



## Does Posh English Sound Attractive?

Li Jiao<sup>1</sup>, Chengxia Wang<sup>2</sup>, Cristiane Hsu<sup>2</sup>, Peter Birkholz<sup>3</sup>, Yi Xu<sup>2</sup>

<sup>1</sup>School of Foreign Languages, Tongji University, China

<sup>2</sup>Department of Speech, Hearing and Phonetic Sciences, University College London, UK

<sup>3</sup>Institute of Acoustics and Speech Communication, Technische Universität Dresden, Germany

2013li\_jiao@tongji.edu.cn, chengxia.wang.15@ucl.ac.uk, cristiane.hsu.11@ucl.ac.uk,  
peter.birkholz@tu-dresden.de, yi.xu@ucl.ac.uk

### Abstract

Poshness refers to how much a British English speaker sounds upper class when they talk. Popular descriptions of posh English mostly focus on vocabulary, accent and phonology. This study tests the hypothesis that, as a social index, poshness is also manifested via phonetic properties known to encode vocal attractiveness. Specifically, posh English, because of its impression of being detached, authoritative and condescending, would more closely resemble an attractive male voice than an attractive female voice. In four experiments, we tested this hypothesis by acoustically manipulating Cambridge-accented English utterances by a male and a female speaker through PSOLA resynthesis, and having native speakers of British English judge how posh or attractive each utterance sounds. The manipulated acoustic dimensions are formant dispersion, pitch shift and speech rate. Initial results from the first two experiments showed a trend in the hypothesized direction for the male speakers' utterances. But for the female utterances there was a ceiling effect due to the frequent alternation of speaker gender within the same test session. When the two speakers' utterances were separated by blocks in the third and fourth experiments, a clearer support for the main hypothesis was found.

**Index Terms:** posh English, vocal attractiveness, formant ratio shift, pitch shift, duration

### 1. Introduction

The word *posh* is a popular English term to refer to the *upper class* [1], including the way they talk, which is typically associated with the Received Pronunciation (RP) [2-3]. Social studies on English accents of different classes have mainly focused on vocabulary and grammar [4-5], although there is also some discussion of pronunciation in terms of choice of consonants, vowels and intonation [3-4, 6]. In this paper, we explore the possibility that even within the same dialect or accent, the level of poshness can vary due to certain phonetic characteristics of the voice.

In this regard, so far there are only a variety of anecdotal and impressionistic descriptions. Some claim that posh speech is much further forward in the mouth, and that the mouth is held tightly [7], or the jaw sticks out and does not move very much [8]. There are also claims that posh speech is slow because posh people are never in a hurry [8]. On the other hand, there is also the impression that posh people are less able to show their emotions in public [9]. Thus, posh speech may sound cold, detached and condescending, which may not be liked by many people [9-11].

The claim that posh speech is more forward in the mouth and the claim that it is cold, detached and condescending are actually contradictory to each other, based on the notion of *frequency code* [12] or *body size projection* [13]. According to this notion, animals, including humans, use natural correlations between body size and the acoustic properties of vocalization as a code in their communication. Other things being equal, a large body size is naturally associated with low pitch, condensed formants and rough voice quality. These qualities are therefore often mimicked or even exaggerated to signal dominance, intimidation and aggressiveness [12, 14-16]. In humans, low pitch, condensed formants and tense voice are used to express anger [13], [17-18] and dominance [12]. Conversely, high pitch, dispersed formants and breathy voice are used to express happiness and friendliness [17-19].

In contrast to the expression of emotion and attitude, vocal attractiveness shows a dichotomy between sexes. Female voices that sound attractive to male listeners have relatively high pitch, wide formant dispersion and breathy quality, which all project a small body size, whereas male voices that sound attractive to female listeners have low pitch and narrow formant dispersion, both projecting a large body size [13]. The only exception to this dichotomy is that an attractive male voice, just like female voice, is also breathy [13].

Assuming that poshness is at least partially signaled by acoustic properties, the question then is, is it spoken more in the front of the mouth as per anecdotal observation, which would resemble an attractive female voice and signal friendliness, or is it spoken with a lengthened vocal tract (hence condensed formants) and lowered pitch, thus signaling dominance, authority and detachment? Given that people, even phoneticians, are generally better at judging social and emotional meanings than making formal phonetic observations [20], the latter is a more likely possibility. To test this prediction, we performed four experiments to obtain perceptual judgments of poshness and attractiveness from resynthesized English utterances that are manipulated in terms of formant dispersion, pitch shift and speech rate.

### 2. Materials and Method

#### 2.1. Stimuli

Stimuli for all four perception experiments were generated through PSOLA resynthesis of natural utterances. Four short statements (in Italics below) produced by one male and one female secondary school students from Cambridge (both 16 years old), were selected from the Intonational Variation in English (IViE) corpus [21] as the base sentences.

1. *You remembered the lilies.*
2. *They are on the railings.*
3. *He is on the lilo.*
4. *You are feeling mellow.*

Each base sentence was first normalized in total duration between the two speakers. Then all the utterances were modified in terms of the following parameters, using the “Change gender” function of Praat [22]:

1. Formant shift ratio (1.1, 1.0 and 0.9) — Expanded (1.1) or compressed (0.9) spectrum to simulate a shortened or lengthened vocal tract.
2. Pitch shift (-3, 0, 3) — Lowered or raised median pitch of the utterance by 3 semitones.
3. Duration [Experiments 1 and 2: (1.0, 1.2 and 1.4); Experiments 3 and 4: (0.8, 1.0 and 1.2)] — Shortened or lengthened total duration of the utterance by the ratios indicated.

This resulted in 4 base sentences \* 2 speakers \* 3 formant ratios \* 3 pitch shifts \* 3 durations = 216 unique stimuli. All stimuli were repeated twice, so that there were 432 trials in total for each of the experiments.

## 2.2. Procedure

All the experiments were run with ExperimentMFC scripts for Praat. Subjects were asked to listen to each utterance and judge how posh (Experiments 1 and 3) or attractive (Experiments 2 and 4) it sounded on a five-level scale. Subjects were seated in front of a computer screen and listened to the stimuli through headphones. Before the real experiments, they received an introduction and went through a short practice to familiarize themselves with the procedure. There was no feedback on how they should judge the stimuli. In each trial the utterance was played only once.

### 2.2.1. Experiments 1 and 2

Fifteen native English speakers (average age: 21.6; 9 females) participated in the first two experiments in a laboratory at the Department of Speech, Hearing and Phonetic Sciences, UCL. All subjects first did the poshness experiment and then the attractiveness experiment. In each experiment, the male and female voices were randomized together.

### 2.2.2. Experiments 3 and 4

These two experiments were conducted to ameliorate a problem arising from the design of Experiments 1 and 2, as will be discussed later in the results section. There were two modifications from Experiments 1 and 2. First, unlike in those experiments, where duration was either lengthened or the original, here duration was manipulated in two directions to generate both a faster and a slower version of each base sentence. Second, and more importantly, the male and female utterances were separated into two blocks rather than randomized together. In each experiment, subjects heard all the male stimuli first and then all the female stimuli. Voices of the same gender were again randomly ordered. Participants were 12 native British English speakers (average age: 24.3; 6 females).

## 3. Results

The 1-5 scale responses were first normalized to [0, 1]. They were then analyzed by four-way Repeated Measures ANOVAs. The independent variables were speaker gender, formant shift ratio, pitch shift and (total) duration.

### 3.1. Experiment 1: Poshness judgment

Figure 1 shows poshness scores as a function of speaker gender and formant shift ratio. There is a strong speaker gender effect ( $F(1, 14) = 55.75, p < 0.0001$ ). Female stimuli were judged to be much posher than male ones (mean score: 0.666 vs. 0.283). There were interactions between speaker gender and formant ratio ( $F(2, 28) = 5.47, p = 0.0099$ ) and between speaker gender and pitch shift ( $F(2, 28) = 6.81, p = 0.0039$ ). There is a three-way interactions between speaker gender, pitch shift and formant shift ratio ( $F(4, 56) = 2.97, p = 0.0272$ ). There is also a marginal three-way interaction between speaker gender, pitch shift and duration ( $F(4, 56) = 2.69, p = 0.0402$ ).

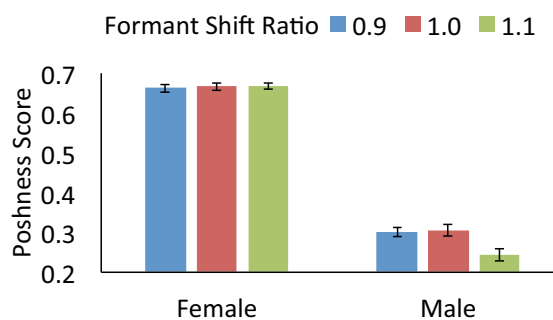


Figure 1: Mean poshness scores as a function of speaker gender and formant shift ratio. The error bars indicate standard errors.

These data show a strong bias in favor of the female speaker. It seems that listeners decided early on in the experiment that the female speaker was posher than the male speaker, and largely ignored the other variations in the stimuli. This is exacerbated by the frequent alteration of speakers within the same block, which allowed the gender effect to mask most of the other effects.

However, a trend could still be seen for the male stimuli in Figure 1: poshness score is the lowest when the ratio is the highest (1.1, which reflects a short vocal tract) and higher when the ratio is the original (1.0) or the lowest (0.9, which reflects a long vocal tract).

### 3.2. Experiment 2: Attractiveness judgment

Figure 2 shows attractiveness scores as a function of speaker gender and formant shift ratio. Similar to Experiment 1, there is a strong preference for the female voice ( $F(1, 14) = 16.97, p = 0.001$ ). Besides that, there is only a significant interaction between speaker gender and formant shift ratio ( $F(2, 28) = 5.38, p = 0.0105$ ).

Nevertheless, a trend can be seen in Figure 2 that male voices with longer vocal tract are heard as more attractive. This trend is in line with previous findings that a narrow

formant dispersion signals a large body size, which makes a male voice attractive [13], [17-18]. The female results did not show a clear trend, however.

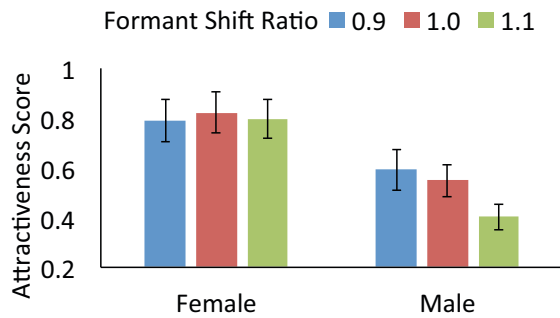


Figure 2: Mean attractiveness scores as a function of speaker gender and formant shift ratio. The error bars indicate standard errors.

### 3.3. Experiment 3: Poshness Judgment (gender blocked)

Free of the masking effect in Experiment 1 due to frequent alternation of male and female voices, results of Experiment 3 show clear main effects of speaker gender ( $F(1, 11) = 20.84, p = 0.0008$ ), formant shift ratio ( $F(2, 22) = 13.28, p = 0.0002$ ), pitch shift ( $F(2, 22) = 4.32, p = 0.0261$ ) and duration ( $F(2, 22) = 5.47, p = 0.0118$ ). Specially, the female voice was judged as more posh than the male voices (mean: 0.582 vs. 0.403). In terms of formant shift ratio, as shown in Figure 3, poshness score is the highest when the ratio is the lowest (0.9), and lowest when the ratio is the highest (1.1). A Student-Newman-Keuls post-hoc test shows that both the 0.9 and 1.0 ratios are significantly different from the 1.1 ratio. This is in accordance with the trend in the male voice of the first experiment in Figure 1. For pitch shift, the Student-Newman-Keuls test shows that lower pitch (-3) received significantly lower scores than the original pitch (0) (0.474 vs. 0.509). Finally, Student-Newman-Keuls test shows that both the original (1.0) and the shortened (0.8) durations have significantly higher scores than the lengthened duration (1.2) (0.51 vs. 0.456, 0.511 vs. 0.456).

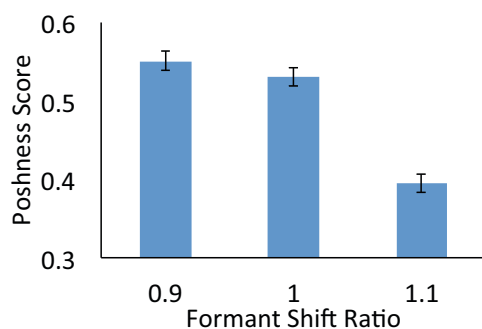


Figure 3: Main effect of formant shift ratio for mean scores of poshness. The error bars indicate standard errors.

Unlike the results of Experiment 1 (Figure 1), although the female utterances are still rated higher than the male ones, there is no interaction between speaker gender and formant shift ratio ( $F(2, 22) = 0.89, p = 0.4262$ ), as can be seen in the upper panel of Figure 4. In other words, for both the male and female speakers, the longer the vocal tract, the posher the voice sounds.

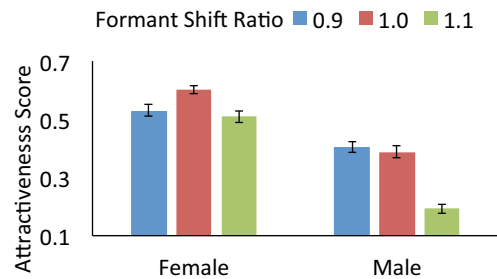
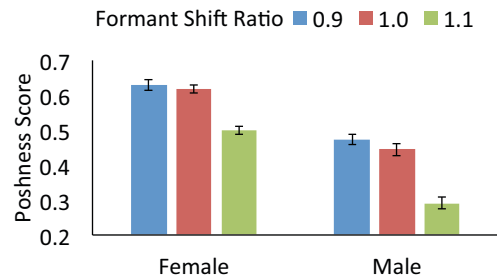


Figure 4: Upper panel: Mean poshness scores as a function of speaker gender and formant shift ratio. Lower panel: Mean attractiveness scores as a function of speaker gender and formant shift ratio. The error bars indicate standard errors.

There are also interactions between speaker gender and duration ( $F(2, 22) = 4.64, p = 0.0209$ ) (upper panel, Figure 5), formant ratio and pitch shift ( $F(4, 44) = 5.50, p = 0.0011$ ) (upper panel, Figure 6) and pitch shift and duration ( $F(4, 44) = 5.29, p = 0.0014$ ).

### 3.4. Experiment 4: Attractiveness Judgment (gender blocked)

Similar to Experiment 3, a four-way Repeated Measures ANOVA shows main effects of speaker gender ( $F(1, 11) = 16.77, p = 0.0018$ ), formant shift ratio ( $F(2, 22) = 8.22, p = 0.0022$ ), pitch shift ( $F(2, 22) = 5.45, p = 0.0119$ ) and duration ( $F(2, 22) = 11.82, p = 0.0003$ ). Specifically, the female voice is heard as more attractive than the male voice (0.546 vs. 0.327). Student-Newman-Keuls post hoc tests show significantly higher scores for formant ratios of 0.9 and the original ratio than for the 1.1 ratio (as can be seen in the lower panel of Figure 4), higher scores for pitch shifts of -3 and the original pitch than for the +3 shift (0.447 vs. 0.404, 0.46 vs. 0.404), and higher scores for duration ratios of 0.8 and the original (1.0) than for the 1.2 ratio (0.481 vs. 0.38, 0.45 vs. 0.38).

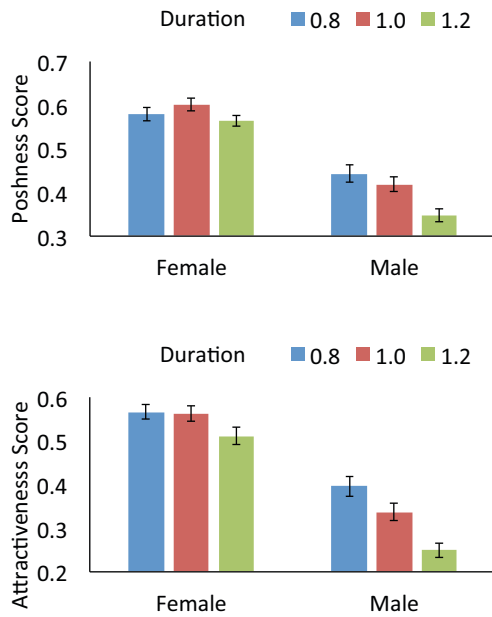


Figure 5: Upper panel: Mean poshness scores as a function of speaker gender and duration. Lower panel: Mean attractiveness scores as a function of speaker gender and duration. The error bars indicate standard errors.

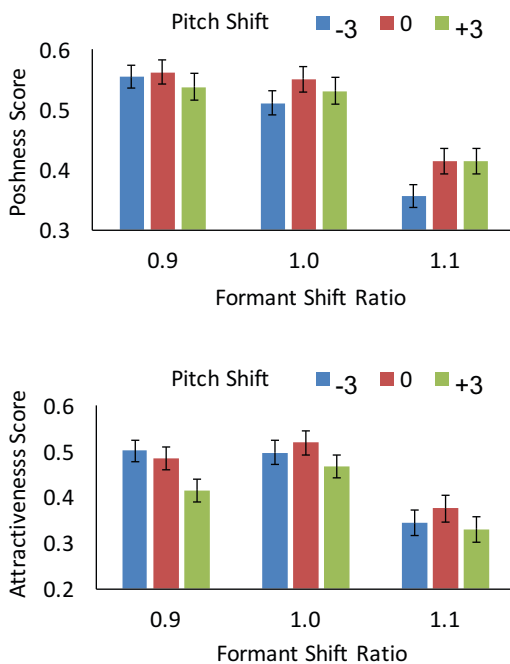


Figure 6: Upper panel: Mean poshness scores as a function of formant shift ratio and pitch shift. Lower panel: Mean attractiveness scores as a function of formant shift ratio and pitch shift. The error bars indicate standard errors.

There are two-way interactions between speaker gender and formant shift ratio ( $F(2, 22) = 10.58, p = 0.0006$ ) (lower panel, Figure 4), speaker gender and duration ( $F(2, 22) = 6.75, p = 0.0052$ ) (lower panel, Figure 5), formant shift ratio and pitch shift ( $F(4, 44) = 4.13, p = 0.0063$ ) (lower panel, Figure 6), formant shift ratio and duration ( $F(4, 44) = 2.60, p = 0.0488$ ). There is also a three-way interaction among speaker gender, formant shift ratio and duration ( $F(4, 44) = 7.53, p = 0.0001$ ).

## 4. Discussion and Conclusion

To test the hypothesis that poshness can be displayed not only through vocabulary and accent, but also through non-segmental phonetic parameters that are related to vocal attractiveness, we conducted four perception experiments. The results of the first two experiments were masked by a strong perceptual bias towards the female voice. Nevertheless, the male voice showed a trend that a lengthened vocal tract was heard both as posh and attractive. The next two experiments, which were free of the masking effect, showed clearer patterns. Both the male and female voices sounded more posh when there was a narrower formant dispersion, hence longer vocal tract. What is particularly significant is that female poshness is in the same direction as male poshness, which is the opposite of an attractive female voice [13]. This confirms our prediction that regardless of gender, poshness is more in line with an attractive male voice.

The confirmation of our initial hypothesis indicates that poshness can indeed be partially conveyed through phonetic dimensions that are involved in the expression of emotion [16-18], attitude [19] and vocal attractiveness [13], which all make use of the body size projection principle [14, 12]. This principle allows us to interpret the meanings associated with the posh accent, and how it relates to vocal attractiveness. The small formant dispersion associated with poshness found in the present study indicates that sounding posh is to project a large body size, which generates an impression of superiority and dominance. In a separate study, we have further found that poshness is also perceptually associated with a breathy quality for both male and female voices, just like attractiveness [23]. The breathiness is interpreted as a means to moderate the level of arrogance that might be associated with superiority in a posh accent.

Unlike popular impressions mentioned in the introduction, however, poshness is found to be associated with faster rather than slower speech rates, as seen in Figure 5. For pitch, despite the significant main effects, there is no consistent directions, as can be seen in Figure 6,

As a caveat, the present findings by no means deny the role of the other factors discussed in the literature, including dialectical differences, vocabulary, grammar and intonation. Future studies can compare the relative contributions of all these factors.

## 5. Acknowledgements

We thank Mia and Judy for their vivid mimicry of posh English, which inspired this study.

## 6. References

*Tutorial and Research Workshop on Experimental Linguistics, June 19-22, Heraklion, Crete, Greece, Proceedings, 2017.*

- [1] “posh,” *Paperback Oxford English Dictionary*. Oxford University Press, Oxford, p. 559, 2012.
- [2] J. Spencer, “Received Pronunciation: Some problems of interpretation,” *Lingua*, vol. 7, pp. 7-29, 1957.
- [3] A. Agha, “The social life of cultural value,” *Language & Communication*, vol. 23, no. 3-4, pp. 231-273, 2003.
- [4] A. S. Ross, “Linguistic class-indicators in present-day English,” *Neuphilologische Mitteilungen*, vol. 55, no. 1, pp. 20-56, 1954.
- [5] K. Fox, *Watching the English*, London: Hodder and Stoughton Ltd, 2004.
- [6] L. Van Buuren, “Margaret Thatcher's pronunciation: An exercise in ear-training,” *Journal of the international phonetic association*, vol. 18, no. 1, pp. 26-38, 1988.
- [7] “How to sound posh - Part one,” *YouTube*, 2015. [Online]. Available: [https://www.youtube.com/watch?v=Nj0Rh\\_\\_1kDw](https://www.youtube.com/watch?v=Nj0Rh__1kDw). [Accessed: 11- Mar- 2017].
- [8] “How to Speak POSH English,” *YouTube*, 2013. [Online]. Available: <https://www.youtube.com/watch?v=BLYPhCR4S1M>. [Accessed: 12- Mar- 2017].
- [9] “Why Posh People are Cold,” *YouTube*, 2014. [Online]. Available: <https://www.youtube.com/watch?v=rKAGTKAwgqE>. [Accessed: 12- Mar- 2017].
- [10] C. Howse, “When you sound too posh to earn the dosh,” *Telegraph.co.uk*, 2012. [Online]. Available: <http://www.telegraph.co.uk/news/uknews/theroyalfamily/9197850/When-you-sound-too-posh-to-earn-the-dosh.html>. [Accessed: 12- Mar- 2017].
- [11] J. Cheshire and V. Edwards, “Schoolchildren as sociolinguistic researchers,” *Linguistics and Education*, vol. 3, no. 3, pp. 225-249, 1991.
- [12] J. J. Ohala, “An ethological perspective on common cross-language utilization of F<sub>0</sub> of Voice,” *Phonetica*, vol. 41, pp. 1-16, 1984.
- [13] Y. Xu, A. Lee, W. L. Wu, X. Liu, and P. Birkholz, “Human vocal attractiveness as signaled by body size projection,” *PLoS ONE*, vol. 8, no. 4, p. e62397, 2013.
- [14] E. S. Morton, “On the occurrence and significance of motivation-structural rules in some bird and mammal sounds,” *The American Naturalist*, vol. 111, no. 981, pp. 855-869, 1977.
- [15] W. T. Fitch, “Acoustic exaggeration of size in birds via tracheal elongation: comparative and theoretical analyses,” *Journal of Zoology*, vol. 248, pp. 31-48, 1999.
- [16] D. Reby, K. McComb, B. Cargnelutti, C. Darwin, W. T. Fitch and T. Clutton-Brock, “Red deer stags use formants as assessment cues during intrasexual agonistic interactions,” *Proceedings of the Royal Society B: Biological Sciences*, vol. 272, no. 1566, pp. 941-947, 2005.
- [17] S. Chuenwattanapranithi, Y. Xu, B. Thipakorn and S. Maneewongvatana, “Encoding emotions in speech with the size code,” *Phonetica*, vol. 65, no. 4, pp. 210-230, 2009.
- [18] Y. Xu, A. Kelly, and C. Smillie, “Emotional expressions as communicative signals,” *In Prosody and Iconicity*. S. Hancil and D. Hirst. Philadelphia: John Benjamins Publishing Company, pp. 33-60, 2013.
- [19] L. Noble and Y. Xu, “Friendly speech and happy speech – Are they the same?” *in ICPHS 2011-17th International Congress of Phonetic Sciences, August 17-21, Hong Kong, Proceedings, 2011*, pp. 1502–1505.
- [20] P. Ladefoged, *Three areas of experimental phonetics*, London: Oxford University Press, 1967.
- [21] E. Grabe, F. Nolan, and K. J. Farrar, “IViE - a Comparative transcription system for intonational variation in English,” *in ICSLP'98-5<sup>th</sup> International Conference on Spoken Language Processing, November 30-December 4, Sydney, Australia, Proceedings, 1998*, vol.4, pp. 1259-1262.
- [22] P. Boersma, “Praat, a system for doing phonetics by computer,” *Glott International*, vol. 5, pp. 341-345, 2001.
- [23] L. Jiao, C. Wang, C. Hsu, P. Birkholz and Y. Xu, “Posh accent and vocal attractiveness in British English,” *in ExLing2017: 8<sup>th</sup>*