

语句边界韵律特征的话轮提示功能

熊子瑜

xiongzuyu@163.com

中国社会科学院语言研究所

摘要*

所谓话轮提示功能，就是指通过一定的手段来标明话轮进程状态，标示说话人的话语是即将结束还是未完待续。有效利用各种手段来标明话轮进程状态是话轮转换有序进行的必要保证。会话交际中存在几类能够标示话轮进程状态的因素，如上下文语境信息、语句边界韵律特征信息和句法语义信息等。

本文以电话录音材料为分析对象，利用 Praat 软件对 1190 个句子的话轮状态进行了标注。标注过程分三种条件进行：一是由一位专业的标音员在有上下文语境的条件下对每个句子的话轮状态进行听辨标注，二是由 5 位受试者在脱离语境的条件下对这些语句的话轮状态进行听辨标注，三是由 2 位具有一定语言学背景的受试者根据这些语句的文字材料对其话轮状态进行标注。在第一种条件下，标注者可以借助的信息包括语境、语音和句法语义三类；在第二种条件下，标注者可以借助的信息只有语音和句法语义两类；在第三种条件下，标注者可以借助的信息只有句法语义一类。然后我们利用 SPSS 软件对这三类标注结果进行了统计分析，以期在一定程度上了解这三类因素的话轮提示功能以及它们之间的交互作用，并在此基础上将语句边界韵律特征的话轮提示功能分离出来。

统计结果表明，在这三类因素中，句法语义因素是提示语句话轮进程状态的首要因素，所起的提示作用最大，语音次之，语境再次之。在识别语句的话轮状态时，仅依靠句法语义信息对纯文本孤立句能够达到 66% 的正确率，这表明某些句法语义特征具有一定的话轮提示功能；在有声条件下对离境化孤立句的听辨实验能够达到 84.5% 的正确率，与纯文本孤立句条件下的正确率相比提高了将近二十个百分点，这表明某些语音韵律特征有助于帮助准确判别语句的话轮进程状态，因此具有一定的话轮提示功能；从离境化孤立句的正确率为 84.5% 可以推导出上下文语境也同样具有一定的话轮提示功能。

此外，统计结果还表明，在识别语句的话轮进程状态时，不同句法语义类型的语句对语音韵律和语境信息的依赖程度不同，询问类的语句对语音韵律和语境信息的依赖程度相对较低，即便是没有语音韵律和语境信息的帮助也能获得较高的正确识别率，而叙述类的语句对语音韵律和语境信息的依赖程度相对较高，需要借助一定的语音韵律和语境信息才能正确判别其话轮进程状态。

一 话轮及话轮转换

有效利用各种手段来标明话轮进程状态是话轮转换有序进行的必要保证。语句边界韵律特征的话轮提示功能，就是指通过语句边界韵律特征来标明话轮进程状态，标示说话人的话语是即将结束还是未完待续。

话轮由一个或多个语句组成，根据语句在话轮中的位置可以将语句分做两类，终端语句和非终端语句，终端语句处于话轮结束位置，其后一般会发话轮转换，非终端语句处于非话轮结束位置，其后一般不发话轮转换。

语句类型和语句在话轮中的位置之间存在着一定的关联，如疑问句通常处于话轮结束位置，其后通常会发话轮转换。但语句类型对语句在话轮中的位置并没有强制的约束力，各种类型的语句既有可能处于话轮非结束位置，也有可能处于话轮结束位置。比如说疑问句尽管通常具有强制转移话轮的作用，但它在实际语流中仍然可能处于话轮中间，在其末尾处也可能不发话轮转换。例如：

例 [——取自 973 语料库/d007. wav]

01 B: [H-X-S]{大}{约}(LE .48)||—会|就到|好吗?|

02 A: [S-F]啊|(UC)啊|

03 A: [H-X-T]您(LE .36)|那您|大概|几点能到?

04 A: [H-X-S]八点半|以前|能到吗?|

05 B: [H-S-K]八点半|以前|我没吃饭呢||。

06 A: [H-T-J] (S-T 1.1)那您看|{您}{那}(LE .37)|什么|时候能到?|

07 A: [H-S-K]因为|我们|现在|房间|挺紧的|。

08 A: [H-S-K]回头|您在|不来了哈||。

第 03 行和第 06 行的语句在句法上都属于疑问句，疑问句通常具有强制转移话轮的作用，其末尾一般都会发话轮转换，但在此处，说话人没有给对方回答问题的机会，而是又追加一些话语，然后才将话轮转移给对方。在语流中，听话人虽然听明白了对方所提的问题，但也没有急于插话回答对方的问题，而是等着对方把话说完，这表明听话人当时完全明白对方还不准备结束话轮。那么是什么因素使得听话人明白了这一点，并因此克服了句法因素对话轮进程的提示作用？

二 话轮提示功能的听辨实验

会话交际中存在某些能够标示话轮进程状态的因素，如上下文语境信息、韵律特征信息和句法语义信息等，本文进行了一系列的听辨实验，期望在一定程度上了解这些因素的话轮提示功能以及它们之间的交互作用。

1 听辨方法

在判断语句是否处于话轮结束位置时，听音人可以借助多方面的信息，如上下文语境信息、韵律特征信息、句法语义信息等，为了在一定程度上考察这些信息在判断过程中所起的不同作用，我们的听辨实验按照三种方法进行。第一种方法是在有上下文语境信息条件下的听辨标注，在这种条件下听音人能够得到所有信息；第二种方法是孤立句条件下的听辨标注，在这种条件下听音人除了得不到上下文语境信息外，其他信息（包括韵律和句法语义等方面）都能够得到；第三种方法是纯文本条件下的判断标注，在这种条件下听音人既不能得到上下文信息，也不能得到语音韵律信息，只能得到句法语义信息。

2 听辨过程

听辨过程按照上述的三种方法分别进行，下面逐一加以简要说明：

1) 上下文语境中的听辨标注

由一位专职标音人员根据原始录音材料对 973 自然口语语料中的话轮、语句以及话轮转换方式等情况进行了标

* 本文宣读于第十届全国当代语言学研讨会，2004 年 12 月于重庆。

注。她根据实际录音材料，将话轮和语句从实际语流中切分出来，并标明语句在话轮中的位置，如果其末尾发生了话轮转换则标为“1”，表示该语句处于结束位置，如果其末尾没有发生话轮转换则标为“0”，表示该语句处于非结束位置。话轮转换的方式分为正常转换、提前转换和延迟转换三种类型。在话轮延迟转换中，尽管有时没有发生真正的话轮转换，但因其实际上出现了说话人要转移话轮的意图，因此也把它看作发生了话轮转换。在话轮提前转换中，因发生了话轮叠接，很难提取正确的基频、能量、时长等韵律特征数据，因此在实际分析中这些话轮叠接暂且不考虑。

在有语境的听辨条件下，标注人员可以了解交际双方有关话轮转换的全部信息，因此所标注的数据应该能够比较真实地反映语句在话轮中的实际位置。通过这一标注，可以获得语句在话轮中的实际位置数据。

2) 离境化孤立句的听辨标注

为了考察语境对话轮转换过程的提示作用，并尽可能分离出语境的作用，本研究将一个个语句从语流中提取出来，进行了孤立句听辨标注，以排除语境因素对听辨判断的影响。在 973 语料中，根据标注信息提取出所有语句，然后将音节数目小于 4 或大于 14 的语句全部剔除，保留音节数目在 4 个到 14 个之间的语句，再在其中剔除那些基频数据不完整和/或音段标注不明确的语句（这样做的目的在于保证后面韵律分析数据的有效），这样还剩下 1190 个语句。在这些单元中，有彼此连贯的、有不连贯的，有男声、有女声，有服务员的话语、有客人的话语，有询问、有回答，有语义结构比较完整的、也有不太完整的，整个说来，基本反映了实际的交际状况，没有针对语料内容和形式做专门的筛选。

为了尽可能地消除语境对听辨判断的干扰，在听辨时，还将这些语句按照随机原则排序，链接到 HTML 格式的文件上，并生成一份标明文件顺序的 EXCEL 表单，在听辨时，听音人只要双击 HTML 文件上的声音图标，就会自动打开所链接的声音，然后由耳机输入到听音人耳朵中，对同一个声音听辨的次数不做限制，只要听音人自认为能够做出判断为止，听音人做出合适的判断后可以在连带的 EXCEL 表单上做记录。

实验由 5 位听音人进行，这些听音人都是汉族，听力正常，能说比较标准的普通话。听辨时，听音人做如下判断：在所听到的语句之后，是否会发生话轮转换（即换一个说话人）。如果认为会发生话轮转换就在 EXCEL 表单的相应声音文件名之后打 1 分，否则打 0 分。在开始听辨之前，对这 5 位听音人进行了简短的训练。由于听辨的内容很多，所需时间较长，因此听辨过程分开进行，将听辨材料分发给听音人，让他们自己掌握听辨进度，为了尽可能避免听音人因困乏造成精力不集中而引起误判，整个听辨过程可以根据听音人的喜好分做多次进行。在听辨结束后，统一收回听音人的听辨结果。

为了和实际位置进行比较，将 5 位听音人的听辨结果进行了归总处理，同一个项目上的 5 位听音人的分值累加。每个语句都有 5 个得分（分别为 0 或 1），加在一起有 6 种可能性：0-5，若总得分为 0-2 的则表明多数听音人认为其后不会发生话轮转换，就归为非结束位置 0，若得分在 3-5 的则表明多数听音人认为其后会发生话轮转换，就归为结束位置 1。

通过离境化孤立句的听辨实验，可以得到语句在离境化情况下的听辨位置数据，把这些数据同有语境条件下的听辨实验数据进行比较，可以在一定程度上概括出上下文语境对话轮进程状态的提示作用，另外还可以用来分析离境化孤立句听辨结果的效度如何。

3) 纯文本孤立句的判别标注

为了在一定程度上分离句法和韵律对听辨结果的相对

贡献率，由两位具有一定语言学知识的学生在无语音的情况下仅根据文本所提供的信息对语句在话轮中所处位置做出判断，判断分 3 个等级：一般只能处于话轮结束的语句得 1 分；一般只能处于非结束位置的语句得 0 分；既可能处于话轮结束又可能处于话轮非结束位置的语句得 2 分。因为分析人员较少，为了在一定程度上排除因分析人员的个体差异对判断结果所造成的影响，在此没有强制做非此即彼的二分判断。

为了和上两种听辨结果进行比较，对两位分析人员的标注数据进行了归总处理，如果两个分析人都认为只能处于非结束位置的，则计为非结束位置，得分为 0，如果两个分析人都认为只能处于结束位置的，则计为结束位置，得分为 1，否则认为是无法判断（既可处于结束位置又可能处于非结束位置），得分为 2。

通过纯文本孤立句的判别实验，可以得到语句在纯文本孤立句条件下的判别位置数据，把这个数据同离境化孤立句的听辨实验数据进行比较，可以在一定程度上概括出语音韵律特征对话轮进程状态的提示作用，另外还可以用来分析语句的句法语义信息对话轮进程状态的提示作用。

三 结果分析

经过听辨实验后，得到三种条件下的判别结果数据，下面将这些数据放在一起进行比较分析，以考察这几种因素的话轮提示功能。

1 上下文语境中的听辨结果分析

根据在有上下文语境条件下听辨标注的结果统计出这些语句在话轮中的分布情况，数据见下表：

	频数	比率
非结束位置	324	27.2%
结束位置	866	72.8%
合计	1190	100.0%

从中可以看出，在这选取的 1190 个语句中，72.8% 的处于话轮结束位置，其后发生了话轮转换，只有 27.2% 的处于非话轮结束位置，其后不发生话轮转换。通过这些数据还可以看出，在电话语料中的话轮通常简短，绝大部分只含有一个语句，这使得大部分语句处于话轮结束位置，其后发生话轮转换。

本文将这 1190 个语句的言语行为类型归结三大类型：叙述、命令和询问。下面按照语句的言语行为类型来统计了它们在话轮中的分布位置：

句子类型 * 实际位置 Crosstabulation

		实际位置		合计	
		非结束位置	结束位置		
句子类型	陈述句	频数	249	394	643
		比率	38.7%	61.3%	100.0%
	祈使句	频数	12	36	48
		比率	25.0%	75.0%	100.0%
	疑问句	频数	63	436	499
		比率	12.6%	87.4%	100.0%
合计		频数	324	866	1190
		比率	27.2%	72.8%	100.0%

从上表数据可以看出不同言语行为类型的语句在话轮中的分布位置存在明显的差异。其中疑问句绝大部分处于话轮结束位置，比率为 87.4%，也就是说 87.4% 的疑问句之后都发生了话轮转换，比 72.8% 的平均水平高出了 15 个百分点；而陈述句处于话轮结束位置上的几率只有 61.3%，也就是说只有 61.3% 的陈述句之后发生了话轮转换，比 72.8% 的平均水平低了 10 个百分点。从这一点看，不同言语行为类型的语

句在话轮中所处位置具有不同的倾向性。

2 离境化孤立句的听辨结果分析

下面根据离境化孤立句条件下 5 位听音人的听辨数据进行统计,首先将离境化孤立句条件下的听辨结果和有上下文语境条件下的听辨结果进行比较。从下表数据可以看出,在 324 个非结束语句中,有 258 个被正确感知为非结束位置,正确率为 79.6%;在 866 个结束语句中,有 748 个被正确感知为结束位置,正确率为 86.4%。两项综合起来,共有 1006 个语句的位置被正确听辨出来,听辨正确率平均为 84.5%。这表明,在离境化的情况下听音人根据语句自身的各方面信息(语音、语义、句法等)可以对其在话轮中所处的位置做出正确率为 84.5%的判断,判断错误率为 15.5%。这表明,上下文语境对人们感知语句在话轮中所处的位置来说具有重要的作用。

实际位置 * 听辨位置 Crosstabulation

		听辨位置		合计	
		非结束位置	结束位置		
实际位置	非结束位置	频数	258	66	324
		比率	79.6%	20.4%	100.0%
	结束位置	频数	118	748	866
		比率	13.6%	86.4%	100.0%
合计		频数	376	814	1190
		比率	31.6%	68.4%	100.0%

下面按照语句的言语行为类型分别统计了它们在听辨位置上的分布概率,以考察语句的言语行为类型对听音人听辨判断的影响,数据如下:

句子类型 * 听辨位置 Crosstabulation

		听辨位置		合计	
		非结束位置	结束位置		
句子类型	陈述句	频数	296	347	643
		比率	46.0%	54.0%	100.0%
	祈使句	频数	14	34	48
	比率	29.2%	70.8%	100.0%	
疑问句	频数	66	433	499	
	比率	13.2%	86.8%	100.0%	
合计		频数	376	814	1190
		比率	31.6%	68.4%	100.0%

从上表数据可以看出在离境化孤立句条件下,不同言语行为类型的语句在听辨位置上存在明显的差异。有 86.8% 的疑问句被听辨为处于话轮结束位置,有 70.8% 的祈使句被听辨为处于话轮结束位置,而只有 54.0% 的陈述句被听辨为处于话轮结束位置。从这一点看,在离境化孤立句条件下,对不同言语行为类型的语句在话轮中所处位置的听辨判断,听音人具有不同的倾向性,听音人倾向于认为疑问句和祈使句之后更可能会发生话轮转换,而陈述句之后发生话轮转换的概率则相对要低一些。这表明在听音人看来,语句的言语行为类型与其在话轮中的所处位置之间存在着一定的关联。

从前面的统计结果可以知道,与实际位置相比,有 184 个语句的听辨位置是判断错误的,占语句总数的 15.5%。下面分析了听辨判断错误和语句的言语行为类型之间的关系,以考察哪种类型的语句更容易正确判断,数据如下:

句子类型 * 听辨位置正确率 Crosstabulation

		听辨位置正确率		合计	
		判断正确	判断错误		
句子类型	陈述句	频数	524	119	643
		比率	81.5%	18.5%	100.0%
	祈使句	频数	38	10	48
	比率	79.2%	20.8%	100.0%	
疑问句	频数	444	55	499	
	比率	89.0%	11.0%	100.0%	
合计		频数	1006	184	1190
		比率	84.5%	15.5%	100.0%

数据表明,在这 184 个听辨位置判断错误的语句中,属

于陈述句的有 119 例,占陈述句总数的 18.5%;属于祈使句的有 10 例,占祈使句总数的 20.8%;属于疑问句的有 55 例,占疑问句总数的 11.0%。这表明在离境化孤立句条件下,疑问句被错判的概率相对较低。也就是说在识别一个语句是否处于话轮的结束位置时,不同类型的语句对语境的依赖程度不同,疑问句依赖语境的程度相对较低,在离境化孤立句条件下通常也能取得较高的正确识别率,而陈述句依赖语境的程度则相对较高,在离境化孤立句条件下的识别正确率会明显下降。

3 纯文本孤立句的判别结果分析

下面根据纯文本孤立句条件下 2 位分析人员的判别数据进行统计,首先将纯文本孤立句条件下的听辨结果和有上下文语境条件下的听辨结果进行比较,数据如下:

文本位置 * 实际位置 Crosstabulation

		实际位置		合计	
		非结束位置	结束位置		
文本位置	无法判断	频数	225	456	681
		比率	33.0%	67.0%	100.0%
	结束位置	频数	57	402	459
		比率	12.4%	87.6%	100.0%
非结束位置	频数	42	8	50	
	比率	84.0%	16.0%	100.0%	
合计		频数	324	866	1190
		比率	27.2%	72.8%	100.0%

从上面数据可以看出,在分析人员做出结束位置和非结束位置的两项判断中,正确率都相当高,分别是 87.6% 和 84.0%,这表明确实存在某些比较明显的句法和语义信息帮助分析人员做出准确判断。但分析人员能做出这两种判断的整个比例却很低,合计 509 个,占总语句数目的 40% 左右,还有将近 60% 的语句,仅依据自身的句法语义信息无法判断它在话轮中的所处位置。为了进一步考察哪些句法语义特征能够帮助分析人员做出准确判断,下面考察语句在句法和语义上的特征与其判别结果之间的关系,数据如下:

句子类型 * 文本位置 Crosstabulation

		文本位置			合计	
		无法判断	结束位置	非结束位置		
句子类型	陈述句	频数	581	13	49	643
		比率	90.4%	2.0%	7.6%	100.0%
	祈使句	频数	28	20		48
	比率	58.3%	41.7%		100.0%	
疑问句	频数	72	426	1	499	
	比率	14.4%	85.4%	.2%	100.0%	
合计		频数	681	459	50	1190
		比率	57.2%	38.6%	4.2%	100.0%

结果表明,大部分疑问句都被标记为处于话轮的末尾位置,也就是说在分析人员看来疑问句之后通常会发生话轮转换,疑问句具有强制转移话轮的作用。还有一些要求对方提供姓名、电话等身份信息的祈使句,如“告我一下客人的名字”、“留下您名字电话”之类,分析人员也认为在其后通常会发生话轮转换。另外某些表示假设关系的语句,如“您要是现在订的话”、“您要是住的时间很长的话”之类表述上不够完整的叙述性语段几乎都被认为不会处于话轮末尾位置。还有一些属于话头性质的叙述性成分,如“不如此吧”、“另外的房间的话”、“我想想啊”之类,也被认为不会出现在话轮末尾位置。除此之外,还有某些语义上不足的成分,如“我想问一下那个”、“我想打听一下”,也被认为不会处于话轮的末尾位置。这表明在分析人员看来,不同句法语义类型的语句在话轮中所处的位置确实具有一定的倾向性,这也说明确实存在某些句法语义信息,在实际过程中起到标示语句在话轮中所处位置的作用。

虽然具有某些句法语义特征的语句(如询问类)在话轮中的位置表现出了一定的倾向性,但绝大部分语句(如叙述类)在话轮中的位置仅依据其句法和/或语义信息是很难完全确定的。因此尽管句法语义信息在判断语句所处的位置起着重要的标示作用,但总得来看,其作用是有限的。如果把

“无法判断”部分完全当作识别错误来处理的话，那么依据文本信息预测的综合正确率只有37.3%，这表明仅依据句法和语义信息来判定该语句应处位置的正确几率是很低的。另外，如果把“无法判断”部分按照随机分布正确和错误几率强制判断其所处位置的话，则综合正确率也只能上升到66%左右。这与有语音信息条件下的正确率84.5%相比，相差近20个百分点，这表明语音韵律特征在识别语句所处位置时起着重要的作用，能够在一定程度上消除模棱两可的情况，使得听音人对大多数语句的判断具有很明显的倾向性。

四 结论

通过以上分析，可以看出影响准确识别语句在话轮中所处位置的因素可以分为句法、语义、语音以及语境等几个方面，在这几个因素中，句法和/或语义因素是基础，所起的作用最大，语音次之，再其次是语境。另外，对于具有不同句法语义特征的语句来说，语音和语境所起的提示作用是不同的，如疑问句的分布位置具有明显的倾向性，大多处于话轮结束位置，因此即便没有语音和语境信息也可以获得较高的正确识别率，而陈述句的分布位置没有明显的倾向性，既可处于结束位置也可处于非结束位置，因此在语流中要想准确识别其位置，只能更多的依赖于语境信息和语音信息。

通过以上分析，可以得出以下结论：

1. 在识别语句所处位置时，纯文本孤立句条件下能够达到66%的正确识别率，这表明某些句法语义特征具有重要的话轮提示功能；离境化孤立句的听辨实验中达到了84.5%的正确识别率，与纯文本孤立句条件下的判别准确率相比提高了将近二十个百分点，这表明某些语音韵律特征具有重要的话轮提示功能；从离境化的正确识别率只有84.5%可以推导出上下文语境也具有一定的话轮提示功能。

2. 在识别语句所处位置时，不同句法语义类型的语句对韵律和语境的依赖程度不同，询问类的语句对韵律和语境的依赖程度低，即便是没有韵律和语境信息的帮助也能获得较高的正确识别率，而叙述类的语句对韵律和语境的依赖程度高，没有韵律和语境信息的帮助很难获得较高的正确识别率。

参考文献 略

The cuing function of prosodic features of sentence boundary in turn-taking

XIONG Ziyu

xiongziyu@163.com

Institute of linguistics, CASS

Abstract

To smooth the process of turn-taking, it is necessary to use some means for cuing the turn-state of an underway sentence. One can use many indicators in a conversation such as context, prosodic features, syntactic and semantic, expression and gestures.

In this paper, we used PRAAT program to label the turn-state of 1190 sentences, which were extracted from a phone-conversation corpus. The labeling process is as following: an object firstly labeled the turn-states of these sentences by listening in a original context; secondly five objects labeled them by listening out of the original context; thirdly two objects labeled the turn-states of these sentences only by text and out of the original context. In the first condition, the object could utilize the context information, the prosodic information, and the syntactic and semantic information to distinguish the turn-states of these sentences. In the second condition, the objects could utilize only the prosodic information and the syntactic and semantic information. In the third condition, the objects could utilize only the syntactic and semantic information. Based on the results of labeling, we used SPSS program to analyze the functions of prosodic features in cuing the turn-state of the sentences.

The statistic result shows that the importance sequence to cue turn-state is the syntactic and semantic information, the phonetic information, and the context information. When discriminating the turn-state of a sentence without sound and out of the original context, the correct rate is about 66%, showing that the syntactic and semantic information has some function in cuing turn-taking. The correct rate for sentences out of context is 84.5% in the perception experiment. The improvement of correct rate tells that some phonetic information helps identify turn-state so that it has some function in cuing turn-taking. This result can also show the function of context information in cuing turn-taking.

Moreover, the result shows that the different types of syntax and semantics tie to the phonetic information and the context information in a different degree when identifying turn-state. Sentences for inquiry can get a high correct rate though there is no phonetic information and context information. But sentences for statement need them to identify turn-state correctly.