Repositório ISCTE-IUL

Voice changes meaning:

The role of gay- vs. straight-sounding voices in sentence interpretation

Abstract

Utterances reveal not only semantic information, but also information about the speaker's social category membership, including sexual orientation. In four studies (N = 345) we investigated how the meaning of what is being said changes as a function of the speaker's voice. In Studies 1a/1b, gay- and straight-sounding voices uttered the same sentences. Listeners indicated the likelihood that the speaker was referring to one among two target objects varying along gender-stereotypical characteristics. Listeners envisaged a more "feminine" object when the sentence was uttered by a gay-sounding speaker, and a more "masculine" object when the speaker sounded heterosexual. In Studies 2a/2b, listeners were asked to disambiguate sentences that involved a stereotypical behaviour and were open to different interpretations. Listeners disambiguated the sentences by interpreting the action in relation to sexual-orientation information conveyed by voice. Results show that the speaker's voice changes the subjective meaning of sentences, aligning it to gender-stereotypical expectations.

Keywords

Voice changes meaning: The role of gay- vs. straight-sounding voices in sentence interpretation

Communication implies an exchange of messages between source and recipient (Griggin, 2009), yet such messages are often open to different interpretations. To facilitate the exchange, recipients look for cues that help them to elaborate and interpret the message. In verbal communication, one of these cues is the speaker's voice, which carries information about the person speaking (see Kreiman, 1997). For instance, when a person says "*my dog is running in the park*", the type of dog the speaker is referring to is not specified. The dog could be of any breed, dimension, or color, and listeners may depict the dog differently depending on who is speaking, as they may have expectations about the kind of dog the speaker likes. Hence, the speaker's voice becomes a key element in message interpretation. Understanding what the listeners infer from messages is fundamental to comprehend how communication and interlocutors' relationship evolve.

People use social categorization and stereotypes to deal with the great amount of information they are exposed to (Macrae, Milne, & Bodenhausen, 1994). In this process, other individuals are often categorized as members of social groups based on minimal cues, including their voices (Dragojevic, 2018; Fasoli, Maass, & Sulpizio, 2016). In this way, they become associated with stereotypes referring to the social category to which they belong (Macrae & Bodenhausen, 2001). The present research examined whether social information conveyed by speakers' voice impacts message interpretation. In particular, we tested whether a speaker's voice indicative of his/her sexual orientation (henceforth SO) shapes the representation of a target object mentioned in the message (Studies 1a/1b) and facilitates the disambiguation of ambiguous messages in line with stereotypical expectations associated with the speaker (Studies 2a/2b).

Message Interpretation

In everyday communication, most messages are open to multiple interpretations (Edwards, 1998). According to a model proposed by Bodie, Worthington, Imhof, and Cooper (2008), the listening process includes a decoding phase, in which recipients not only process the literal meaning of the words, but also assign a specific connotation to the message (Edwards, 2011; Imhof, 2010). This process may involve expectations and knowledge about who conveyed the message, about contextual factors, such as the words used, and about the speaker's prosodic features and speech sound (Bodie, 2011; Bodie et al., 2008).

On the one hand, to interpret or disambiguate messages, people rely on elements that are present in the message, such as nouns, pronouns, verbs or entire clauses. For instance, causality in a sentence is interpreted differently depending on the type of verb used (Garvey & Caramazza, 1974) and an agent is held more responsible when the sentence is presented in active rather than passive form (Platow & Brodie, 1999). Similarly, both the definitional gender (e.g., king) and stereotypic gender (e.g. nurse) of nouns impact sentence disambiguation processes by slowing down comprehension in cases of mismatches (e.g., Kreiner, Sturt, & Garrod, 2008). Studies using written scenarios have further shown that behaviors are interpreted according to the stereotypes associated with the actor. For instance, the same behavior is interpreted differently, depending on whether it is performed by a man or a woman (Edwards, 1998), a Black or White individual (Sagar & Schofield, 1980), or by someone of high or low socioeconomic background (Darley & Gross, 1983). Messages are also interpreted literally or as sarcastic depending on the speaker's group membership because stereotypical expectations are used to disambiguate the speaker's intent (Pexman & Olineck, 2002).

On the other hand, in spoken communication, message interpretation can also be affected by "contextual factors" such as the speaker's voice characteristics and prosodic patterns

(Badzinski & Gill, 1994; Bodie et al., 2008; Imhof, 2010; Schafer, Carlson, Clifton, & Frazier, 2000). Prosodic features are those characteristics of speech (e.g., pitch, intonation, pause, etc.) that go beyond the phonemes and that listeners interpret, often unconsciously (Lehiste & Lass, 1976). For instance, listeners use voice pitch and intonation to interpret the speaker's intentions, assertiveness, and persuasiveness (Apple, Streeter, & Krauss, 1979; Fernald, 1989). Voice sound is associated with mental representations as shown by the fact that non-existing words said with a happy, rather than a sad tone, were more likely associated with happy than sad target pictures (Nygaard, Herold, & Namy, 2009), and ambiguous words are processed in line with the voice emotional tone (Nygaard & Lunders, 2002). Moreover, prosodic features are used to solve the inconsistencies between vocal intonation and message content (Badzinski, 1991; Bugental, 1974; Cohen, Douaire, & Elsabbagh, 2001) and to interpret ambiguous messages as either serious or sarcastic (Glenwright, Parackel, Cheung, & Nilsen, 2014).

Vocal cues can also convey the speaker's social identity, especially when other, more explicit information about the person is unavailable. The idea that phonetically cued information about the speaker influences comprehension is at the center of Sumner, Kim, King, and McGowan's (2014) *dual-route approach to speech perception*. According to this model, one route involves the actual encoding of words as elements that have a lexical meaning that is processed when listening to someone communicating. This route represents the "lexical representation" of the message. The other route involves encoding of acoustic patterns leading to a "social representation". Specifically, since listeners have learned and stored in their mind that specific acoustic patterns refer to different social groups (gender-based, nationality-based, etc.), such information is activated when listening to a message. Thus, according to the dual route model, listeners simultaneously extract linguistic and social information from speech. For

instance, pitch is processed simultaneously as a linguistic cue and as signaling speaker gender (Kaganovich, Francis, & Melara, 2006). Importantly, the two processes determine comprehension in an interactive fashion, or, as Sumner et al. (2014, p. 10) state "comprehension is the composite of social and linguistic activation".

Although the model was initially developed for speech perception, it provides a useful general framework to explain the process we predict to occur in message interpretation. This is in line with Badzinski and Gill's (1994) argument that the interpretation of features such as words and sounds is integrated with the information coming from the source and the context, with the aim of listeners to achieve a coherent interpretation. There is multiple evidence supporting this "*social weighted encoding*" process: For instance, accent, prosodic patterns, TH-fronting in English or sibilant /s/ can be taken as vocal markers of typical street-talk/city-slang, social status, and SO, respectively (Campbell-Kibler, 2011; Levon, 2015; Mack & Munson, 2012; Pharao, Maegaard, Møller, & Kristiansen, 2014). Similarly, when judging non-standard accented speakers, listeners use accent and intonation to guess the speaker's origin and nationality (Dragojevic, Berglund, & Blauvelt, 2018; Gnevsheva, 2016). Hence, voice sound conveys information that leads listeners to categorize speakers as socially rooted within social groups. This activates social knowledge, which in turn affects message decoding. In fact, words stored in memory as terms usually uttered by women are recognized faster if said by a woman than a man (see King & Sumner, 2015; for age see Walker & Hay, 2011; for regional accent, see Sumner & Kataoka, 2013), listeners recall better stereotypical national names if they are uttered by speakers with a congruent accent (Senior, Hui, & Babel, 2018), and facilitation effects occur when there is a match between the speakers' voices and listeners' expectations about their nationality

(McGowan, 2015). Moreover, different semantic associations are elicited by the same words spoken by a male or female speaker (King & Sumner, 2015).

Neuroscience research has also provided evidence that the speaker's voice affects message elaboration at the brain level. Lattner and Friederici (2003) found that messages with stereotypical content spoken by male and female speakers elicited a P600 – an event-related component, namely a modulation of electrical brain activity signaling difficulty in message elaboration – when the message content and the speaker's gender were incongruent. Moreover, Van Berkum, van den Brink, Tesink, Kos, and Hagoort (2008) showed that when the speaker's voice is incongruent with message content, there is a difficulty in integrating the two types of information at the brain level and making sense of the content, as shown by a higher activation of N400, which is usually observed when processing words that are semantically anomalous or inconsistent.

Together, these findings suggest that speaker's voice is taken into consideration in message elaboration. However, most prior research has focused on how vocal features conveying group membership inhibit (or facilitate) comprehension, by demonstrating either "surprise" reactions or slowdowns in reading and comprehension when encountering mismatching or incongruent information. Yet, to our knowledge, research has rarely examined how the speaker's social identity conveyed by voice may impact *message interpretation*. Only a small body of literature has shown shifts in interpretation of meaning as a function of speaker's voice. For instance, accent, speech rate, and age-sound all affect the interpretation of the exact same message (Giles, Henwood, Coupland, Harriman, & Coupland, 1992). We extend this work by testing the interplay between voice signaling SO and gender stereotypes related to the message content. We suspect that stereotypes play an important role in such contexts. To reach a "coherent"

interpretation, message content is likely to be interpreted in line with speaker stereotypes and this may have consequences for stereotype maintenance and stigmatization.

Sexual Orientation and Gender Stereotypes

Stereotypes represent oversimplified generalizations about characteristics associated with a group and its members that help individuals navigate the complexity of the world (Macrae et al., 1994). Communication often involves stereotypes. On the one hand, individuals have stereotypical expectations about their interlocutors. Such stereotypes "guide" communication and help individuals to anticipate and adapt to the interaction (Burgoon, 1993). On the other hand, stereotypical, rather than counter-stereotypical messages, are the most likely to be communicated (Lyons & Kashima, 2003). Message interpretation is an intrinsic part of interpersonal and intergroup communication that affects how interlocutors relate to each other (Giles, 2016) and process information (Edwards, 2011). It is therefore important to understand whether and when messages are interpreted according to stereotypes, as this defines the relationship and interaction between interlocutors belonging to different groups. Also, by interpreting messages in line with stereotypes, individuals may contribute to their persistence as shared knowledge since they reiterate them (Devine, 1989). This can be particularly important for those groups that are targets of negative stereotyping and stigma such as sexual minorities (Herek & McLemore, 2013).

Gay and lesbian individuals are targets of stereotyping and stigmatization, even when their SO is not explicitly disclosed but "detected" by others from minimal cues such as voice (Fasoli et al., 2016). An increasing body of research has shown that listeners categorize speakers' SO according to how they sound (Gaudio, 1994; Linville, 1998; Smyth, Jacobs, & Rogers, 2003; Munson, 2007), even when they speak a foreign language (Sulpizio et al., 2015; Sulpizio, Fasoli, F., Antonio, R., Eyssel, F., Paladino, M. P., & Diehl, 2019), talk for a few seconds (Mack &

Munson, 2012), or utter meaningless syllables rather than conveying a meaningful message (Tracy, Bainter, & Satariano, 2015). Sometimes this categorization process happens to be accurate, whereas at other times, it occurs on the basis of stereotypical ideas of how gay men and lesbian women are believed to sound (for an overview see Fasoli et al., 2016; Rule, 2017). As a matter of fact, men who produced fronted /s/, have a soft voice, and sound feminine (Mack & Munson, 2012; Munson, 2007), and women who have a low-pitched and masculine sounding voice (Munson, 2007; Van Borsel, Vandaele, & Corthals, 2013) are likely to be categorized as gay and lesbian, respectively. Research has also shown that different acoustic features are associated with gay-sounding voices. For instance, the sibilant /s/, fundamental frequency and vowel space dispersion and speaking rate affect perceived SO (Sulpizio et al., 2015). Importantly, Munson (2007) has shown that different acoustic cues predict perception of SO and voice gender typicality (masculinity/femininity).

Since vocal cues lead to SO categorization, they also elicit stereotype-based inferences about speakers. Although multiple stereotypes about gay men and lesbian women exist, these two sexual minorities are often stereotyped along gender characteristics. Gender Inversion theory (Kite & Deaux, 1987) suggests that gay men and lesbian women are seen as similar to their opposite gender. For instance, gay men are perceived as warm but lacking in competence whereas lesbian women are rated as competent but less warm (Vaughn, Teeters, Sadler, & Cronan, 2017; but see Fiske, Cuddy, Glick, & Xu, 2002). Gender and SO stereotypes go hand in hand even when voice is considered. Listeners often judge gay-sounding speakers as more feminine, and lesbian-sounding speakers as more masculine than straight-sounding male and female speakers, respectively (Munson, 2007) and such stereotyping extends to personality traits and personal interests (Fasoli, Maass, Paladino, & Sulpizio, 2017).

Although research on voice-based identification of SO is expanding rapidly, the link between gay/straight voices and the content of the conveyed message has received little attention. Some researchers have investigated whether message content may affect perception of speakers' SO. Gaudio (1994) examined whether a speaker reading out an "emotional/dramatic" (i.e., a gay character's monologue) rather than a neutral text would be more likely to be perceived as gay. Results showed no influence of message content on SO judgments. Subsequently, Smyth, Jacobs, and Roger (2003) compared perceptions of speakers' SO when they were speaking spontaneously as opposed to when they were reading a scientific or a dramatic text. Surprisingly, they found that speakers were perceived as more gay- and feminine-sounding when reading a scientific text. The authors suggested that, perhaps, the formal content of the message activated the stereotype of gay men being formal and educated. Also, personal information such as self-reported hobbies can be taken as cues of SO (Rieger, Linsenmeier, Gygax, Garcia, & Bailey et al., 2010). Thus, there is some evidence that the content of the message may affect the perception of the speaker as gay or straight.

To our knowledge, no research has addressed the opposite path, that is, whether SO conveyed by voice influences message interpretation. If voice is taken as a SO cue, it is likely to affect communication: when heterosexuals imagine themselves talking with gay men, they expect specific communication schemas to emerge (Hajek & Giles, 2005). It is therefore likely that what is being said by gay-sounding voices is interpreted according to such expectations. We expect here that people align their interpretation of the message with the stereotypes that are activated by a gay/lesbian vs. straight sounding voice. Since gay men and lesbian women are predominantly stereotyped along gender dimensions (e.g., masculinity and femininity; see Blashill & Powlishta, 2009), in this research we focused on gender stereotypes. Thus, we aim to

provide first evidence that SO vocal cues trigger gender stereotype-based message interpretation and, thus, support future examination of other sexual stereotypes.

Overview

The main aim of this research is to show that vocal information that cues speakers' SO changes the interpretation of what is being said because vocal information revealing social information (age, gender, regional background, etc.) is integrated with lexical information. Thus, we predicted that people would interpret the message in line with the stereotypes related to the social category conveyed by the speaker's voice, so as to match their stereotypical expectations.

We used two different methods to test this general hypothesis. In Studies 1a and 1b, participants listened to single sentences (such as "*My dog runs in the park"*) uttered by both gay- or straight-sounding speakers. Participants were then shown pairs of pretested photos (e.g., a Doberman and a Chihuahua) and asked to indicate which of the two images better reflected the type of dog the speaker was referring to. We expected that participants would judge gay-sounding speakers as more likely to refer to a feminine target (Chihuahua) than when the speaker sounded straight (Hypothesis 1). The rationale underlying this prediction is that the speaker's voice conveys information about his SO and activates associated stereotypes, which in turn channel the interpretation of the message in line with these stereotypes. Since small dogs such as Chihuahuas are stereotypically associated with women and gay men, this may be the dog that listeners envisage when listening to a gay-sounding man referring to his dog. This first set of studies is expected to show that listeners form different images of the same statement as a function of the vocal SO information.

In the second set of studies (Study 2a and Study 2b), the same general hypothesis was tested using a different paradigm. We used ambiguous sentences that have been largely

employed in the psycholinguistic literature to study sentence comprehension (see John & McClelland, 1990). Here, we considered ambiguous sentences (e.g. "*I met my [male] friend with the ballet shoes*") in which the critical action or object (in the case, "*the ballet shoes*") could either refer to the speaker or to the other protagonist ("*the friend*"). We also varied the voice of the speaker, expecting listeners to use the vocal information to disambiguate the sentence. We hypothesized that listeners would resolve the ambiguity based on vocal information in the direction implied by the speakers' voice. In the example cited here, the ballet shoes would be more likely to be seen as the property of the speaker when the voice sounded gay rather than straight because gay men are stereotypically perceived as doing stereotypically feminine sports and liking dancing. Hence, stereotypically gay/lesbian behaviors mentioned in an ambiguous sentence would be more likely interpreted as performed by the speaker if he/she sounds gay/lesbian (Hypothesis 2).

Study 1: Does Voice Impact the Presumed Object Being Referred to?

The first two studies examined whether vocal features associated with speaker SO impact how listeners perceive generic objects referred to in a sentence. That is, do listeners envisage different exemplars when the general category referred to in a sentence is said by a straight- vs. gay-sounding voice? In the studies below, the sentence referred to a general category of objects (e.g., a dog or a car) without specifying an exact exemplar. Participants were required to select the likely exemplar of the general category being referred to in the sentence. Based on pretesting, the exemplars were selected so as to be perceived as stereotypical of men or women, and also of gay or straight men (see Kite & Deaux, 1987). We predicted that listeners would integrate vocal and semantic information so as to create a coherent impression. Thus, a gay-sounding voice was expected to evoke different exemplars of the same category (for instance, a Chihuahua rather

than Doberman, and a small car such as a Smart rather than a SUV) than a straight-sounding voice.

Study 1a

Method

Participants. Participants were Portuguese psychology students who took part in the study in exchange for credit. Of the 37 participants, one was excluded because the participant self-identified as gay. The final sample consisted of 36 Portuguese participants (5 males, M_{age} = 19.50, SD = 3.47) who, on average, knew approximately 4 gay men (M = 4.31, SD = 3.87).

Speakers. Variations among speakers' voices exist and speakers can sound gay or straight regardless of how they identify (see Sulpizio et al., 2015). However, for the purpose of this and the following studies, we selected speakers who self-identified as gay/lesbian or straight and whose voices were perceived as accurately signaling their gay/lesbian or straight SO, respectively. SO voice-based judgments are often triggered by voice stereotypes and SO-related acoustic cues (see Fasoli et al., 2016). Hence, the gay- and straight-sounding voices used in this research were likely to reflect stereotypical ideas of how speakers of different SO sound.

Two gay and two straight male speakers were selected from a pool of pre-existing recordings that included the sentences uttered in the current study. A pilot test (N = 22) on the four voices assured that they were perceived as gay- or straight-sounding. Speakers were judged on a scale from 1 (*exclusively heterosexual*) to 6 (*exclusively homosexual*). Straight speakers were judged as more heterosexual (M = 2.67, SD = .63) than gay speakers (M = 3.66, SD = .65), $t(21)$ = 5.15, p = .001.

Sentences. Speakers uttered four sentences, each referring to one of 4 different targets, namely a dog, a car, a profession, and a sport (see Table 1).

Pictures. Pictures referring to four target categories, namely dogs, cars, sports, and professions were selected on the basis of pretests ($N = 39$). Participants indicated how stereotypical each target was of men, women, gay and straight men (e.g., "How typical is it that a [man] has this car?"). Answers were provided on a scale from 1 (*not at all*) to 7 (*completely*). Pairwise *t*-tests were performed between ratings of different targets. Two targets of each category were selected in order to represent targets more stereotypical of males and of females (all *t*s > 4.57, *p*s < .001). We also compared whether the targets were perceived as stereotypical of gay and straight men, respectively. The feminine targets were perceived as more stereotypical of gay than straight men, and vice versa for the masculine targets (all *t*s > 2.29, *p*s < .04). The same objects that were seen as stereotypical of women were also seen as stereotypical of gay men, whereas those stereotypical of men were also rated to be stereotypical of heterosexual men. The target objects were the following (*dog*: Chihuahua vs. Doberman; *car*: Smart vs. SUV; *job*: stylist vs. mechanic; *sport*: dance vs. football).

Procedure. Participants completed the study in the lab. Before putting on headphones, participants were informed that the study was about how people interpret messages. No reference to speakers' SO was made. After consenting to take part in the study, they were instructed to listen to audio recordings of sentences, all of which included a self-reference such as "my" or "I" ("my dog is running in the park"), uttered by four (2 gay and 2 straight) speakers. Participants listened to the same sentences 4 times, each time uttered by one of the 4 speakers. Hence, participants listened to a total of 16 audio recordings (4 sentences X 4 speakers) in a randomized order. After listening to each audio recording, participants indicated the likelihood that the speaker was referring to one or the other of two (a masculine and a feminine) targets mentioned in the sentence (i.e., a dog, a car, a profession, or a sport), which were presented in pictures

positioned at the end points of a 6-point Likert scale. The position of the feminine and masculine target on the left/right pole of the scale was counterbalanced across participants. After completing this task, participants were asked to provide basic demographic information (age, gender, nationality, SO), report the number of gay men they knew, and if they had encountered technical problems in listening to the audio files. In addition, they indicated how many of the 4 speakers they had listened to were gay by indicating a number from 0 to 4.

Results

Participants' ratings were re-coded such that the higher the score, the more participants chose the masculine target. Since the pattern of results was the same across the four items (dog, car, profession, sport), we calculated the reliability of the 8 sentences (4 for each speaker) uttered by the two gay- ($\alpha = .52$) and two straight-sounding speakers ($\alpha = .60$) and averaged the ratings in two indexes, one for gay and straight speakers respectively.

A 2 (Speaker SO: gay vs. straight) repeated-measure ANOVA was performed on participants' ratings. As predicted, participants were less likely to envisage the masculine object when the speaker was gay-sounding ($M = 3.50$, $SD = .64$) than when he was straight-sounding ($M = 3.99$, $SD = .62$), $F(1,36) = 16.92$, $p < .001$, $\eta_p^2 = .32$. [1]

Study 1b

Study 1b aimed at replicating the findings of the previous study using different voices and increasing the listener sample.

Method

Participants. Fifty-five Portuguese participants took part in the study. After excluding those who self-identified as gay/bisexual ($n = 5$), the final sample consisted of 50 participants (11 males, $M_{age} = 22.41$, $SD = 4.00$). On average, they knew 6 gay men ($M = 5.87$, $SD = 3.70$).

Speakers. Two gay and two straight speakers (different from the ones in Study 1a), whose voices conveyed the SO they identified with, were selected on the basis of a pretest. Pretest participants ($N = 17$) judged 18 speakers' SO both on a dichotomous choice (gay vs. straight) and on a Likert scale from 1 (*exclusively heterosexual)* to 7 (*exclusively homosexual*), after the speakers uttered each of the 4 target sentences. The two gay speakers were correctly categorized as gay by 70.6% of participants, whereas the two straight speakers were correctly categorized as straight by 88.2% of participants. Ratings on the Likert scale confirmed this categorization, such that gay speakers ($M = 4.55, SD = .87$) were rated on average as more gay than straight speakers ($M = 2.12, SD = .77$), $t(15) = 7.68, p < .001$ and both were collocated at the correct side of the scale.

Procedure. The procedure was identical as that of Study 1a, with the exception that the study was run online by providing the study link via email or posting it on social networks.

Results

Responses were recoded so that higher scores meant a greater likelihood of choosing the masculine target, and were averaged across objects (dog, car, profession, sport). This was done separately for the 8 sentences uttered by the gay speakers ($\alpha = .73$) and for the 8 sentences uttered by the straight speakers ($\alpha = .67$). The same analysis as in Study 1a was then conducted. As predicted, participants were more likely to indicate the more feminine objects when the speaker uttering the sentence sounded gay ($M = 3.15, SD = .73$) than when he sounded straight ($M = 4.10, SD = .64$), $F(1,54) = 76.19, p < .001, \eta_p^2 = .58$.[2]

Discussion

The two studies confirmed that people use gender stereotypes associated with straight and gay individuals to interpret sentences and to envisage the objects alluded to. In particular, depending on whether the voice sounded gay or straight, they activated different exemplars of the object

category the speaker was referring to. They envisaged more feminine objects (for instance a

Chihuahua and dancing) or associated the masculine object less with the speaker when the voice

sounded gay, despite the fact that the sentences were identical in both cases. Thus, the exemplars

that became activated and available when a general category was alluded to in a sentence

changed in line with the speaker's voice and with the SO-related stereotypes associated with that

voice. This supports Sumner et al.'s (2014) contention that social information revealed by vocal

cues is integrated with semantic content. Importantly, voices are not only spontaneously

interpreted as reflecting the speaker's SO, but they also guide the interpretation of what is being

referred to.

Study 2: Ambiguous Sentences. Whom is the Speaker Referring to?

In the second set of studies we investigated the same general question with a different

methodology. We presented ambiguous sentences that may either be interpreted as referring to

the speaker or to another person mentioned in the sentence. We hypothesized that listeners, in

order to interpret the sentences, would use vocal cues to resolve the ambiguity. Take the example

of the sentence "*I met my [male] friend with the ballet shoes".* To understand who has the ballet

shoes (the speaker or the friend), vocal information may come in handy. If the voice sounds gay,

people may disambiguate the sentence envisaging the shoes being worn by the speaker because

gay men are stereotyped as liking dancing and ballet. On the contrary, if the speaker sounds

straight, listeners may believe that the man rather than the speaker has the ballet shoes. Besides

changing the stimuli in Study 2, we also extended our research by examining both male and

female voices.

Study 2a

Method

Participants. 60 Portuguese participants accessed and completed the online survey. Four

participants did not identify as heterosexual and one as a non-native speaker. Hence, the final

sample consisted of 55 heterosexual Portuguese participants (20 males; M_{age} = 23.29, SD = 5.60).

Speakers. Two gay/lesbian and two straight speakers of each gender were selected. Male

speakers were the same as those of Study 1b, whereas 4 female speakers were selected on the

basis of a previous study (N = 56) where participants judged speakers' SO on a scale from 1

(*exclusively heterosexual*) to 7 (*exclusively homosexual*). The two lesbian speakers (M = 3.34,

SD = 1.21) were judged as more lesbian-sounding than the two straight speakers (M = 2.33, SD =

1.09), t(55) = 5.35, p < .001.

Sentences. A set of 5 sentences for male and 5 for female speakers were created and pretested.

Each sentence involved an action (e.g., buying something) and two targets (i.e., the speakers and

another person). The sentences were formulated in a way that both the speaker and the other

person could have performed the action (see Table 1) and subjected to two pretests. The first

pretest (N = 28; all Portuguese) served to select sentences with the highest possible level of

ambiguity. Participants read a list of sentences and indicated for each which of the two people

had most likely performed the action, using a scale from 1 (*the speaker*) to 7 (*the other person*)

where 4 indicated that both targets were equally likely. On the basis of this pretest, we selected

two sentences for male (buying a vase, using ballet shoes) and two for female speakers (buying a

motorcycle, drinking a beer) that differed reliably from both scale endpoints and that were as

close as possible to the scale midpoint. Although there was an overall tendency to attribute the

actions slightly more to the speaker than to the other person, means were relatively close to the

scale midpoint (for male speakers: M = 3.22, SD = 1.34, t-test against the midpoint: t(26) = -3.02,

$p = .006$; for female speakers: $M = 3.74$, $SD = .68$; t-test against the midpoint: $t(26) = -1.97$, $p = .06$).

A second pretest (n = 19) served to assure that the actions were perceived as more stereotypical of gay/lesbian than of straight speakers. All 10 actions were rated on scales from 1 (*more typical of straight men*) to 7 (*more typical of gay men*) and on scales from 1 (*more typical of straight women*) to 7 (*more typical of lesbian women*). The two ambiguous sentences for male speakers (see pretest 1) were indeed perceived as stereotypical of gay men ($M = 5.34$, $SD = 1.04$, t-test against the midpoint: $t(18) = 5.62$, $p < .001$) and the two ambiguous sentences for female speakers as stereotypical of lesbian women ($M = 4.52$, $SD = .75$, t-test against the midpoint: $t(18) = 3.04$, $p = .007$; for sentences see Table 1).

Procedure. Participants were recruited online. Before starting the study, participants were informed that the study was about how people interpret messages and that it involved listening to audio files. Then, they were presented with a consent form and they formally agreed in taking part in the study. Next, participants listened to one speaker at a time. Sentences for male speakers contained a target action/object that was stereotypically gay-related (e.g., "Yesterday, my friend Mike and I went shopping for an interior design vase*"*), whereas sentences for female speakers involved a stereotypically lesbian-related target action/object (e.g., "Yesterday, Katia and I were at the shop to buy a motorcycle"). Each speaker uttered the two ambiguous sentences selected for his/her gender. Participants listened to 16 sentences in total: 8 sentences uttered by 2 gay and 2 straight male speakers and 8 sentences uttered by 2 lesbian and 2 straight female speakers. Thus, participants listened to the same sentence four times. The order of speaker and sentence presentation was randomized across participants. Participants were first asked to indicate who was performing the action by answering a scale from 1 (*The speaker*) to 6 (*Name of the other*

person mentioned in the sentence). Subsequently, participants completed a two-item measure (α = .95) asking to report how similar they considered themselves to the speaker and how much they identified with the speaker, respectively. Answers were provided on a scale from 1 (*not at all*) to 5 (*completely*). This measure was introduced for exploratory purposes unrelated to the main hypothesis and results reported in Supplementary Materials. Participants then indicated how much they would have liked to perform each of the actions mentioned in the sentences (i.e., buying a vase, using ballet shoes, buying a motorcycle, drinking a beer) and completed a thermometer measure ("Please indicate how favorable and warm you feel toward each group listed below", from 0 to 100) that was included to test attitudes toward different groups, namely heterosexual men, heterosexual women, bisexual men, bisexual women, gay men and lesbian women. Finally, participants indicated their gender, age, SO, native language, and completed two items in which they were asked to indicate the number (from 0 to 4) of gay/lesbian speakers they had listened to. Finally, they were debriefed and thanked.[3]

Results

A 2(Speaker SO: gay/lesbian vs. straight) x 2 (Speaker Gender: male vs. female) x 2 (Participant gender: male vs. female) mixed ANOVA with repeated measures on the first two variables was performed. Ambiguity ratings served as dependent variable, with higher scores indicating a greater likelihood that the male speaker was seen as performing the "gay" action and the female speakers as performing the "lesbian" action. The only reliable effect that emerged was a main effect for Speaker SO, $F(1,52) = 4.27$, $p = .044$, $\eta_p^2 = .08$. Listeners were more likely to attribute the ambiguous action to the speaker when s/he sounded lesbian/gay (*M = 4.44, SD = .88*) than when s/he sounded straight (*M = 4.13 SD = 1.14*). Given that the actions were stereotypically gay or lesbian, this pattern is perfectly in line with the predictions.

Study 2b

A limitation of Study 2a is that sentences were created to merely test the disambiguation

hypotheses when male and female speakers uttered objects/actions that were stereotypically

incongruent for their gender but associated with gay men or lesbian women. Thus, for instance,

male speakers uttered sentences involving feminine/gay-related objects/actions (e.g., ballet

shoes) but none including masculine/straight-related objects/actions. To overcome this limit, in

our final study, male and female speakers were uttering sentences that involved both

stereotypically congruent or incongruent objects/actions (e.g., male speakers uttering sentences

about football is an example of gender stereotypically congruent object/action whereas shopping

for an interior design vase is considered gender stereotypically incongruent object/action). As in

our previous study, we predicted that listeners would interpret the sentences as referring to the

speaker (rather than the other protagonist) when voice and content were congruent (e.g.

gay/lesbian voice and stereotypically gay/lesbian behavior or straight voice and stereotypically

straight behavior), but not when voice and content were incongruent (e.g. gay/lesbian voice and

stereotypically straight behavior). Hence, we predicted that voice features (gay/lesbian vs.

straight sounding) would channel the disambiguation in the direction of the assumed SO. A

secondary aim of Study 2b was to extend our previous research to a different cultural and

linguistic context, namely Italy.

Method

Participants. In total, 204 Italian participants completed the survey (129 males, 75 females, M_{age}

= 31.94, SD = 18.78). In contrast to our previous studies, the sample contained a considerable

subsample of sexual minority participants (n = 31), the majority of whom identified as

gay/lesbian (n = 19) or bisexual (n = 6). We therefore maintained sexual minority participants in our analyses, using the participants' SO as predictor variable.

Speakers. Speakers were selected on the basis of two pretests, one for male and one for female speakers (N = 95 in total), in which participants rated 18 male and 14 female speakers' SO on a scale from 1 (*exclusively heterosexual*) to 7 (*exclusively homosexual*). We selected two speakers for each gender whose voice was perceived as sounding accordingly with the SO they self-identified with. The two gay speakers (M = 5.81, SD = 1.43) were rated as more gay-sounding than two straight male speakers (M = 1.74, SD = .93; $t(55)$ = 16.11, $p < .001$). Also, the two lesbian speakers (M = 4.14, SD = 1.76) were rated as more lesbian-sounding than the two straight female speakers (M = 2.41, SD = 1.31; $t(39)$ = 5.34, $p < .001$).

Sentences. Sentences were selected on the basis of two pretests (N = 72 in total, all Italian students). Each pretest included 12 sentences. Overall, participants rated the ambiguity of the sentences on a scale from 1 (action performed by the speaker) and 7 (action performed by the other person mentioned in the sentence) with 4 indicating an equal likelihood that the action was performed by the speaker or by the other person. Moreover, each behavior mentioned in the ambiguous sentences was judged in terms of gender stereotypicality (from 1 = typical of men to 7 = typical of women) and SO stereotypicality (from 1 = typical of straight men/women to 7 = typical of gay men/lesbian women), namely in terms of stereotype associated to gender and SO groups. Based on both pretests we selected 8 sentences as stimulus materials for the main study. The selected sentences were all perceived as ambiguous as they did not differ from the scale midpoint (one sample t-test, ts < -1.75, ps $> .09$). Moreover, half of the actions were perceived as more gender stereotypical and as more stereotypical of straight men/women (one sample t-test comparing means to scale midpoint, ts > 6.45, ps $< .001$). The other half were perceived as more

gender counter-stereotypical and as more stereotypical of gays/lesbians (one sample compared to

scale midpoint t-test, ts > 3.34, ps < .002; for sentences see Table 1).

Procedure and materials. The procedure was identical to that of Study 2a, except for the

following characteristics. First, the study was run in Italian rather than Portuguese. Second, given

that participants listened to a greater number of sentences and, hence, the experimental task

required more time, we did not assess identification with the speaker. Participants listened to 8

pretested sentences. Two of these described stereotypically straight behaviors, one stereotypical

of straight men (i.e., playing soccer) and the other of straight women (i.e., taking dance lessons).

Of the remaining six sentences, half described behaviors stereotypical of gay men (i.e., getting an

eccentric haircut, not wanting a stable relationship, buying a bracelet), the other half were

behaviors stereotypical of lesbians (i.e., getting a short haircut, having a pick-up, getting a

tattoo).

Moreover, all sentences were ambiguous, such that it was unclear whether the speaker was

referring to him- or herself or to the other person (e.g., "I went to meet the guy with the pick-

up"). Depending on content, sentences were either uttered by female or male speakers. The

participants' task was to indicate on a 7-point whether the speaker was referring to him- or

herself or to the other person (e.g. "Who has the pick-up, the speaker or the guy?"). Before

ending the survey, participants reported their age, gender, SO, number of gay/lesbian friends, and

how many gay and lesbian voices they had listened to (from 0 to 4).[5]

Results

Reactions to the two behaviors that were counter-stereotypical of gays and lesbians were

reversed-coded such that in all cases, high values indicated a stereotype-congruent

disambiguation of the sentence. We averaged the ratings for each group of speakers varying in

terms of gender and SO. Next, a 2 (Speaker SO: gay/lesbian vs. straight) x 2 (Participant Gender: male vs. female) x (Participant SO: sexual minority vs. straight) ANOVA with repeated measures on the first variable was conducted on the disambiguation ratings. The analysis revealed only a main effect for Speaker SO, $F(1, 202) = 6.85$, $p = .028$, $\eta_p^2 = .02$. As predicted, listeners disambiguated the sentences based on speakers' vocal information, assuming that the speaker was referring to him/herself more when content and SO conveyed by voice matched ($M = 4.50$, $SD = 1.09$) than when it did not ($M = 4.23$, $SD = 1.09$). Disambiguation was not modified by the participants' gender nor by the participant's SO when added to the analyses.

A different way to look at the same data is to calculate an index of voice-content match and one for voice-content mismatch. For the voice-content match index, we averaged ratings for straight and gay/lesbian speakers associated with stereotypically gender congruent and incongruent behaviors, respectively. The opposite was done for the voice-content mismatch index. A 2 (Voice-content: match vs. mismatch) x 2 (Participant Gender: male vs. female) x 2 (Participant SO: sexual minority vs. straight) ANOVA with repeated measures on the first variable was performed. A main effect of Voice-content, $F(1,200) = 4.60$, $p = .033$, $\eta_p^2 = .02$, showed that listeners disambiguated the sentences based on speakers' vocal information assuming that the speaker was referring to him/herself more when content and SO conveyed by voice matched ($M = 4.59$, $SD = 1.27$) than when it did not ($M = 4.19$, $SD = 1.45$). Disambiguation was not modified by the participants' gender or SO ($Fs < 2.17$, $ps > .14$).

Discussion

The findings of Study 2 are conceptually in line with the previous set of studies, showing that people use vocal cues to disambiguate sentences. The same ambiguous statements were attributed to the speaker (rather than to the other person mentioned in the sentence) when the

described action was stereotypically in line with the SO conveyed by the speakers' voice. Thus, confronted with ambiguous sentences, people rely on vocal cues (in this case telling of SO) to make sense of what is being said. The same pattern of results emerged when we considered actions stereotypical of gay men and lesbians (Study 2a), and when we included behaviors stereotypical of straight men and women (Study 2b). Although there was an overall tendency to believe that the speaker referred to him/herself, this assumption was stronger when vocal features matched the sentence (gay/lesbian-sounding voice and gay/lesbian-stereotypical behavior or straight-sounding voice and heterosexual behavior) than when it did not.

General Discussion

In verbal communication, the speaker's voice is likely to affect how the message is interpreted. Across four studies, we demonstrate that the speaker's voice influences the way listeners envisage a specific object mentioned in the message (Studies 1a/1b) and the way they disambiguate sentences and infer that a specific behavior was or was not performed by the speaker (Studies 2a/2b). In the first set of studies, the target object mentioned by a gay-sounding speaker was more likely to be imagined as a stereotypically feminine object. For instance, the dog referred to by a gay-sounding speaker was imaged as a Chihuahua rather than as a Doberman. Similarly, in the second set of studies, a stereotypically gay/lesbian behavior was more likely to be attributed to the speaker (than to another person mentioned in the ambiguous sentence) if the speaker sounded gay/lesbian. Thus, a gay-sounding speaker is more likely to be perceived as the one performing a feminine action (e.g., buying a design vase) and a lesbian-sounding speaker is more likely to be engaged in a stereotypically masculine action (e.g., buying a motorcycle). In both studies, the message was interpreted in line with the stereotypes related to the social category conveyed by the speaker's voice, suggesting that voice is an important cue

affecting message interpretation. Thus, listeners used oversimplified and stereotypical knowledge about SO to make inferences about what the speaker was saying.

These findings contribute to the current literature in several respects. First, this research tested and complemented Sumner et al.'s dual-route model of speech perception. Our data support the idea that speakers process both content and social information revealed by voice and that the two are integrated into a single, coherent impression. In our studies, speakers' voices did affect how listeners perceived the message confirming that speech perception goes beyond the mere encoding of semantic information. While the existing literature has already demonstrated that voice-message incongruence hampers message comprehension (e.g., Dragojevic, 2018) and leads to difficulties in processing the message content at the brain level (Lattner & Friederici, 2003; Van Berkum et al., 2008), the present research supports the idea that voice also matters in how listeners explicitly interpret a message that is open to various interpretations. This finding is quite different from prior work and the context within which the dual-route model has been developed, showing that reading time and sentence comprehension slowdown in the presence of voice-content incongruence. What our studies show is that vocal and semantic information are spontaneously integrated into a coherent impression and that there is a shift in meaning that goes well beyond a simple time delay in comprehension.

Second, we extend the previous literature – which tested the impact of "marked" social categories such as gender and age on message elaboration (see Giles et al., 1992) – by examining a more subtle and ambiguous social category, namely SO (Rule, 2017). In particular, we showed that SO-related vocal cues can affect communication in similar ways as other more clearly defined social categories inferred from voice do. Moreover, we extend previous work on

message interpretation based on written stimuli where actors were explicitly introduced as member of a given category (e.g., man/woman; see Edwards, 1998).

Third, rather than testing whether stereotypical message content influences listeners' judgments of speakers' SO (Gaudio, 1994; Smyth et al., 2003), we investigated the opposite route. We showed that voice conveying SO triggers stereotypical interpretation of messages. Hence, our studies provide evidence for the path from voice to message interpretation, but not for the opposite path, namely whether what is being said impacts the perception of the speakers' identity (e.g., stereotypes in message content may well trigger inferences of the speaker's SO). Whether this model works in both directions or is only unidirectional remains an important question for future research.

Finally, it is worth noting that, in the last study, we did not find any difference between sexual minority and straight participants. Although listeners' SO was tested only in one study and with a small sample of SO minority participants, our data seem to suggest that message interpretation was the same for all listeners. This may be explained by the fact that all individuals have internalized shared stereotypes (Devine, 1989). Hence, regardless of whether they explicitly endorse them or not, such stereotypes may be automatically active in people's minds when interpreting a message and communicating with others. However, additional research on larger samples of diverse SO are needed before definite conclusions can be drawn about the generality of the observed shifts in meaning.

Limitations and Future Directions

This research has some limitations. First of all, we tested our hypotheses on a limited and carefully selected sample of voices that was somehow ideal given that voices were clearly perceived as straight and gay/lesbian. Hence, it would be important in the future to use larger and

more representative voice samples. Second, a replication of these studies with additional materials would be valuable. The sentences we used in our studies were limited and may not have been optimal in some respects. For instance, the ambiguous sentences in Study 2 always included a reference to a person of the same sex as the speaker, making it possible that the other person was assumed to be the speaker's partner. In our data, this possibility does not appear to have invalidated the pattern given that straight-sounding voices produced a different interpretation of content despite the fact that they referred to same-sex others in the sentences. However, future studies may want to control for inferred relationships between the speaker and the other person mentioned in the sentence.

Third, it should also be noted that the effects observed here were overall small. This could reflect listeners' attempt to avoid appearing prejudiced and blatantly interpreting messages in line with stereotypes (see Devine, 1989). This would suggest that message interpretation reflects listeners' use of stereotypes activated by voice-inferred social categories. In line with this reasoning, it would be important to test whether endorsement of gender roles, stereotype acceptance, and prejudice moderate the observed effects. Indeed, low stereotype endorsement and positive attitudes toward sexual minorities, along with saliency of societal anti-prejudice norms, could motivate listeners to control their behaviors (see Devine, 1989), including message interpretation. Furthermore, the strengths of the effects may also reflect the strength of the association between the speakers' voice characteristics and the message content. There are remarkable variations in the way a speaker sounds gay, lesbian or straight (Sulpizio et al., 2015). The stronger the 'gay accent' is, the more likely it is that the listener would see a match between the speakers' sexual identity and the stereotypical message content. Studies involving speakers

27

whose voices vary on the 'SO spectrum' would allow to test how this variation affects message interpretation.

With regard to the generalizability of our findings, it would be interesting to extend our results to other gay-related stereotypes or to other types of voices. Indeed, we limited our research to gender stereotyping and gender-inverted attribution (Blashill & Powlishta, 2009; Kite & Deaux, 1987), while ignoring stereotypes attributed to sexual minorities that are gender unrelated (e.g., gay men being promiscuous, immoral, or highly acculturated). Future research could test, for instance, whether gay-sounding speakers are perceived as more acculturated and whether messages are interpreted along this stereotypes (e.g., if a speaker states "my education is important to me", would listeners expect a higher education for a gay- than a straight-sounding speaker?). Similarly, posh-sounding speakers may lead listeners to interpret messages in line with stereotypes related to social status (e.g., imagining they are talking about an expensive rather than a cheap car). Thus, in principle, shifts in subjective meaning may occur for any kind of stereotype and for any kind of voice able to trigger them.

Moreover, it would be important to test potential practical implications of the phenomenon under investigation, assuming that the way individuals process messages may affect their beliefs and behaviors. Prior research has shown that biased communication constitutes one way in which stereotypes are perpetuated (Lyons & Kashima, 2003). We argue here that interpreting a message in stereotype-congruent ways may have a similar function. In line with Summer et al.'s (2014) dual route model, our findings suggest that the encoding of spoken words becomes qualitatively different when integrated with phonetically-cued social information. If listeners align their interpretation of message content with socio-acoustic cues indicative of category membership, then the stereotypes about that category are likely to persist. In the case of sexual minorities,

interpreting message content in line with gender and sexual stereotypes implies that these stereotypical overgeneralizations will hardly be challenged. A "neutral" statement such as *my dog is running in the park* turns into a stereotype-confirming piece of information as soon as the image of a Chihuahua is evoked by acoustic cues that suggest a gay SO. Similarly, grammatically ambiguous self-disclosures such as *I went to meet the guy with the pick-up truck*, become stereotype confirmations (*lesbians love trucks*) when they are interpreted as referring to the speaker rather that to *the guy.* Thus, regardless of the communicative intentions of the speaker, the very meaning of what s/he said may be reinterpreted in light of subtle vocal cues. Together, our findings suggest that socio-acoustic cues may feed into a stereotype-confirming process either by guiding imaging or by disambiguating equivocal messages. It is already known that sounding gay causes avoidance and discrimination (Fasoli et al., 2017). The novel contribution of the current research lies in the fact that vocal cues also affect message interpretation in a way that is likely to bolster existing stereotypes.

Finally, and most importantly, it remains to be understood how and when the integration of content and social information takes place. Although our findings are in line with Sumner et al.'s (2014) dual-route approach to speech perception, they are silent as to the exact stage of information processing at which the integration occurs. We believe that the most challenging questions for future research are to understand (a) whether social and semantic information exert a mutual (rather than uni-directional) influence, (b) how the two sources of information are integrated, (c) how early during information processing this integration takes place, and (d) whether this integration occurs outside of people's awareness. For the moment, we can only conclude that socially relevant vocal cues do change the interpretation of sentence. In sum, our

29

findings contribute to the existing knowledge on communication and stereotypes, by illustrating that vocal cues and the inferred social categories influence message interpretation.

Notes

[1] The same pattern of results emerged for the items concerning car (M_{gay} = 2.99, SD = 1.02 vs. $M_{straight}$ = 4.17, SD = .93; $t(35)$ = -5.54, p < .001), job (M_{gay} = 3.16, SD = 1.09 vs. $M_{straight}$ = 3.85, SD = 1.09; $t(35)$ = -3.25, p = .003), sport (M_{gay} = 4.24, SD = 1.14 vs. $M_{straight}$ = 4.69, SD = .72; $t(35)$ = -2.53, p = .016), but fell short of significance for dog (M_{gay} = 3.24, SD = .87 vs. $M_{straight}$ = 3.63, SD = .85; $t(35)$ = -1.69, p = .099). Also, participants reported to have listened to at least two gay speakers (30.6% of participants), one (44.4% of participants), or none (25% of participants).

[2] Significant differences emerged across all items: car (M_{gay} = 3.06, SD = .97 vs. $M_{straight}$ = 4.18, SD = .80; $t(49)$ = -6.37, p < .001), profession (M_{gay} = 3.05, SD = 1.10 vs. $M_{straight}$ = 3.96, SD = .95; $t(49)$ = 3.25, p < .001), sport (M_{gay} = 3.64, SD = 1.00 vs. $M_{straight}$ = 4.45, SD = .91; $t(49)$ = -4.78, p < .001), and dog (M_{gay} = 2.86, SD = .95 vs. $M_{straight}$ = 3.63, SD = .99; $t(49)$ = -6.09, p < .001). Also, the majority of participants (44%) correctly thought that 2 speakers were gay, 16% recognized only 1 gay speaker and 38% of the participants thought that none of them was gay.

[3] Participants reported a slightly positive attitude toward bisexual and homosexuals (M = 56.38, SD = 26.45) on the thermometer measure. The majority of them recognized one (27.3%) or two (40.0%) gay speakers, and one (34.5%) or two (21.8%) lesbian speakers. Overall, they were not interested in performing any of the actions (M = 2.47, SD = .68; $t(1, 54)$ = -5.82, p < .001).

[4] Participants knew on average 7 homosexuals (M = 6.89, SD = 6.59) and recognized one (32.3%) or two (36.0%) speakers as gay, and one (27.3%) or two (26.8%) speakers as lesbian.

References

Apple, W., Streeter, L. A., & Krauss, R. M. (1979). Effects of pitch and speech rate on personal attributions. *Journal of Personality and Social Psychology*, *37*, 715-727. doi:10.1037/0022-3514.37.5.715

Badzinski, D. M. (1991). Children's cognitive representations of discourse: Effects of vocal cues on text comprehension. *Communication Research, 18*, 715-736. doi:10.1177/009365091018006001

Badzinski, M. D., & Gill, M. M. (1994). Discourse features and message comprehension. In S. A. Deetz (Ed.), *Communication yearbook 17* (pp. 301-332). Thousand Oaks, CA: Sage.

Blashill, A. J., & Powlishta, K. K. (2009). Gay stereotypes: The use of sexual orientation as a cue for gender-related attributes. *Sex Roles*, *61*, 783-793. doi:10.1007/s11199-009-9684-7

Bodie, G. D. (2011). The understudied nature of listening in interpersonal communication: Introduction to a special issue. *The International Journal of Listening*, *25*, 1-9. doi:10.1080/10904018.2011.536462

Bodie, G. D., Worthington, D., Imhof, M., & Cooper, L. O. (2008). What would a unified field of listening look like? A proposal linking past perspectives and future endeavors. *The International Journal of Listening*, *22*, 103-122. doi:10.1080/10904010802174867

Bugental, D. E. (1974). Interpretations of naturally occurring discrepancies between words and intonation: Modes of inconsistency resolution. *Journal of Personality and Social Psychology, 30*, 125-133. doi:10.1037/h0036654

Burgoon, J. K. (1993). Interpersonal expectations, expectancy violations, and emotional communication. *Journal of Language and Social Psychology*, *12*, 30-48. doi:10.1177/0261927X93121003

Campbell-Kibler, K. (2011). Intersecting variables and perceived sexual orientation in

men. *American Speech*, *86*, 52-68. doi:10.1215/00031283-1277510

Cohen, H., Douaire, J., & Elsabbagh, M. (2001). The role of prosody in discourse

processing. *Brain and Cognition*, *46*, 73-82. doi:10.1016/S0278-2626(01)80038-5

Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of*

Personality and Social Psychology, *44*, 20-33. doi:10.1037/0022-3514.44.1.20

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled

components. *Journal of Personality and Social psychology*, *56*, 5-18.

Dragojevic, M. (2018). Language attitudes. In H. Giles & J. Harwood (Eds.), *Oxford Research*

Encyclopedia of Intergroup Communication (Vol. 2, pp. 35-57). New York, NY: Oxford

University Press.

Dragojevic, M., Berglund, C., & Blauvelt, T. K. (2018). Figuring out who's who: The role of

social categorization in the language attitudes process. *Journal of Language and Social*

Psychology, *37*, 28-50. doi:10.1177/0261927X17706942

Edwards, R. (1998). The effects of gender, gender-role, and values on the interpretation of

messages. *Journal of Language and Social Psychology*, *17*, 52-71.

doi:10.1177/0261927X980171003

Edwards, R. (2011). Listening and message interpretation. *The International Journal of*

Listening, *25*, 47-65. doi:10.1177/0261927X980171003

Fasoli, F., Maass, A., & Sulpizio, S. (2016). Communicating the "invisible": The case of sexual

orientation. In H. Giles & A. Maass (Eds.), *Advances in intergroup communication* (pp. 193-

204). New York, NY: Peter Lang Publications.

Fasoli, F., Maass, A., Paladino, M. P., & Sulpizio, S. (2017). Gay-and lesbian-sounding auditory

cues elicit stereotyping and discrimination. *Archives of Sexual Behavior*, *46*, 1261-1277.

doi:10.1007/s10508-017-0962-0

Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the

melody the message?. *Child Development*, 1497-1510. doi:10.2307/1130938

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype

content: Competence and warmth respectively follow from perceived status and competition.

Journal of Personality and Social Psychology, 82, 878-902. doi:10.1037/pspa0000163

Garvey, C., & Caramazza, A. (1974). Implicit causality in verbs. *Linguistic Inquiry, 5*, 459-464.

Gaudio, R. P. (1994). Sounding gay: Pitch properties in the speech of gay and straight

men. *American Speech*, *69*, 30-57. doi:10.2307/455948

Giles, H. (Ed.). (2016). *Communication accommodation theory: Negotiating personal*

relationships and social identities across contexts. Cambridge, UK: Cambridge University

Press.

Giles, H., Henwood, K., Coupland, N., Harriman, J., & Coupland, J. (1992). Language attitudes

and cognitive mediation. *Human Communication Research*, *18*, 500-527. doi:10.1111/j.1468-

2958.1992.tb00570.x

Glenwright, M., Parackel, J. M., Cheung, K. R., & Nilsen, E. S. (2014). Intonation influences

how children and adults interpret sarcasm. *Journal of Child Language*, *41*, 472-484.

doi:10.1017/S0305000912000773

Gnevsheva, K. (2016). Beyond the language: listener comments on extra-linguistic cues in

perception tasks. *Language Awareness*, *25*, 257-271. doi:10.1080/09658416.2016.1197932

Griggin E. M. (Ed.). (2009). *A first look at Communication Theory*. New York, NY: Mc Graw Hill.

Hajek, C., & Giles, H. (2005). Intergroup communication schemas: Cognitive representations of talk with gay men. *Language and Communication*, *25*, 161-181. doi:10.1016/j.langcom.2005.01.002

Herek, G. M., & McLemore, K. A. (2013). Sexual prejudice. *Annual Review of Psychology, 64*, 309-333. doi:10.1146/annurev-psych-113011-143826

Imhof, M. (2010). What is going on in the mind of the listener? The cognitive psychology of listening. In A. Wolvin (Ed.), *Listening and human communication in the 21st century* (pp. 97-126). Oxford, UK: Blackwell.

John, M. F. S., & McClelland, J. L. (1990). Learning and applying contextual constraints in sentence comprehension. *Artificial Intelligence*, *46*, 217-257. doi:10.1016/0004-3702(90)90008-N

Kaganovich, N., Francis, A. L., & Melara, R. D. (2006). Electrophysiological evidence for early interaction between talker and linguistic information during speech perception. *Brain Research*, *1114*, 161-172. doi:10.1016/j.brainres.2006.07.049

King, E., & Sumner, M. (2015). Voice-specific effects in semantic association. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 37th annual conference of the Cognitive Science Society* (pp. 1111-1116). Austin, TX: Cognitive Science Society.

Kite, M. E., & Deaux, K. (1987). Gender belief systems: Homosexuality and the implicit inversion theory. *Psychology of Women Quarterly*, *11*, 83-96. doi:10.1111/j.1471-6402.1987.tb00776.x

Kreiman, J. (1997). Listening to voices: Theory and practice in voice perception research. In K. Johnson & J. Mullenix (Eds.), *Talker variability in speech processing* (pp. 85-108). San Diego, CA: Academic Press.

Kreiner, H., Sturt, P., & Garrod, S. (2008). Processing definitional and stereotypical gender in reference resolution: Evidence from eye-movements. *Journal of Memory and Language*, *58*, 239-261. doi:10.1016/j.jml.2007.09.003

Lattner, S., & Friederici, A. D. (2003). Talker's voice and gender stereotype in auditory sentence processing–evidence from event-related brain potentials. *Neuroscience Letters*, *339*, 191-194. doi:10.1016/S0304-3940(03)00027-2

Lehiste, I., & Lass, N. J. (1976). Suprasegmental features of speech. In N. Lass (Ed.), *Contemporary issues in experimental phonetics* (pp. 225-239). New York, NY: Academic Press.

Levon, E. (2015). Integrating intersectionality in language, gender, and sexuality research. *Language and Linguistics Compass, 9*, 295-308. doi:10.1111/lnc3.12147

Linville, S. E. (1998). Acoustic correlates of perceived versus actual sexual orientation in men's speech. *Folia Phoniatrica et Logopaedica*, *50*, 35-48. doi:10.1159/000021447

Lyons, A., & Kashima, Y. (2003). How are stereotypes maintained through communication? The role of stereotype sharedness. *Journal of Personality and Social Psychology*, *85*, 989-1005. doi:10.1037/0022-3514.85.6.989

Mack, S., & Munson, B. (2012). The influence of /s/ quality on ratings of men's sexual orientation: Explicit and implicit measures of the 'gay lisp' stereotype. *Journal of Phonetics*, *40*, 198-212. doi:10.1016/j.wocn.2011.10.002

Macrae, C. N., & Bodenhausen, G. V. (2001). Social cognition: Categorical person

    perception. *British Journal of Psychology*, *92*, 239-255. doi:10.1348/000712601162059

Macrae, C. N., Milne, A. B., & Bodenhausen, G. V. (1994). Stereotypes as energy-saving

    devices: A peek inside the cognitive toolbox. *Journal of Personality and Social*

    Psychology, *66*, 37-47. doi:10.1037/0022-3514.66.1.37

McGowan, K. B. (2015). Social expectation improves speech perception in noise. *Language and*

    Speech, *58*, 502-521. doi:10.1177/0023830914565191

Munson, B. (2007). The acoustic correlates of perceived masculinity, perceived femininity, and

    perceived sexual orientation. *Language and Speech, 50*, 125-142.

    doi:10.1177/00238309070500010601

Nygaard, L. C., Herold, D. S., & Namy, L. L. (2009). The semantics of prosody: Acoustic and

    perceptual evidence of prosodic correlates to word meaning. *Cognitive Science*, *33*, 127-146.

    doi:10.1111/j.1551-6709.2008.01007.x

Nygaard, L. C., & Lunders, E. A. (2002). Resolution of lexical ambiguity by emotional tone of

    voice. *Journal of Memory and Cognition, 30*, 583-593.

Pexman, P. M., & Olineck, K. M. (2002). Understanding irony: How do stereotypes cue speaker

    intent?. *Journal of Language and Social Psychology*, *21*, 245-274.

    doi:10.1177/0261927X02021003003

Pharao, N., Maegaard, M., Møller, J. S., & Kristiansen, T. (2014). Indexical meanings of [s+]

    among Copenhagen youth: Social perception of a phonetic variant in different prosodic

    contexts. *Language in Society*, *43*, 1-31. doi:10.1017/S0047404513000857

Platow, M., & Brodie, M. (1999). The effects of linguistic voice on evaluations and attributions

of ingroup and outgroup members. *Asian Journal of Social Psychology*, *2*, 187-200.

doi:10.1111/1467-839X.00033

Rieger, G., Linsenmeier, J. A., Gygax, L., Garcia, S., & Bailey, J. M. (2010). Dissecting

"gaydar": The role of masculinity–femininity. *Archives of Sexual Behavior*, *39*, 124-140.

doi:10.1007/s10508-008-9405-2

Rule, N. O. (2017). Perceptions of sexual orientation from minimal cues. *Archives of Sexual

Behavior*, *46*, 129-139. doi:10.1007/s10508-016-0779-2

Sagar, H. A., & Schofield, J. W. (1980). Racial and behavioral cues in black and white children's

perceptions of ambiguously aggressive acts. *Journal of Personality and Social

Psychology*, *39*, 590-598. doi:10.1037/0022-3514.39.4.590

Schafer, A., Carlson, K., Clifton jr, H., & Frazier, L. (2000). Focus and the interpretation of pitch

accent: Disambiguating embedded questions. *Language and Speech*, *43*, 75-105.

doi:10.1177/00238309000430010301

Senior, B., Hui, J., & Babel, M. (2018). Liu vs. Liu vs. Lick? Name influence on voice recall.

Applied Psycholinguistics, *39*, 1117-1146. doi:10.1017/S0142716418000267

Smyth, R., Jacobs, G., & Rogers, H. (2003). Male voices and perceived sexual orientation: An

experimental and theoretical approach. *Language in Society*, *32*, 329-350.

doi:10.1017/S0047404503323024

Sulpizio, S., Fasoli, F., Maass, A., Paladino, M. P., Vespignani, F., Eyssel, F., & Bentler, D.

(2015). The sound of voice: Voice-based categorization of speakers' sexual orientation within

and across languages. *PloS One*, *10*, e0128882. doi:10.1371/journal.pone.0128882

Sulpizio, S., Fasoli, F., Antonio, R., Eyssel, F., Paladino, M. P., & Diehl, C. (2019). Auditory

gaydar: Perception of sexual orientation based on female voice. *Language and Speech.*

Advanced online publication. doi:10.1177/0023830919828201

Sumner, M., & Kataoka, R. (2013). Effects of phonetically-cued talker variation on semantic

encoding. *The Journal of the Acoustical Society of America*, *134*, EL485-EL491.

doi:10.1121/1.4826151

Sumner, M., Kim, S. K., King, E., & McGowan, K.B. (2014). The socially weighted encoding of

spoken words: a dual-route approach to speech perception. *Frontiers in Psychology*, *4*, 10-13.

doi:10.3389/fpsyg.2013.01015

Tracy, E. C., Bainter, S. A., & Satariano, N. P. (2015). Judgments of self-identified gay and

heterosexual male speakers: Which phonemes are most salient in determining sexual

orientation?. *Journal of Phonetics*, *52*, 13-25. doi:10.1016/j.wocn.2015.04.001

Van Berkum, J. J., Van den Brink, D., Tesink, C. M., Kos, M., & Hagoort, P. (2008). The neural

integration of speaker and message. *Journal of Cognitive Neuroscience*, *20*, 580-591.

doi:10.1162/jocn.2008.20054

Van Borsel, J., Vandaele, J., & Corthals, P. (2013). Pitch and pitch variation in lesbian

women. *Journal of Voice*, *27*, 656e13-656e16. doi:10.1016/j.jvoice.2013.04.008

Vaughn, A. A., Teeters, S. A., Sadler, M. S., & Cronan, S. B. (2017). Stereotypes, emotions, and

behaviors toward lesbians, gay men, bisexual women, and bisexual men. *Journal of

Homosexuality*, *64*, 1890-1911. doi:10.1080/00918369.2016.1273718

Walker, A., & Hay, J. (2011). Congruence between 'word age' and 'voice age' facilitates lexical

access. *Laboratory Phonology*, *2*, 219-237. doi:10.1515/labphon.2011.007

Table 1. *Sentences Across Studies*

Speaker Gender	Sentences of Study 1a/1b	
	Portuguese	*English*
Male speakers	O meu cão está a correr no parque	my dog is running in the park
	Ontem levei o meu carro ao mecânico	yesterday I took my car to the mechanic
	Consegui finalmente um novo emprego	I finally got a new job
	Faço desporto há muitos anos	I have been doing sports for many years
	Sentences of Study 2a	
	Portuguese	*English*
Male Speakers	Ontem eu e o meu amigo Miguel fomos a uma loja de decoração de interiores à procura de um vaso	Yesterday, my friend Mike and I went shopping for an interior design vase
	Surpreendi o meu amigo com umas sapatilhas de ballet	I surprised my friend with the ballet shoes
Female Speakers	Ontem eu e a Cátia fomos a um stand comprar uma mota	Yesterday, Kate and I went to a shop to buy a motorcycle
	Antes de irmos acampar, eu e a minha amiga fomos beber uma cerveja	Before going camping, my friend and I met to drink a beer
	Sentences of Study 2b	
	Italian	*English*
	<u>Gender stereotypical actions/objects</u>	
Male Speakers	A causa della partita di calcio, il mio amico ed io ci siamo salutati e abbiamo preso strade diverse	Because of the football match, my friend and I said bye-bye and took different routes

Female Speakers	Ho incontrato il mio amico con le scarpe da ballo	I met my friend with the ballet shoes
	Gender counter-stereotypical action/object	
Male Speaker	Io e Paolo siamo andati dal mio parrucchiere preferito per un taglio di capelli colorato ed eccentrico	Paolo and I went to my favorite hairdresser for a colorful and eccentric haircut
	Non volendo una relazione stabile, io e mia sorella abbiamo discusso a lungo	Not wanting a stable relationship, my sister and I discussed a lot).
	Io e mio fratello Giovanni siamo andati al centro commerciale per comprare un bracciale	My brother Giovanni and I went to the shopping centre to buy a bracelet
Female Speaker	Paola ed io siamo andate dal mio parrucchiere preferito per tagliare i capelli corti	Paola and I went to my favorite hairdresser to get a short haircut
	Io e Maria abbiamo guardato il braccio con il tatuaggio	Mary and I looked at the arm with the tattoo
	Per discutere una serie di questioni importanti, sono andata a incontrare il ragazzo con il furgoncino	To discuss a number of important issues, I went to meet the guy with the pick-up truck

1

Author biographies

Fabio Fasoli is a Lecturer in Social Psychology at the University of Surrey. His research interests regard social communication, gaydar, LGBT discrimination, homophobic language, and humour.

Anne Maass is Professor of social psychology at the University of Padua, Italy. Her main research interests concern (a) the intersection between language and social cognition, (b) gender stereotypes and stereotype threat and (c) the psychology of economic inequality.

Rachel Karniol is Professor Emerita of Social Development in the School of Psychological Sciences at Tel Aviv University. Her expertise is at the crossroads of social and developmental psychology, with a focus on processes in understanding other people, empathy, moral development, and gender development.

Raquel António is a PhD candidate at ISCTE-IUL on the topic of homophobic bullying and bystanders' behaviors. Her master's degree is in Social and Organizational Psychology. She also worked as research assistant in three research projects on intergroup relations and discrimination.

Simone Sulpizio is assistant professor in the Faculty of Psychology at the Vita-Salute San Raffaele University. His research activity is mainly focused on two main research lines that are: The Neurobiological and cognitive mechanisms underlying language, and how social and emotional factors may interact and change language processing.