

# 天津津南方言擦音的声学特性研究

马富榆

天津师范大学外国语学院, 天津

收稿日期: 2023年8月21日; 录用日期: 2023年9月7日; 发布日期: 2023年9月20日

## 摘要

声学参数在擦音研究中是不可或缺的, 以往研究基本是使用谱重心和分散度两个声学参数构建擦音格局, 并且鲜少涉及方言擦音的研究。本文探究了天津津南方言的5个擦音的8个声学参数(相对时长、谱顶点、谱矩、相对振幅、后接元音第二共振峰起点频率)与其语音特征的关系, 旨在补充天津方言的研究。方差分析和判别分析结果显示: 8个声学参数可以有效地区分不同的5个发音位置、2个后接元音以及发音人性别。

## 关键词

天津津南方言, 擦音, 声学参数, 统计

# Acoustic Characteristics of Fricatives in Tianjin Jinnan Dialect

Fuyu Ma

School of Foreign languages, Tianjin Normal University, Tianjin

Received: Aug. 21<sup>st</sup>, 2023; accepted: Sep. 7<sup>th</sup>, 2023; published: Sep. 20<sup>th</sup>, 2023

## Abstract

Acoustic parameter is indispensable in fricative study, and the previous studies mainly build fricative pattern by using spectral peak location and dispersion. What's more, the study on fricatives of dialects is rare. This study investigates the relationship between eight acoustic parameters (normalized duration, spectral peak location, spectral moment, normalized amplitude and F2 onset frequency) and phonetic features for five fricatives of Tianjin Jinnan dialect. The aim is to complement the study of Tianjin dialect. The analysis of variance and the discriminant analysis show that these 8 acoustic parameters can effectively distinguish 5 places of articulation, the following 2 vowels and the gender of the speakers.

## Keywords

Tianjin Jinnan Dialect, Fricative, Acoustic Parameter, Statistical Analysis

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

语音中有着许多声学参数, 这些声学参数承载着不同的信息, 不同语音的声学参数有所不同。相对于元音来说, 辅音的研究较少。针对辅音的研究多是构建辅音格局, 如冉启斌[1]使用谱重心和分散度两个声学参数作为二维坐标, 构建了擦音格局。利用声学参数的方法探究擦音最早是由 Svantesson [2]提出的, 他利用了三个参数来构建擦音格局。后来 Jongman [3]扩大了声学参数的数量, 使用时长、振幅、轨迹方程、频谱能量分布等 10 个声学参数对英语中的擦音进行了分析, 发现多个声学参数可以区分开来英语擦音的发音位置。李善鹏和顾文涛[4]使用 9 个声学参数探究普通话 6 个塞擦音的声学参数和塞擦音发音位置、送气状态和后接元音等的关系, 统计结果显示这 9 个声学参数可以对塞擦音的语音特征进行区分。除此之外, 汉语方言的擦音研究较少。

本文参考了以上的研究方法, 以天津津南方言的 5 个擦音(f, s, sh, h, x)为研究对象, 在 praat 软件上获取了相对时长、谱顶点、谱重心、分散度、偏度、峰度、相对振幅和后接元音第二共振峰起点频率等 8 个声学参数。除谱矩(谱重心、分散度、偏度和峰度)使用脚本外, 其余参数均为手动获取。之后使用 SPSS 软件中的方差分析、判别分析两种统计方法, 探究这些声学参数与发音位置、后接元音和发音人性别之间的关系, 以期这些声学参数可以区分方言中不同的擦音, 进一步加深对津南方言的认识。

## 2. 语料说明及录制

天津津南方言有五个擦音, 分别为唇齿音 f/f/, 齿腭音 s/s/, 卷舌音 sh/s/, 硬腭音 x/c/, 软腭音 h/x/ (后文中使用拼音形式), 它们的发音位置依次靠后。将五个擦音分别后接元音 a 和 i, 其中 i 是高元音, a 是低元音, 由此组成 10 个音节。当 x 后接 a 时, 为齐齿呼, 因此添加韵头 i 变为 xia。此外, 当 f 和 h 后接 i 时, 难以发音, 因此将其变为 ei, 即 fei 和 hei。

同时, 为了去除声调的影响, 每个音节读四个声调。对于其中没有对应汉字的音节, 采取拼音读法, 发音人在录音前提前训练, 以保可以自然地发音。

发音人为四男四女, 均为天津津南本地人, 在此长期生活。每个字由八位发音人重复录制两遍, 共计 5 (擦音)\*2 (后接元音)\*8 (发音人)\*4 (声调)\*2 (录制次数) = 640 个音节。录音均在发音人家中进行, 使用 RecForge 软件录制, 采样率为 44,000 赫兹。发音人提前熟悉字表, 保证录制时每个音节正确地读出。

## 3. 声学参数定义

### 3.1. 相对时长

摩擦段开始于波形图中开始有波形起伏的地方, 结束于元音周期波开始的地方。此外, 提取整个音节的时长, 计算摩擦段在整个音节中所占的比重, 由此得出摩擦段相对时长。

## 3.2. 频谱能量分布

表征频谱能量分布的声学参数有谱顶点和谱矩。其中谱矩包括谱重心、分散度、偏度和峰度。

### 3.2.1. 谱顶点

谱顶点代表频谱能量的上限，是摩擦段中点处振幅最高处的频率。

### 3.2.2. 谱矩

本文使用脚本<sup>1</sup>提取摩擦段中间 80% 部分的谱重心、分散度、偏度和峰度这四个声学参数。对其一定义：谱重心其实是平均值，用来表示频谱能量的集中区域；分散度即方差，表示频谱能量分布的离散程度；偏度是指频谱倾斜，如偏度为正，说明能量向负倾斜，即能量集中在低频区域；峰度表征能量的平坦或尖峰的分布，如峰度为正，则能量分布有尖峰，如峰度为负，则能量分布较为平坦。

## 3.3. 相对振幅

提取摩擦段的相对振幅。首先利用 praat 中的 query 选项提取摩擦段的均方根振幅值，利用公式将其转化。之后，提取元音段的振幅，元音段振幅为元音振幅最大点所在周期及前后相邻周期的振幅平均值。由摩擦段振幅减去元音段振幅，我们就得到了相对振幅值。<sup>2</sup>

## 3.4. 后接元音 F2 起始频率

后接元音的第二共振峰会携带辅音发音位置的信息，是衡量辅音的重要参数。

## 4. 方差分析

利用 SPSS 26.0 的三因素方差分析。以发音位置、后接元音、性别为固定因子，八个声学参数分别为因变量。统计结果如表 1 所示。

Table 1. The ANOVA result

表 1. 方差分析结果

	发音位置	后接元音	发音人性别	发音位置 × 后接元音	发音位置 × 发音人性别
相对时长	F (4,640) = 50.35, p < 0.001	F (1,640) = 31.34, p < 0.001	F (1,640) = 3.42, p = 0.065	F (4,640) = 4.81, p = 0.001	F (4,640) = 0.81, p = 0.518
谱顶点	F (4,640) = 5383.98, p < 0.001	F (1,640) = 2.03, p = 0.154	F (1,640) = 52.36, p < 0.001	F (4,640) = 2.24, p = 0.063	F (4,640) = 36.92, p < 0.001
谱重心	F (4,640) = 425.73, p < 0.001	F (1,640) = 0.148, p = 0.701	F (1,640) = 16.66, p < 0.001	F (4,640) = 1.94, p = 0.103	F (4,640) = 4.39, p = 0.002
分散度	F (4,640) = 115.16, p < 0.001	F (1,640) = 19.16, p < 0.001	F (1,640) = 49.65, p < 0.001	F (4,640) = 1.15, p = 0.330	F (4,640) = 1.94, p = 0.103
偏度	F (4,640) = 62.05, p < 0.001	F (1,640) = 0.002, p = 0.964	F (1,640) = 29.83, p < 0.001	F (4,640) = 2.77, p = 0.026	F (4,640) = 1.23, p = 0.295
峰度	F (4,640) = 4.85, p = 0.001	F (1,640) = 0.02, p = 0.888	F (1,640) = 4.53, p = 0.034	F (4,640) = 1.09, p = 0.362	F (4,640) = 1.14, p = 0.338
振幅	F (4,640) = 441.33, p < 0.001	F (1,640) = 108.26, p < 0.001	F (1,640) = 49.54, p < 0.001	F (4,640) = 3.63, p = 0.006	F (4,640) = 6.95, p < 0.001
元音 F2 起始频率	F (4,640) = 194.93, p < 0.001	F (1,640) = 320.40, p < 0.001	F (1,640) = 164.21, p < 0.001	F (4,640) = 6.53, p < 0.001	F (4,640) = 1.59, p = .0175

<sup>1</sup> 谱矩提取使用张明辉老师制作的脚本。

<sup>2</sup> 感谢李善鹏老师提供的相对振幅提取方法。

## 4.1. 相对时长

表 2 是摩擦段相对时长的均值。

**Table 2.** The mean of normalized duration (%)

**表 2.** 相对时长的均值(%)

	f		s		sh		x		h	
	a	i	a	i	a	i	a	i	a	i
女性	30.2	31.1	33.8	40.2	34.2	40.1	34.4	39.9	28.5	27.4
男性	28.5	31.9	35.9	39.3	32.4	38.3	33.7	37.9	26.2	24.8

不同发音位置的时长具有显著差异。唇齿音 f (30.5%) 和软腭音 h 的相对时长(26.8%)比齿腭音 s (37.4%)、卷舌音 sh (36.3%)、硬腭音 x (36.5%) 的短。

不同后接元音的时长具有显著差异。擦音后接 i 时的时长(35.1%)稍长于后接 a 时(31.8%)。

发音位置和后接元音具有显著的交互作用。齿腭音 s、卷舌音 sh、硬腭音 x 在的时长在后接元音 i 时更长。

## 4.2. 频谱能量分布

表 3 是谱顶点、谱重心、分散度、偏度和峰度的均值。

**Table 3.** The mean of spectral energy distribution

**表 3.** 频谱能量分布的均值

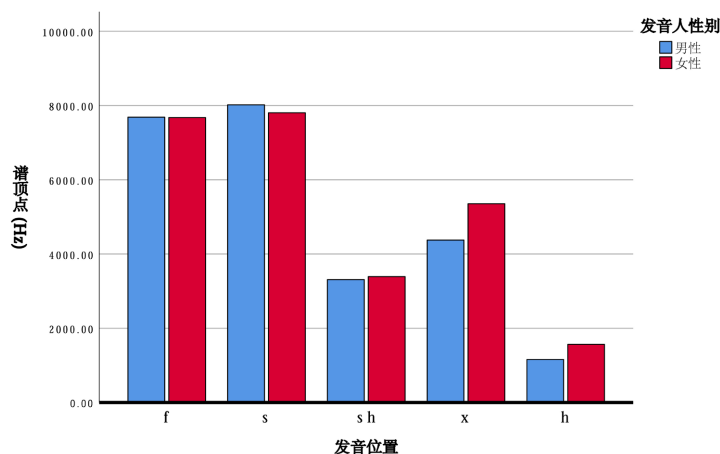
擦音	后接元音	谱顶点(Hz)	谱重心(Hz)	分散度(Hz)	偏度	峰度
f	a	7627.1	4926.0	2243.1	0.29	0.91
	i	7734.7	4771.1	2271.6	0.42	0.58
s	a	7822.0	6778.9	1264.6	-0.31	1.56
	i	7999.5	6551.3	1461.9	0.11	0.37
sh	a	3411.9	3685.8	1524.9	1.94	6.12
	i	3287.9	3592.3	1738.9	1.43	2.46
x	a	4825.1	5587.7	1298.8	0.50	1.16
	i	4903.1	5638.3	1401.0	0.41	0.15
h	a	1360.3	2649.3	1411.4	1.59	6.01
	i	1365.5	2946.6	1554.4	2.21	11.21

### 4.2.1. 谱顶点

不同发音位置的谱顶点具有显著差异。唇齿音 f、齿腭音 s、卷舌音 sh、硬腭音 x、软腭音 h 之间均有显著性差异。其中，齿腭音的谱顶点(7910.7 Hz)最高，软腭音的谱顶点(1362.9 Hz)最低。

不同性别发音人的谱顶点具有显著的主效应。女性的擦音谱顶点(5157.7 Hz)比男性的谱顶点(4909.7 Hz)更高。

发音位置和性别之间具有显著的交互作用。男性与女性在擦音谱顶点的差异，体现在齿腭音 s、硬腭音 x 和软腭音 h 上。下直方图图 1 体现了发音位置与性别之间的谱顶点分布。



**Figure 1.** The distribution of spectral peak location between the place of articulation and the gender of speaker  
**图 1.** 发音位置和性别之间的谱顶点分布

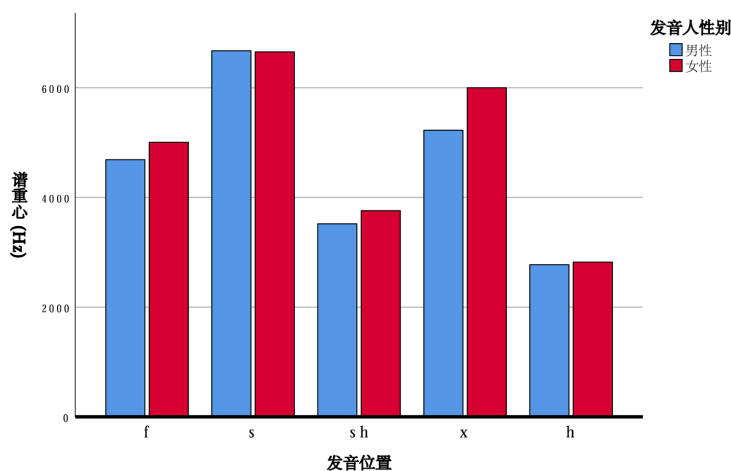
#### 4.2.2. 谱矩

##### 1) 谱重心

不同发音位置的谱重心具有显著差异。其中齿腭音 s 的谱重心(6665.13 Hz)最高，软腭音 h 的谱重心(2797.95 Hz)最低。

性别具有显著的主效应。女性的谱重心(4848.78 Hz)比男性的谱重心(4576.71 Hz)更高。

发音位置和性别之间具有显著的交互效应。女性发唇齿音 f 和硬腭音 x 时的谱重心比男性的谱重心更高。下直方图图 2 体现了发音位置和性别之间的谱重心分布。



**Figure 2.** The distribution of center of gravity between the place of articulation and the gender of speaker  
**图 2.** 发音位置和性别之间的谱重心分布

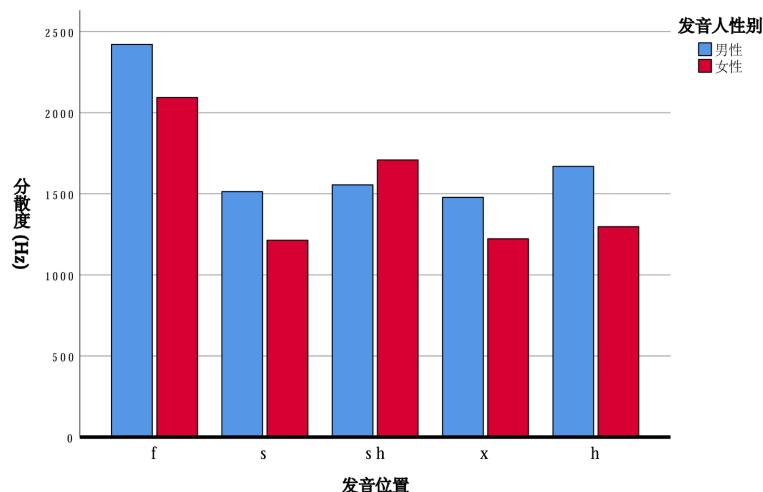
##### 2) 分散度

不同发音位置的分散度具有显著差异。唇齿音 f (2257.3 Hz)和卷舌音 sh (1631.9 Hz)均与齿腭音 s (1363.2 Hz)、硬腭音 x (1349.9 Hz)、软腭音 h (1482.93 Hz)之间存在显著性差异。唇齿音 f 的分散度最大，它的发音位置最靠前。

不同后接元音的分散度具有显著差异。擦音后接 i 时的频谱能量分散度(1685.6 Hz)比后接 a 时的分散度(1548.6 Hz)更高。

不同发音人性别的分散度具有显著差异。男性的分散度(1727.3 Hz)比女性的分散度(1506.8 Hz)更高。

发音位置和性别之间具有显著的交互作用。除了卷舌音 sh 以外,其他的擦音都是男性的分散度大于女性。对于女性来说,唇齿音 f 与卷舌音 sh 都与其他音有显著差异。而对于男性来说,只有唇齿音 f 与其他音有显著差异。下直方图图 3 体现了发音位置和性别之间的分散度分布。



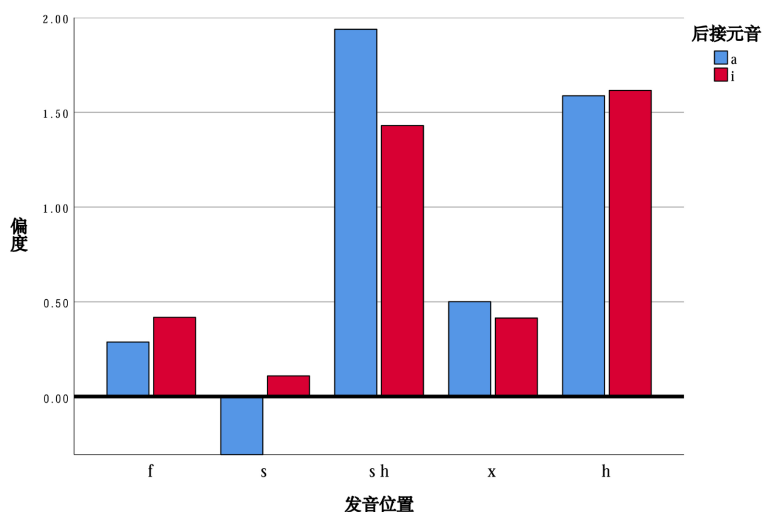
**Figure 3.** The distribution of dispersion between the place of articulation and the gender of speaker  
**图 3.** 发音位置和性别之间的分散度分布

### 3) 偏度

不同发音位置的偏度具有显著差异。唇齿音 f (0.35)和齿腭音 s (-0.10)、卷舌音 sh (1.68)、软腭音 h (1.60)之间均有显著差异,其中卷舌音 sh 的偏度最大,说明能量集中更集中在低频;齿腭音 s 的偏度为负,能量更集中在高频。

不同发音人性别的偏度具有显著差异。男性的偏度比女性的偏度大,表示男性的频谱能量更集中在低频。

发音位置和后接元音具有显著的交互作用。其中卷舌音 sh 后接不同元音时差异最大。下直方图图 4 体现了发音位置和后接元音之间的偏度分布。



**Figure 4.** The distribution of skewness between the place of articulation and the following vowel  
**图 4.** 发音位置和后接元音之间的偏度分布

## 4) 峰度

不同发音位置的峰度具有显著差异。其中软腭音 h (8.61)与唇齿音 f (0.75)、齿腭音 s (0.98)之间具有显著差异，软腭音 f 的峰度最高。

不同发音人性别的峰度具有显著差异。男性的峰度(4.55)比女性的峰度(1.56)更高。

## 4.2.3. 相对振幅

表 4 是相对振幅的均值。

Table 4. The mean of standardized amplitude (dB)

表 4. 相对振幅的均值(dB)

	f		s		sh		x		h	
	a	i	a	i	a	i	a	i	a	i
女性	-22.6	-21.0	-13.7	-9.3	-10.5	-7.8	-13.7	-10.5	-21.8	-19.4
男性	-23.0	-22.0	-14.7	-12.2	-13.6	-12.3	-14.3	-10.9	-22.2	-21.1

不同发音位置的相对振幅具有显著差异。其中卷舌音 sh 的振幅(-11.1 dB)与其他擦音都有显著差异，但唇齿音 f (-22.2 dB)与软腭音 h (-21.1 dB)之间以及唇齿音 s (-12.5 dB)与硬腭音 x (-12.3 dB)之间没有显著差异。卷舌音 sh 的相对振幅最高。

不同后接元音的相对振幅具有显著差异。擦音后接 i 时的振幅(-14.7 dB)显著高于后接 a 时的振幅(-17.0 dB)。

不同发音人性别的相对振幅具有显著差异。女性的振幅(-15.0 dB)高于男性的振幅(-16.6 dB)。

发音位置和性别之间具有显著的交互作用。对于齿腭音 s、卷舌音 sh 以及软腭音 h，振幅的男女差异显著，男性振幅低于女性振幅。图 5 体现了发音位置和性别之间的相对振幅分布。

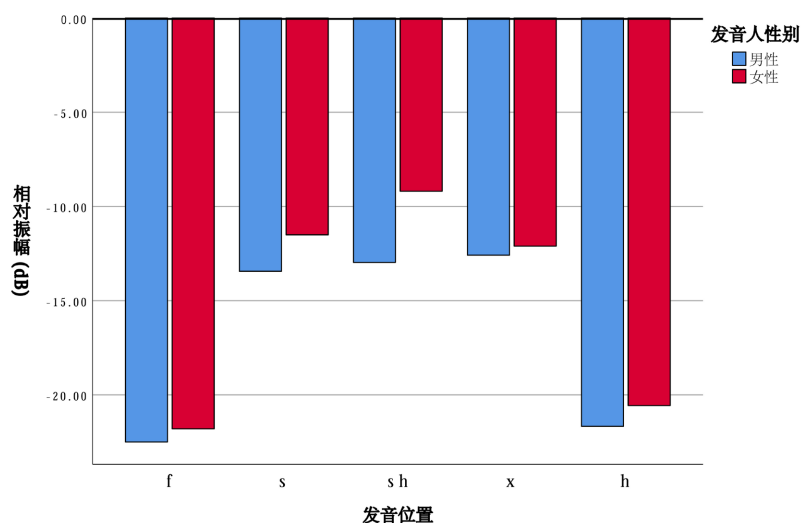


Figure 5. The distribution of standardized amplitude between the place of articulation and the gender of speaker

图 5. 发音位置和性别之间的相对振幅分布

发音位置和后接元音具有显著的交互作用。五个不同发音位置的擦音后接 i 时的振幅都显著高于后接 a 时的振幅。下直方图图 6 体现了发音位置和后接元音之间的相对振幅分布。

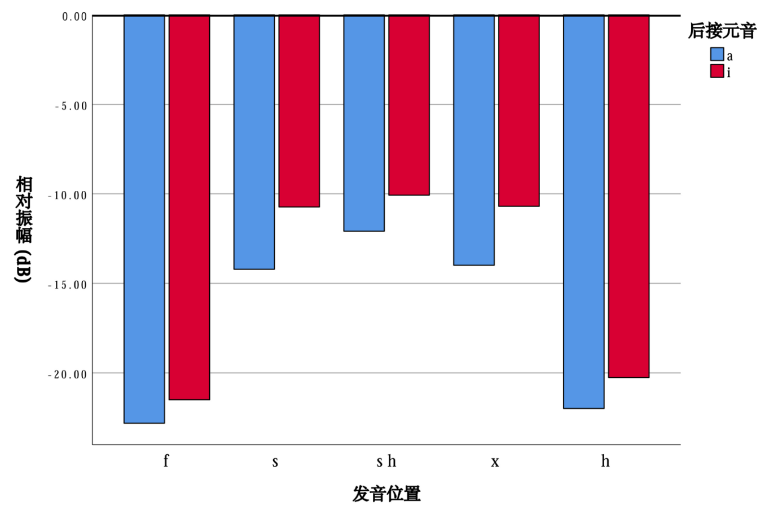


Figure 6. The distribution of standardized amplitude between the place of articulation and the following vowel  
图 6. 发音位置和后接元音之间的相对振幅分布

#### 4.2.4. 后接元音 F2 起始频率

表 5 是后接元音第二共振峰起始频率的均值。

Table 5. the mean of F2 onset frequency (Hz)

表 5. 后接元音 F2 起始频率的均值(Hz)

	f		s		sh		x		h	
	a	i	a	i	a	i	a	i	a	i
女性	1290.5	1743.2	1718.3	2029.7	1846.3	2400.5	2418.5	2570.6	1636.9	1928.0
男性	1117.8	1556.4	1336.8	1785.9	1540.8	2119.2	2004.0	2295.8	1287.7	1713.7

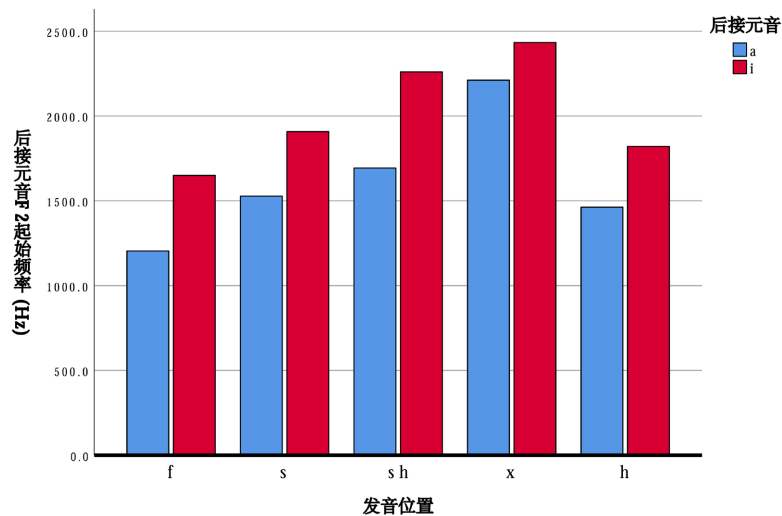


Figure 7. The distribution of F2 onset frequency between the place of articulation and the following vowel  
图 7. 发音位置和后接元音的 F2 起始频率之间的分布

不同发音位置的 F2 具有显著差异。除齿腭音 s (1717.7 Hz)与软腭音 h (1641.6 Hz)之间没有显著差异，其余均差异显著。硬腭音 x 的后接元音 F2 起始频率(2322.2 Hz)最高。



不同后接元音的 F2 具有显著差异。后接元音为 i 的 F2 起始频率(2014.3 Hz)高于后接元音为 a 的 F2 起始频率(1619.8 Hz)。

不同发音人性别的 F2 具有显著差异。女性发音人的 F2 (1958.25 Hz)比男性发音人的 F2 (1675.8 Hz) 更高。

发音位置和后接元音具有显著的交互作用。其中卷舌音后接元音 i 和后接元音 a 时的差异最大。上直方图图 7 是发音位置和后接元音的 F2 起始频率之间的分布。

## 5. 判别分析

判别分析结果显示, 时长、振幅等八个声学参数对擦音五个发音位置的辨别率是 96.4%。其中唇齿音 f 和软腭音 h 的辨别率最高, 高达 100%。硬腭音 x 的辨别率最低, 为 83.6%, 有 16.4% 误判为卷舌音 sh。

八个声学参数对两个后接元音的辨别率是 73.8%。对元音 a 的辨别率是 74.4%, 对元音 i 的辨别率是 73.1%。

对发音人性别的辨别率是 73.0%。女性的辨别率(75.9%)比男性的辨别率(70.0%)更高。

## 6. 讨论与总结

根据方差分析和判别分析, 我们可以得出以下结论。

第一, 发音位置对八个声学参数都具有显著的主效应, 但是可以完全区分五个发音位置只有谱重心和谱顶点。从谱顶点和谱重心的均值来看, 发音位置越靠后, 频谱能量越低, 比如软腭音 h。

第二, 可以区分两个后接元音的声学参数是相对时长、振幅、元音 F2 起始频率和分散度这四个声学参数。其中特点是当擦音后接高元音 i 时, 相对时长更长, 振幅更高, 频谱能量分散度更大, 元音 F2 起始频率更高。

第三, 可区分发音人的性别的声学参数是相对振幅、元音 F2 起始频率、谱顶点、谱重心、分散度、偏度和峰度。其中观察发现, 谱顶点、谱重心、相对振幅和后接元音 F2 起始频率为女性更高; 分散度、偏度和峰度为男性更高。

最后, 方差分析和判别分析的结果综合显示, 相对时长、频谱能量分布、相对振幅, 元音 F2 起始频率等八个声学参数可以用来区分天津津南方言擦音的五个发音位置、两个后接元音以及发音人性别。

## 参考文献

- [1] 冉启斌. 辅音声学格局研究[J]. 当代外语研究, 2011(9): 12-16.
- [2] Svantesson, J. (1986) Acoustic Analysis of Chinese Fricatives and Affricates. *Journal of Chinese Linguistics*, **14**, 53-70.
- [3] Jongman, A., et al. (2000) Acoustic Characteristics of English Fricatives. *The Journal of the Acoustical Society of America*, **108**, 1252-1263. <https://doi.org/10.1121/1.1288413>
- [4] 李善鹏, 顾文涛. 普通话塞擦音的声学特性研究[J]. 清华大学学报(自然科学版), 2016, 56(11): 1202-1208