

汉语疑问句句重音感知初探

柳雪飞¹, 李雅², 陶建华², 许小颖³

- (1. 中国社会科学院 语言研究所, 北京 100732;
2. 中国科学院 自动化研究所, 模式识别国家重点实验室, 北京 100190;
3. 北京师范大学 文学院, 普通话水平培训测试研究中心, 北京 100875)

摘要: 本文通过实验研究的方法探讨了疑问句的常规句重音分布情况, 通过感知标注实验, 对疑问句句重音与句法结构之间的关系进行研究, 初步总结出疑问句句重音的分布规则, 并将是非疑问句句重音的位置与陈述句句重音位置进行了对比, 初步探讨了疑问句句重音自身的特点, 研究结果有助于进一步拓展语音合成等工作的思路。

关键词: 句重音; 句法; 疑问句;

Abstract: The paper discusses the default stress distribution of the interrogatives via syntax structure, and investigates the relationship between the stress and the syntax structure by a perception experiment. The ultimate goal of this research is generating the default stress distribution of interrogative sentences in order to offer some elicitation to the TTS system.

Key Words: stress; syntax; interrogative sentences

1 引言

汉语重音研究一直是语言学界和言语工程学界研究的重点, 也一直是制约语音合成技术进一步发展的关键因素。而且, 由于汉语是声调语言, 重音的感知受到声调、语调等多种因素的共同影响, 这使得汉语重音的研究成为语音分析和言语工程的难点。针对陈述语气的汉语重音研究已经有一定的进展[6][7][8][9], 在汉语疑问句的重音研究方面, 虽然也有一些研究成果[18], 但总体上还需要进一步的深入探索。从语言学的角度看, 汉语疑问句句重音的研究是实现汉语韵律与句法结构、信息结构接口的重要方面, 通过对语句重音指派规则及预测的研究可以将语音、语法以及语用有机地联系起来; 从言语工程的角度来看, 加强对疑问句句重音的研究将有助于扩大语音合成、口语人机交互的应用范围。

针对汉语疑问句句重音问题, 本文试图从语法的角度对疑问句句重音与句法结构之间的关系进行探讨, 以期对疑问句句重音在言语工程上的使用提供一些思路。由于句重音的位置可能因不同的人或不同的上下文环境而异, 因此, 为了简化工作, 文章主要探讨疑问句中常规句重音的位置, 其它情况则不再讨论。

2 前人研究

关于句法——重音的关系已有不少学者从不同的角度进行了不同程度的研究, 一般来说, 可以分为两种主要观点: 一种是以乔姆斯基(Chomsky)和赫勒(Halle)[1]为代表的纯句法重音观。他们认为, 一旦确定了句法结构, 重音的指派就是完全自动的。也就是说, 我们完全可以根据句法结构来预测语句重音的位置, 持这一观点的研究包括: 利博曼(Liberman)和普润斯(Prince)[2]、王丹、程宗军、郑波等[3], 袁野[4], 杨文昌[5], 董洪杰[6], 王丹和杨玉芳[7], 王韞佳、初敏和贺琳[8]等。另一种是以Bolinger[11]为代表的纯语用的重音观。他们认为句子的重音不能由句法或词法来解释, 它与语义或语用有着更大的关系, 是语义和情绪上的凸显, 持这种观点的学者不多, 主要有Anne Culter[12]和曹剑芬[14]等。

通过前面论述可以看出, 多数学者认为可以从句法结构来预测重音位置的。本人经过分析认为, 虽然句法结构和重音位置并不总是完全一致的, 但是, 它们确实有着比较密切的关系。有的学者认为, 句法重音与词重音和短语重音在总体上是基本一致的, 可以从句法结构关系推测出句重音的位置, 如, 主谓式结构中一般是谓语重, 述宾结构中, 宾语重, 述补

结构中，补语重，偏正式结构中，一般是偏重[4、5、8]。但是，句重音与词重音和短语重音之间又存在着一些复杂的情况，并不总是一致的，比如，偏正式合成词在短语层面，一般是偏重，而在语句层面却不完全是偏重[5、8]，句重音有时落在偏部，有时落在正部。因此，传统重音理论所强调的偏正结构“偏重”的重音分布规律实际上是一个短语层面的规律。

也有少数学者对汉语陈述句的语句重音与句法结构的关系进行了研究，他们认为汉语语句重音的分布和表现形式主要有以下几个特点：(1) 语句重音具有明显的后置倾向，主谓结构重读谓语中的主要动词，在有宾语的句子中倾向于落在宾语部分；(2) 在宾语部分，句重音倾向于落在定语上；(3) 句子谓语较复杂，句重音则倾向于落在表性状和程度的状语或者表状态和程度的补语上[5、8、10、13、15]。黄伯荣、廖序东[15]进一步提出，语句中表疑问和指示的代词也可以重读，王丹、程宗军[2]等虽然也认为句法成分位置越靠后，被重音知觉的程度就越大，但是，他们与上述学者的观点略有所不同，他们认为状中结构的正部比偏部的重音知觉程度更大，并列结构中的后部的重音知觉程度更大；定中结构也是正部比偏部重音知觉大。

由此可以看出，国内虽然有不少学者对语句的句法结构与重音之间的关系进行了研究，但他们之间的观点存在着一定的分歧，并不是完全一致的，而且，他们研究的大多是词重音、短语重音，涉及到句重音也仅仅局限于对陈述句的研究，疑问句几乎是一个空白。日本学者 Kitagawa Y.[16]对日语中特殊疑问句与韵律之间的关系进行了研究。他认为，在理解和分析特殊疑问句时，应当将韵律和句法结合起来考虑。李爱军、王海波[17]从声学的角度分析了现代汉语中陈述句和疑问句的句重音，但是，他们的研究都没有涉及到疑问句中句法结构与句重音之间的关系问题。

本文以大规模语料库为基础，旨在研究疑问句句法结构与句重音之间的关系，为口语对话系统的进一步研究提供一些有益的线索。本文的研究思路是：首先对疑问句句语料进行感知标注，然后按照四大句型（是非疑问句、特指疑问句、选择疑问句以及正反疑问句）进行分类总结，归纳出疑问句句重音的规则，找出语法结构与句重音之间的关系，并将疑问句句重音位置与陈述句进行对比，总结出疑问句句重音自身的特点。

以下论述分为三个部分：第三部分：简单介绍一下疑问句句法结构到句重音的实验研究及实验结果；第四部分：对实验结果进行分析，总结出疑问句句重音的指派规则，并与陈述句进行对比；第五部分：结语以及今后的工作。

3. 基于感知的重音标注实验及结果

3.1 实验过程

为了可以得到更准确的疑问句句法结构与句重音之间的关系，从而提高合成语音的质量，我们进行了一次重音感知的标注实验。根据实验目的，本文共选择了 400 句疑问句作为测试文本，其中，是非疑问句大约有 116 句；特指问句大约有 234 句；选择问句大约有 40 句，正反问句大约有 10 句。各类疑问句所占的比例如图 1 所示：

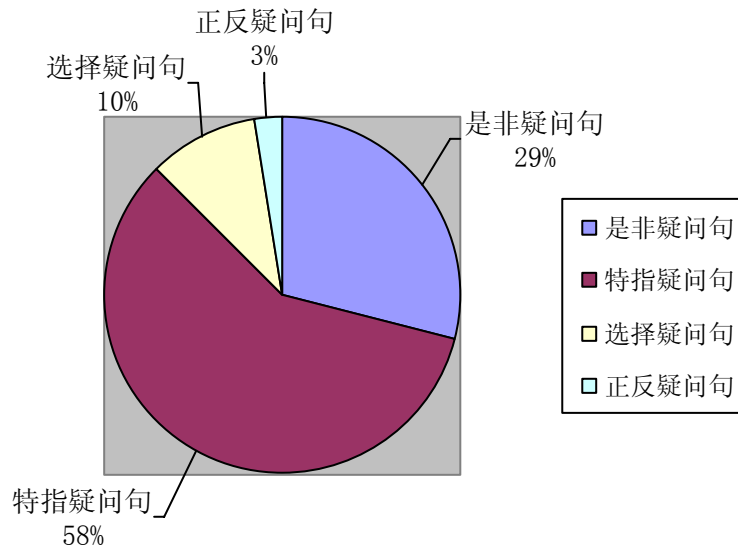


图 1: 实验中疑问句的类型分布

语料选取完毕之后, 我们首先对这些疑问句进行了以词为单位的切分, 然后进行多人感知的标注实验。在这次实验中, 我们共邀请了 20 位标注人进行标注, 这 20 位标注人都是北京师范大学的在读研究生, 所学的专业都是汉语言文字学。他们平时都讲流利的普通话。由于他们的知识水平都比较高, 而且, 参与标注的人数也比较多, 因此, 我们认为这次标注的结果能够反映一定的规律。

考虑到他们自身的语言知识积累会对标注结果产生影响, 因此, 我们要求他们尽量以自然发音风格读出, 而且, 在标注过程中不能参考他人的标注结果, 也不能进行讨论, 以第一语感为准。由此, 在实验过程中, 这 20 位标注人主要根据自身的感知进行判断, 选出句子中认为需要重读的地方, 并记录下来。由于疑问句一般都较短, 因此一句话中被感知为重音的个数一般不超过两个 (较长的句子可以有两处重音), 并对每个句子的重音位置进行统计。

3.2 实验结果

这次实验的结果比较理想, 400 句疑问句中标注结果基本一致的有 348 句, 占整个语料的 71% 左右, 这说明标注的一致性还是比较高的。具体来说, 特指疑问句的一致性最高, 20 人都将特指疑问词标为句重音; 其次为选择疑问句, 20 人中有 18 个人将句重音标在前一选择肢上; 再次为是非疑问句, 由于是非疑问句情况比较复杂, 所以一致性相对前二者较差, 但是较为简单的句式, 一致性还是比较高的; 最后是正反疑问句, 只有一句 20 人的标注结果一致。

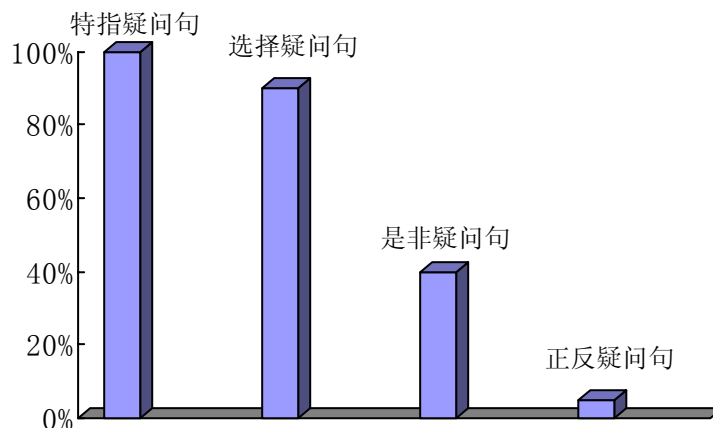


图 2: 各类句子标注一致性比较

我们对这次的重音感知结果进行了更细致的分析，得到了更具体的疑问句重音指派规则。具体如第 4 节所述。

4 句法到重音的指派规则及原因

4.1 是非疑问句

是非疑问句在整个语料库中所占的比例比较大，大约占整个语料库的 31%。从整个语料库的构成上来看，是非疑问句的情况最为复杂，具体可以分为以下几类进行讨论：

(1) 结构简单的主谓句或主谓宾句，重音一般落在谓语或宾语上。如，*你相信吗？我能看见海吗？* 分别重读“相信”和“海”。

按照信息结构的理论，句法位置越靠后的成分，负载的信息量越大，越容易被重读。

(2) 结构复杂的主谓宾句，即宾语前有较长的定语的情况

在这种情况下，重读一般要分两种情况：

a) 重读定语或定语的某一部分，此时的定语一般是刻划中心语的内涵的，如，*你具备侦探才能吗？* 重读“侦探”

b) 重读中心语，这时的定语一般为人称代词、数词以及描写事物外延的词，如，*有谁看见我的帽子吗？对不起，我能为您买一份饮料吗？你还有读者文摘吗？* 分别重读“帽子”、“饮料”和“文摘”。

两类句子重读位置的不同可以归结为定语类型的不同，前一类的定语属于“确定指称” (specifying a reference) 的一类；后一类的定语属于“刻划概念” (characterizing a concept) 的一类[19]。确定指称的修饰成分关涉中心语的外延，用来确定中心语所指为何物、与何相联系。其修饰语往往是中心语以外的事物，包括表领属的成分、冠词等。因此，这类的修饰语所含的信息量较低，一般不需重读；刻划概念的修饰语则定义、深化中心语的内涵，更加细致地呈现中心语某方面的性质，其修饰语往往是中心语的某方面性质。因此，这类修饰语的信息含量较高，一般需要重读。

(3) 若句子的宾语前出现多重定语，一般重读离中心语较远的那个定语（代词作定语除外）。如，*你想把别人的最后一件衬衫也剥掉吗？* 重读“最后”。从信息的角度来讲，离中心语越远的词，其区别性越大，包含的信息量也就越多，因此，也就越容易突出重读。

(4) 若句子的宾语较简单，在谓语前出现表示时间、程度、强调的状语时，一般需要重读该状语。如，*成功的代价一定是神经衰弱吗？您说今夜就去吗？你的父亲认为一生中最大的成就是什么？* 分别重读“一定”、“今夜”、“最大”。谓语较复杂，包含的信息量也就越多，因此，重音一般会落在谓语的某一成分上，由于状语是修饰突出谓语的，因此与谓语相比，状语包含的信息量相对较多，故重音一般落在状语上。

(5) 兼语句中，重音一般落在兼语上。如，由**国华**来完成吗？让**树庆**去种菜吗？重音分别落在“国华”和“树庆”上。

对于兼语句来说，除去兼语外的其他成分对于说话双方来说是已知的，是旧信息；兼语部分对说话双方来说则是信息，是信息焦点，故需要重读。

4.2 特指疑问句

特指疑问句的情况比较简单，20位标注人都将语句重音标在特指疑问词上。而且特指疑问句在整个语料库中所占的比例最大（占57%左右），标注的一致性也最好，如，**你**和**什么人**住在一起呢？你**匆匆忙忙**到**哪儿**去？他**为什么**那么穷？重音分别落在“什么”、“哪儿”和“为什么”上。若特指疑问词并不表示疑问，语句重音则落在句子的其它成分上。由于重音一般是提示信息焦点的，因此，信息焦点所在的位置一般就是重音所在的位置。大多数学者认为，特指疑问句中的特指疑问词是问话人最想知道的部分，因此，自然而然就成为语句的信息焦点[22]，因此，特指疑问句中的语句重音一般落在表示信息焦点的特指疑问词上。

虽然20位标注人都将句重音标在特指疑问词上，但在自然口语对话中，实际的句重音并非都落在特指疑问词上，而是落在句子的其它成分上，只有说话人在表示强调或惊奇的一些情况下，才会重读特指疑问词。20位标注人之所以将语句重音标在特指疑问词上，是因为他们都是语言学专业的学生，在标注过程中会无意识地运用一些语言学的知识来判断。但是，从可接受度上来看，将句重音标在特指疑问词上并非是不可接受的，而且，从计算机语音合成的角度来看，我们应以语法标注为主。对于特殊疑问句句重音为何落在句子的其他地方，我们会在今后的工作中加以研究和讨论。

4.3 选择疑问句

选择疑问句占整个语料库的8%，标注的一致性也比较好，20位标注人大多倾向于将语句重音标在前一个选择肢上。如，是**欧洲**时间呢还是**当地时间**？是**宽恕**了女儿还是**盛怒**未息？**上海**好还是**北京**好？重音分别落在“欧洲”、“宽恕”和“上海”上。

这种现象可以从说话者的心理倾向上来分析。说话者在让听话者做出选择时，一般是先说出自己期望得到的回答，然后再说出另一个选择项，让听话者选择，即，说话者已有一个心理预期。从实现性上来看，前一选择肢实现的可能性比后一选择肢要大，是说话者有意强调的内容，故一般接受重音。

4.4 正反疑问句

语料中正反疑问句的数量不多，只占整个语料库的3%，而且一致性也比较差。这一方面是由于切分词的标准的不统一，有的是按照语法词切分的，有的是按照韵律词切分的，导致正反疑问词“是不是”、“能不能”、“有没有”等在有的句子中被切分开来，在有的句子中则作为一个标注整体；另一方面是由于正反疑问句的句子大都较长，如“极美的女人是不是都不会说话呢？”“你是不是觉得我这架小飞机好看呀？”句子较长结构较复杂，标注的一致性就会较差。由于重音的标注主要是根据自己的语感来确定的，主观性很强，句子较长结构较复杂，每个人感知的重点就会不一样，相应地一致性也会较差。

4.5 陈述句与疑问句句重音对比

我们在前期的标注工作中，初步研究了陈述句的句重音位置，发现陈述句的句重音实际上多数集中在句首词语上[23]，这与多数学者认为的陈述句的语句重音具有明显的后置倾向有较大的分歧，而且，从我们的标注经验看，句重音与呼吸句群的分布更相关一些，每个句子中各个呼吸句群起点的韵律词通常是换气的开始，因此，这些韵律词容易发音突显，获得重音。以呼吸句群为单位，各个呼吸句群内的重音分布与上述学者总结的规律比较吻合。

由于是非疑问句是在原来陈述句的基础上加上“吗”、“吧”、“啊”等疑问语气词形成的或是直接加上疑问语调形成的，这类疑问句与陈述句的关系最为密切，加强对这类疑问句的研究也具有十分重要的意义。通过与陈述句的对比，我们发现此类疑问句句重音的位置与

陈述句句重音的位置差别不是很大，主要的区别集中在以下几个方面：

(1) 通过对陈述句句料标注结果的观察，我们发现大多数陈述句的句重音落在句首词语上，句首获得重音的概率要明显高于其他位置，这个特点与句子长度无关。这类句子大约占整个标注语料的 57%，其它的句重音位置出现在句子的各个成分上，相对地，句末词一般不获重音，没有明显的后置倾向。而疑问句的句重音则有明显的后置倾向，占整个标注一致的语料的 58% 左右，而且在标注一致的语料中，特殊疑问句所占的比例很大（占 57% 左右），而特指疑问词多出现在句首和句中。

我们认为，陈述句句首的韵律词一般处于较长的韵律短语中，在开始讲话时，一般会获得充足的能量和较高的基频，相对地，就会获得语句重音。而疑问句的语气则比较舒缓，加之句末语调的上扬，使得疑问句后半部分的音域整体扩大，因此，句重音多落在句末位置。这里所指的句首、句末不一定指句子的第一个韵律词或最后一个韵律词，也有可能是第二个韵律词。

(2) 关于述补结构，在陈述句中，有的学者倾向于补语重[2、5、13、15]，而王韞佳、初敏、贺琳[3，9]则认为是等重或述语重，但是，在疑问句中并没有表现出这种分歧，如果句子结构简单，则一般是谓语重，若句子复杂，则句子重音则落在其它的成分上，补语一般不获得重音。

(3) 关于定语与宾语孰重的问题。疑问句中，需要分两种情况来讨论，若定语是“确定指称”类的定语，一般要重读中心语；若定语属于“刻划概念”类的，则需要重读定语。在陈述句中，则没有表现出这种区别，只有在定语是数量词、人称代词时，才会重读中心语或句子的其它成分，其他情况则一律重读定语，例如：**深褐色**的海面显得凝重。**秋天**的海平添了几分充实与忙碌。分别重读“深褐色”和“秋天”。

5 讨论和进一步工作

文章通过重音感知标注实验探讨了句法与疑问句句重音的关系。在此实验之前我们还做了一个疑问句句重音预测实验，创建了一个句法到重音的映射模型（具体的实验设计及过程参见文献[18]），但是句法到重音的映射实验的结果很不理想，通过观察，我们发现导致这两个实验结果产生差异的原因实际上是句法之外的一些因素：

(1) 重音预测模型是完全按照赋予的句法规则进行计算的，在遇到比较复杂的情况时不会灵活处理，而人为标注则比较灵活，可以随时根据变化的情况做出选择。如，两个词按照规则都需要重读时，计算机的处理就比较呆板，没有多余的额外规则可供选择，而人可以根据语感或者说话的需要进行调整，选择需要重读的词语，并且，人们的语感大多数情况下是一致的；

(2) 在对语料中的句子进行以词为单位的切分时，没有固定的标准可供遵循，有时是按照语法词切分，有时则是按照韵律词切分，导致词的切分比较混乱，比如，“有没有”、“能不能”、“是不是”等在一些句子中被切分开，在一些句子中则作为一个整体，还有一些虚词，如“的”、“地”、“得”、“了”的切分问题，在有的句子中被单独切分开，而在有的句子中则被并到前一个词上，这些情况的出现可能会导致预测结果的不一致，但是对人为的标注结果则影响不大；

重音感知标注实验主要是人为标注，在一般情况下，人们的语感基本上一致的，而且参与标注的学生都是语言学专业的学生，专业的素养也使得他们具有较一致的语感，因此，这个实验的结果比较理想。虽然重音预测实验的结果并不理想，但是，感知实验标注的一致性和可接受度还是相对较好的，计算机的预测主要是以总结出来的规律为主，在两种重音位置都可以接受的情况下，以符合规律的重音位置为准，预测的准确率应该会提高，也就是说，我们可以根据句法结构来预测重音的位置。

进一步的研究将考虑结合言语工程的研究，建立一个有效的疑问句下重音预测模型，

在模型的构建中,将对疑问句按照句型进行分类,并对每类归纳出的规则进行预测,在预测中分别加入文本上下文信息,以其提高疑问句句重音的预测效果[]。

6 参考文献

- [1] Chomsky ,N.& Halle,M. The Sound Pattern of English[M]. N.Y.:Haper&Row.
- [2] 王丹,程宗军,郑波等. 普通话常规重音分布规则的研究[J].应用声学.2007, 26(1):46-54.
- [3] 王韞佳、初敏、贺琳.汉语焦点重音和语义重音分布的初步实验研究[J].世界汉语教学.2006 (2): 86-97.
- [4] 袁野. Stress—XP 理论与汉语短语重音研究[J].华文教学与研究.2010 (4) :69-74.
- [5] 杨文昌.汉语的词重音与句重音[J].现代语文(语言研究) .2007 (10) .
- [6] 董洪杰.汉语语句重音在分析[J].西安文理学院学报(社会科学版) .2010,13 (3) .
- [7] 王丹,杨玉芳. 自然语言中焦点与重音关系的研究进展[J].陕西师范大学学报(哲学社会科学版), 2004,33 (4) .
- [8] 王韞佳,初敏,贺琳. 汉语语义重音分布的初步研究[J].第六届全国现代语音学学术会议, 2003:256-261.
- [9] 王韞佳,初敏,贺琳. 汉语语句重音的分类和分布的初步实验研究[J].心理学报,2003,35(6): 734-742.
- [10] 郑献芹.谈重音[J].安阳师范学院学报, 2005 (6): 63-65.
- [11] Bolinger,D. Accent is predictable(if you are a mind reader)[J].Language, 1972.
- [12] Anne Culter. Semantics,Syntax and Sentence accent[J],85-91.
- [13] 曹建芬.基于语法信息的汉语韵律结构预测[J].中文信息学报.2003,17 (3) :41-46.
- [14] 曹建芬. 从语音合成看汉语的重音结构[J].纪念吕叔湘先生百年诞辰国际学术研讨会, 2004.
- [15] 黄伯荣,廖序东.现代汉语[M].北京: 高等教育出版社, 1997:126-127.
- [16] Kitagawa Y., "Prosody, Syntax and Pragmatics of Wh-questions in Japanese"[J], English Linguistics 2005, 22(2):302-346.
- [17] Aijun Li, and Haibo Wang. "Friendly Speech Analysis and Perception in Standard Chinese"[J], INTERSPEECH 2004, 1-4.
- [18] Ya Li, Xuefei Liu, Xiaoying Xu and Jianhua Tao. Assign Stress for Interrogative Sentences via Syntax Structure Mapping, Speech Prosody 2012.
- [19] 张敏.认知语言学与汉语名词短语[M].北京: 中国社会科学出版社, 1998: 264-274.
- [20] 端木三.重音、信息和语言的分类[J].语言科学.2007,6 (5): 3-14.
- [21] 陈虎. 自然语言的重音分布及其语义解释—西方研究综述[J].现代外语(季刊), 2003, 26 (1): 94-103.
- [22] 尹洪波. 现代汉语疑问句焦点研究[J].江汉大学学报(人文科学版), 2008,27 (1): 92-96.
- [23] 卢颖超,李雅,陶建华,许小颖.连续语流中重音标注和大规模重音语料库建设[J].清华大学学报(自然科学版), 2011,51 (9) :1271-1275.

(原载《中国语音学报》第四辑, 2013)