Animal Behaviour 130 (2017) 47–55

Contents lists available at ScienceDirect

Animal Behaviour

journal homepage: www.elsevier.com/locate/anbehav

Tennis grunts communicate acoustic cues to sex and contest outcome

Jordan Raine*, Katarzyna Pisanski, David Reby

Mammal Vocal Communication and Cognition Research Group, School of Psychology, University of Sussex, Falmer, U.K.

ARTICLE INFO

Article history:
Received 13 April 2017
Initial acceptance 29 May 2017
Final acceptance 12 June 2017

MS. number: 17-00342R

Keywords:
dominance
fundamental frequency
nonverbal vocalizations
pitch
tennis grunts
vocal communication

Despite their ubiquity in human behaviour, the communicative functions of nonverbal vocalizations remain poorly understood. Here, we analysed the acoustic structure of tennis grunts, nonverbal vocalizations produced in a competitive context. We predicted that tennis grunts convey information about the vocalizer and context, similar to nonhuman vocal displays. Specifically, we tested whether the fundamental frequency (F0) of tennis grunts conveys static cues to a player's sex, height, weight, and age, and covaries dynamically with tennis shot type (a proxy of body posture) and the progress and outcome of male and female professional tennis contests. We also performed playback experiments (using natural and resynthesized stimuli) to assess the perceptual relevance of tennis grunts. The F0 of tennis grunts predicted player sex, but not age or body size. Serve grunts had higher F0 than forehand and backhand grunts, grunts produced later in contests had higher F0 than those produced earlier, and grunts produced during contests that players won had a lower F0 than those produced during lost contests. This difference in F0 between losses and wins emerged early in matches, and did not change in magnitude as the match progressed, suggesting a possible role of physiological and/or psychological factors manifesting early or even before matches. Playbacks revealed that listeners use grunt F0 to infer sex and contest outcome. These findings indicate that tennis grunts communicate information about both the vocalizer and contest, consistent with nonhuman mammal vocalizations.

Despite their ubiquitous use across ages, sexes, contexts and cultures, human nonverbal vocalizations remain under-investigated. In fact, aside from laughter (see Bryant et al., 2016; Scott, Lavan, Chen, & McGettigan, 2014) and infant cries (Lingle, Wyman, Kotrba, Teichroeb, & Romanow, 2012 for review), human nonverbal vocalizations (such as moans, sighs, roars, screams and grunts) have received little attention, especially from a functional and evolutionary perspective.

Indeed, most research on human nonverbal vocalizations has focused on their classification according to emotional content (e.g. Belin, Fillion-Bilodeau, & Gosselin, 2008; Lima, Castro, & Scott, 2013; Sauter, Eisner, Calder, & Scott, 2010), while overlooking their potential to convey indexical cues about the vocalizer such as age, sex, body size and social dominance. Such cues are typically present in the vocal signals of nonhuman mammals, and function to mediate interactions in social and sexual contexts (Briefer, 2012; Taylor, Charlton, & Reby, 2016). Human nonverbal vocalizations probably predate language, and appear homologous in structure and function with nonhuman vocalizations (e.g. laughter Davila-Ross, Owren, & Zimmermann, 2010; Pisanski, Cartei, McGettigan, Raine, & Reby, 2016; infant distress vocalizations Lingle et al., 2012). As such, they may constitute a relatively direct link between animal and human vocal systems. Investigating their production, control and perception may therefore provide a unique window into the origins and evolution of human vocal behaviour (Pisanski, Cartei, et al., 2016).

Here, we examined whether the acoustic structure of tennis grunts, a nonverbal vocalization produced during a competitive contest, contains functionally relevant and perceptible cues. While there are clear qualitative differences between tennis matches and nonhuman mammal contests (tennis players do not voluntarily yield to dominant competitors, and there are more clearly defined rules and endpoints), animal contests often follow ritualized rules and structures, during which competitors produce signals that contain static and dynamic information about their respective physical condition and motivation (e.g. ungulates: Jennings & Gammell, 2013). Thus, tennis matches provide a potentially useful model to examine whether similar information is communicated in human competitive interactions.

Investigations of the function of tennis grunts have so far focused on their distracting quality to opponents (Farhead & Punt, 2015; Sinnett & Kingstone, 2010), and their enhancement of ball

* Correspondence: J. Raine, School of Psychology, Pevensey 1 Building, University of Sussex, Falmer BN1 9QH, U.K.
E-mail address: j.w.raine@sussex.ac.uk (J. Raine).

http://dx.doi.org/10.1016/j.anbehav.2017.06.022

velocity without increasing oxygen cost (e.g. O'Connell, Hinman, Hearne, Michael, & Nixon, 2014). No previous study has examined their possible communicative value. We hypothesized that tennis grunts are competitive nonverbal vocalizations homologous to those produced by nonhuman animals in agonistic contexts. Accordingly, we predicted that grunts would contain static and dynamic cues to anatomical and physiological traits of the vocalizer. As vocalizations produced in a competitive context, tennis grunts may be affected by changes in players' physiological and psychological states (e.g. arousal and stress, which correlate positively with fundamental frequency, F0, Briefer, 2012; Pisanski, Nowak, & Sorokowski, 2016). Thus, we predicted that grunt F0 would correlate with the outcome of competitive contests (i.e. tennis matches). Finally, we predicted that listeners would be able to use these static and dynamic cues to make functionally relevant inferences about both the tennis player and the match.

Recent research generalizing the source-filter model of speech production (Fant, 1960) to vertebrate vocal signals has highlighted the function of F0 (corresponding to the rate of vocal fold vibration, affecting perceived pitch) and formant frequencies (resonances of the supralaryngeal vocal tract, affecting perceived timbre) in communicating various static and dynamic cues in nonhuman mammal vocalizations (see Briefer, 2012; Taylor et al., 2016). For example, sexually selected calls communicate F0-based cues to dominance (e.g. male deer groans: Liu et al., 2016; Vannoni & McElligott, 2008) and formant cues to body size (red deer, *Cervus elaphus*: Reby et al., 2005; Australian sea lions, *Neophoca cinerea*: Charrier, Ahonen, & Harcourt, 2011; dogs, *Canis lupus familiaris*: Taylor, Reby, & McComb, 2010). However, whether fundamental and formant frequencies also communicate similar information in human nonverbal vocalizations remains to be determined.

To address this, we investigated the acoustic structure of tennis grunts produced by male and female tennis players during professional matches. Because tennis grunts are relatively high-pitched vocalizations characterized by a low spectral density, formant frequencies were poorly defined and difficult to both perceive and measure (Pisanski, Fraccaro, Tigue, O'Connor, & Feinberg, 2014; Ryalls & Lieberman, 1982). We therefore focused our analyses on the mean F0 of grunts. We recorded the sex, height, weight and age of the vocalizers (static cues), the type of tennis shot accompanying the grunt (forehand/backhand/serve) and the outcome of the given match (vocalizer won/lost) (dynamic cues). We then investigated whether tennis grunts have perceptual and functional relevance via playback experiments using a separate sample of natural and F0-resynthesized tennis grunts.

We predicted that (1) females would produce tennis grunts with higher F0 than males due to sexual dimorphism in human vocal folds and F0 (Taylor et al., 2016; Titze, 1994); (2) F0 would not indicate height, weight or age in our sample of young adult players due to the weak relationship between speaking F0 and body size within the sexes in human speech (Pisanski, Fraccaro, Tigue, O'Connor, Röder, et al., 2014 for meta-analysis) and other mammal vocalizations (Taylor et al., 2016 for review), and based on the relative stability of F0 in human speech after puberty (Fouquet, Pisanski, Mathevon, & Reby, 2016); (3) postural differences between shot types would affect F0 due to the influence of dynamic biomechanical constraints on vocal production mechanisms (Fitch & Hauser, 1995; Titze, 1994); and (4) the F0 of vocalizations occurring during match losses would be higher than during match wins. This fourth prediction stems from the aforementioned negative relationship between F0 and dominance, and evidence that F0 increases under stress, distress and arousal in both humans (Pisanski, Nowak, et al., 2016) and nonhuman mammals (Briefer, 2012 for review). Finally, we predicted that F0 cues in tennis grunts would influence listeners' attributions of vocalizer sex and match outcome.

METHODS

Analysis of Tennis Grunts

Within-subject variables

Of the top 30 players in the world of each sex at the time of data collection, 90% of male players and 76% of female players produced tennis grunts during matches. Of these, 50% of male players and 63% of female players grunted sufficiently frequently to allow for behavioural observation (i.e. during at least 75% of tennis shots) and were therefore eligible for inclusion in our data sample. A further 40% of males and 13% of females grunted less frequently (e.g. only in later stages of tournaments and/or during important points), and were therefore not eligible for inclusion in our data sample. Thus, from these top 30 players, we identified seven males and seven females who consistently grunted when hitting both serves and groundstrokes (a shot executed after the ball bounces once) within and between matches.

Using PRAAT 5.3.62 DSP package (Boersma & Weenink, 2014) and Boom 2 software (Global Delight Technologies, 2014), we extracted 367 tennis grunts from direct audio output of television footage of 50 matches provided by the International Tennis Federation and the IMG Sport Video Archive. Recordings were saved as WAV files at 44.1 kHz sampling frequency and 16 bit amplitude resolution. For each grunt, we coded shot type (forehand/backhand/serve; the most common shot types, Johnson & McHugh, 2006), the duration of the match at the point of grunt production (hereafter ‘match progress’, expressed as a percentage of total match duration), and match outcome (win/loss). For each match, two grunts per set were recorded for males and three for females. This resulted in an equal number of grunts per match for each sex, as women play best-of-three sets while men play best-of-five. Grunts were sampled at equally spaced time intervals across the duration of each set. Within each match, an equal number of forehands, backhands and serves were recorded. Within each vocalizer, we recorded a roughly equal number of grunts from match wins and losses. Within the constraints of the limited number of televised matches available to us, we matched wins and losses as closely as possible in terms of tournament stage, to control for the potential effect of match importance on physiological and psychological state.

Between-subject variables

To test whether player sex, height, weight and age predicted grunt F0, we conducted an additional between-subject analysis. Because F0 variation among serve vocalizations in the within-subjects data set was relatively small (see Fig. 1a), we randomly selected two serve vocalizations from each of the 14 players described above, and extracted 92 serve vocalizations from an additional 23 male and 23 female professional tennis players (mean age $\pm$ SD $= 25.09 \pm 0.42$ years), to achieve an adequate sample size of 30 players per sex and 120 serve vocalizations. We chose serve vocalizations because they are always produced at the start of rallies, from an initially stationary position where posture is most standardized. Player sex, height, weight and age data were acquired from www.atpworldtour.com and www.wtatennis.com. We also extracted 10 s prematch interview clips to measure each player's mean speaking F0.

Acoustic analysis

We extracted mean F0 using a dedicated processing script in PRAAT (Boersma & Weenink, 2014). We systematically inspected each extracted pitch contour and verified it using a narrow band spectrogram displaying the first 2000 Hz of the signal. Erroneous pitch values (e.g. octave jumps) were manually corrected. Fifty