Supplementary Material 1

Speakers' Selection

In both countries, speakers were selected from pilot tests where participants were asked to listen to gay and straight speakers (20 speakers in Italy and 16 speakers in UK) while pronouncing neutral sentences (e.g., the dog runs in the park). Two gay and two straight speakers were chosen in each country so that their voices conveyed a gay and straight sexual orientation, respectively. In Italy (n = 15), gay speakers ($M = 4.53$, $SD = 1.06$) were perceived as more gay sounding than straight speakers ($M = 2.53$, $SD = 1.27$; $t(14) = 3.98$, $p = .001$). In the UK (n = 59), the gay speakers ($M = 3.97$, $SD = 1.20$) were perceived as more gay sounding than the straight speakers ($M = 2.68$, $SD = 1.10$; $t(58) = 6.31$, $p < .001$).

Supplementary Material 2

Surprise

The same analysis as for the other dependent variables was performed on perceived surprise. Participants were overall more surprised to listen to a message that included a DGL ($M = 4.18$, $SE = .13$) than an a CGL ($M = 2.82$, $SE = .12$; $F[1, 268] = 56.93$, $p < .001$, $\eta^2_p = .17$). Higher surprise was also reported in the other-labelling ($M = 3.77$, $SE = .13$) than self-labelling condition ($M = 3.23$, $SE = .13$; $F[1, 268] = 8.54$, $p = .004$, $\eta^2_p = .03$). A significant interaction between Label and Target, $F(1, 268) = 11.77$, $p = .001$, $\eta^2_p = .04$, indicated that a message containing a DGL elicited always high surprise in our listeners regardless of the target ($M_{self} = 4.22$, $SE = .18$ vs. $M_{other} = 4.24$, $SE = .18$; $p = .74$). On the contrary, participants were more surprised to listen to a message including a CGL in the other- than self-labelling condition ($M_{self} = 2.25$, $SE = .17$ vs. $M_{other} = 3.39$, $SE = .18$; $p < .001$). Hence, the less surprising situation was when a person self-labelled with a CGL.

Moreover, a significant main effect of Country, $F(1, 268) = 23.90$, $p < .001$, $\eta^2_p = .04$, showed that, overall, British participants ($M = 3.94$, $SE = .12$) reported to be more surprised than Italian participants ($M = 3.06$, $SE = .14$) when listening to the messages.

Supplementary Material 3

Perception of Speakers across Conditions and Countries

A 2(Speaker: gay vs. straight) x 2(Target: self vs. other) x 2(Label: category vs homophobic) x 2(Country: English vs. Italian) repeated measure ANCOVA as in the main analyses was performed on perceived gender typicality, SO voice sound, and SO categorisation. Pairwise comparisons (Bonferroni correction) were applied for significant interactions.

Gender Typicality. A marginal Speaker X Target x Label, $F(1, 268) = 3.81$, $p = .05$, $\eta^2_p = .01$, indicated that overall the gay speakers were always perceived as less masculine that the straight speakers (all ps < .001). However, the gay speakers were perceived as particularly less masculine when they self-labelled with a category label (ps < .016; see Table below). No other differences emerged.

	Category label (CGL)		Homophobic label (DGL)	
	Self-labelling	*Other-labelling*	*Self-labelling*	*Other-labelling*
Gay speakers	4.20 (.15)a	4.73 (.16)ac	4.82 (.16)ac	4.51 (.16)ac
Straight speakers	5.98 (.12)b	5.93 (.12)b	6.12 (.12)b	5.85 (.12)b

Note. Means that share the same letter are not significantly different from each other

Sexual Orientation Voice Sound. A significant main effect of voice, $F(1, 268) = 29.48$, $p < .001$, $\eta^2_p = .10$, and of Country, $F(1, 268) = 7.61$, $p = .006$, $\eta^2_p = .03$, were qualified by a significant interaction between these two factors, $F(1, 268) = 70.14$, $p < .001$, $\eta^2_p = .21$. Pairwise comparisons showed that in both countries the gay speakers were perceived as more gay-sounding than the straight speakers (ps < .001), but that Italian participants had more extreme ratings than British participants. Italians rated the gay speakers as more gay-sounding ($M_{Italians} = 4.89$, $SE = .11$ vs $M_{Britons} = 3.82$, $SE = .09$; $p < .001$) and the straight

speakers as more straight sounding ($M_{Italians}$ = 2.67, SE = .11 vs $M_{Britons}$ = 3.01, SE = .10; p = .005) that the British participants did.

Moreover, a main effect of Label, $F(1, 268)$ = 8.96, p = .003, η^2_p = .03, indicated that overall participants rather participants as sounding more gay when they used a CGL (M = 3.80, SE = .08) rather than a DGL (M = 3.45, SE = .08).

Sexual Orientation Categorisation. A significant main effect of voice, $F(1, 267)$ = 6.63, p = .01, η^2_p = .02, and of Country, $F(1, 267)$ = 31.69, p < .001, η^2_p = .11, were qualified by a significant interaction between these two factors, $F(1, 267)$ = 30.50, p < .001, η^2_p = . 10. In both countries the gay speakers were more likely to be categorised as gay than the straight speakers (ps < .001). The likelihood of categorising the gay speakers as gay was higher in Italy (M = 1.54, SE = .06) than in the UK (M = .89, SE = .05; p < .001), whereas no difference on the categorization of straight speakers (M_{Italy} = .67, SE = .07 vs. M_{UK} = .59, SE = .05; p = .30) emerged.

Moreover, a significant 3-way interaction between Label, Target, and Country, $F(1, 267)$ = 5.79, p = .02, η^2_p = .02, showed that participants were more likely to categorise the speakers as gay when they self-labelled than when they labelled others (all ps < .006), except for when Italian participants listened to speakers using a CGL (p = .44). However, this pattern of result is not influenced by the speakers' SO.

	Category label		Homophobic label	
	Self-labelling	*Other-labelling*	*Self-labelling*	*Other-labelling*
Italy	1.28 (.10)a	1.17 (.10)a	1.26 (.10)a	.72 (.10)c
United Kingdom	1.15 (.08)a	.61 (.08)b	.75 (.08)c	.42 (.08)b

Note. Means that share the same letter are not significantly different from each other