Voice pitch as a sign of a speaker's identity

M. Grazia Busà
mariagrazia.busa@unipd.it

Abstract. This paper discusses the pitch component of the human voice: how it is produced, how it is used, what can affect it and how it can influence people's opinions and reactions. The paper focusses on how pitch is unique to every individual, and how it reveals an individual's personal and socio-cultural facts and attitudes. The paper also shows how speakers can, more or less consciously, manipulate their pitch for particular purposes: to appear more attractive, powerful or simply to show membership in a socio-linguistic community.

1. Introduction: Our voice tells people who we are

Speakers' voice plays a fundamental role in communication. It provides linguistic information, it allows us to distinguish between types of sentences (e.g., questions vs. statements), to emphasize parts of discourse (e.g., 'It is *John* who won the price' (implied: and not Mike)) and disambiguate ambiguous sentences (e.g., 'flying planes can be dangerous' -in which 'planes' can either be the grammatical subject of the phrase, or the object of the verb 'flying'). Voice also allows us to express our intentions, e.g., whether we speak to reproach, ask or joke, and communicates our feelings and moods by revealing whether we are happy, sad, or rather bored and detached. At the same time, our voice also reveals our personality and identity by providing information about who we are, where we are from, what is our age, our education, our social status, how powerful we are or we feel, what is our health status (whether we are tired, drunk, under the weather). In other words, in addition to linguistic meanings, our voice carries information about ourselves, how we feel and think, and how we stand in society (Ladefoged 1996, 2006).

As listeners, we are able to tell speakers apart and identify speakers' individual characteristics on the basis of several voice features, including intonation, tone of voice, voice quality, pitch, rhythm, speaking rate, etc. These are all subsumed under the heading of *prosody*, the field that studies the

linguistic properties of speech above the segments (Lehiste 1996; Kent, Read 2002; Ladd 1996; Ladefoged 2006). Prosody reflects speakers' personal and socio-cultural identity, as is summarized in the quote below.

> The combination of these features [i.e., intonation, tone of voice, voice quality, pitch, rhythm, speaking rate[1]] in spoken language allows listeners to recognize a sample of speech as belonging to a particular talker, and prevents others from reproducing his or her speech pattern precisely. These fundamentals can be seen as marking us out as unique individuals.
> (Llamas, Watt 2009: 1)

Voice pitch (or pitch) plays a particularly relevant role in marking speakers' identity. This paper reviews why pitch is crucial for speakers' identification, what are some of the main factors that can cause variations in speakers' pitch, and how these variations contribute to shaping the perception of individuals' personal and socio-cultural identities.

2. Fundamental frequency

Everyone's voice is different, the main reason being that voice is linked to people's anatomy. In speech, sounds are produced by the air coming from the lungs, passing through the larynx, and then going out into the nose and mouth (the so-called *vocal cavity* or *vocal tract*). Inside the larynx are two bands of muscle and tissue, called *vocal folds* (or *vocal chords*). Elastic and aerodynamic forces produced by the air pressure coming from the lungs cause the vocal folds to spring open and close. Such periodic opening and closing movements of the vocal folds are referred to as vibrations of the vocal folds. The rate of vibration of the vocal folds is different for each individual, because it is related to the individual's anatomical characteristics (i.e., being short, tall, slim, overweight has an effect of the size and thickness of the folds). However, contextual factors (such as age, emotional involvement, language and register) also contribute to the production of differences in vocal fold vibrations.

The frequency of an individual's vocal fold vibration is called *fundamental frequency*, or *F0*. Fluctuations in F0 create what in linguistics is

[1] Author's addition.

called *intonation*, that is the voice going up or down when we speak. Different configurations of vocal folds vibrations result in different voice qualities. So, through modifications of the vocal folds speakers can produce voice that has different characteristics, for example whisper, breathiness, murmur, creakyness.

Typically, children have the shortest vocal tracts, smallest larynxes and shortest vocal folds; men have the longest vocal tracts, largest larynxes and longest vocal folds; women have vocal tracts, larynxes and vocal folds of sizes that are in-between those of children and men (Titze 1989). Because lighter and shorter vocal folds can vibrate at higher frequencies, children have F0s with highest frequencies, and men lowest (Peterson, Barney 1952). To give a rough estimate, the average F0 values range around 120-130 Hz for men, around 220-225 Hz for women, and around 265 Hz for children (Cruttenden 1997; Hillenbrand *et al.* 1995; Peterson, Barney 1952). This explains why F0 is the key object of investigation in studies focusing on sex-related vocal properties (Hillenbrand, Clark 2009).

An individual's F0 is reflected in all sounds produced by that individual. In particular, vowels have frequencies that are multiples of F0. This is very important because, from an acoustic point of view, all speakers produce sounds differently from any other individual. Our brain can recognize the traits in each speaker's voice that contribute to creating his/her own particular identity. At the same time, our brain can also abstract from the acoustic variability generated by different individuals and perform categorisations that are essential for linguistic knowledge, which is why we can recognize linguistic units even though these are in fact acoustically very different.[2]

3. Voice Pitch

The perceptual correlate of F0, that is, how we perceive speakers' voice (as high or low) is called *pitch*. Each individual has a characteristic, habitual pitch, but also the ability to speak at higher or lower pitch levels. The range of variations between an individual's maximum and minimum pitch levels is called *pitch*

[2] Categorization is the process that allows us to make sense and conceptualize the world we live in. It allows us to recognize commonalities in our experiences and concepts, and to classify them for some specific purposes. It is one of the basic cognitive principles at the basis of conceptual and linguistic organization (adapted from: https://en.wikipedia.org/wiki/Categorization).

range. Speakers modulate their pitch levels depending on the communicative situation, which may affect the speakers' mood, motivation and speaking effort (Hirschberg, Benus, Gravano 2007). For example, speakers can modulate their pitch to show expressiveness, and be perceived more attractive, aggressive or charismatic; they can use falsetto or creaky voice to signal their socio-cultural identity (Eckert 2008; Borkowska, Pawlowski 2011; Hughes, Mogilski, Harrison 2014). As a case in point, it has become trendy among young, 'upwardly mobile' American and British women to speak using a *creaky voice*, a phenomenon known in linguistics phonetics as *vocal fry*. This is produced by compressing the vocal folds, which reduces the airflow through the larynx and the frequency of vibrations, causing speech to sound "creaky" (Yuasa 2010; Wolk, Abdelli-Beruh, Slavin 2012). This phenomenon has been observed widely in college-aged women in the US and UK, but it is popular among VIPs as well. Among the stars that use this kind of speech are Zooey Deschanel, Kim Kardashian, Britney Spears, Scarlett Johansson (e.g., Shaw, Crocker 2015).

Pitch is a very important parameter for the identification of a speaker's emotional state (other important paramenters being loudness and speech rate). This can be explained, again, on physiological grounds: emotions cause an alteration of the muscles that are responsible for the respiratory and phonatory activities during speech; on their turn, these muscle alterations cause modifications in the (subglottal) air pressure mediating the vibration of the vocal folds in the larynx; an increase in pressure causes the vocal folds to vibrate at an increased frequency, producing a higher pitch (Van den Berg, Zantema, Doornenbal 1957). Listeners are generally able to determine a speaker's emotional state by processing the speaker's variations in pitch (as well as rate and loudness) over time (Hirschberg, Benus, Gravano 2007). For example, pitch that is less varied, i.e., more monotonic, tends to be associated with sad or neutral emotions, while pitch that is more varied, i.e., is characterized by highs and lows, tends to be associated with fear, anger, or happiness. As for speech rate, slow speech is associated with sadness, while fast speech tends to be classified as anger, fear or neutral (Breitenstein, Van Lancker, Daum 2010).

It has been shown that men and women process emotions differently, with possible effects on the use of emotional (non-verbal) information on language

comprehension and production, though in ways that are still not fully understood (Besson, Magne, Schon 2002; Schirmer and Simpson 2007). It is claimed that men's speech has a narrower pitch range and is less dynamic than women's speech; conversely, women's speech is considered more high-pitched and varied. This may be related to the differences in the processing of emotions in the two sexes (Loveday 1981; Henton 1989, 1995). However the evidence in support of women's having a more 'emotional' pitch than men is not uncontroversial. Henton (1989, 1995) found no significant difference in pitch dynamism between male and female speakers, while Daly and Warren (2001) found evidence of measurable and reliable speaker-sex differences in both pitch range and dynamism.

3.1 The effects of body size

As seen in section 2, some voice characteristics are related to the speakers' body size. In nature, animals' vocalizations reflect their overall body mass. More specifically, F0 is inversely related to the mass of the vibrating membrane (either vocal folds in mammals or syrinx in birds), which correlates with body mass. Large animals' membranes vibrate al lower frequencies, producing lower F0s, than smaller animals that produce more high-pitched vocalizations (Morton 1977; Ohala 1984).

Animals are very resourceful in their use of F0 modulations. That is, they can make their voice sound rough and low in pitch to sound aggressive, but can emit sounds that are tone-like and high-pitched to sound submissive or non-threatening. Taking dogs as an example, they growl when they want to appear threatening and whine or yelp to appear submissive. It appears advantageous for animals to use a lower pitch to give the vocal impression of being bigger than the actual size: projecting a bigger size scares away the enemy and, in parallel, makes males more attractive for females (Goldstein 1980; Feinberg *et al*. 2005). Conversely, using a high-pitched voice to try to appear submissive could be considered "a form of infant mimicry", having the effect of pacifying possible aggressors: "for obvious reason, natural selection has left most species with a very strong inhibition against harming conspecific infants" (Ohala 1984: 5). Thus, animals' ability to modulate their F0 in potentially threatening situations

can be explained on biological and evolutionary grounds, that is, as a strategy to reduce the risks of being killed by aggressors in the wild (Morton 1977). This advantage may have driven genetic change in some species, and caused the larynx to drop in adult males to exaggerate the acoustic effects of body size, as for example in the case of the red deer (Fitch, Reby 2001).

In animals, this particular use of F0 paralles the use of visual signals that are made with the body and are intended to convey a message. Animals use visual signals when they are competing with other animals for resources: Predators on the prowl, preys seeking protection, males showing sexual maturity and attractiveness, all make themselves look bigger or try to appear as large as possible. For example, cobras have specialized muscles in the neck that let them flare out a "hood" to signal when they are ready to attack or are excited. Cats that are startled arch their backs and stick their hair up to look bigger. Toads inflate their bodies by gulping air and stand high on all fours to create the illusion of larger size. Lions' manes are a sign of sexual maturity and health, but are also a sign of potency; their size and density makes lions appear larger and more powerful, scaring away possible opponents. Peacocks attrack females with their beautiful sapphire and emerald plumage, making their tails highly elongated.

In humans, fundamental frequency is not unequivocally related with body size. Graddol and Swann (1989) found no statistically significant correlation between body size and F0 excursion, while (Yuasa 2008) showed that F0 correlates with the speaker's body size, elastic properties of vocal tract tissue and vocal fold length. Pisanski *et al.* (2014) ironically quote the case of the *Wizard of Oz*, who had a low, powerful and threatening voice that made him sound very fearful, but who turned out to be a little and insignificant man. However, even though there is no robust, exceptionless, direct relationship between voice parameters and size, humans, like animals, tend to associate a low voice with a body of large size. In fact, it has been suggested that voice pitch and formants interact in complex ways to affect voice perception and that pitch may help listeners perceive size accurately (Pisanski et al 2014).

Size projection through vocal and visual signals in humans is not as clear as in animals, but it does take place to some extent. Angry speakers lower their

pitch and raise their voice volume to sound bigger and more powerful while making themselves look as big as possible by using wide arm gestures, spaced legs, etc. Conversely, happy speakers higher their pitch and understate their body size to communicate appeasement (Ohala 1984). This kind of behaviour may have been advantageous for the human speaker and may have driven genetic change, causing the male larynx to drop during puberty (Morton 1977; Fitch, Reby 2001; Putz 2004).

3.2 The effects of age

Voice also conveys information about age. People are surprisingly good at determining a speaker's age range just by hearing their voice (Ryan 1973; Ptacek, Sander 1966). This is because people's voice changes with age, and so does pitch. This was tested in several longitudinal studies. For example, Stoicheff (1981) analysed the F0 and standard deviation of 111 female subjects aged between 20 and 82, and found that in the postadolescence through early adulthood speakers' F0 remained relatively stable, to decrease in the 50-59 age range, and to remain then stable at older ages. The group in the 50-59 age range also showed an increase in intra-subject variability of F0. The changes in F0 values observed in this group were interpreted as "due to changes in the vocal fold mass, and to a decreased laryngeal control over fundamental frequency in postmenopausal adults" (Stoicheff 1981: 437).

Comparable results were obtained by de Pinto and Hollien (1982) who recorded 11 Australian women at the age of 18-25 years and then the same subjects 35 years later. The results showed that the subjects in their 50s had F0 values that were about 40Hz lower than they had in their 20s. A similar study was conducted on male speakers (Hollien, Ship 1972) and showed that the men's F0 varied with age: the subjects in their 20s had F0s ranging around 119 Hz, those in their 40s around 107 Hz, and those in their 80s around 146 Hz.

In an individual's life span, F0 changes are systematic and due to the "maturational development of the vocal apparatus as well as the development of cognitive and linguistic skills" (Cooper, Sorensen 1981). These changes, which are broadly predictable, are shown in Fig. 1 (from Baken, Orlikoff 2000). Children, boys and girls, have a high fundamental frequency, with small

differences in the two sexes. For males, F0 drops dramatically during puberty, and continues to decrease until men are around 40, due to physical modifications of the larynx; it raises slightly after 40, through and past the 80s due to changes in the musculature of the vocal tract. Women's voice lowers continusouly thoughout their lives, with menopause being a contributing factor for the lowering of F0, but it raises slightly after 80 (Cooper, Sorensen 1981; Cruttenden 1997).

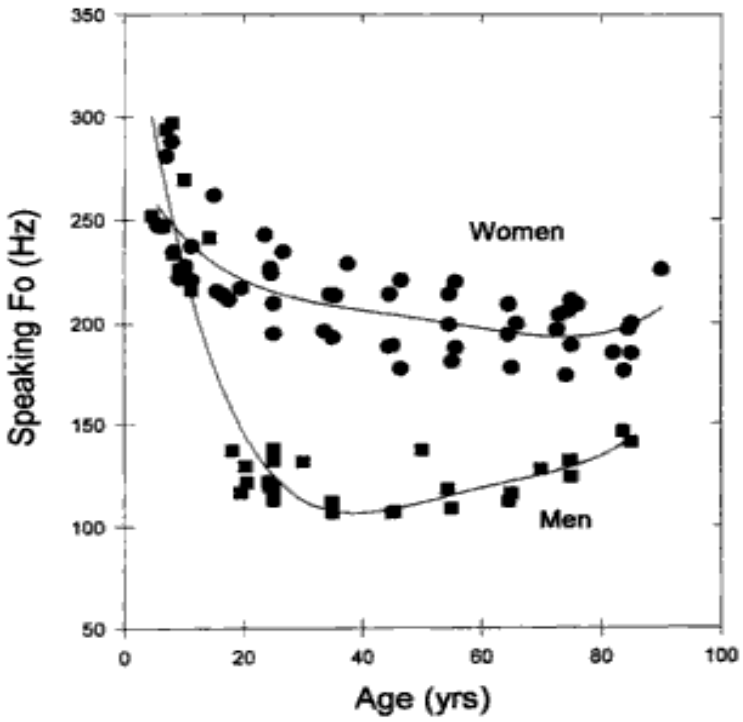

Figure 1. General trend of F0 across male and female life span (from Baken, Orlikoff 2000: 174).

3.3 Cultural and Social variations in use and interpretation of pitch

Pitch may be culturally and socially stigmatised. In general, lower pitch tends to be associated with masculinity and confidence while higher pitch is considered more feminine and emotional –probably as a generalization of the fact that men typically have lower pitch than women. Lower pitch is also generally associated with other masculine traits such as dominance, while higher pitch is associated with female traits such as submissveness (Borkowska, Pawlowski 2011).

These largely-shared interpretations of vocal speech patterns may cause speakers to deliberately vary their pitch depending on the effect they want to make on people. In the US, lower pitch is considered more prestigious and

admirable than higher pitch, both as concerns one's own implicitly perceived status, and the status of those one is talking to. In other words, speakers know how to adjust their picth if they are in higher power positions, and recognize higher-power position people based on their pitch (Cheng *et al*. 2016, Leongomez *et al*. 2017). Both men and women tend to select male and female leaders with lower voices, suggesting that having lower-pitched voices may be critical for obtaining positions of leadership (Klofstad, Anderson, Peters, 2012). Also, women that want to appear more professional, or dominant, tend to lower their pitch and use a creaky voice (Yuasa 2010; Wolk, Abdelli-Beruh, Slavin 2012; Lowen 2017).

The issue of what factors may influence female speakers in deciding to lower their pitch when speaking in public is well expressed in this quote:

> Pitch can be a problem for all public speakers, whatever their sex, because it tends to rise when someone is nervous or is speaking louder than usual –both of which are likely to happen in speeches. For women the problem is more acute because the natural pitch of their voices has a higher starting point than is the case for men, with the result that it cannot rise as far before reaching a level at which it sounds 'shrill'.
> This might not matter but for the fact that high pitch tends to be strongly associated with emotional or irrational outbursts - a deeply rooted cultural assumption that probably derives from (and is sustained by) the screams of each new generation of infants. The fact that the sound of a woman raising her voice is more likely to be negatively evaluated as 'shrill' or 'screeching' is probably at the heart of a source of irritation that's familiar to many professional women, namely the tendency of male colleagues to accuse them of 'overreacting' whenever they become involved in arguments (Atkinson 2009).

Public figures are aware that their language projects their identity. Women are under constant scrutiny, and particularly in politics. Female politicians often believe that, to succed in a man-dominated world, they should downplay typically feminine traits, including their voice. A female politician who was aware of the effects of her voice on the public was Margaret Thatcher. At the start of her career Ms. Thatcher took lessons to make her voice sound firmer and more powerful. As a result, she managed to lower her pitch remarkably (about 46 Hz), and slow down her speech to achieve greater credibility and impact (Atkinson 1984).

Another female politician who showed awareness of the voice power in shaping the public perception of her identity was Hillary Clinton. When Clinton ran for the US presidency in the 2008 and 2016 political campaigns, she was careful to avoid using a highly-varied pitch (i.e., with a lot of ups and downs) –a characteristic that is associated with emotional and female speech. Instead, she emphasised stressed words, which is not associated with emotion but rather with powerful speech (Reeve 2015). Clearly, she wanted to appear strong and determinate rather than emotional. However, her voice was the object of harsh criticism in both campaigns. Some of the adjectives that were used to describe it were: flat, screaming, shrieking, angry, bitter, grating, non-conversational, screechy, blaring, loud (Brinlee 2016; Chozick 2016). But this kind of criticism was probably the result of extreme scrutiny for a woman politicians's voice: "Men are supposed to be assertive, loud, and competitive. Women are supposed to be soft-spoken, cooperative, and helpful. "No matter who's saying something, a man or a woman, they're being judged on their language via their gender"" (Fought in Reeve 2015).

3.4 Cross-linguistic and cross-cultural variations in use and interpretation of pitch

Languages differ in their use and interpretation of pitch. This accounts for many differences in the way languages 'sound' to non-native speakers. For example, as reviewed in Busà (this volume), German speakers speak with a lower pitch level and less variation in pitch than English speakers. Because of these differences in pitch, the former may sound monotonic to the latter and, viceversa, the latter may sound overexcited to the former (Mennen 2006; Mennen, Schaeffler, Docherty 2008, 2012). German speakers also have lower pitch level and less pitch variation than speakers of Bulgarian and Polish; on their turn, the speakers of these two languages have higher and more varied pitch than English speakers (Andreeva *et al*. 2015). Language-related differences in pitch seem to be preserved by bilinguals when speaking the two languages. For example, bilingual speakers of English and Japanese use a higher pitch in Japanese than in English (Todaka 1993; Yamazawa, Hollien 1992; Graham 2013).

Variations in use and interpretation of pitch are interwined with language- and culture-specific norms, for example in relation to gender. For instance, comparing national varieties of English, American women seem to use a larger vocal range than British women, while American men employ a more limited vocal range than British men (Graddol, Swann 1989 in Kroløkke, Søresen 2005). Even within the same language, different linguistic communities (whether they share the same language or speak different languages) tend to be characterized by different pitch range and pitch variation (Dolson 1994). For example, in New York City Puerto Rican girls were found to use rather high pitch, while American women used low pitch (Luchsinger, Arnold 1965). Different dialects of the same language may also differ in the use F0 (e.g., Torgerson 2005; Deutsch, Le, Shen, Henthorn 2009). This is because individuals are influenced by the pitch levels of speech that they perceive in a given linguistic community and shape their pitch accordingly so as to fit the prevailing range in that community (Dolson 1994; Deutsch *et al.* 2009).

Long-term exposure to pitch patterns shapes not only speakers' realization of pitch but also conditions speakers' expectations as to what should be regarded as appropriate in certain linguistic communities. In other words, it leads speakers/listeners to formulate quality judgments of pitch range. This may account for some differences in the use and interpretation of pitch that are gender- and culture-related. Such differences may be wider when comparing linguistic communities that diverge significantly in cultural heritage, lifestyle and social expectations, as exemplified below.

Japan and the United States "exemplify completely opposite sets of cultural norms and consequently offer excellent sample pools by which to compare and contrast abiding social variables" (Yuasa 2008: 8). If the American society stands for values such as individualism, leadership and high competition, the Japanese society is very attached to tradition, honor and respectability, with relatively pronounced gender roles and unbalanced economic conditions for women (Wikipedia, *Women in Japan*, 9.09.2017). In the most conservative Japanese culture, women are expected to follow several patterns of modesty, tidiness, courtesy, compliance, and self-reliance (ibid.).

The socio-cultural differences in gender roles in Japan are reflected in the speakers' use of pitch: high pitch is associated with politeness and deference, and Japanese women are expected to use high pitch as a sign of dependency and modesty. In fact, Japanese women have been reported to use a pitch range that is consistently higher than that of American (Ohara 1999), British (Loveday 1981) and Dutch women (van Bezooijen 1995), thus reflecting an idea of submissiveness. On the other hand, Japanese men seem to have considerably lower pitch values than British males (Loveday 1981), projecting an idea of strength and dominance. These socio-cultural differences in the use of pitch also frame its interpretation: Japanese listeners (males and females) perceive high-pitch voice in female speakers as being more attractive than Dutch listeners do (van Bezooijen 1995).

3.5 Other factors affecting voice pitch

Among the many variables that affect speakers' voice pitch a couple more are worth mentioning, as they relate to style, speakers' first or second language spoken, and charisma.

3.5.1 Style

It is well known that speakers can modify their speech characteristics depending on the particular purpose or social setting of their communication, a phenomenon that in linguistics is known as *register*. Voice pitch is one of the features that speakers can modify when they vary their registers. For example, spontaneous speech is characterized by higher and more varied pitch than read speech (Daly, Warren 2001; Garrido 1993). Also, speakers' pitch is more varied in informal than in formal contexts, and in both cases it is more varied than in read speech (Zipp, Dellwo 2011).

Speakers also vary their pitch to convey different pragmatic meanings. For example, because high pitch is associated with perceived subdominance (see section 3.3), it is also associated with politeness (Ohala 1984; Gussenhoven 2002) or friendliness (Chen, Gussenhoven, Rietveld 2004 on Dutch and English; Ohara 2001 on Japanese) [though some research shows that low pitch is correlated with the usage of polite register (Winter, Grawunder 2011, 2012;

Brown *et al*. 2014 on Korean) and that high pitch may also be used to signal impoliteness (e.g. Stadler 2007 on German and New Zealand English; Nadeu, Prieto 2007 on Catalan)]. As a result, speakers who aim to appear nicer, more polite and friendlier can modulate their pitch to make it sound higher, while people who keep a low pitch in certain circumstances may be perceived as rude or unfriendly.

3.4.2 Effects of speakers' L1 pitch on L2

As seen in section 3.4, languages differ in their use and interpretation of pitch. Like any other L1 feature, L1 pitch is likely to influence and be transferred to the L2, with effects on way in which L2 is produced and interpretation. It is possible that L2 pitch contributes, as part of the L2 speakers' prosody, to what is perceived as 'foreign accent', and that it may have an impact on L2 speakers' comprehensibility and intelligibility (Anderson-Hsieh, Johnson, Koehler 1992; Kormos, Dénes 2004; Munro, Derwing 2001; Munro 2008; Wennerstrom 2000).

However, it is more likely that L2 pitch may be associated with paralinguistic meanings that are not intended by L2 speakers, but that are interpreted in ways that L2 speakers are unaware of. For example, there are indications that L2 speech may be characterised by a narrower pitch range and less pitch variation than L1 speech. This was found for several L1-L2 interactions, including Russian spoken by Finns (Ullakonoja 2007), English spoken by Swedish, Italian and Japanese speakers (Hincks 2004, 2005; Busà, Urbani 2011; Urbani 2012; Aoyama, Guion 2007; Graham 2013), and Spanish spoken by Italians (Orrico *et al.* 2016). It is possible that L2 speakers focus on conveying meaning in the L2 through lexical, grammatical, semantic and segmental information at the expense of prosody (Jenkins 2002), making their pitch rather unvaried and sounding monotonic (and monotonous).

This may have important pragmatic consequences, as it may affect the image the speaker projects of him/herself. In either L1 or L2, limited pitch variation tends to be associated with lack of speaker's liveliness (see below). For L2 speakers, lack of liveliness in speech may generate the seeds for prejudice or negative stereotyping, and thus increase the odds of L2 speakers' social or professional discrimination (Busà 2010). As reviewed above (section 3.4), the

differences in pitch patterns do seem to affect the way German and British speakers sound to each other.

3.4.3 Liveliness and charisma

Pitch range has been used as a measure of speaker's perceived *liveliness* (Hinks 2004, 2005; Mennen *et al*. 2008, Traunmüller, Eriksson 1995) –though this too is subjected to variations depending on language and socio-cultural/socio-phonetic factors, as reviewed above. Typically, a voice that is heavily inflected or has a wide pitch range will sound animated; a voice that has a narrow pitch range will sound monotonic.

The use of a lively voice, that is a voice that varies in pitch or intonation, is particularly emphasized for speaking in public. Public speaking training guidelines advise speakers to pause before moving to a new chunk of information, and then start the new piece of information by raising the pitch. This lets listeners orient themselves when following the information flow, and helps them maintain their attention on the message. Varied pitch (i.e. intonation characterized by highs and lows) also makes the speaker sound less monotonous to the ear and prevents listeners from *habituating* to the speaker's rhythm of speech, giving them reason to wander with their thoughts. *Habituation* is a technical term that describes a decrease in response to a stimulus after repeated presentations (from Wikipedia, sub voce, 30.08.2017). In nature, it is an adaptive behavior, and one of the simplest and most common forms of learning since it allows people to "tune out" non-essential stimuli and focus on the things that really demand attention. For example, people's attention may be drawn by an unusual loud noise in their surroundings when they first hear it, but then tend to habituate to it, and disregard it. Thus, the constant repetition of a stimulus changes the efficacy of that stimulus: The more we hear it, the less we notice it. In a way, it becomes uninteresting to our brains. That is why, to sound interesting, speech should be characterised by a varied rather than a monotonic pitch.

Charisma is a well known but ill-defined term. It has to do with the ability to attract, charm, and influence people. The characteristics of a charismatic person are also not well defined, but there is little doubt that a charismatic

speaker needs to have a charismatic voice. As reviewed in section 3.1, speakers with lower pitch frequency are perceived as more 'dominant' or 'attractive' (Collins 2000; Feinberg, et al. 2005) while speakers with higher pitch frequency are perceived as more 'subordinate' (Collins, Missing 2003). However, this is also subjected to socio-cultural interpretation. A distinction has been made between institutionally and impulsively oriented cultures (Gordon, 1989; Turner, 1976). In institutionally oriented cultures, strong control and regulation of emotions are required to fulfil institutional roles and standards. An example is provided by the Japanese culture. In impulsively oriented cultures, there is a lower regulation and control in public and institutional contexts. An example is the Italian culture. According to D'Errico *et al*. (2013), charismatic people are the ones that express emotions coherently with the socio-cultural expectations of emotion expression of their culture (D'Errico *et al*. 2013). This explains why what could be perceived as too passionate (for example as a result of a highly-varied pitch) in certain Asian contexts might be perceived as mellow in southern European cultures. But socio-cultural differences in interpretation of charisma may also subtler and not so easy to pick up. In their analysis of how French and Italian speakers perceive the charismatic voice D'Errico *et al*. (2013) found that Italians are more likely to trust a speaker who uses normal or high pitch and make short pauses, while French prefer speakers who have normal to low pitch and make long pauses.

4. Conclusion

Voice plays a relevant role in the representation of speakers' identity by providing information about sex, age, body size, emotional and physiological state, regional varieties, native language, and speaking styles. All of this is done in parallel with the transmission of linguistic information.

An important feature of our voice, our pitch determines how we are perceived by others: presumptions about our intelligence, desires, determination, confidence and influence are all inferred from how we speak. Thus, our use of pitch has an effect on our social and professional relationships. Our ability to modify pitch can also be used to create our own identity profile through the

construction of a specific persona to try to apper more powerful, successful and attractive and influence others through our voice.

This paper has reviewed some main facts about voice pitch and discussed how speakers with no impairment are capable to produce a wide range of pitch excursions, depending on their mood, motivation and effort. In their use and interpretation of pitch they follow the norms of the socio-cultural and linguistic communities where they live, and are able to modify their pitch in many ways to become part of a specific community.

Thus, if the eyes are the mirror of the soul, pitch is the reflection of our personality.

References

ANDERSON-HSIEH J., JOHNSON R., KOEHLER K. (1992), "The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure", *Language Learning*, 42, 529-555.

ANDREEVA B., MÖBIUS B., DEMENKO G., ZIMMERER F., JÜGLER J. (2015), "Linguistic measures of pitch range in Slavic and Germanic Languages", *Proceedings Interspeech 2015*, 968-972.

ATKINSON M. (1984), *Our Masters' Voices: the Language and Body Language of Politics*, London, Routledge.

ATKINSON M. (2009), "Margaret Thatcher and the evolution of charismatic woman: Part I. Cultural and vocal challenges. http://maxatkinson.blogspot.com/2009/01/margaret-thatcher-and-creation-of.html (1.09.2017).

AOYAMA K., GUION, S.G. (2007), "Prosody in Second Language Acquisition: An acoustic analysis on duration and F0 range", In Bohn O.-S., Munro M.J. (eds), *The Role of Language Experience in Second-Language Speech Learning - In Honor of James Emil Flege*, Amsterdam, John Benjamins, 281-297.

BAKEN R.J., ORLIKOFF R.F. (2000), *Clinical Measurement of Speech and Voice*, San Diego, Singular Publishing Group.

BESSON M., MAGNE C., SCHON D. (2002), "Emotional prosody: sex differences in sensitivity to speech melody", *Trends in Cognitive Sciences,* 6(10), 405-407.

BORKOWSKA B., PAWLOWSKI B. (2011), "Female voice frequency in the context of dominance and attractiveness perception", *Animal Behaviour*, 82(1), 55-59.

BREITENSTEIN C., VAN LANCKER D., DAUM I. (2001), "The contribution of speech rate and pitch variation to the perception of vocal emotions in a German and an American sample", *Cognition and Emotion,* 15(1), 57-79.

BRINLEE M. (2016), "8 sexist comments about Hillary Clinton's voice during the 2016 election that'll make you want to scream". https://www.bustle.com/articles/160884-8-sexist-comments-about-hillary-clintons-voice-during-the-2016-election-thatll-make-you-want-to. (1.09.2017).

BROWN L., WINTER B., IDEMARU K., GRAWUNDER S. (2014), "Phonetics and politeness: Perceiving Korean honorific and non-honorific speech through phonetic cues", *Journal of Pragmatics*, 66, 45-60.

BUSÀ M.G. (this volume), "Representing others: The role of non-verbal language. With special reference to Italian speakers of English".

BUSÀ M.G (2010), "Effects of L1 on L2 pronunciation: Italian prosody in English". In Gagliardi A., Maley A. (eds.), *ILS, ELF, Global English: Teaching and LearningProcesses, Linguistic Insights: Studies in Language and Communication*, Bern, Peter Lang, 200-228.

BUSÀ M.G., URBANI M. (2011), "A cross linguistic analysis of pitch range in English L1 and L2", *Proc. 17th International Conference of Phonetic Sciences (ICPhS)*, Hong Kong, 380-383.

CHEN A., GUSSENHOVEN C., RIETVELD T. (2004), "Language-specificity in the perception of paralinguistic intonational meaning", *Language and Speech* 4, 311-350.

CHENG J.T., TRACY J.L., HO S., HENRICH J. (2016), "Listen, follow me: Dynamic vocal signals of dominance predict emergent social rank in humans", *Journal of experimental psychology: general*, 145(5), 536.

CHOZICK A. (2016), "Hillary Clinton raises her voice, and a debate over speech and sexism rages". https://www.nytimes.com/2016/02/05/us/politics/hillary-clinton-speeches-sexism.html. (1.09.2017).

COLLINS S.A. (2000), "Men's voices and women's choices", *Animal Behaviour,* 60, 773-780.

COLLINS S.A., MISSING, C. (2003), "Vocal and visual attractiveness are related in women", *Animal Behaviour, 65,* 997-1004.

COOPER W.E., SORENSEN J.M. (1981), *Fundamental Frequency in Sentence Production*, New York, Springer Verlag.

CRUTTENDEN A. (1997), *Intonation* (2nd ed.), Cambridge, Cambridge University Press.

DALY N., WARREN P. (2001), "Pitching it differently in New Zealand English: Speaker sex and intonation patterns", *Journal of Sociolinguistics*, 5(1), 85-96.

DE PINTO O., HOLLIEN H. (1982), "Speaking fundamental frequency characteristics of Australian Women: Then and now", *Journal of Phonetics*, 10, 367-375.

D'ERRICO F., SIGNORELLO R., DEMOLIN D., POGGI I. (2013), "The perception of charisma from voice. A cross-cultural study", *Proceedings of IEEE Conference on Affective Computing and Intelligent Interaction (ACII 2013),* Geneva, Switzerland, September 2-5, 2013.

DEUTSCH D., LE J., SHEN J., HENTHORN T. (2009), "The pitch levels of female speech in two Chinese villages", *Journal of the Acoustical Society of America*, 125 (5), 208- 2013.

DOLSON M. (1994), "The pitch of speech as a function of linguistic community", *Music Perception*, 11, 321-331.

ECKERT P. (2008), "Variation and the indexical field", *Journal of Sociolinguistics,* 12, 453-76.

FEINBERG D.R., JONES B.C., LITTLE A.C., BURT D.M., PERRETT D.I. (2005), "Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices", *Animal Behavior*, 69, 561-568.

FITCH W.T., REBY D. (2001), "The descended larynx is not uniquely human", *Proceedings of the Royal Society, Biological Sciences,* 268(1477), 1669-1675.

GARRIDO J.M. (1993), "Analysis of global pitch contour domains at paragraph level in Spanish reading text", *Working Papers. Lund University, Department of Linguistics*, (Proceedings of an ESCA Workshop on Prosody, september 27-29, 1993, Lund, Sweden), 41,104-107.

GOLDSTEIN U. (1980) *An articulatory model for the vocal tracts of growing children*. Ph.D. dissertation, Massachusetts Institute of Technology.

GORDON S.L. (1989), "Institutional and impulsive orientations in selective appropriating emotions to self". In Franks D.D., McCarthy D. (eds.), *The Sociology of Emotions: Original Essays and Research Papers,* Greenwich, CT, USA, JAI Press, 115-136.

GRADDOL D., SWANN J. (1989), *Gender voices*, Oxford, Basil Blackwell.

GRAHAM C. (2013), "Revisiting f0 range production in Japanese-English simultaneous bilinguals", *Annual Report of UC Berkeley Phonology Lab* 110-125.

GUSSENHOVEN C. (2002), "Intonation and interpretation: Phonetics and Phonology". In Bel B., Marlien I. (eds.), *Speech Prosody 2002: Proceedings of the First International Conference on Speech Prosody. ProSig and Université de Provence Laboratoire Parole et Language*, Aix-en-Provence, 47-57.

HENTON C. (1989), "Fact and fiction in the description and female and male pitch", *Language and Communication,* 9(4), 299-311.

HENTON C. (1995), "Pitch dynamism in female and male speech", *Language and Communication,* 15(1), 46-61.
HILLENBRAND J.M., GETTY L.A., CLARK M.J., WHEELER K. (1995), "Acoustic characteristics of American English vowels", *Journal of the Acoustical Society of America*, 97, 3099-3111.
HILLENBRAND J.M., CLARK M.G. (2009), "The role of F0 and formant frequencies in distinguishing the voices of men and women", *Attention, Perception & Psychophysics,* 71 (5), 1150-116.
HINCKS R. (2004), "Processing the prosody of oral presentations", *Proc. InSTIL/ICALL2004 – NLP and speech technologies in advanced language learning systems*, 63-69.
HINCKS R. (2005), "Measuring liveliness in presentation speech", *Proc. INTERSPEECH 2005 - Eurospeech, 9th European Conference on Speech Communication and Technology,* 01/2005.
HIRSCHBERG J., BENUS S., GRAVANO A. (2007), "Prosody, emotions, and...'whatever'", *Proc. Interspeech*, Antwerp, 2629-2632.
HOLLIEN H., SHIP T. (1972), "Speaking fundamental frequency and chronological age in males", *Journal of Speech and Hearing Research,* 15, 155-159.
HUGHES S.M., MOGILSKI J.K., HARRISON M.A. (2014), "The perception and parameters of intentional voice manipulation", *Journal of Nonverbal Behavior*, 38(1), 107-127.
JENKINS J. (2000), *The Phonology of English as an International Language*, Oxford, Oxford University Press.
KENT R.D, READ C. (2002), *The Acoustic Analysis of Speech*, San Diego, Singular Publishing Group.
KLOFSTAD C.A., ANDERSON R.C., PETERS S. (2012), "Sounds like a winner: voice pitch influences perception of leadership capacity in both men and women", Proc. Royal Society B, 279, 2698–2704.
KORMOS J., DÉNES M. (2004). "Exploring measures and perceptions of fluency in the speech of second language learners", *System,* 32, 145-164.
KROLØKKE C., SØRENSEN A.S. (2005), *Gender Communication Theories and Analyses*, New York, Sage.
LADD D.R. (1996), *Intonational Phonology*, Cambridge, Cambridge University Press.
LADEFOGED P. (1996), *Elements of Acoustic Phonetics*, (2nd ed.), Chicago, University of Chicago Press.
LADEFOGED P. (2006), *A Course in Phonetics*, Boston, Thomson.
LEHISTE I. (1996), "Suprasegmental features of speech". In Lass N.J. (ed.), *Principle of Experimental Phonetics*, Saint-Louis, Mosby, 226-244.
LEONGÓMEZ J.D., MILEVA V.R., LITTLE A.C., ROBERTS S.C. (2017), "Perceived differences in social status between speaker and listener affect the speaker's vocal characteristics", *PloS one*, 12(6), e0179407.

LLAMAS C., WATT D. (eds.) (2009), *Language and Identities*, Edinburgh, Edinburgh University Press.
LOWEN L. (2017), "Do women with lower voice pitch have more authority, achieve greater success?", https://www.thoughtco.com/women-lower-voice-pitch-authority-success-3533843. (17.05. 2017).
LOVEDAY L. (1981), "Pitch, politeness and sexual role", *Language and Speech*, 24 (1), 71-89.
LUCHSINGER R., ARNOLD G. (1965), *Voice-Speech-Language*, London Constable&Co Ltd.
MENNEN I. (2006), "Phonological and phonetic influences in non-native intonation". In Trouvain J., Gut U. (eds.), *Non-native Prosody: Phonetic Descriptions and Teaching Practice*, Berlin, Mouton De Gruyter, 1-18.
MENNEN I., SCHAEFFLER F., DOCHERTY G. (2008), "A methodological study into the linguistic dimensions of pitch range differences between German and English", *Proc. Speech Prosody*, University of Campinas, 527-530.
MENNEN I., SCHAEFFLER F., DOCHERTY G. (2012), "Cross-language difference in f0 range: a comparative study of English and German", *Journal of the Acoustical Society of America*, 131 (3), 2249-2260.
MORTON E.W. (1977), "On the occurrence and significance of motivation-structural rules in some bird and mammal sounds", *American Naturalist,* 111, 855-869.
MUNRO M.J. (2008), "Foreign accent and speech intelligibility". In Hansen Edwards J.G., Zampini M.L. (eds.), *Phonology and second language acquisition,* Amsterdam, John Benjamins, 193-218.
MUNRO M.J., DERWING T. (2001), "Modelling perceptions of the comprehensibility and accentedness of L2 speech: The role of speaking rate", *Studies in Second Language Acquisition,* 23, 451-468.
NADEU M., PRIETO P. (2011), "Pitch range, gestural information and perceived politeness in Catalan", *Journal of Pragmatics*, 43, 841-854.
OHALA J.J. (1984), "An ethological perspective on common cross-language utilization of F0 of voice", *Phonetica*, 41, 1-16.
OHARA Y. (1999), "Performing gender through voice pitch: a cross-cultural analysis of Japanese and American English". In Pasero U., Braun F. (eds.), *Perceiving and Performing Gender*, Opladen/Wiesbaden, Westdeutscher Verlag, 105-116.
OHARA Y. (2001), "Finding one's voice in Japanese: A study of the pitch levels of L2 users". In Pavlenko A., Brackledge A., Piller I., Teutsch-Dwye M. (eds.), *Multilingualism, Second Language Learning, and Gender*, New York, Mouton de Gruyter, 231-254.
ORRICO R., CATALDO V., SAVY R., BARONE L. (to appear), "Transfer, fossilisation and prosodic drift in Foreign Language Learning", *Proceedings AISV 2016*.

PETERSON G., BARNEY H.L. (1952), "Control methods used in a study of the vowels", *Journal of the Acoustical Society of America*, 24, 175-184.
PISANSKI K., FRACCARO P.J., TIGUE C.C., O'CONNOR J.J.M., FEINBERG D.R. (2014), "Return to Oz: Voice pitch facilitates assessments of men's body size", *Journal of Experimental Psychology: Human Perception and Performance*. Advance online publication.http://dx.doi.org/10.1037/a0036956
PTACEK P.H., SANDER E.K. (1966), "Age recognition from voice", *Journal of Speech, Language, and Hearing Research*, 9, 273-277.
PUTZ D.A. (2004), *Female Choice, Male Dominance, and the Evolution of Low Voice Pitch in Men.* PhD Thesis. University of Pittsburgh. http://www.anthropology.pitt.edu/node/565 (1.09.2017).
REEVE E. (2015), "Why do so many people hate the sound of Hillary Clinton's voice?", https://newrepublic.com/article/121643/why-do-so-many-people-hate-sound-hillary-clintons-voice. (1.09.2017).
RYAN W.J. (1973), "Predictors of Age in the Male Voice", *The Journal of the Acoustical Society of America*, 53(1), 345.
SCHIRMER A., SIMPSON E. (2007), "Brain correlates of vocal emotional processing in men and women". In Izdebski K. (ed.), *Voice and Emotion*, San Diego, Plural Publishing, 75-86.
SHAW F., CROCKER V. (2015), "Creaky voice as a stylistic feature of young American female speech: An intraspeaker variation study of Scarlett Johansson", *Lifespans & Styles: Undergraduate Working Papers on Intraspeaker Variation*, 1(3), http://journals.ed.ac.uk/lifespansstyles, Volume 1, article 3.
STADLER S. (2007), *Multimodal (Im)politeness: The Verbal, Prosodic and Non-Verbal Realization of Disagreement in German and New Zealand English*, Hamburg, Verlag Dr. Kovac.
STOICHEFF M.L. (1981), "Speaking fundamental frequency characteristics of nonsmoking female adults", *Journal of Speech and Hearing Research*, 24, 437-441.
STRAND E. (2000), *Gender Stereotype Effects in Speech Processing*, PhD dissertation, Ohio State University.
TITZE I.R. (1989), "The physics of small-amplitude oscillation of the vocal folds", *Journal of the Acoustical Society of America*, 83, 1536-1552.
TODAKA Y. (1993), *A cross-language study of voice quality: bilingual Japanese and American speakers*. Doctoral dissertation, University of California, Los Angeles.
TORGERSON R.C. (2005), *A comparison of Beijing and Taiwan Mandarin tone register: An acoustic analysis of three native speech styles*. Master's thesis, Provo (Utah), Brigham Young University.
TURNER R.H. (1976), "The real self: From institution to impulse," *American*

Journal of Sociology, 81, 989-1016.

TRAUNMÜLLER H., ERIKSSON A. (1995), "The perceptual evaluation of *F0* excursions in speech as evidenced in liveliness estimations", *Journal of the Acoustical Society of America*, 97 (3), 1905-1915.

ULLAKONOJA R. (2007), "Comparison of pitch range in Finnish (L1) and Russian (L2)*", Proc. 16th ICPhS,* Saarbrücken, 1701-1704.

URBANI M. (2012), "Pitch range in L1/L2 English. An analysis of F0 using LTD and linguistic measures". In Busà M.G., Stella A. (eds.), *Methodological Perspectives on L2 Prosody, Papers from ML2P 2012*, Padova, Cleup, 79-83.

VAN BEZOOIJEN R. (1995), "Sociocultural aspects of pitch differences between Japanese and Dutch women", *Language and Speech,* 38, 253-265.

VAN DEN BERG J.W., ZANTEMA J.T., DOORNENBAL P. (1957), "On the air resistance and the Bernoulli effect of the human larynx", *Journal of the Acoustical Society of America,* 29(5), 626-631.

WENNERSTROM A. (2000), "The role of intonation in second language fluency". In Riggenbach H. (ed.), *Perspectives on Fluency,* Ann Arbor, MI, University of Michigan Press, 102-127.

WINTER B., GRAWUNDER S. (2011), "The polite voice in Korean: searching for acoustic correlates of contaymal and panmal". In Sohn H., Cook H., O'Grady W., Serafim L., Cheon S. (eds.), *Japanese/Korean Linguistics*, Stanford, CSLI, 19, 419-431.

WINTER B., GRAWUNDER S. (2012), "The phonetic profile of Korean formal and informal speech registers", Journal of Phonetics, 40(6), 808-815.

WOLK L., ABDELLI-BERUH N.B., SLAVIN D. (2012), "Habitual use of vocal fry in young adult female speakers", *Journal of Voice*, 26(3), e111–e116.

YAMAZAWA H., HOLLIEN H. (1992). "Speaking fundamental frequency pattern of Japanese women", *Phonetica*, 49, 128-140.

YUASA P.I. (2008), *Culture and Gender of Voice Pitch. A Sociophonetic Comparison of the Japanese and Americans*, London, Equinox.

YUASA P.I. (2010), "Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women?", *American Speech* 85(3), 315-337.

ZIPP L., DELLWO V. (2011), "Reading-Speech-Normalization: A method to study prosodic variability in spontaneous speech", *Proc. 17th International Conference of Phonetic Sciences (ICPhS)*, Hong Kong, 2328-2331.

