

句法边界的韵律学表现^{*}

杨玉芳

(中国科学院心理研究所 北京 100012)

1996 年 4 月 10 日收到

1996 年 9 月 6 日定稿

摘要 本文研究朗读语句中不同等级的句法边界与附近音节的韵律学参数和边界处停顿时长之间的系统关系。结果看到,边界前音节的时域和频域参数随边界等级的系统变化。在时域,边界前音节的时长和停顿之和随边界等级提高几乎是线性增长;边界前音节时长随边界等级的变化是双向的,在短语边界处达到最大;停顿在大的句法边界处增长很快;在音节内部辅音时长和能量峰值的归一化位置也随边界等级有系统变化。在频域,边界前音节基频均值随边界等级提高而下降,音域逐渐收缩。这些结果将为连续言语合成、识别和理解系统中处理语句句法结构和语音的关系提供实验依据。

PACS 数 43.70

Prosodic cues to syntactic boundaries

YANG Yufang

(*Institute of psychology, The Chinese Academy of Sciences Beijing 100012*)

Received Apr. 10, 1996

Revised Sept. 6, 1996

Abstract The systematic differences of temporal and frequency parameters of syllable sets at the word, word group, phrase, clause and sentence boundaries were investigated. It was found that (1) temporal parameters of the syllable immediate before the boundaries change significantly, but do not for syllables in other positions, (2) for syllables before boundaries the sum of duration and following pause increase linearly with the level of boundaries, (3) the duration of the syllable increases at first before word group and phrase boundaries, and then decreases before clause and sentence boundaries, (4) the pause duration increases dramatically as the level of boundary is higher, (5) the average values of fundamental frequencies decrease as boundary level is higher.

引言

在过去几十年中,汉语普通话音节和词的合成和识别研究取得相当大的进展。在此基础上研制连续言语合成、识别和理解系统,关键问题之一是要研究和利用大的语言单位和结构层次上具有的语音现象和规律。这其中包括语流中音位之间的协同发音和音变规则,更重要的还有超音段层面的现象,即言语的韵律学特征。

^{*} 国家自然科学基金资助项目

韵律特征是言语的组成部分, 具有重要的语言功能, 是把言语切分成不同层次的结构成分的组织框架。在国际上, 语音工程界已将韵律特征广泛运用于语音系统中, 研制了作为文语转换系统一部分的以句法分析为基础的韵律特征生成系统¹, 和作为识别和理解系统一部分的以语音分析为基础的韵律特征检测和标记系统²。在国内, 韵律特征还没有得到充分的注意和应用。在把韵律特征运用于汉语语音系统之前, 需要研究它与语句的句法结构、语境和语义的关系。

本文用朗读语句为语料, 研究语句产生过程中发话人如何用韵律学参数表现语句中各种句法单位的边界。具体地说, 就是在一定的语言环境中, 观察和寻找不同等级的句法边界与附近音节的韵律学参数和边界处停顿时长之间的系统关系。结果将为连续语言合成和识别中处理语句句法结构和语音的关系提供实验依据。

1 研究方法

1.1 方法的考虑

本实验中采用的研究方法是一种聚合分析方法, 兼有经验和统计的性质。

1.2 语句编制

在语句中音节的时长和其后的停顿不仅取决于它在句法和韵律结构中的位置, 还和许多因素有关, 如音节的固有时长和音高, 音节在语句中的序位, 语句的长度等。为了能准确地找出句法边界等级与音段的韵律学参数和停顿之间的关系, 语料编制过程中考虑了如下的限制:

(1) 所考察的句法边界类型有词中音节边界、词界、词组边界、短语边界、分句边界和句子边界, 分别用边界等级 1 至 6 作标记。

(2) 所考察的句法边界尽可能与韵律结构成分边界保持一致。

(3) 语句长度接近汉语口语平均句长, 实词为 7 个左右, 总词数为 10 个左右。

(4) 六种边界等级分别放在六个字数相同或相近的语句中考察, 边界出现在相同的音节序位之间, 六个语句构成一个语句组。

(5) 边界前两个音节和边界后一个音节在六个语句中保持一致。

(6) 为更好地考察音域问题, 边界前音节的声调类型在语句组之间做到平衡。对每个声调类型, 有两组语句。

这样, 共设计了八组语句, 每组六句, 共 48 个语句。全部语句在附录中给出。

1.3 录音、采样和分析

共有六位发音人, 四男二女, 为北京广播学院播音系的教师和进修生。要求用播音时的朗读风格读出。用 Nagra I V-S 型录音机录音。然后通过线性输出在 IBM PS/2 微机上采样, 采样率为 10 kHz, AD/DA 为 12 bit, 低通滤波截止频率为 4.0 kHz。每位发音人读的每个语句都单独存入一个样本文件中, 形成 288 个文件。对样本文件中所要观察的边界附近的音节通过 ILS 软件作 LPC 分析, 提取基频和共振峰值。

1.4 参数选择和测量

在时域, 选择五个参数: 音节时长 + 停顿时长 ($S + P'$), 音节时长 (S'), 停顿时长 (P'), 辅音时长 (C') 和能量峰值位置 (EP' , 指从音节开始到能量峰值间的时间间隔), 单位是 ms。在频域, 测量音节声调曲线特征段的频率, 单位是 Hz。

测时域五个参数时, 对照波形、基频曲线和共振峰结构图确定测量位置, 程序将时长数据记入数据文件。测基频曲线时, 对照基频曲线和共振峰结构图, 取基频曲线的特征段。测的同时作时域上的归一化。等间隔抽取十个样点, 并记入数据文件。

1.5 数据预处理

为了排除语速和发音人性别等引起的差异, 更好地观察句法边界和韵律学参数之间的关系,

在作统计分析前, 对数据作了预处理。

对时域参数作了下述几步预处理:

(1) 测六位发音人所发的每个语句的总时长 (Sen)。

(2) 计算音节时长 + 停顿时长、音节时长和停顿时长在语句中占的相对比例 (分别用 $S+P$ 、 S 和 P 表示) 以及辅音在音节 (Sy) 中占的相对比例和能量峰值位置在音节 (Sy) 中的归一化位置 (分别用 C 和 EP 表示)。计算公式:

$$S + P = (S + P') / Sen$$

$$S = S' / Sen$$

$$P = P' / Sen$$

$$C = C' / Sy$$

$$EP = EP' / Sy$$

(3) 以处于二字词首字位置 (边界等级为 1) 的音节的时域参数作标准, 计算其它边界等级上时域参数的相对值 (对边界作归一化处理)。

为减小男女声之间的差异, 并考虑音域分析的方便, 对声调基频值作对数转换。转换公式³:

$$Deg = 5 * \log_2(F/F_0)$$

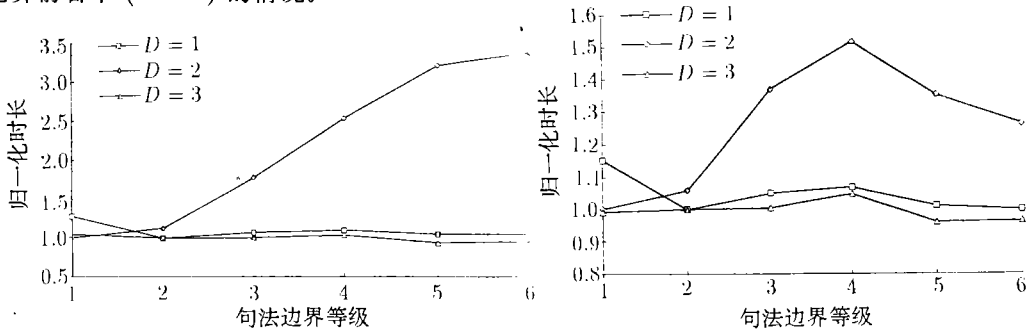
其中: $F_0 = 55$ Hz, 对男声; $F_0 = 80$ Hz, 对女声。

2 统计分析和结果

2.1 时域

为了考察边界附近音节的时域参数随边界等级变化的趋向和统计显著性, 取句法边界等级 ($C=1 \sim 6$) 和语句组 ($B=1 \sim 8$) 作自变量, 取五个时域参数作因变量, 分别对边界附近三个音节 (边界前第二个音节, $D=1$; 边界前音节, $D=2$; 边界后音节, $D=3$) 作 MNOVA 分析。表 1 和表 2 分别给出边界音节 ($D=1, 2, 3$) 的五个时域参数在六种边界等级上的平均值及统计检验结果。

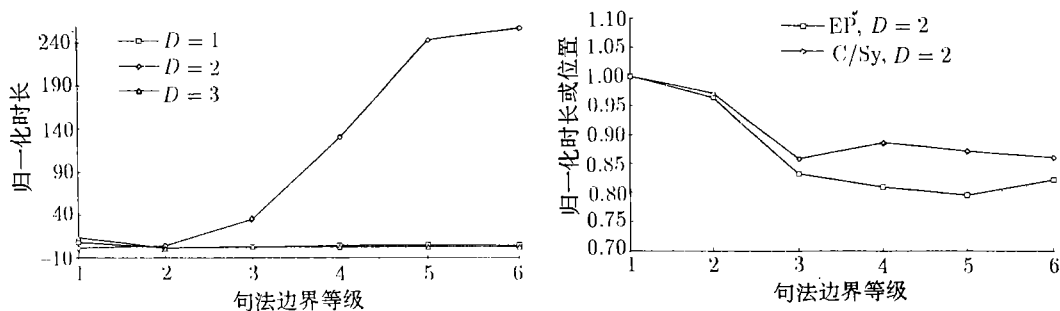
图 1 给出了作归一化处理后的音节 + 停顿 (a)、音节 (b)、停顿 (c)、辅音时长和能量峰值位置 (d) 与边界等级的关系。其中音节序位 ($D=1, 2, 3$) 为参数。对辅音时长和能量峰值仅考虑边界前音节 ($D=2$) 的情况。



(a) 音节 + 停顿时长 ($S + P$) 与句法边界等级的关系

(b) 音节时长 (S) 与句法边界等级的关系

图 1



(c) 停顿时长 (P) 与句法边界等级的关系

(d) 辅音时长 (C) 和能量峰值归一化位置 (EP) 与句法边界等级的关系

图 1

表 1 音节时域归一化参数与边界等级的关系

音节位置 (D)	边界等级 (C)	时 域 参 数					
		S + P	S	P	C'	C	EP
1	1	1.274	1.153	7.689	2.480	3.533	0.889
	2	1.000	1.000	1.000	1.000	1.000	1.000
	3	1.066	1.050	2.352	1.580	1.875	0.992
	4	1.096	1.068	3.908	1.268	1.337	1.004
	5	1.046	1.010	4.789	1.433	1.615	0.949
	6	1.045	0.998	4.210	1.436	2.710	0.999
2	1	1.000	1.000	1.000	1.000	1.000	1.000
	2	1.122	1.059	3.679	0.957	0.964	0.971
	3	1.780	1.370	34.931	1.097	0.832	0.858
	4	2.538	1.517	130.250	1.163	0.809	0.886
	5	3.220	1.352	243.278	1.084	0.795	0.871
	6	3.388	1.263	257.168	1.066	0.821	0.860
3	1	1.044	0.991	13.152	1.638	1.524	1.017
	2	1.000	1.000	1.000	1.000	1.000	1.000
	3	1.002	1.004	2.713	6.092	5.201	1.243
	4	1.030	1.047	2.407	4.160	4.328	1.058
	5	0.933	0.959	2.458	4.398	4.584	1.209
	6	0.952	0.964	2.388	4.310	3.871	1.222

表 2 音节时域归一化参数与边界等级关系的统计检验

音节位置 (D)	指 标	时 域 参 数					
		S + P	S	P	C'	C	EP
1	F(4,200)	2.4640	1.8812	2.1671	1.0864	1.0775	0.2182
	P	0.0456	0.1140	0.0730	0.3645	0.3735	0.9265
2	F(5,240)	140.96	27.1365	88.1540	2.8739	6.6477	2.1468
	P	0.0000	0.0000	0.0000	0.0153	0.0000	0.0599
3	F(4,200)	1.9579	1.6201	1.5904	2.6208	2.5407	1.4186
	P	0.1013	0.1694	0.1771	0.0355	0.0286	0.2279

从表 1 和表 2 可以看出，对于边界前音节 (D = 2)，所考察的五个参量都随边界等级发生系统变化。而且，除了能量峰值 (EP) 外，这些变化都在统计上达到显著。边界前音节时长和其后的停顿时长之和随边界等级单向增长，且在边界等级 2 ~ 5 之间几乎是线性地增长。音节时长随边界等级的变化不是单向的，在等级 2 ~ 4 是线性地增长；在等级 4，即短语边界处达到最大值，约为在二字词首字位置时 (C = 1) 的时长的 1.5 倍；在边界等级 5 ~ 6 处是下降的。停顿也是随边界等级单向地增长，在等级 5 ~ 6 处增加很快。辅音时长随边界等级的变化趋向是一样的，在

短语边界处达到最大值,为词首位置时的 1.16 倍,比全音节延时比例小。辅音在音节中的比例是随边界等级下降的,且变化达到统计显著性。能量峰值在音节中的归一化位置与边界等级有关,但变异不显著。

对于边界前第二个音节 ($D = 1$),在第一句中为二字词末字,在第二至六句中为首字,如果排除为末字的情况,仅考虑其作为首字的情况,则五个时域参数均与边界等级无关。就是说,边界前第二个音节其时长及音节内部时间结构都不随边界等级变化。

边界后的音节 ($D = 3$),在第一句中为二字词末字,在二至六句中为首字。在作为首字时 $S + P$ 、 S 和 P 都与边界等级无关,辅音在音节中占的比例有变化且变化显著,但与边界等级之间不是单向的关系。

2.2 频域

在频域,考察了句法边界附近音节的声调基频曲线特征段的频率均值和音域与句法边界等级间的关系。

在统计分析前,先对缺失值进行处理,测得的 864 条基频曲线中有 42 条部分样点为 0,占 4.9%;有 19 条全部样点为 0,占 2.2%。这是目前基频提取技术水平所限制的,不论采取那种基频提取方法,都不可能完全避免。对于第一种情况,若有一半以上样点为 0,且为 0 的样点十分集中,则按整条曲线为 0 处理;若为 0 的样点少于一半,且为 0 样点和非 0 样点是穿插的,则用线性内插值代替缺失值。对于整条曲线为 0 的情况,则用同一发音人在同样的边界等级 (C 相同) 和位置 (D 相同) 上调类相同的音节的基频曲线代替。这样处理引起的误差有两个,一是替补音节和基频缺失音节音位组成不同,因而固有音高不同;二是两者相邻音节调类不同,可能使基频曲线的过渡部分不同。由于音节固有音高引起的误差不会太大,我们在测量中取的是基频曲线的特征段,相邻音节调类的差别引起的误差可以排除。因此,这里采用的补缺办法应该是可以接受的。

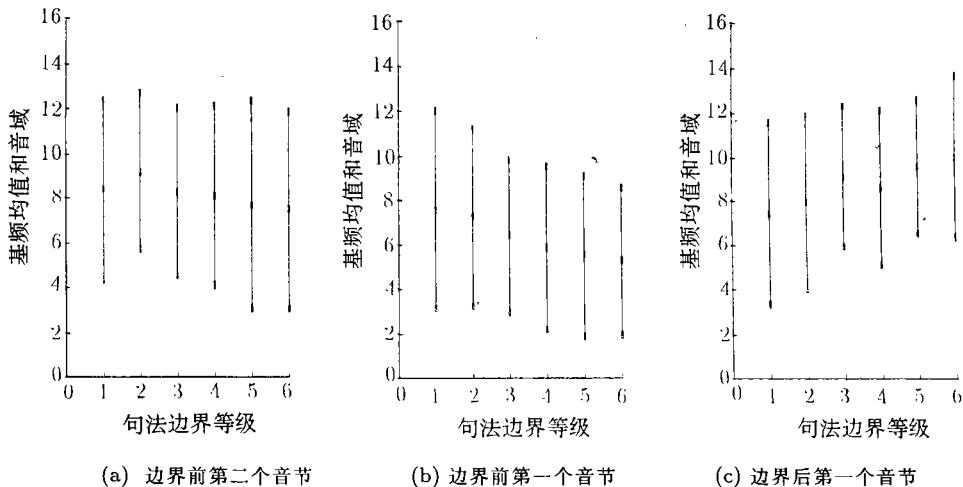


图 2 基频均值和音域与边界等级的关系

为了对基频与边界等级的关系有定量描述,计算了在六种边界等级上,对发音人平均的音节基频均值和标准差。然后取边界等级和声调类型作自变量和基频均值作因变量;分别对 $D = 1, 2, 3$ 三种序位音节的基频均值作 ANOVA 分析,表 3 给出计算结果。图 2 给出了 $D = 1, 2, 3$ 时基频均值和音域与边界等级的关系。其中音域定义为 $M \pm 2SD$ (4 倍的标准差)。

结果表明,对于边界前音节 ($D = 2$),基频均值随句法边界等级提高而下降,音域随边界等级提高而缩小。边界前第二个音节 ($D = 1$),基频均值随边界等级提高下降,而音域逐步加宽。边界后音节,基频均值和音域在不同等级的边界处有变化,但不是单向的。

表 3 音节基频均值和音域与边界等级的关系

音节位置 (D)	边界等级 (C)	平均值 (M)	标准差 (SD)	音域 ($M \pm 2SD$)
1	1	8.3545	2.0189	4.3167 - 12.3923
	2	9.1787	1.7586	5.6615 - 12.6959
	3	8.2557	1.8949	4.4659 - 12.0455
	4	8.0676	2.0043	4.0590 - 12.0762
	5	7.6799	2.3221	3.0357 - 12.3241
	6	7.4621	2.1972	3.0677 - 11.8565
2	1	7.5649	2.2244	3.1161 - 12.0137
	2	7.2209	1.9923	3.2363 - 11.2055
	3	6.3786	1.7172	2.9442 - 9.8130
	4	5.8525	1.8415	2.1695 - 9.5355
	5	5.4791	1.7985	1.8821 - 9.0761
	6	5.2431	1.6545	1.9341 - 8.5521
3	1	7.4399	2.0543	3.3313 - 11.5485
	2	7.9388	1.9649	4.0090 - 11.8686
	3	9.0758	1.8521	5.3716 - 12.7800
	4	8.5977	1.7495	5.0987 - 12.0967
	5	9.5619	1.5128	6.5363 - 12.5875
	6	9.9832	1.8271	6.3290 - 13.6374

3 讨论

本实验用朗读语句为语料,研究了语句句法结构确定之后,发音人如何用韵律学参数表现不同等级的句法边界。根据得到的实验结果,有以下几点值得讨论。

实验中对边界附近三个音节的韵律学参数随边界等级变化进行了观察和分析。音节时长和停顿发生系统变化的只有边界前音节。边界前第二个音节和边界后的音节的时长在作为二字词首字时不发生变化。这一结果与 Wightman⁴ 对英语所作的研究结果是一致的。说明,汉语和英语一样,句法边界对音节时长的影响仅限于边界前音节。这里着重讨论一下边界前音节时域参数随边界等级产生的变化。

音节和其后的停顿的总时长随边界等级单向增长,在词以上的边界等级间几乎是线性增长的。其中句子边界增长比较慢,可能是由于有些发音人是把两个句子当作两个分句处理的。

音节时长随句法边界等级变化是双向的,在短语边界处达到最大值。这种变化规律说明,在汉语中音节延时的确是表现句法边界强度的手段,但它的表现能力有限。因为在语流中音节延时不可能是无限的。另一方面可能还说明,音节延时不仅表示句法成分或语言单位之间的分割程度,还表现语流的连续。在分句和句子末尾,当一个相对完整的语句结构完成时,音节延时反而减小。

边界处的停顿随边界等级增长很快,特别是从短语边界到句子边界。说明停顿在表现句法边界强度中,特别是大的句法边界,是一种强有力的手段。

有人认为⁵,汉语不同于英语,表现句法边界的主要手段是停顿而不是延时。从本实验结果看,特别是从延时和停顿的绝对值的变化看,两者的贡献在不同的边界等级上是不一样的。为了说明这一问题,图 3 给出了音节 + 停顿,音节和停顿的时长绝对值与边界等级的关系。从中可以看出,在表现词和词组边界时,以延时为主;在表现短语边界时,二者的作用相当;在表现分句和句子边界时,主要靠停顿。这与用英语所作的研究结果是一致的。

实验中还观察到,边界音节的辅音随边界等级的系统变化。尽管变化的幅度小于整个音节,却达到了统计显著性。这与用英语作的研究结果不同。已经证明²,英语中边界延时仅限于边界前

音节的元音核和尾辅音。在这两种语言之间为什么有这样的差异, 值得进一步研究。

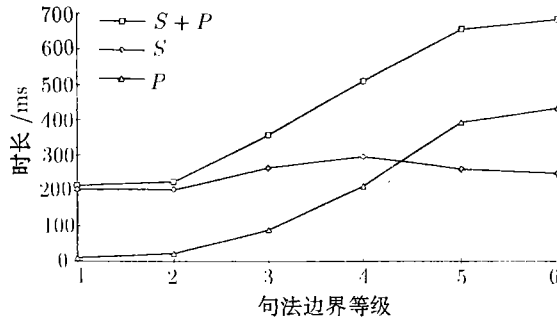


图 3 音节 + 停顿时长、音节时长和停顿时长与边界等级的关系

频域测量和分析结果表明, 边界前两个音节基频均值都随边界等级提高下降。但是音域的变化则不同。边界前音节 ($D = 2$) 的音域随边界等级提高逐渐收缩, 而边界前第二个音节 ($D = 1$) 的音域有逐渐扩张的趋势, 边界后音节 ($D = 3$) 基频均值和音域的变化与边界等级之间没有系统的关系。边界前两个音节音域的变化趋势可以用语音对比规则加以解释。

参 考 文 献

- 1 沈 炯. 北京话声调的音域和语调. 林 焱, 王理嘉. 北京语音实验录. 北京: 北京大学出版社, 1985, 73—130
- 2 Wightman C W, Shattuck-Hufnagel S, Ostendorf M, Price P J. Segmental durations in the vicinity of prosodic phrase boundaries. *J. Acoust. Soc. Am.*, 1991, **91**(3): 1707—1717
- 3 Rossi M. A model for predicting the prosody of spontaneous speech (PPSS MODEL), *Speech Communication*, 1993, **13**: 87—107.
- 4 Compbell N. Automatic detection of prosodic boundaries. *Speech Communication*, 1993, **13**: 343—354
- 5 Shen S N. The use of prosody in disambiguation in Mandarin Chinese. *Phonetica*, 1993, **50**: 261—271

附录: 语句表

- 第一组
1. 小李同志提高工的愿望将很快变为现实。
 2. 增加产量提高工效是同样重要的两个方面。
 3. 改变产量提高工效下降的状况是当务之急。
 4. 日元汇率提高攻破了美国贸易保护的防线。
 5. 经济实力提高, 工业和农业的发展是基础。
 6. 市场物价提高。工人的经济收入有所增加。
- 第二组
1. 在夜间行驶的非机动车辆必须严加控制。
 2. 王老师讲授的飞机动力这门课极受欢迎。
 3. 新购买的波音飞机动力装置属世界一流。
 4. 研制这种新型飞机动用了大量人力物力。
 5. 学员学习驾驶飞机, 动机和目的必须明确。
 6. 这是一架老式飞机。动力系统仍完好无损。
- 第三组
1. 小学生参观茶叶工厂的时间还没有确定。

2. 米勒里把观察夜空当作一种爱好和习惯。
3. 连续不断观察液体颜色变化是实验要求。
4. 这次现场观察业已收到意想不到的成效。
5. 根据我的观察，夜行车辆多半是出租汽车。
6. 冯利主张观察。叶辉主张立即采取措施。

- 第四组
1. 那班的学生专攻园林和建筑设计学。
 2. 李晓阳经常去官源林场购买月季花。
 3. 孩子们到天坛公园林荫绿地做游戏。
 4. 整洁的玉渊潭公园邻近钓鱼台宾馆。
 5. 建造现代街心公园，林木花草不可少。
 6. 北京有座北海公园。林明从来没去过。

- 第五组
1. 上星期日本家叔叔回上海市探亲。
 2. 新型的日本家电占据了世界市场。
 3. 当今的日本家电市场已近于饱和。
 4. 发达的日本加快了电子产品更新。
 5. 为反对日本，加拿大派出援华医疗队。
 6. 老徐在日本。家庭其他成员都在中国。

- 第六组
1. 白山的松江水流时而湍急时而缓慢。
 2. 注意观察江水流向是防洪抗洪的要求。
 3. 这里嘉陵江水流速度达每秒十米。
 4. 一部分长江水流向洞庭湖和渤海湖。
 5. 如果堵住江水，刘阳河水源将会枯竭。
 6. 他们引调江水。刘家峡电站受到影响。

- 第七组
1. 大吊车搬运动力设备既安全又省时。
 2. 李康凯的运动力学研究取得重要突破。
 3. 有关地壳运动历史演变的著作很少见。
 4. 坚持体育运动利于青少年的健康成长。
 5. 练习举重运动，力量和技巧同样重要。
 6. 我爱球类运动。立新喜欢田径和游泳。

- 第八组
1. 他们从两广贩运水果取得很大经济收益。
 2. 普通话的广泛运用有利于经济文化交流。
 3. 这本内容广泛韵味十足的文集十分畅销。
 4. 讨论深入广泛孕育出了新的构思和见解。
 5. 知识越是广泛，运用起来越是得心应手。
 6. 读书应该广泛。运用应该灵活而有创新。