言学, 词汇语义学。

会话(Conversation)是对话的语用表现形式, 本文特指其为以文本形式记录的对话内容, 大致可分为通过录音转写保存的面对面自然会话和互联网环境下通过键盘输入形成的网络会话(Internet Relay Conversation, IRC)两大类。对话行为(Dialog Act, DA)描述了会话过程中话语消息(Utterance)所发挥的交互作用, 主要包括语义内容(Semantic Content)和交际功能(Communicative Function)两大部分内容。其中语义内容主要涉及话语消息相关的实体、关系、动作和事件内容; 交际功能则主要关注对话参与者在理解了话语消息后使用其语义内容对其信息状态进行更新的行为方式。它们大致相当于 Searle(1969)提出的言语行为(Speech Act)理论中使用的言内之义和言外之义两个概念<sup>[22]</sup>, 但其内涵和外延都有了新的扩展。从这个意义上说, 对话行为分析可以看成是一种浅层的对话理解, 可以作为进行更深入的对话理解模型探索的一个合适切入点。

从 1997 年起, 国外在 DA 描述体系和标注语料库开发方面进行了大量研究。提出的典型 DA 描述体系有: DAMSL(Dialogue Act Markup in Several Layers)<sup>[1]</sup>, DIT(Dynamic Interpretation Theory)<sup>[2]</sup>, ISO 24617-2<sup>[12]</sup>。构建的典型自然会话 DA 标注库有: 针对电话录音的 SWBD<sup>[13]</sup>、针对会议记录的 ICSI-MRDA<sup>[10]</sup>等。处理的网络会话 DA 标注库包括: 异步会话类的电邮 E-mail 和论坛 Forum 数据等<sup>[14]</sup>; 同步聊天数据<sup>[15]</sup>; 微博类 Twitter 数据<sup>[27]</sup>; 短信服务类数据<sup>[24]</sup>。相对而言, 国内在这方面的研究工作则比较少, 主要集中在针对访谈类语料的 DA 标注体系设计<sup>[26]</sup>和会话结构分析标注<sup>[25]</sup>方面, 在自然会话和网络会话的 DA 分析标注方面还有许多研究工作亟待开展。

本文在汉语日常会话的 DA 标注体系方面进行了初步探索, 提出了一套汉语日常会话 DA 标注体系, 其主要特点是: 1) 引入主客观阐述和正反向反应子类描述, 拓展了标注体系的 DA 描写能力; 2) 引入连贯修辞对描述机制, 与内置的 DA 依存对相配合, 形成了更完整的会话结构描述体系; 3) 引入话题线索分析机制, 有效组织会话中的话题变化趋势, 为后续处理网络会话中的话题纠缠问题打下了很好基础。在 500 个日常会话片段上的双人独立标注数据显示, 该体系具有很好的可操作性, 可以适应汉语日常会话中各种常用对话行为描述的应用需求, 为后续进行汉语日常会话的对话行为自动识别和对话理解模型开发打下了很好的基础。

在下面几节中, 我们首先介绍相关研究进展(第 2 节), 然后分析我们的标注规范设计要点(第 3 节), 并给出目前的实验分析数据(第 4 节), 最后是初步结论和后续工作展望。

## 2. 相关研究进展

### 2.1 自然会话的 DA 标注研究

在自然会话研究方面, Allen and Core (1997) 综合前人的研究成果, 提出了一种针对不同对话任务的通用对话行为描述体系 DAMSL<sup>[1]</sup>, 把不同对话行为标记组织成如下 4 大板块: 交际状态(Communicative Status, CS)、信息水平(Information Level, IL)、前指功能(Forward Looking Function, FLF)、后指功能(Backward Looking Function, BLF)。该标记体系设计主要吸收了言语行为<sup>[22]</sup>和毗邻对(Adjacency Pairs)<sup>[16]</sup>理论的描述思想, 采用多维度(Multi-dimension)的标记设计方式, 大大提升了该体系的通用性和灵活性。但由于没有很好地考虑不同信息维度的 DA 标记之间的内容互斥性, 致使在实际应用中会产生 DA 标记的组合爆炸问题, 不利于在此标记体系下进行相应标注语料的机器学习和知识获取研究<sup>[20]</sup>。

Switchboard-DAMSL(SWBD-DAMSL)是与 DAMSL 几乎同步进行的 DA 标注项目<sup>[13]</sup>, 主要针对自由话题的双人口语电话转录会话。首先使用 DAMSL 体系对所有 SWBD 中的 205,000 条话语消息进行了多人独立 DA 标注。通过统计分析, 获取了 220 个不同的 DAMSL 标记组合。然后对这些标记进行了聚类重组, 最终形成了包含 42 个单维度标记的 DA 标记集<sup>[20]</sup>。其主要贡献在于: 通过引入新的排斥性限制, 重构形成了更为简练的单维度 DA 标

记集，大大提升 DA 自动分析模型的处理效率。

ICSI(International Computer Science Institute)会议记录项目主要关注多人对话内容<sup>[10]</sup>。其 DA 标注体系 ICSI-MRDA(Meeting Recorder Dialogue Acts)主要基于 SWBD-DAMSL, 并针对多人对话的描述特点进行了改进和完善, 补充了针对话轮控制(Turn-taking)和不完整话语消息的 DA 描述标记。按照 DA 功能的相似度将所有 DA 标记重组为 12 个标记群, 并进一步把这些标记群归并为两大等级 (tier) T1 和 T2。每条话语消息的基本标注规则是: 选择 1 个 T1 标记和 0~n 个 T2 标记, 即 T1 (^ T2)\*。这样, ICSI-MRDA 的 DA 标记又重新回到 DAMSL 的多维度描述形式, 但其中 DA 功能描述的深度和广度都有了新的提升。

Popescu-Belis(2003)对 ICSI-MRDA 标注数据进行了统计分析, 加强了对两大描述等级 T1 和 T2 内部及两者之间的 DA 描述功能互斥性分析控制, 重新归纳构建了 MALTUS (Multidimensional Abstract Layered Tagset for Utterances) 标记集。据此最终形成 29 个不同的 DA 标记覆盖所有的 ICSI-MRDA 标注数据, 并在两者之间建立直接的映射关系: ICSI-MRDA → MALTUS<sup>[20]</sup>。

针对日常会话中话语消息固有的多功能性(multi-functionality)描述特点<sup>[5]</sup>, DA 标记集设计经历了从多维度组合(DAMSL)到单维度标记(SWBD), 再到多维度组合(ICSI-MRDA)和单维度标记(MALTUS)的螺旋式发展历程。研究人员希望能根据不同会话的描述特点, 在理论完备性和实际易操作性之间寻找到一个较好的平衡点。这里的处理关键是对 DA 描述维度的准确定义。

针对 DAMSL 在 DA 描述维度定义上的模糊性问题, Bunt(2006)提出了一种更为准确的 DA 维度定义方法<sup>[3]</sup>, 总结归纳了 11 个 DA 描述维度。据此, 可以在 DA 描述的多功能性和多维度标记间, 通过如下的语义相关原则建立起一一对应关系: 任意两个交际功能所赋予的描述维度或者是互斥的, 或者存在逻辑蕴含关系<sup>[3]</sup>。Ptukhova & Bunt (2009)在三个会话标注数据上实际验证了这些 DA 维度之间的相互独立性<sup>[18]</sup>。

在此基础上, Bunt(2009a)设计了一套新的基于动态解释理论(Dynamic Interpretation Theory, DIT)<sup>[2]</sup>的 DA 标注体系 DIT++<sup>[4]</sup>。把所有交际功能描述 DA 标记组织成两大部分: 1) 通用目的功能 (General-Purpose Functions, GPFs) 标记; 2) 特定维度功能 (Dimension-Specific Functions, DSFs) 标记。DIT++已在多个会话标注库上得到应用, 包括: DIAMOND 指令型双人德语会话库; AMI 面向任务的多人英语会话库; OVIS 信息搜寻类人机会话库等。欧盟项目 LIRICS 也使用了精简的 DIT++标注集。其中在德语和英语标注库上三人独立标注的高一致性显示了 DIT++标记的实用价值<sup>[17]</sup>。

从 DAMSL 到 DIT++, 经过十多年的发展, 集成各个对话行为标注集的描述特点, 逐步形成了一个 DA 标注领域的国际标准: ISO 24617-2<sup>[12]</sup>。Bunt et. al(2010)介绍了该标准草案的基本内容<sup>[7]</sup>, 细化了两类 DA 依存关系描述方法。Bunt et. al(2012)则对该标准最终版中补充的修辞关系和量子化等修改内容进行了分析说明<sup>[8]</sup>。

近年来, 国内研究者在汉语访谈类语料的对话行为和会话结构分析标注方面也进行了一些研究。杨平(2006)提出了一套针对访谈语料的对话行为分类体系, 分析了访谈问句体现的挖掘信息、组织对话和表达态度等对话行为特点<sup>[26]</sup>。王珊等(2016)采用包含 5 个大类、16 个小类的会话结构描述体系, 对 20 个典型访谈节目片段中的话语消息进行了会话结构的详细标注<sup>[25]</sup>。这些研究与国外在自然会话上的对话行为分析标注形成了很好的互补借鉴效果。

## 2.2 网络会话的 DA 标注研究

从 2010 年起, 随着社交媒体的应用和普及, 出现了大量网络会话语料, 对其中的话语消息进行 DA 分析标注逐渐成为学术界新的研究热点。但其相关研究更关注 DA 自动分析建模, 相应的 DA 标记体系设计则相对比较简单, 大多沿用已有的自然会话 DA 标注体系, 如: DAMSL, DIT++等, 针对 IRC 语料的实际应用需求进行简化压缩, 重点标注 IRC 语料中的

不同话语消息的 GPFs。下面对相关内容进行简要分析说明。

Joty et. al(2011)主要关注 IRC 中的异步会话语料,包括电邮 E-mail 和论坛 Forum 数据等。从 MRDA 标记集中选择了 12 个 DA 标记分别对 40 个电邮线索和 200 个论坛线索数据进行了分析标注。其中的高频 DA 标记包括 Statement, Question, Response, Action motivator, Polite mechanism 等,大多可以归入 DIT++ 的 GPFs 和社会义务管理 SOM 维度中<sup>[14]</sup>。

Kim et. al(2012)主要关注 IRC 中的同步聊天语料,包括话题比较明确的双人客服会话、网络课堂多人讨论和话题随意的日常聊天等。其中话语消息的不规范性和话题纠缠(Topic entanglement)问题比较突出。从 DAMSL 标记集中选择了 14 个 DA 标记对包含 5276 条话语消息的 15 个图书馆论坛聊天组进行了分析标注。其中的高频 DA 标记包括: Statement, Response, Question, Background, Opening, Closing 等,其统计分布特点与 Joty et. al(2011)在异步电邮和论坛语料上得到的数据非常类似<sup>[15]</sup>。

Zarisheva & Scheffler(2015)主要关注 IRC 中的微博类 Twitter 数据。针对 Twitter 中不同内容体裁和描述风格混杂的会话特点,从 DIT++ 标记集中选择了所有通用目的功能 GPFs 和三类特定维度功能:话语结构关系 DSM、社会义务关系 SOM 和自我交际关系 OCM,对 1213 条德语 Twitter 数据进行了 DA 标注和分析<sup>[27]</sup>。

Xue et al. (2016)主要关注 IRC 中的短信服务(Short Message Service, SMS)类数据。针对 SMS 中类似电邮的异步会话描述特点,采用基于 DAMSL 的 FLF-BLF 对话行为关系对和基于 PDTB<sup>[21]</sup>的修辞连贯关系对相结合的 SMS 会话结构描述体系,对 44 个 SMS 数据文件进行了对话行为和会话结构的分析标注<sup>[24]</sup>。

### 3. 标注规范设计

本节基于上节对前人研究的综述分析,针对汉语日常会话描述特点,总结形成了一个汉语对话行为分析标注规范。后续主要内容包括:标注单位选择,标注体系设计,具体实例分析。有关该标注规范的详细内容,可参阅参考文献[29]。

#### 3.1 标注单位选择

给定一个日常会话片段,进行对话行为标注的第一步是选择合适的标注单位。综合前人的研究成果,在标注单位选择方面一般可分为以下几个处理层次:

- 1) 话轮单位(Turn Unit):指对话过程中某个对话参与者临时控制话语权,在一定时间内产生或发出的一段信息。一个话轮传递的信息可以是一个词语片段、一个句子或一个复杂段落。考虑到话轮描述内容的复杂性和多样性,研究人员一般不选择它作为 DA 标注单位。
- 2) 话语消息单位(Utterance Unit):指一个话轮中对话参与者发出的一条信息。在文本形式记录的话轮单位中,一般以其中的结句点号(包括句号、问号、感叹号等)作为标记自动切分出来,大致相当于书面独白语篇中的句子(sentence)单位。话语消息是许多 DA 标注体系(如:DAMSL、SWBD、MRDA)选择的基本标注单位。
- 3) 功能片段(Functional Segment):指一个话语消息中可以表征某个交际功能的最小信息单位。在文本形式记录的话语消息中,大多表现为以逗号、分号等分隔点号切分的小句(clause)单位。Bunt(2011)在分析话语消息的多功能特点时,详细定义了这个处理单位<sup>[6]</sup>。它是 DIT++ 和 ISO 24617-2 等近年来新提出的 DA 标注体系选择的基本标注单位。

我们对 DA 标注单位的选择原则是:以话语消息作为基本标注单位,在某些多功能的话语消息上,进一步切分出功能片段来描述的其中各个特殊交际功能,以此作为 DA 标注的最小处理单位。

### 3.2 标注体系设计

在 DA 标记体系设计方面，则采用交际功能和语义内容描述并重的处理原则。在交际功能描述方面，设计了一个核心功能(Core Functions)标记集，提取类似 DIT++中的通用目的功能，对不同类型会话的核心功能进行描述；同时，自由选择组合多个特定维度功能，与核心功能相结合，共同完成相关交际功能的完整分析标注。在语义内容描述方面，设计了一个话题线索(Topic Threads)标记组，将会话片段中描述同一话题内容的话语消息组织在一起。在此基础上，进一步分析各个话语消息/功能片段间的 DA 依存关系和修辞连贯关系，凸显话题线索内部的会话结构表示，初步形成对会话片段中隐含的参与者交际意图和内容主旨的全面分析理解。

对于核心功能标记集，吸收 DAMSL 中的 FLFs、BLFs 和 DIT++中的 GPFs 描述精髓，补充 ISO 24617-2 中的 DA 依存对机制，初步形成如下 5 类核心功能标记：

- 1) 阐述(Statement)类：细分为针对外部客观世界的实体/事件描述(S-1)和针对内部主观世界的认知/评价描述(S-2)两个子类；
- 2) 询问(Question)类：细分为是否问题(Q-1)、选择问题(Q-2)、特指问题(Q-3)和其他问题(Q-4)四个小类；
- 3) 动作(Action)类：细分为说话者承诺完成动作(DO-1)、受话者预期完成动作(DO-2)和会话双方协同完成动作(DO-3)三个小类；
- 4) 回答(Answer)类：对于是否问题(Q-1)，增加 A-1, A-0 两个标记，显性表示‘是/否’答案信息。大部分情况下，可与会话中该回答对应的问题会形成 Q-A 功能依存对；
- 5) 反应(Response)类：细分为同意/接受(R-1)、否定/拒绝(R-2)、中立/评价(R-3)、保留(R-4)等四个小类。大部分情况下，可与上文的阐述或动作类话语消息/功能片段形成 S-R 或 DO-R 反馈依存对；

对于特定维度功能，目前主要选择社会义务管理(SOM)维度，设置打招呼、自我介绍、道歉、感谢、告别、祝贺以及其他等 7 个社会义务行为标记。对于话题线索，采用简化的话题线索标记：T<话题序号>，使用不同自然数表示会话中使用的不同话题内容。对于修辞连贯关系，基本沿用已有的书面独白语篇连贯性描述体系中设计的连接关系标记<sup>[28]</sup>，并通过设置连贯标记显性标注会话中可能的修辞连贯对。

### 3.3 具体实例分析

为方便处理，我们把每个经过 DA 标注处理的会话片段保存在一个矩阵数据表中。矩阵的每行依次保存语篇中的各个话语消息/功能片段及其相应的 DA 分析标注结果。矩阵中的每列从左到右依次保存以下信息：1~3 列为会话基本信息，包括消息编号、参与者 ID 和话语片段等；4 列为话题线索；5~6 列为连贯修辞对描述，包括连贯标记和修辞关系；7~9 列为交际功能描述，包括依存标记、核心功能和特定维度等；最后第 10 列为备注。

下面给出一个具体标注实例，从中可以看出：1) 第 11 条话语消息被人工切分为两个功能片段，便于准确标注两者不同的交际功能：11-1 → 同意反应 R-1，11-2→主观评价 S-2；2) 多功能 DA 标注是常见现象，第 4、5、7、10、13 条话语消息都标注了两个核心功能标记；3) 通过交际功能标注而形成的内置 DA 依存对，我们在‘依存标记’类中使用了不同标记进行显性标注，如：4-5 条消息间形成的反馈依存对 S-2 → R-2（依存标记=2-1）；6-7 条消息间形成的功能依存对 Q-3 → A-1（依存标记=1-1）；4) 连贯修辞对主要反映了话语消息描述的语义内容之间的关联性，如：4-5 条消息内容间形成的因果连接关系 4-1；5) 话题线索标记反映了会话中不同话题的变换内容，如 6-16 条消息共同组成的话题线索组 T2 反映了其中工作技巧话题的描述内容。

消息	参与	话语片段	话题	连贯	修辞	依存	核心	特定	备注
----	----	------	----	----	----	----	----	----	----

编号	者 ID		线索	标记	关系	标记	功能	维度	
4	1-B	我觉得写这个很困难。	T1	1	4-1	2-1	R-2, S-2		
5	1-A	我认为它一点儿都不难。	T1	1	4-1	2-1	R-2, S-2		
6	1-B	你有什么技巧吗?	T2	0			Q-3		
7	1-A	当然有技巧了。	T2	1	1-1	1-1	A-1, S-1		‘是’ 答案
8	1-B	那你传授一下技巧。	T2	1	3-2		DO-2		请求
9	1-A	你得写出预计目标。	T2	1	6-3		DO-2		指导
10	1-B	有了目标才可以知道后面的规划。	T2	1	3-2	2-1	R-1, S-2		
11-1	1-A	是这样的,	T2			2-1	R-1		FS
11-2		你很有悟性。	T2				S-2		
12	1-B	有了目标之后呢?	T2				Q-3		
13	1-A	当然是条件分析。	T2	5	2-1	1-1	A, DO-2		
14	1-B	好的,我记住了。	T2			2-1	R-1		
15	1-A	分析后给出具体的措施。	T2	2	2-1		DO-2		
16	1-B	还有什么?	T2				Q-3		

#### 4. 标注数据分析

基于上述 DA 标注规范,我们选择了 500 个日常会话片段<sup>1</sup>进行了初步的 DA 标注实验。这些会话片段按照话题内容不同组织成工作、饮食、爱情、交通和天气 5 大组,每组包含 100 个会话片段,每个片段包含 20 条话语消息,合计 10000 条消息。我们安排了两名研究生分别独立进行 DA 分析标注,得到两份完整的 DA 标注数据。表 1 显示了其中一份标注数据的汇总统计结果。从中可以看出:

- 1) 在标注过程中对 25%左右的话语消息进行了功能片段的人工切分,显示出目前 DA 标注过程中选择合适标注单位的必要性和重要性。如何设计有效的自动功能片段切分模型是后续研究中需要解决的一个重要 DA 预处理技术。
- 2) 平均每条话语消息/功能片段上标注的核心功能或特定维度功能达到了约 1.1 个,初步显示了日常会话中话语消息的多功能描述特点。其中的核心功能占了 98.3%的绝大多数,基本符合我们的预期判断。
- 3) 在目前标注的 5840 个 DA 依存对中,功能依存和反馈依存所占比例分别为 31.2%和 68.8%。更深入的研究发现,心理距离不同的会话者选择的不同话题会话中,两种 DA 依存对的分布比例会显示出不同差异。心理距离较大的会话者选择的天气、交通等话题会话中,Q-A 功能依存对所占的比例会较小。相关详细内容将另文发表。
- 4) 目前标注的话题线索组平均可覆盖约 13 条话语消息/功能片段,包含约 6 个 DA 依存对。如何发现消息片段和 DA 依存对组合形成话题线索的内在规律,探索日常会话中话题线索的转换机理,将是后续研究中一个亟待解决的核心课题。

表 1 对话行为标注的整体统计数据(交际功能=核心功能/特定维度;DA 依存对=功能/反馈)

<sup>1</sup> 相关原始数据由博彦科技提供,特此表示感谢。

	工作	饮食	爱情	交通	天气	汇总
会话	100	100	100	100	100	500
话语片段	2460	2534	2405	2552	2501	12452
交际功能	2667/72	2730/18	2532/51	2715/51	2694/40	13338/232
话题线索	197	188	177	186	188	936
DA 依存对	372/785	389/803	413/676	349/860	300/893	1823/4017
连贯对	2263	2346	2229	2367	2313	11518

表 2 记录了核心功能及其相应的 DA 依存对的汇总统计数据。从中可以发现：

- 1) 在汉语日常会话中，5 大核心功能标记从大到小的分布比例分别为：阐述 S-32.57%，反应 R-30.15%，询问 Q-13.67%，回答 A-13.67%，动作 DO-9.94%。这与英文电话录音<sup>[13]</sup>和同步聊天库<sup>[15]</sup>中获取的统计数据分布基本一致。对其中使用最多的两类核心功能：阐述 S 和反应 R 的子类标记分析统计发现，客观描述(S-1)和主观评价(S-2)所占比例分别为：64% VS 36%；同意/接受(R-1)、中立/评价(R-3)和否定/拒绝(R-2)所占比例分别为：46% VS 41% VS 12%。而且不同话题的会话片段中的相关分布比例基本一致。如何有效区分核心功能子类，对后续 DA 标记自动识别提出了新的挑战。
- 2) 在 DA 依存对中占比 69%(4017/5840)左右的反馈依存对中，阐述-反应类(S-R)又占了绝大多数(82%以上)。它们与询问-回答类(Q-A)功能依存对共同形成了日常会话中最常用的交际模式。更深入的统计数据发现，Q-A 依存对大多出现在话题线索首部位置作为话题转换标记，而 S-R 依存对则主要出现在话题线索内部作为话题连接手段，而且不同子类的依存对在话题线索首部/内部的分布比例还有较大差异。如何有效利用 DA 依存对与话题线索之间的复杂关联特征，是后续会话结构自动分析中需要重点关注的一个技术难点。

表 2 核心功能和 DA 依存对的汇总统计数据

	工作	饮食	爱情	交通	天气	汇总
S	795	846	799	913	991	4344
R	786	806	677	860	893	4022
Q	372	389	413	349	300	1823
A	372	389	413	349	300	1823
DO	342	300	230	244	210	1326
S-R	604	628	566	713	791	3302
Q-A	372	389	413	349	300	1823
DO-R	181	175	110	147	102	715

基于双人独立标注数据，我们进一步分析了两个标注者针对不同 DA 标注项的宏一致率，得到表 3 的汇总数据。其中宏一致率的计算方法是：首先计算针对每个标注者的一致率，计算公式为：完全一致的标注项总数/该标注者的该类标注项总数，然后计算两个标注者一致率的算术平均值作为宏一致率。相关数据表明：

- 1) 核心功能标记的宏一致率达到了 91% 以上，相应的 DA 依存对宏一致率也达到 86% 左右。而特定维度标记的宏一致率则略低，约为 85% 左右。其主要原因是它们在总体 DA 标注中只占了 1.7% 左右，如果出现几个个人认识偏差而产生的标注差异情况，就会造成最终宏一致率数据的较大波动。总体而言，目前的 DA 标记集设计具有很好的可操作性，可以适应汉语日常自然对话中的行为功能模式描述需求。

- 2) 连贯依存对的宏一致率达到了 92%左右, 表明标注者能够很好地分析理解日常对话中相邻话语消息的内容连贯性。如何深入探索这些连贯依存对与 DA 依存对之间的信息互补性在对话理解中的重要作用, 是后续研究中需要关注的一个问题。
- 3) 话题线索的宏一致率只有 70%左右, 这从一个侧面反映了对日常会话中话题线索的把握具有较高的认知难度, 对后续的自动话题线索识别研究提出了新的挑战。

表 3 两个独立标注者的一致性汇总分析数据

	工作	饮食	爱情	交通	天气	汇总
核心功能	0.895218	0.909128	0.899278	0.929119	0.927096	0.912199
特定维度	0.775576	0.767361	0.88898	0.920368	0.872368	0.849282
DA 依存对	0.826226	0.868955	0.839303	0.874129	0.868245	0.855847
连贯对	0.872069	0.925731	0.925661	0.937077	0.926476	0.917563
话题线索	0.561224	0.705439	0.784468	0.777197	0.691489	0.701582

## 5. 结语

本文综合 DAMSL, DIT++, ISO 24617-2 等目前常用 DA 标注体系的描述精髓, 设计了一套针对汉语日常会话的对话行为标注体系。通过引入主客观阐述和正反向反应子类描述, 拓展了对不同交际功能的描写能力。同时引入话题线索分析机制, 有效组织会话中的话题变化趋势, 为后续处理网络会话中的话题纠缠问题打下了很好基础。在 500 个日常会话片段上进行的 DA 标注实验结果显示, 占比 98.3%的核心功能标记的双人独立标注宏一致率达到了 91%以上, 相应的 DA 标记对宏一致率也达到 86%左右。表明目前的 DA 标记集设计具有很好的可操作性, 可以适应汉语日常会话的行为功能模式描述需求。

在后续研究中, 我们将探索基于目前 DA 标注库的汉语日常会话对话行为和话题线索自动识别技术, 深入分析基于对话行为的会话结构模式与话题线索转换之间的复杂心理认知机理。并把相关 DA 标注体系扩展应用到微博、微信等网络会话语料上, 挖掘互联网话语中的会话结构描述特点, 逐步探索互联网话语的认知计算模型。

## 参考文献

- [1] Allen J, Core M. DAMSL: Dialogue Act Markup in Several Layers (Draft 2.1) [R]. University of Rochester, Rochester, NY. <http://www.cs.rochester.edu/research/cisd/resources/damsl/RevisedManual/>. (1997)
- [2] Bunt, H. Dialogue Pragmatics and Context Specification [A]. In: *Harry Bunt & William Black (eds.) Abduction, Belief and Context in Dialogue* [C]. Benjamins, Amsterdam, pp.81–150. (2000)
- [3] Bunt, H. Dimensions in Dialogue Act Annotation [A]. In *Proceedings Fifth International Conference on Language Resources and Systems (LREC 2006)* [C], Genova. (2006)
- [4] Bunt, H. The DIT++ taxonomy for functional dialogue markup [A]. In D. Heylen, C. Pelachaud, R. Catizone and D. Traum (eds.) *Proceedings of AAMAS 2009 Workshop "Towards a Standard Markup Language for Embodied Dialogue Acts"* [C], Budapest, pp.13-24. (2009a)
- [5] Bunt, H. Multifunctionality and multidimensional dialogue semantics [A]. In *Proceedings of DiaHolmia Workshop on the Semantics and Pragmatics of Dialogue* [C], Stockholm. (2009b)
- [6] Bunt, H. Multifunctionality in Dialogue [J]. *Computers Speech and Language*, 2011, 25: 222-245.
- [7] Bunt, H., J. Alexandersson, et. al. Towards an ISO standard for dialogue act annotation [A]. In *Proc. of the 7<sup>th</sup> International Conference on Language Resources and Systems (LREC 2010)* [C], Malta, pp. 2548–2558. (2010)
- [8] Bunt, H., Alexandersson, et. al. ISO 24617-2: A semantically-based standard for dialogue annotation [A]. In *LREC 2012* [C], pp. 430–437. (2012)

- [9] Clark, H. H. *Using language* [M]. New York: Cambridge University Press. (1996).
- [10] Dhillon, Rajdip, Bhagat, Sonali, Carvey, Hannah, & Shriberg, Elizabeth. Meeting Recorder Project: Dialog Act Labeling Guide [R]. Technical Report TR-04-002. ICSI, Berkeley, CA. (2004).
- [11] Fusaroli, R., Raczaszek-Leonardi, J., & Tylen, K. Dialog as interpersonal synergy [J]. *New Ideas in Psychology*, 2014, 32: 147–157.
- [12] ISO DIS 24617-2. Language resource management – Semantic annotation framework (SemAF) [S], Part 2: Dialogue acts. ISO, Geneva. (2010)
- [13] Jurafsky, D., E. Schriberg, and D. Biasca Switchboard SWBD-DAMSL Shall-Discourse-Function Annotation Coders Manual [R]. [http://stripe.colorado.edu/~sim\\$jurafsky/manual.august.html](http://stripe.colorado.edu/~sim$jurafsky/manual.august.html). (1997)
- [14] Joty, S., Carenini, G., and Lin, C.-Y. Unsupervised modeling of dialog acts in asynchronous conversations [A]. In *IJCAI Proceedings-International Joint Conference on Artificial Intelligence* [C]. (2011).
- [15] Kim, Su Nam, L. Cavedon, and T. Baldwin. Classifying Dialogue Acts in Multi-party Live Chats [A]. In *Proc. of 26<sup>th</sup> Pacific Asia Conference on Language, Information and Computation* [C]. pp. 463-472. (2012).
- [16] Levinson, Stephen C. *Pragmatics* [M]. Cambridge: Cambridge University Press. (1983).
- [17] LIRICS D4.4, Multilingual test suites for semantically annotated data [R]. <http://lirics.loria.fr>. (2007)
- [18] Petukhova, V. and H. Bunt. The independence of dimensions in multidimensional dialogue act annotation [A]. In *Proceedings 2009 NAACL-HLT Conference* [C], Boulder, Co. (2009)
- [19] Pickering, M. J., & Garrod, S. Toward a mechanistic psychology of dialogue [J]. *Behavioral and Brain Sciences*, 2004, 27: 169–190.
- [20] Popescu-Belis, Andrei. Dialogue act tagsets for meeting understanding: an abstraction based on the DAMSL, Switchboard and ICSI-MR tagsets [R]. IM2.MDM-09. University of Geneva. (2003)
- [21] Rashmi Prasad, Nikhil Dinesh, et. al. The Penn Discourse Treebank 2.0 [A]. In *Proc. of the 6th International Conference on Language Resources and Evaluation (LREC 2008)* [C], Marrakech, Morocco. (2008).
- [22] Searle, John R. *Speech Acts* [M]. Cambridge: Cambridge University Press. (1969).
- [23] Traum, D. & S. Larsson The Information State Approach to Dialogue Act Management [A]. In *J. van Kuppevelt & R. Smith (eds.) Current and new Directions in Discourse and Dialogue* [C], Kluwer, Dordrecht. (2003)
- [24] Xue Nianwen, Qishen Su and S. Jeong. Annotating the discourse and dialogue structure of SMS message conversations [A]. *Proc. of the 10<sup>th</sup> Linguistic Annotation Workshop* [C]. pp 180-187. (2016)
- [25] 王珊, 刘锐 谈话节目语料库的构建与会话结构分析 [J]. 《中文信息学报》, 2016, 30(6): 140-146.
- [26] 杨平 面向自然语言处理的会话行为分类研究 [J]. 《语言文字应用》, 2007年8月第3期, pp. 128-135.
- [27] Zarisheva, Elina and T. Scheffler Dialog Act Annotation for Twitter Conversations [A]. In *Proc. Of the SIGDIAL 2015 Conference* [C], Prague, Czech Republic. pp.114-123. (2015)
- [28] 周强. 汉语语篇的连贯性分析标注规范 Ver 2.4 [R]. 清华大学信息技术研究院语音和语言技术中心技术报告. (2016)
- [29] 周强. 汉语会话的对话行为标注规范 Ver 1.0 [R]. 清华大学信息技术研究院语音和语言技术中心技术报告. (2017)

作者联系方式：周强，清华大学信息技术大楼 FIT 1-303，邮编: 100084，电话: 62772432，电子邮箱: zq-lxd@mail.tsinghua.edu.cn