

汉语普通话不同文体朗读时的呼吸重置研究

Breathing-Resets in Reading Aloud the Texts of Different Literature Genres in Mandarin Chinese

谭晶晶 李永宏

Tan Jingjing, Li Yonghong

摘要: 本文目的在于探讨人们在使用汉语普通话朗读不同文体时的呼吸重置特点。本文运用肌电脑电仪和呼吸带传感器测量了发音人在朗读不同文体时呼吸节奏的变化,通过自动标注提取出呼吸重置的幅度和时长,并分析了它们的频度分布,发现在朗读不同文体时,发音人的呼吸节奏有较明显的差别,朗读韵文和非韵文时,呼吸重置分别可以分成2级和3级。实验结果表明,呼吸重置的时长和幅度显著正相关,重置幅度比重置时长更能反映不同文体的差别,也更能反映不同呼吸级别的大小。

Abstract: This paper is a preliminary study on the breathing-reset characteristics in reading aloud texts of different literature genres in Mandarin Chinese, namely, Chinese ancient poetry, song lyrics, novel, prose and news. Breathing signals are recorded by EMG and respiration sensor. The parameters of duration and amplitude of breathing-resets are automatically detected. Some conclusions have been drawn by analyzing the frequency distribution of the duration and amplitude: 1) The breathing rhythm varies with the literature genres. There are two levels of breathing-resets in verse-reading and three in essay-reading. 2) The duration and amplitude are highly-correlated across genres. 3) The amplitude is more useful in indicating the literature genres and the hierarchies of breathing-resets.

关键词: 语音生理学 韵律 呼吸重置 文体

Key words: physiological phonetics, prosody, breathing-reset, literature genres

0 引言

语篇的韵律特征是近年来语音学界研究的一大热点。由于语流的韵律构造是一个复杂的过程,它受到生理、心理、语法等各个方面的制约,因此研究者的角度也各不相同。目前的研究大多集中在音系学、声学和心理学方面,如文[1]、[2]、[3],从生理角度进行研究的则并不多见。在这一方面,Liberman (1967) [4]曾提出过一个“呼吸群 (breath-group)”的理论,认为人们是通过

呼吸群来产生和感知语调的,在郑秋豫 (2005) [5]近年的研究中,也把呼吸看成是划分韵律层级的重要线索。“呼吸群”的理论可以有效地解释语流在实现过程中受到的生理因素的制约,也能有效地解释语流中的音高下倾现象,可是目前完全从生理(呼吸)的角度对汉语语流的韵律问题进行的研究还非常少见。

文[6]利用肌电脑电仪和呼吸带传感器记录了呼吸信号,对新闻朗读中呼吸节奏的变化做了一些分析,试图寻找人们朗读时呼吸节奏的类型,探讨呼吸节奏和韵律结构的关系,发现呼吸结构可以反映人们对表达内容的整体认知规划,同时它也反映了人们生理机能的制约。由于当时实验手段的制约,还不能直接提取呼吸信号的参数,划分呼吸重置级别时,主要依靠人工判断。最近,本文编写了程序来自动提取呼吸重置的时长和重置幅度,并用它对近体诗、词、小说、散文、新闻等不同文体的朗读语料进行了分析,试图寻找人们在朗读不同文体时呼吸节奏的特点。

1 语料说明

1.1 实验语料

本实验选取了1名发音人朗读的近体诗20首、词30阕、小说20篇、散文20篇、新闻40篇,为了尽量减少发音人的断句错误,本文选取的都是发音人比较熟悉的语料。该发音人为北京大学电视台新闻播音员,女,22岁,受过良好的播音训练。

1.2 语料录制

本实验的录音工作在北京大学中文系的隔音室内进行。录音使用的主要设备是澳大利亚PowerLab公司生产的肌电脑电仪,本实验用它

自带的录音软件 Chart5 采集了 4 个通道的信号：1 通道为语音信号，2 通道为噪音（EGG）信号，3 通道为通过 MLT1132 呼吸带传感器采集的呼吸信号，4 四通道闲置，如图 1 所示。

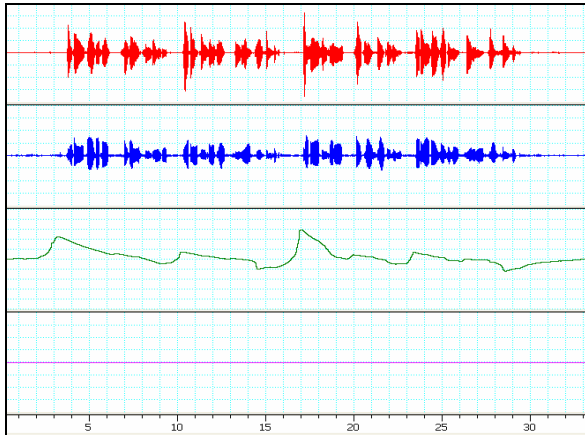


图 1 用 Chart5 采集的四通道信号

2 参数提取

2.1 参数定义

将呼吸信号的变化反映在二维图谱上，横轴是时间，纵轴是由呼吸导致的电压变化（称之为“呼吸曲线”），呼吸曲线上升表示吸气，下降表示呼气，如图 2 所示。

本实验提取的参数主要是呼吸重置的时长和重置幅度，其中重置时长指的是从开始吸气到开始呼气之间经过的时间，即图 2 中波谷和波峰之间的水平距离 T ，重置幅度指的是一次吸气过程中呼吸信号数值的变化幅度，即图 2 中波谷和波峰之间的垂直距离 A 。

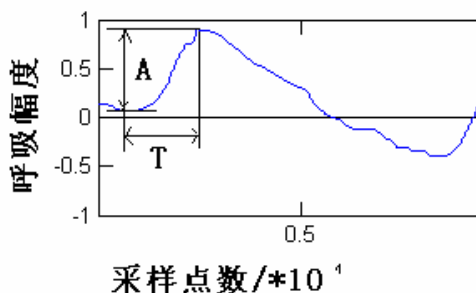


图 2 呼吸重置

2.2 程序说明

本实验采用的呼吸信号处理程序是北京大学中文系语音乐律平台子程序，在 Windows 平台下用 Matlab 编程实现。该软件功能包括：自动或手动对呼吸信号的重置进行标记，包括重置

谷底标记和重置峰值标记，手动标记用不同的标记形式对不同大小的呼吸重置分别进行标记，主要的参数为重置谷底的位置和相对幅值，重置峰值的位置和相对幅值，计算每一个呼吸重置需要的时长和幅度，并保存自己设置的带参数的四通道文件，可以用带参数的四通道读取函数直接读取，标记直接显示在呼吸信号之上，便于复查，最后对标记好的 wav 文件进行批处理，自动读标记数据到 Excel 文件中。具体处理步骤如下：

1}. 考虑到朗读不同文体时呼吸信号幅度和时长的可比性，首先对第 3 通道的呼吸信号进行归一化处理。breathe_signal 为原始采集的呼吸信号。

```
breathe_signal = breathe_signal / max(abs(breathe_signal));
```

2}. 对归一化后的呼吸信号进行平滑滤波。由于采集的原始呼吸信号带有很多细微的高频和部分干扰信号，因此对归一化后的呼吸信号进行了低通平滑滤波，滤波器采用零相位数字滤波 `filtfilt(b,a,x)`，通过将输入数据前向和反向处理，以完成零相位数字滤波。它先将数据按顺序滤波，然后将所得结果逆转后反向通过滤波器，这样得到的序列为精确零相位失真，并使滤波器的阶数加倍。Filtfilt 通过与初始条件相匹配。可使起始和结束阶段的暂态过渡过程最短^[4]。

Filtfilt 差分方程表示为

$$y(n) = b(1)_x(n) + b(2)_x(n-1) + \dots + b(nb+1)_x(n-nb) - a(2)_y(n-1) - \dots - a(na+1)_y(n-na)$$

参数 b 序列的取值越小，滤波后的信号就越接近原始信号，但运算量越大，耗费的时间就越多，一方面考虑对高频信号的滤除，以满足呼吸的处理，另一方面要误差尽量小，方差限制在一定范围内，参数的取值为：

```
b=ones(1,400)/400
```

```
a=1
```

3}. 自动标记波峰和波谷。对平滑后的呼吸信号，利用局部最大法，综合考虑呼吸重置峰值点之间的距离和幅度差、谷值点之间的距离和幅度差，多次迭代循环，尽量消除小的类似于呼吸重置的峰和谷，最终获得能满足要求的呼吸重置的峰值和谷值。程序中还设置了手动修改功能，

如果标记位置有误，可进行手工调整。

4}. 手动标记波峰和波谷。经过对呼吸信号的大量研究，大致根据呼吸重置的幅度大小和时长，确定出正常人说话，一般有3级呼吸，程序中可以对每一级别的呼吸重置和呼吸单元进行标记，也可以作为自动标注的补充手段。

5}. 参数保存。标记打完后，可以按实验室自己的格式保存为带参数的四通道 wav 文件，所有的标记参数都保留在原始 wav 文件的后面，参数按国际 RIFF 标准定义。

6}. 对提取呼吸参数的 wav 文件进行批处理，程序中针对手动和自动标记分别做了2套自动提取参数的程序，能够直接把文件夹下面的所有 wav 文件的标记数据直接读取到 xls 文件中。

3 实验结果

3.1 近体诗¹

本实验选取了五言律诗、七言律诗、五言绝句、七言绝句各5首。由于近体诗的格律严格，结构简单，因此呼吸信号的类型也比较简单。其呼吸信号的特点是：无论是五言诗还是七言诗，每一联(每联包含两小句)开始处都有一个呼吸重置，七言诗则每小句开始处都有一个呼吸重置，小句开始处的呼吸重置比每一联开始处的呼吸重置小。如图3所示。

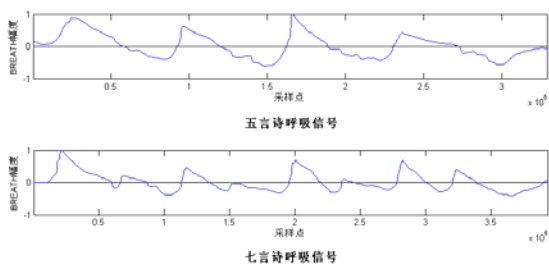


图3 五言诗(上)和七言诗(下)呼吸信号示例

图3为五言律诗《天末怀李白》和七言律诗《登高》的呼吸信号，这2首诗都有4联(共8小句)，每一联开始处都有一个较大的呼吸重置，七言律诗由于每小句字数较多，无法一口气念完一联，因此在每一联的第2个小句开始的时候会先吸一小口气，这样就形成了一个比较小的呼吸重置。

¹格律诗有律诗和绝句之分，律诗分四联八小句，绝句为两联四小句，每小句5或7个字，即“五言”、“七言”。

从图3中可以看出呼吸信号的重置幅度有大有小，为了探讨大小不同的呼吸重置幅度的频度分布，本文自动提取了20首近体诗中所有的呼吸重置时长和重置幅度(共计119处)，在SPSS12.0中将它们画出直方图，如图4、5所示。

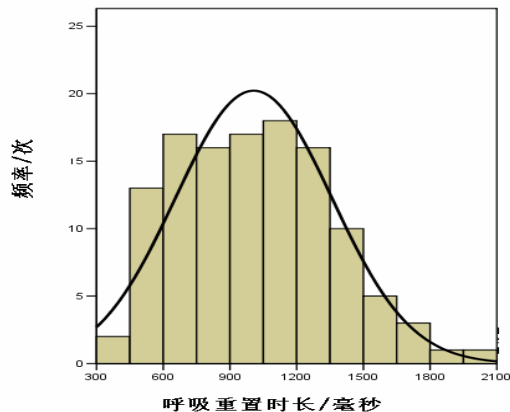


图4 近体诗重置时长频度分布图

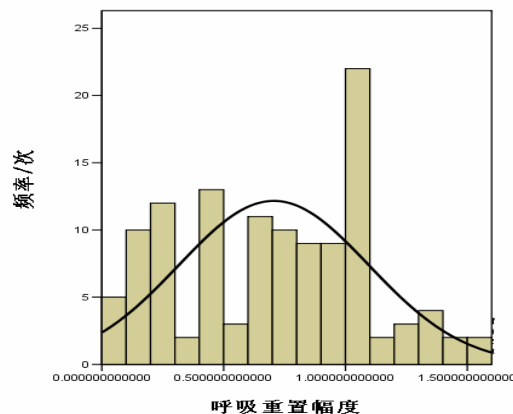


图5 近体诗重置幅度频度分布图

从图4、5中可以看到，近体诗重置时长的频度分布基本符合正态分布，重置幅度的分布则在0.50和1.00处分别出现了2个较明显的峰值，尤其是1.00处频度达到20以上，这说明近体诗的呼吸重置大致可以分成2级。这种频度分布可以从近体诗的呼吸节奏特点中得到解释：每首诗都会有大呼吸重置，但是只有七言诗有小呼吸重置，因此大呼吸重置的数量就要比小呼吸重置的数量多。

3.2 词²

本次实验选取了10个不同词牌，每个词牌选取3首词，共30阙词，涵盖小令、中调和长

²词按字数多少分为小令(58字以下)、中调(59-90字)、长调(91字以上)。词按“阙”计，一首词为一阙，有些词分为上下两段，称为“上阙”、“下阙”。

调。词的呼吸信号有如下特点：相同词牌的三阙词呼吸节奏完全相同，每阙词的上阙和下阙开始处都有一个较大的呼吸重置，除长调的呼吸重置幅度大致可以分成大、中、小 3 级之外，其余词的呼吸重置大致可以分成大、小 2 级。如图 6 所示：

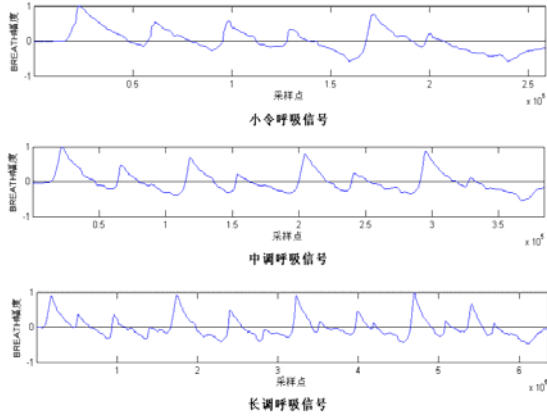


图 6 小令（上）、中调（中）、长调（下）呼吸信号示例

我们将 30 阙词中所有的呼吸重置时长和重置幅度（共 229 处）画出如图 7、8 所示的频率直方图，从图中可以看出，词中的呼吸重置时长和重置幅度都有 2 个较明显的峰，尤其是重置幅度的频率分布图中 2 个峰（分别出现在 0.3 和 1.0 处）基本对称，这说明将词中的呼吸重置分为 2 级是比较合理的。

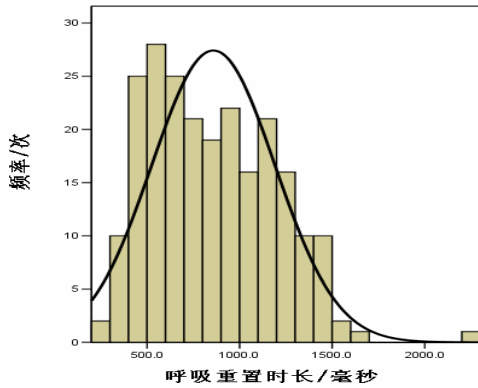


图 7 词重置时长频率分布图

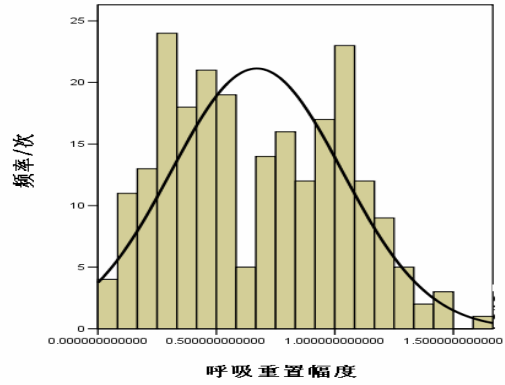


图 8 词重置幅度频率分布图

3.3 小说

本次实验从可以作为现代汉语典范的白话小说中选取了 20 个完整自然段作为实验材料，每段大约 250 字。小说呼吸信号的特点是：每段小说中会出现 1~3 个较大的呼吸重置、若干个中等大小的呼吸重置以及很多小呼吸重置，如图 9 所示。

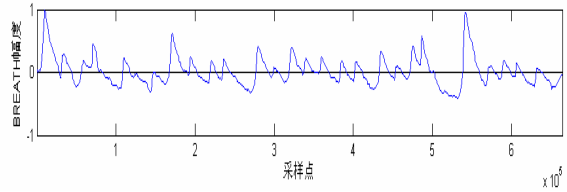


图 9 小说呼吸信号示例

从图 10、11 中可以发现：小说的呼吸重置时长和重置幅度的频率分布呈偏态分布，重置时长在 500ms 处有一个高峰、此后频率急剧下降，重置幅度在 0.20 处有一个高峰，重置幅度的频率下降的趋势与重置时长相比显得比较和缓，0.60~1.20 之间有一段比较稳定。这种频率分布也反映了小说大呼吸重置少、小呼吸重置多的特点。

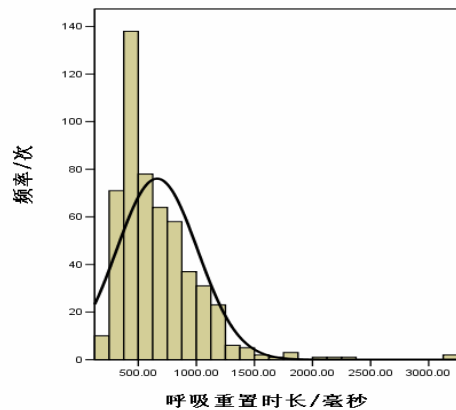


图 10 小说重置时长频率分布图

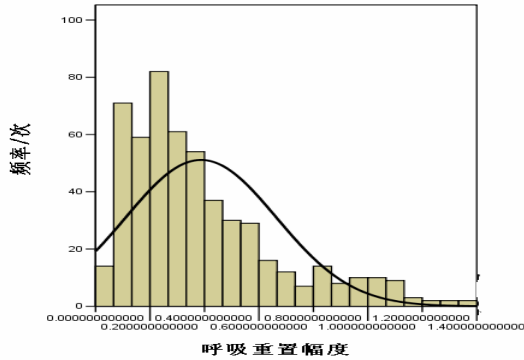


图 11 小说重置幅度频度分布图

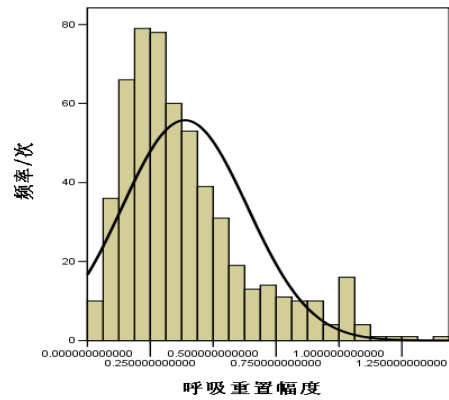


图 14 散文重置幅度频度分布图

3.4 散文

本实验选取的 20 段散文来自于可以作为现代汉语典范的白话散文，字数也都控制在 250 字左右。散文呼吸信号的特点和小说非常相似：每段散文中会出现 1~3 个较大的呼吸重置、若干个中等大小的呼吸重置以及很多小呼吸重置，如图 12 所示。

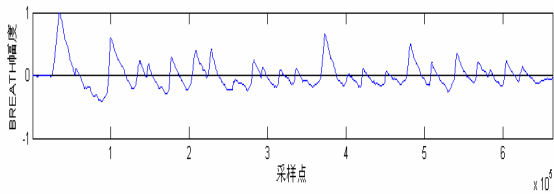


图 12 散文呼吸信号示例

散文呼吸重置时长和重置幅度的频度分布也是偏态分布（如图 13、14 所示），重置时长在 300ms 处有一个高峰，频度高达 180 余次，之后时长频度迅速下降。重置幅度则在 0.12~0.40 之间分布非常集中，之后在 1.00 处又有一个小峰。这种分布说明朗读散文时发音人的呼吸节奏也呈现出大呼吸重置少、小呼吸重置多的特点。

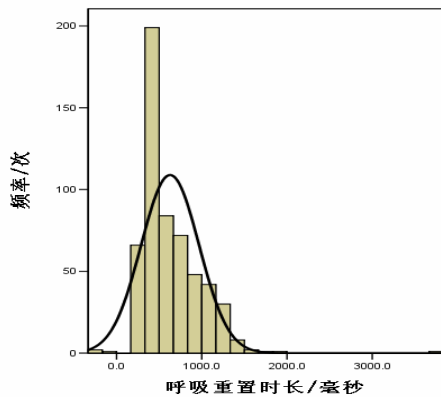


图 13 散文重置时长频度分布图

散文和小说的呼吸信号从重置时长和重置幅度的分布频度方面来看非常相似，只是在具体的频度峰值上有细微的差别，这可能是因为在具体的频度峰值上有细微的差别，这可能是因为散文和小说从广义的文学体裁上来说都属于和“韵文”相对应的“散文”，因此朗读的时候在节奏上会有相似之处，在朗读小说和散文时能感受到它们之间的区别，可能有其他方面因素（如音高、语速甚至语义）的影响。

3.5 新闻

新闻的文章结构有其固定的格式，发音人在朗读新闻时也容易形成相对固定的呼吸节奏模式。文[6]曾分析过该发音人朗读 40 篇新闻的呼吸节奏，认为在新闻朗读中呼吸节奏可以分为 3 级呼吸单元，分别和自然段、复句和分句相对应，相应的也有 3 级大小不同的呼吸重置，如图 15 所示：

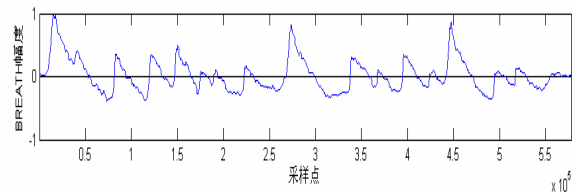


图 15 新闻呼吸信号示例

在本次实验中，本文自动提取了这 40 篇新闻中共 671 处的呼吸重置时长和重置幅度，并将其画成了频度分布直方图，如图 16、17 所示：

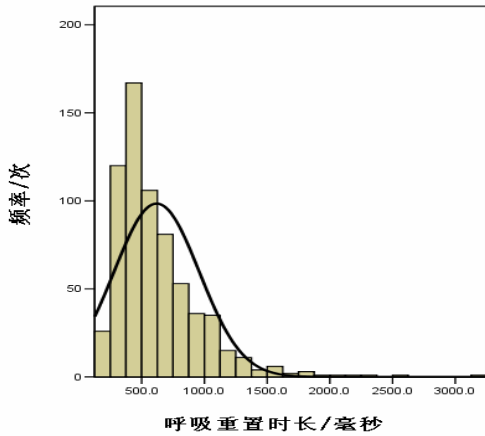


图 16 新闻重置时长频度分布图

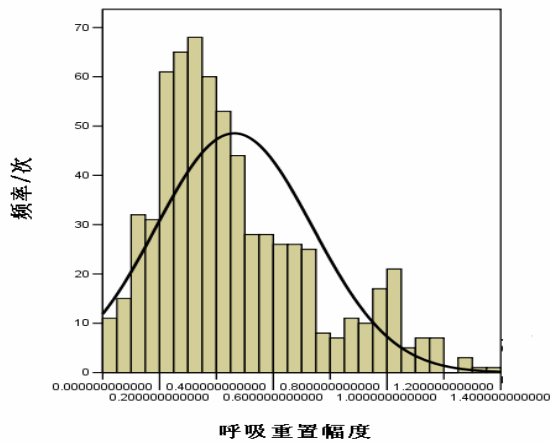


图 17 新闻重置幅度频度分布图

从图 16 中可以看到，新闻的呼吸重置时长呈明显的偏态分布，在 400ms 处有一个高峰，随后频度分布曲线急剧下降，重置时长频度的这种分布和小说、散文非常相似。

从图 17 中可以看到新闻的呼吸重置幅度在 0.30 处有 1 个很明显的高峰，在 1.00 处也有 1 个比较明显的小峰。新闻和小说、散文的不同之处在于，新闻大呼吸重置的数量比散文、小说多，另外，新闻的重置幅度在 0.50~0.80 之间有一段比较稳定的频度分布，这一段大致对应着中等的呼吸重置，这一级的呼吸重置数量比散文、小说多。这说明之前将新闻的呼吸重置分成 3 级是有道理的。

3.6 呼吸重置时长和重置幅度的关系

本文运用 SPSS12.0 对呼吸重置时长和重置幅度作了相关分析，得到如下结果：

Correlations				
			duration	amplitude
Kendall's tau_b	duration	Correlation Coefficient	1.000	.517(**)
		Sig. (2-tailed)	.	.000
		N	2108	2108
	amplitude	Correlation Coefficient	.517(**)	1.000
		Sig. (2-tailed)	.000	.
		N	2108	2108
Spearman's rho	duration	Correlation Coefficient	1.000	.692(**)
		Sig. (2-tailed)	.	.000
		N	2108	2108
	amplitude	Correlation Coefficient	.692(**)	1.000
		Sig. (2-tailed)	.000	.
		N	2108	2108

** Correlation is significant at the 0.01 level (2-tailed).

表 1 重置时长和重置幅度的相关性统计结果

重置时长和重置幅度之间的 Kendall 和 Spearman 等级相关系数分别为 0.517 和 0.692，说明二者之间是正相关关系，重置时长越长，重置幅度越大。经过双尾 t 检验，发现其显著性水平为 0.01，也就是说重置时长和重置幅度之间的正相关性非常显著。

3.7 各种文体重置幅度的共同点

尽管各种文体呼吸重置幅度的频度分布各有其特点，但是仔细观察，也可以从中找到一些共性：

1) 虽然频度大小不一，但各种文体的呼吸重置幅度在 1.00 处都有一个小高峰。这是因为一段语流中大呼吸重置一般有 1~3 个，从数量上来看比小呼吸重置少很多。呼吸信号一般都是从 0 开始的，接着就达到整段语流呼吸信号的最高值。而我们的呼吸信号在进行归一化处理的时候，是把整段语流的呼吸信号最大值定为 1，因此每段语流基本都会出现一个幅度为 1 的重置，这样一来 1.00 的出现频率就会比邻近的数值多一些，在统计某种文体的重置幅度频度时，就会在 1.00 处出现一个高峰。

2) 重置幅度在 1.00 两侧都会有少数分布，这可能反映了一级呼吸重置的不同类型。通过观察呼吸信号，发现如果语流中有一句话很长，在说完这句话的时候发音人就会出现呼气过度的情况，这时呼吸信号就是负值，在这句话之后如果有一个大呼吸重置，其重置幅度就有可能超过 1.00。此外，一般位于语流中间位置的大呼吸的重置幅度不会大于位于语流最初的重置幅度，因此此时的重置幅度就小于 1.00。上述 2 种情况不如 1) 多见，所以频度上就会小一些。

3) 从二维图谱上的呼吸信号来看, 可以看到比较明显的 1、2、3 级呼吸重置, 但是从重置幅度的频度分布图上, 2 级和 3 级重置之间的界限则不太清楚。

上述分析表明, 不同文体朗读时的呼吸节奏从宏观上来看是具有一致性的。

4 结论和余论

通过分析, 本文对发音人朗读不同文体的语料时的呼吸节奏变化有了初步的了解, 并得到了以下结论:

1) 从宏观上看不同文体的呼吸重置幅度有一定的一致性: 呼吸信号的重置时长和重置幅度之间显著正相关。1 级呼吸重置的分布一致性较高, 2 级呼吸重置和 3 级呼吸重置之间的界线不是很清楚。

2) 不同文体的呼吸节奏有较明显的差异。宏观上, 近体诗、词等韵文的呼吸节奏比较简单, 其呼吸重置的大小大致可以分成两级。小说、散文、新闻等非韵文的呼吸节奏较复杂, 其呼吸重置可以大致分成三级。具体到韵文、非韵文内部, 不同文体朗读时的呼吸重置时长和幅度也有细微的不同。

3) 从数量上看, 小说、散文、新闻的大呼吸重置 (1 级呼吸重置) 最少, 小呼吸重置 (3 级呼吸重置) 最多。

4) 呼吸信号的重置幅度比重置时长更能反映不同文体的差别, 也更能反映不同呼吸级别的大小。

目前, 对语流中的呼吸节奏研究还刚刚起步, 有很多问题等着我们去探索。例如, 我们知道人们正常情况下只有呼气的时候才能说话, 吸气的时候不能说话, 因此呼吸信号在二维图谱中的上升段对应的就是语音信号的静音段。但是在对呼吸信号的波峰和波谷进行自动标注的时候我们发现了这样的情况:

图 18 是五言律诗《望岳》的语音、嗓音、呼吸信号。在自动标注呼吸信号的时候, 程序按照局部最大值的办法, 把波谷的标记打到了 A 的位置, 但对照语音和嗓音信号可以看到一直要到 B 的位置, 发音人才把“阴阳割昏晓”这句

诗朗诵完。A、B 之间的呼吸信号是上升的, 但这期间发音人一直在发音, 这种和我们已有的知识不符的情况也许是因为这次测量的是呼吸时胸腔的变化, 而此时发音人使用了腹式呼吸, 腹部的力量将横膈膜往上顶, 导致了胸腔扩大, 反映在图谱上就成了呼吸信号上升的假相。这种发音时胸腹部呼吸的交互作用就是一个很值得研究的问题。

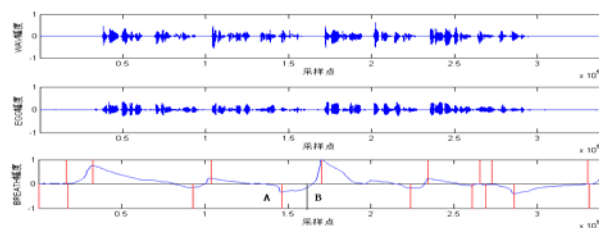


图 18 胸呼吸和腹呼吸交互作用示例

另外, 呼吸信号和语音信号、嗓音信号等声学参数之间有着什么样的关系? 根据呼吸信号划分出来的韵律单元和人们感知到的韵律单元是否一一对应? 呼吸节奏和语法结构之间有什么样的关系? 呼吸节奏是否能反映不同发音人的个人特征? 能否根据呼吸信号建立出不同风格的韵律模型用于语音合成? 只有对呼吸节奏作更深入的研究, 才能揭示这些问题的答案。

参考文献

- [1] 王洪君. 普通话中节律边界与节律模式、语法、语用的关联, 《语言学论丛》, 2002, 26: 279-300
- [2] 熊子瑜韵律单元边界特征的声学语音学研究, 《语言文字应用》, 2003, 2: 116-121
- [3] 王蓓、杨玉芳、吕士楠汉语韵律层级边界结构的声学相关物, 《声学学报》, 2004, 29 (1): 29-36
- [4] Liberman P. Intonation, Perception, and Language. Cambridge, Massachusettes. The MIT Press, 1967
- [5] Tseng C, Pin S, Lee Y, et al. Fluent speech prosody: Framework and modeling. Speech Communication. 2005, 46: 284-309
- [6] 谭晶晶. 新闻朗读的呼吸节奏初探. 孔江平. 第七届中国语音学学术会议论文集, 电子版
- [7] 陈亚勇. Matlab 信号处理详解, 人民邮电出版社 2001
- [8] 卢纹岱. SPSS for Windows 统计分析 (第 3 版). 北京. 电子工业出版社. 2006