

Sociophonetic factors of speakers' sex differences in Voice Onset Time: a Florentine case study

State of the art. Voice Onset Time (VOT) is one of the most studied subsegmental features in relation to the social variables of speech communities. In particular, studies on the relationship between speaker's sex and VOT length have recently shifted from exegetical approaches based on universal sex-determined physiological features to a growing attention on the local sociophonetic realities. Swartz 1992 [1] first noticed a systematic tendency for female English speakers showing greater VOT values for voiceless stops. In the following years, many scholars confirmed [2, 3, 4, 5, 6, 7, 8] Swartz's pattern: that led to a series of explanations of the phenomenon seen as a universal of speech production (i.e. men's wider supraglottal space, women's shorter and stiffer vocal folds [1, 9, 3, 4, 5, and earlier in 10, 11] and the female tendency to enhance speech accuracy [4]). Two factors broke this established point of view:

A) Studies on this relationship, focused on developmental age [12, 6], showed that male English speakers around the age of 9 have a dramatic increase in VOT lengths, simultaneous to a drop of f_0 values. These changes, not shown by the female counterparts, are related to sex-determined physical growth. VOT values become adultlike around the age of 11: from the English male speakers' point of view, this is an *inhibition* of a physical feature, determined by sociophonetic factors.

B) Starting with Oh 2011 [13], VOT analysis in non-English-speaking communities showed male speakers with greater values (e.g. for the unaspirated voiceless series [14, 15], and also [16]). Oh explicitly underlined the necessity to dig into the local sociolinguistic structures to find satisfying explanations to the experimentally acquired VOT patterns.

However, to our knowledge, no one ever tried to link the findings to questions that are, or should be [17], central in the sociophonetic science, i.e.: what type of indexicality VOT values hold? What is speakers' evaluative perception of the trait? Is there a structurally motivated reason to the selection of a specific VOT-sex pattern?

The research. In order to obtain a synchronic sociophonetic description of the so-called “gorgia enfatica” (the allophonic presence of voiceless aspirated stops in strong position in the vernacular variety of Florence [18], [19]), we linked the quantitative analysis of a production test to the qualitative analysis of a perceptual dialectological session about the subjective evaluation of the trait. Our three-step protocol consisted in: 1) session of read speech: story-telling task and 60 sentences built on a selection of target words and written in a morphosyntactically marked variety of Regional Tuscan Italian; 2) “pseudo-matched guise”: listening and evaluative task of three digitally reproduced couples of sentences pronounced by the researcher (native speaker), spectrographically checked to differ only by VOT durations (\pm 30 ms.); 3) open-ended interview based on a deductive outline: informal conversation led by the researcher to gradually dig into the metalinguistic competence of the interlocutor. The interviewed subjects were 24, all born and living in Florence, divided in M VS F (sex variable), I VS II VS III (age variable) and G VS NG (Graduates VS NonGraduates, class variable distinction based on Hudson's idea of “sophistication” [20]). Through manual annotation of spectrograms we collected the VOT values of 1041 [k], 2287 [t] and 738 [p], discerning the following contexts: C-, -CC-, -sC-, -C:- (with particular attention to syntactic doubling). We considered valid for our analysis only the VOTs preceded by a release burst [21]. We also took account of many of the prosodic factors that have a direct positive or negative influence on VOT values, such as initial strengthening, prepausal lengthening and those related to emotional speech (pitch rising and hyperarticulation). The raw data showed a general tendency $M > F$, correlated to the production of clearly perceptible aspirated allophones. ANOVA tests were conducted to analyse the statistical significance of the interactions between our data and the social variables. $M > F$ was confirmed as statistically significant ($p < 0,05$) for $k-$, $-Ck-$, $p-$, $-p-$, $-sp-$ and all the contexts containing [t]. Speaking of the class variable, all the contexts containing [k] and [t]

showed a statistically significant tendency $NG > G$, confirming the traditional distribution of the “gorgia enfatica” (dating back to the *AIS* [22]). No significant interactions were found with the age variable, according to recent studies on this relationship [23]. We also noticed that some of the micro-categories of speakers showed greatly significant differences in terms of VOT lengths, i.e. $M I G > F I G$ ($p = 5,01^{-21}$), $M II NG > M II G$ ($p = 6,03^{-21}$) and $M III NG > F III NG$ ($p = 9,59^{-16}$). This peculiar distribution found a suggestive explanation with the perceptual sections of the interviews. The matched guise experiment found identifications of the aspirated allophones as Florentine in the frequency order of $[k] > [t] > [p]$ and, speaking of the social variables, of $III > I > II$. The voice using the aspirated allophones was often considered manlier than the other, and more appropriate in informal situations (mainly groups of male friends, but also family figures: fathers, grandfathers, husbands...). During the open-ended interviews we decided to explore the folk terminology relative to Florentine dialectal traits, in order to naturally propose a metalinguistic reflection on the “gorgia enfatica”. We noticed that metalinguistic sensibility decreased with the aging ($I > II > III$) and increased with the level of instruction ($G > NG$). Particularly relevant for our research was the recurrence of a referential ambiguity in the use of the word “aspirazione”. People tend to speak of “aspirazione” in relation to both weak fricative allophones $[h, \theta, \varphi]$ and strong aspirated ones $[k^h, t^h, p^h]$, suggesting that the term generally refers to an acoustic impression of consonantal segments containing continuous traits. Guided pronunciation of aspirated allophones was a key element of the reflection about the “gorgia enfatica”, that opened the way to detailed description of the phenomenon, especially by I. The “gorgia enfatica” is considered virile, archaic and rustic: these traits hold a positive connotation among the young males, but find a firm rejection among the female counterparts. This distribution suggests that the “gorgia enfatica” is considered a *covert prestige* trait, opposed to the generally accepted “gorgia”. Young males took the “gorgia enfatica” as an important element of their “youth language” from a noticeable distribution of the strong allophones among the elders: they are not only conscious of the trait, but also explicitly aware of it. From this point of view, our third greatly significant opposition ($M II NG > M II G$) represents the axis of the generational transmission of the trait, showing the possibility of *retrenchment* phenomena. The “gorgia enfatica” can thus be considered a *marker* of manliness in the Florentine dialectal system. What about the reason behind the choice of the aspirated allophone itself? The presence of aspirated allophones for voiceless plosives in languages with phonematic oppositions in terms of voicing is typologically unmarked, as shown by Keating's “polarization principle” [24]: in fact, voiceless aspirated plosives were found in varieties of the entire Italian peninsula [25, 26]. However, the clear distribution and indexical value of the trait in Florence pulls the “gorgia enfatica” far from general issues. One first observation is that aspirated allophones are perceived as non-standard and, following the sex/prestige pattern, more prone to be considered as manly. But there is probably more. As we saw about the terminological confusion in the use of the word “aspirazione”, the “gorgia enfatica” can be born as an *exaggeration* of the normal “gorgia”, creating allophones for the unaspirated plosives also in strong positions. This interpretation follows the typical definition of a *covert prestige* trait, used to form a sense of local identity exclusive to a certain group. *Exaggeration* of dialectal traits considered particularly representative of the local identity is described in Del Puente 1995 [27] for the metaphony in Naples (with striking similarities with our distribution) and is suggested in Craffonara 1979 [28] for the palatalization of $[ka]$ and $[ga]$ in the Dolomitic Ladin. Moreover, aspirated allophones in strong positions are typical of erroneous imitation of the Florentine dialect by foreigners, that tend to ironically *exaggerate* the normal “gorgia”, as already noticed by Franceschi 1969 [29]. We can conclude saying with the words of Johnson 2006 [30] that “people (perhaps especially men) perform gender”: but the mechanics of this performance can only be found with a sociophonetic analysis of the local speech community.

Duccio Piccardi
Ph.D. student
University of Pisa

Bibliography:

- [1] Swartz, Bradford L., *Gender difference in voice onset time*, in "Perceptual and Motor Skills" 75, 1992, pp. 983-992.
- [2] Ryalls, John - Zipprer, Allison - Baldauff, Penelope, *A preliminary investigation of the effects of gender and race on Voice Onset Time*, in "Journal of Speech and Hearing Research" 40(3), 1997, pp. 642-645.
- [3] Whiteside, Sandra P. - Irving, Caroline J., *Speakers' sex differences in voice onset time: Some preliminary findings*, in "Perceptual and Motor Skills" 85 (2), 1997, pp. 459-463.
- [4] Whiteside, Sandra P. - Irving, Caroline J., *Speaker's sex differences in voice onset time: A study of isolated word production*, in "Perceptual and Motor Skills" 86, 1998, pp. 651-654.
- [5] Koenig, Laura L., *Laryngeal Factors in Voiceless Consonant Production in Men, Women, and 5-Year-Olds*, in "Journal of Speech, Language and Hearing Research" 43, 2000, pp. 1211-1228.
- [6] Whiteside, Sandra P. - Henry, Luisa - Dobbin, Rachel, *Sex differences in voice onset time: a developmental study of phonetic context effects in British English*, in "Journal of the Acoustical Society of America" 116 (2), 2004, pp. 1179-1183
- [7] Robb, Michael - Gilbert, Harvey - Lerman, Jay, *Influence of gender and environmental setting on voice onset time*, in "Folia Phoniatica et Logopaedica" 57 (3), 2005, pp. 125-133.
- [8] Morris, Richard J. - McCrea, Christopher R. - Herring, Kaileen D., *Voice onset time differences between adult males and females: Isolated syllables*, in "Journal of Phonetics" 36, 2008, pp. 308-317.
- [9] Titze, Ingo R., *Principles of Voice Production*, Englewood Cliffs, 1994.
- [10] Smith, Bruce L., *Effects of Place of Articulation and Vowel Environment on "Voiced" Stop Consonant Production*, in "Glossa" 12(2), 1978, pp. 163-175.
- [11] Pickett, James M., *The sounds of speech communication: A Primer of Acoustic Phonetics and Speech Perception*, Austin 1980.
- [12] Whiteside, Sandra P. - Marshall J., *Developmental Trends in Voice Onset Time: Some Evidence for Sex Differences*, in "Phonetica" 58, 2001, pp. 196-210.
- [13] Oh, Eunjin, *Effects of speaker gender on voice onset time in Korean stops*, in "Journal of Phonetics" 39, 2011, pp. 59-67.
- [14] Peng, Jui-Feng - Chen, Li-mei - Lee, Chia-Cheng, *Voice Onset Time of Initial Stops in Mandarin and Hakka: Effect of Gender*, in "Taiwan Journal of Linguistics" 21:1, 2014, pp. 63-80.
- [15] Cheng, Ming-Chung, *Voice Onset Time of Syllable-Initial Stops in Sixian Hakka: Isolated Syllables*, in "Journal of National Taiwan Normal University" 58 (2), 2013, pp. 193-227.
- [16] Madhu, Sudarshan Reddy. B. – Mahendra, Kumar. N. - Sreedevi, N., *Voice Onset Time across Gender and Different Vowel Contexts in Telugu*, in "Language in India" 14:12, 2014, pp. 252-263.
- [17] Candea, Maria - Trimaille, Cyril, *Introduction. Phonétique, sociolinguistique, sociophonétique: histoires parallèles et croisements*, in "Langage et société" 151, 2015, pp. 7-25.
- [18] Castellani, Arrigo, *Precisazioni sulla gorgia toscana*, in "Boletim de Filologia" XIX, 1960, pp. 242-262.
- [19] Giannelli, Luciano - Savoia, Leonardo Maria, *L'indebolimento consonantico in Toscana. I/II*, in "Rivista Italiana di Dialettologia" II/IV, 1978/1979-1980, pp. 23-58/38-101.
- [20] Hudson, Richard A., *Sociolinguistics*, Cambridge 1996².
- [21] Marotta, Giovanna, *Non solo spiranti. La gorgia toscana nel parlato di Pisa*, in "L'Italia Dialettale" 62, 2001/2004, pp. 27-60.
- [22] Jaberg, Karl - Jud, Jacob, *Sprach- und Sachatlas Italiens und der Südschweiz (=AIS)*, Zofingen 1928-1940.
- [23] Singh, Rita - Keshet, Joseph – Gencaga, Deniz – Raj Bhiksha, *The relationship of voice onset time and voice offset time to physical age*, in *Proceedings of The 41st IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2016)*, Shanghai, China, March 20-25, 2016.
- [24] Keating, Patricia A., *Phonetic and phonological representation of stop consonant voicing*, in

“Language” 60, 1984, pp. 286-319.

[25] Izzo, Herbert John, *Tuscan & Etruscan*, Toronto 1972.

[26] Stevens, Mary - Hajek, John, *Post-aspiration in standard Italian. Some first cross-regional acoustic evidence*, in *Proceedings of the 11th Annual Conference of the International Speech Communication Association. ISCA*, 2010, pp. 1557-1560.

[27] Del Puente, Patrizia, *La metafonia napoletana. Un tentativo di analisi sociolinguistica*, in “L’Italia dialettale” 58, 1995, pp. 49-67.

[28] Craffonara, Lois, *Zur Palatalisierung von CA und GA in den Sellatälern*, in “Ladinia” III, 1979, pp. 69-93.

[29] Franceschi, Temistocle, *Il principio dell’esagerazione come criterio di ricerca linguistica*, in “Archivio Glottologico Italiano” LIV, 1969, pp. 49-105.

[30] Johnson, Keith, *Resonance in an exemplar-based lexicon: The emergence of social identity and phonology*, in “Journal of Phonetics” 34, 2006, pp. 485-499.