The vocal repertoire of preterm infants: Characteristics and possible applications

Vanessa André, Virginie Durier, Severine Henry, Fouad Nassur, Jacques Sizun, Martine Hausberger, Alban Lemasson

The vocal repertoire of preterm infants: characteristics and possible applications

Vanessa André[1], Virginie Durier[1], Séverine Henry[1], Fouad Nassur[1], Jacques Sizun[2, 3], Martine Hausberger[1], Alban Lemasson[1*]

[1] Univ Rennes, Normandie Univ, CNRS, EthoS (Ethologie animale et humaine) - UMR 6552, F-35000 Rennes, France.

[2] Centre Hospitalier Régional Universitaire, Pôle de la Femme, de la Mère et de l'Enfant, Brest, France.

[3] Université de Bretagne Occidentale, Faculté de Médecine et des Sciences de la Santé, Laboratoire de Neurosciences de Brest - EA 4685, Brest, France

* Corresponding author: +332 23 23 68 20, alban.lemasson@univ-rennes1.fr, Station Biologique de Paimpont, 35380, Paimpont, france

Acknowledgments

We warmly thank the babies and their parents who accepted to participate in this study, as well as the NICU staff for their help. We also thank Ann Cloarec for helping with the writing in English and Nicolas Dollion, Basile Griois and Emmanuelle Martin for their implication in data collecting and processing. This project emerged after discussions among members of the research network GIS ''Cerveau-Comportement-Societé''.

Funding

Financial support was provided by the CNRS, the University of Rennes 1 and the University hospital centre of Morvan at Brest.

Authors declare no conflict of interest.

Vanessa André: Conceptualization, Methodology, Investigation, Formal analysis, Writing; **Fouad Nassur**: Formal analysis; **Alban Lemasson, Martine Hausberger, Virginie Durier, Jacques Sizun**: Conceptualization, Methodology, Formal analysis, Writing, Project administration, Funding acquisition

Abstract

We investigated infants' capacities to express themselves orally at very early developmental stages. Most reports focus on crying when in pain or hungry. We evaluated young preterm infants' spontaneous vocal production in non-painful contexts. We identified a vocal repertoire composed of nine types of vocalisations. High-pitched sounds were associated with relaxed postures, implying a positive valence, whereas long low-pitched vocalisations, associated more with grimaces and muscle tensions, appeared to have a more negative valence. Infants' vocalisations were useful indicators of their internal state in two situations (when exposed to clothing constraints and environmental noises).

Key words: prematurity, newborns, vocalisations, comfort, discomfort

1. Introduction

Young infants experience language, start uttering articulated sounds around 7 months old (Oller, 1980) and produce their first words around 8-12 months (Bates, 1979). However, infants vocalise long before (Scheiner, 2002). During their first months they produce so-called cries, laughs, squeals, moans and growls (Young and Decarie, 1974; Laufer and Horii, 1977; Koopmans-van Beinum and van der Stelt, 1986). In fact, infants pay attention to adults' vocal stimulations (Stevenson et al., 1986) and appear to be communicating with their caregivers who react to these vocalisations (Zeskind and Collins, 1987; Dessureau et al., 1998). Parents and infants engage in vocal interactions respecting turn-taking rules (Stevenson et al., 1986; Ginsburg and Kilbourne, 1988; Reissland and Stephenson, 1999), as early as 32 weeks post conception in preterm infants (Caskey et al., 2014).

However, little is known about very young infants' vocalisations. Although newborns vocalise spontaneously (Michelsson et al., 1996) by producing acoustic structures shaped by prenatal experience with their native language (Mampe et al., 2009), few studies describe the diversity of sounds emitted during the first days/weeks of life. Most acoustic studies of young infants focused on crying during painful or uncomfortable procedures (Owens and Todt, 1984; Gaspardo et al., 2008; Gustafsson et al., 2013; Kelly et al., 2017; Koutseff et al., 2017). These vocalisations are intense strident productions, with a high pitch and abrupt frequency modulation (Porter et al., 1986; Hadjistavropoulos et al., 1994; Branco et al., 2007). Newborns' cries can encode the level of arousal, becoming more strident as the pain increases (Hadjistavropoulos et al., 1994; Bellieni et al., 2004) or more aperiodic in distress contexts (Koutseff et al., 2017).

Studies of older infants suggest a possible acoustic encoding of the emotional valence of their internal state. Distress vocalisations (pain, hunger or fussing cries) of two to six-month-old infants display a higher frequency than "non-cry" vocalisations (Fuller and Horii, 1986) and can be distinguished by ear (Brennan and Kirkland, 1982). A few studies of several-months old infants confirmed that different types of vocalisations are associated with positive (*e.g.* laughs or squeals) and negative (*e.g.* cries or wails) emotions (Young and Decarie, 1974; Kent and Murray, 1982; Scheiner, 2002). One study of full-term infants a few hours after birth highlighted acoustic differences during invasive (intramuscular injection) or non-invasive (touching, rubbing) procedures, the former being associated with longer and higher-pitched sounds (Grunau et al., 1990). Further studies of infants' vocalisations are needed to understand their spontaneous vocal expression of various internal states better.

Animal bioacoustics' studies show a link between a caller's internal state, its physical impact on the muscles and the acoustic structures, with possible parallels with humans (Scherer and Kappas, 1988; Briefer, 2012). The "Motivation-Structural rules hypothesis" predicts that low-pitched and atonal vocalisations are produced when being hostile, while callers produce high-pitched and tonal sounds when frightened or appeasing while approaching in a friendly manner (Morton, 1977). Indeed, the acoustic patterns (*i.e.* repetitive structures, abrupt frequency modulations, very high pitch) of sounds typically associated with alarm or distress have a direct (unpleasant and attention-getting) effect on receivers (Rendall and Owren, 2009). Both emotional valence and arousal intensity can be encoded acoustically. For example, mammals associate increase of arousal with increases of mean fundamental frequency, energy spectrum and number of repeated units (Briefer et al. 2015a, Lemasson et al. 2012 and 2015), whereas call duration and fundamental frequency variations may code for emotional valence (Briefer et al., 2015a and 2015b). Call rates and vocal sequence length also increase following stress (Marx et al., 2001; Cooper and Vierck, 1986; Lemasson et al., 2010 and 2012).

With the recent increase of premature births, improving preterm infants' well-being and identifying visible and reliable welfare indicators has become a major societal challenge. Vocalisations are interesting candidates (Van Beek et al., 1994) and understanding their "meaning" is essential. Indeed, even though preterms can have difficulties expressing discomfort with body movements (Craig et al., 1993; Gibbins et al., 2008), they vocalise frequently (Van Beek et al., 1994; Gaspardo et al., 2008), even before their expected term age (Caskey et al., 2011). This study focused on preterm infants' vocal productions in non-painful resting contexts. We aimed to establish possible relationships between vocal rate and acoustic structures related to their internal state. First, we established the vocal repertoire of preterm infants based on vocal recordings and evaluated its relevance *via* acoustic measurements and the analysis of associated behavioural expressions. Second, we tested the reliability of vocalisations as indicators of an infant's emotional state by focusing on their rate of emission in relation to clothing constraints and background noises, contexts known to impact infants' comfort (Darcy et al., 2008; Lasky and Williams, 2009; Durier et al., 2015). We hypothesised that physical constraints would trigger an increase of acoustic structures supposedly associated with discomfort, and that human-related noises would stimulate infants' vocal production (attempts to communicate, social facilitation).

2. Methods

Protocols, approved by the Brest and Rennes Regional and University Hospital Centre ethical committees, followed the Helsinki declaration. Neonates were included after their parents' informed consent had been received with ensured anonymity.

2.1 Participants

The study took place in January-April 2011. Parents received a document explaining the experiment, *i.e.* infants video recorded while resting in bed, at least 15 min after any treatment. Participants were 10 preterm infants with no brain injury and no breathing apparatus (6 girls and 4 boys born at 31.4 ± 1.9 s.d. weeks, NB1-NB10) observed between 35-38 weeks post-conception (age: 32.7 ± 16.6 days). Six infants had clothing constraints (CC: pyjamas and sleep-sack) and four did not (NCC: bodysuit and light wrapping), as defined by Durier et al. (2015). They were tested at the level II unit of the Brest University Hospital NICU that follows NIDCAP developmental care guidelines (Westrup, 2007).

2.2 Procedure

The infants were filmed, between two feeding periods, with an infrared camcorder (Sony HDR-XR200) connected to a microphone (Dell® Latitude-D600, Sample rate 44100Hz). Ten minutes of the video recording of each subject were analysed. These periods were chosen randomly under the condition that they included different stages of the Prechtl's scale (1974), that is moments when asleep (eyes closed, calm or small movements) and awake (eyes open, agitated or not).

2.3 Analyses of behaviours and vocalisations

All the infants' behaviours (Table 1) were scored every second using 1/0-sampling (Altmann, 1974). To test the reliability of our scoring, we asked a second rater to blindly resample 420 scores (i.e. 4 babies, one score every 30s) using the same repertoire and obtained an inter-rater agreement of 92% (Cohen's Kappa = 0.86). Each vocalisation was extracted from the videos using VLC software for subsequent analysis with ANA® software (Richard, 1991; resampling 11kHz /16bit). In parallel, all the occurrences of environmental noises were scored and categorised as human (talking, laughs, coughs...) or material (scope beep, closing a door, object falling…) noises.

The vocalisations were classified based on audio-visual inspections of spectrograms (as in animal bioacoustics: McCowan and Reiss, 1997; Datta and Sturtivant, 2002; Lemasson and Hausberger, 2011). Acoustic measurements were done for all tonal vocalisations (Fig.1). We noted whether a vocalisation was emitted alone or as part of a sequence (series of vocalisations separated from one another by silent breaks ≤ 1s, Fig. 2).

2.4 Statistical analyses

To confirm that the vocal types identified audio-visually had an acoustic validity, a principal component analysis (PCA) was run on the acoustical data (log transformed). The coordinates on the first two axes were then extracted for use in a mixed linear model (with individual identity as random factor to account for the unbalanced subjects' contributions). Then, after checking the residual plots, a Manova and its subsequent Anova compared the different vocal types applying Bonferroni correction.

In order to assess the potential association of vocal types with the infants' internal state, we performed a first exploratory factorial component analysis (FCA) with all vocalisations and associated behaviours. It revealed a link between vocal types and particular behaviours, *i.e.* body tension or facial expressions. We ran a final FCA concentrating on these behaviours. Post-hoc tests were used to further confirm the results. Wilcoxon (N≥6) or Sign tests (N=5) were used for testing intra-individual variations in the proportions of facial expression and body tension according to vocal types.

Mann-Whitney tests were used for comparing groups: impact of age, sex and time of day (morning *versus* afternoon) on the infant's emission of the different vocal types and vocal patterns (isolated *versus* sequence). Similarly, Mann-Whitney tests were used to compare the characteristics of vocal production between infants with (CC) and without (NCC) clothing constraints. Finally, Spearman correlation tests evaluated relationships between the type of environmental noise and vocal production.

3. Results

3. 1. Vocal repertoire of preterm infants at rest

We recorded a total of 1550 vocalisations: 155 isolated vocalisations and 291 vocal sequences that included 4.80 +/- 3.95 s.d. consecutive vocalisations. Audio-visual inspection of the vocalisations enabled us to propose a vocal repertoire composed of nine vocal types (Fig. 3). Some vocalisations (6.7%) could not be classified due to poor recording quality. First, vocalisations were classified as either tonal (with fundamental frequency), atonal (noisy sound) or intermediate (partially tonal and atonal). Tonal, intermediate and atonal vocalisations represented 44.0%, 8.5% and 47.5% of our recordings, respectively. Atonal as well as intermediate vocalisations were each divided into two vocal types, according to their duration leading to four vocal types: AS (Atonal Short), AL (Atonal Long), IS (Intermediate Short) and IL (Intermediate Long). Tonal sounds were divided according to their fundamental frequency and duration yielding five vocal types: TLS (Tonal Low-pitched Short), TLM (Tonal Low-pitched Medium), TLL (Tonal Low-pitched Long), THS (Tonal High-pitched Short) and THL (Tonal High-pitched Long).

The acoustic parameters of all 635 tonal vocalisations (22 - 133 per infant) were measured (Table 2). A PCA analysis, extracting 60.31% of the inertia on axis 1 and 23.75% on axis 2, was computed on these parameters and confirmed the acoustic distinctiveness of the five pre-identified tonal vocal types (Fig. 4a,b). Fundamental frequencies and duration explained most of the distribution along axis 1 and axis 2 respectively (Fig. 4c): the first axis opposed TLL, TLM and TLS to THL and THS, and the second axis opposed TLL, TLM and THL to TLS and THS.

Each vocal type was emitted by 86% (+/- 22) of the infants and all infants emitted isolated vocalisations and vocal sequences (Table 3). No effects of sex (Mann-Whitney test, Ngirls=6, Nboys=4, 2<U<10.5, P>0.05), gestational age (threshold for "very preterm" birth: less *versus* more than 32 weeks, Mann-Whitney test, Nless=4, Nmore=6, 6<U<12, P>0.05) or time (morning *versus* afternoon, Mann-Whitney test, Nmorning=3, Nafternoon=7, 2<U<10.5, P>0.05) on the vocal type or vocal patterns produced could be observed.

3.2. Towards acoustic indicators of comfort and discomfort?

The FCA run with the most relevant behaviours (body tensions and facial expressions, Fig. 5a,b) yielded a total inertia of 64.75 % for the first two axes. Axes 1 and 2 (respectively characterized by facial expressions and body tensions) opposed on one side long and medium low-pitched tonal vocalisations (TLL, TLM) and on the other side short and long high-pitched tonal vocalisations (THS, THL) and atonal short (AS) sounds.

Post-hoc analyses confirmed that infants presented more facial expressions (Wilcoxon tests, Lips: Z=2.201, P=0.028; Eyes: Z=2.201, P=0.028) and were tenser (arm distance from the head: Z=2.201, P=0.028; Sign test, arm tension: Z=2.236, P=0.025) when emitting relatively tonal low-pitched long (TLL) and medium (TLM) vocalisations than when emitting tonal high-pitched short (THS) and long (THL) or atonal short vocalisations (AS). Since 82% of our sampled facial expressions had a negative valence (Young and Décarie, 1974; Steiner, 1979; Ganchrow et al., 1983), it appears that the tenser the infants, the more they used long and low-pitched vocalisations.

3.3. Characteristics of vocalisations in relation to practical situations

Infants with clothing constraints (CC), known to be in a relatively uncomfortable situation (Durier et al., 2015), emitted more vocalisations (Mann-Whitney test, $N1_{CC}$=6, $N2_{NCC}$=4, U=1; P=0.019) and presented vocal sequences composed of a greater number of vocalisations (Mann-Whitney test, $N1_{CC}$=6, $N2_{NCC}$=4, U=1; P=0.019) than infants with no clothing constraints (NCC). Thus, the more the infants felt uncomfortable, the more they vocalised. Moreover, infants with clothing constraints emitted a higher proportion of low-pitched medium (TLM) and long (TLL) tonal, and long atonal (AL) vocalisations, whereas no such difference was observed for the other vocal types (Mann-Whitney test, Z=2.530; P=0.010; Z=2.239, P=0.019; Fig. 6). Conversely, infants with no clothing constraints emitted higher proportions of atonal short (AS) vocalisations (Z=2.239, P=0.019; Fig. 6). Thus, the more the infants felt uncomfortable, the more they emitted long vocalisations, notably atonal and low-pitched ones.

The higher the number of human noises, the more the infants produced isolated vocalizations (Spearman correlation test, P=0.036, R=0.665, Fig. 7a), whereas the exact opposite was observed for material noises (P=0.058, R=-0.616). They also emitted less vocal sequences when the number of material noises was high (P=0.047; R=-0.638) (but not for human noises: P=0.080, R=0.578) (Fig. 7b). The types of vocalizations produced also differed according to the types of surrounding background noises. There were more atonal vocalisations overall (P=0.002, R=-0.842), and particularly more atonal short (AS, P=0.002, R=-0.845), but also tonal high-pitched short (THS) vocalisations (P=0.028, R=-0.689) produced when material noises were rare. No other vocal type appeared affected.

4. Discussion

To our knowledge, this is the first description of young preterm infant vocalisations emitted before the age of 38 weeks post-conception, in a calm (resting) context. Infants produced a broad range of vocalisations falling into nine vocal types distinguishable acoustically and non-randomly associated with behaviours. High-pitched sounds were associated with less facial movements and relaxed postures and thus appeared to have a positive valence. In contrast, long low-pitched vocalisations were associated with tensed arms and facial expressions associated with discomfort and therefore appeared to have a more negative valence. These vocalisations could be useful indicators of infants' internal state in practical situations: their number, their pattern of emission (isolated *versus* sequence) and the vocal types produced varied according to the presence of constraints imposed by clothing as well as the type of background noises. Constraining clothes were associated with higher rates of vocal sequences and long low-pitched vocalisations. Also, whereas human-associated noises triggered an increase of vocal production, material noises inhibited vocal production.

Our study underlines the rich vocal production of preterm infants, composed of subtle differences between vocal types that are more complex than the well-described distress cries (Porter et al., 1986; Hadjistavropoulos et al., 1994; Branco et al., 2007). Although gradation does exist, we identified stereotyped vocal types shared by several infants (independent of sex and time). Thus, long before they babble (Oller, 1980), preterm infants are able to vocalise flexibly (Michelsson et al., 1996; Caskey et al., 2011). The first non-verbal vocalisations could be premises of vocalisations later called "cries" (TLL/TLM), "squeals" (THS/THL), "moans" (TLS) or "growls" (IL / IS) that were reported to appear during the first 2 months (Young and Decarie, 1974; Laufer and Horii, 1977; Koopmans-van Beinum and van der Stelt, 1986). Further acoustic comparisons are required to secure this conclusion. Our study confirms Caskey et al. (2011)'s report documenting that preterm infants begin to emit primitive vocalisations before their expected term age.

As predicted, the acoustic structure of these preterm infants' vocalisations could be associated with either positive or negative emotions experienced by callers. Shinya et al. (2014) showed that differences in cries' fundamental frequency of young preterm infants were not caused by body size but rather by neurophysiological states. Our findings support studies that described different types of vocalisations of several-month-old infants associating "squeals" with an overall positive valence and "cries" with a more negative valence (Young and Decarie, 1974; Kent and Murray, 1982; Scheiner, 2002). The importance of sound duration was confirmed when the production of atonal sounds of infants with and without clothing constraints was compared, as the former emitted longer and the latter shorter atonal sounds. All in all, these findings converge with earlier findings on the acoustic encoding of emotions by humans and animals: (1) low-pitched vocalisations would be associated with "hostile" or chronic discomfort contexts, and slightly higher frequencies emitted by calm and relaxed individuals (Morton, 1977; Scherer and Kappas, 1988; Portfors, 2007; Taylor et al., 2009), and (2) sound duration would increase with arousal intensity (Lemasson et al. 2012). However, these relationships are not linear as very high-pitched sounds for example are typical of intense acute distress or discomfort (Sébe et al., 2012). More general recording contexts should now be explored to evaluate in more depth the relative importance of tonal *versus* atonal, high- *versus* low-pitched and long *versus* short structures for coding emotional valence and arousal intensity by infants.

Beyond the acoustical structures, we found that vocal rate varied with context and could be another indicator of internal state. Indeed, clothing constraints increased vocal production, particularly of vocal sequences, indicating a high level of arousal, often associated with negative emotional states (Cooper and Vierck, 1986; Lemasson et al., 2010 and 2012).

Moreover, perception of human noises increased infants' vocal production whereas material noises inhibited it. We argue that infants vocalised in an interactive way when their social environment was appropriate. A previous study underlined that the more adults talked, the more infants vocalised in response (Caskey et al., 2011). Here, the infants produced more single calls when hearing human sounds, which may leave time for a response (a first step towards turn-taking interactions). Besides, the infants may have identified that when hearing human voices, a potential help/comfort source was present, and thus increased their vocalisations, whereas material noises may have been associated with potential impending disagreeable care, inducing a decline of vocal production and particularly of their positive valence vocal production.

We acknowledge that our sample size is limited (10 infants, 10 min each) and that more infants, maternity types and recording contexts are needed to be allowed to make more definitive conclusions. However, our results confirm that infants' vocalisations can be reliable indicators of their internal state and therefore highlight the fact that carers need to pay more attention to the vocal expression of newborns.

References

Altmann, J. (1974). Observational study of behavior: sampling methods. *Behaviour*, *49*(3), 227-266.

Bates, E. (1979). On the emergence of symbols: Ontogeny and phylogeny. In *Children's language and communication: The Minnesota Symposium on Child Psychology Vol. 12*, Collins, A. (ed.). Hillsdale, Erlbaum associated, 121-157.

Bellieni, C. V., Sisto, R., Cordelli, D. M., & Buonocore, G. (2004). Cry features reflect pain intensity in term newborns: an alarm threshold. *Pediatric research*, *55*(1), 142-146.

Branco, A., Fekete, S. M., Rugolo, L. M., & Rehder, M. I. (2007). The newborn pain cry: Descriptive acoustic spectrographic analysis. *International journal of pediatric otorhinolaryngology*, *71*(4), 539-546.

Brennan, M., & Kirkland, J. (1982). Classification of infant cries using descriptive scales. *Infant Behavior and Development*, *5*(2-4), 341-346.

Briefer, E. F. (2012). Vocal expression of emotions in mammals: mechanisms of production and evidence. *Journal of Zoology*, *288*(1), 1-20.

Briefer, E. F., Maigrot, A. L., Mandel, R., Freymond, S. B., Bachmann, I., & Hillmann, E. (2015a). Segregation of information about emotional arousal and valence in horse whinnies. *Scientific reports*, *4*.

Briefer, E. F., Tettamanti, F., & McElligott, A. G. (2015b). Emotions in goats: mapping physiological, behavioural and vocal profiles. *Animal Behaviour*, *99*, 131-143.

Caskey, M., Stephens, B., Tucker, R., & Vohr, B. (2011). Importance of parent talk on the development of preterm infant vocalizations. *Pediatrics*, *128*(5), 910-916.

Caskey, M., Stephens, B., Tucker, R., & Vohr, B. (2014). Adult talk in the NICU with preterm infants and developmental outcomes. *Pediatrics*, *133*(3), e578-e584.

Cooper, B. Y., & Vierck, C. J. (1986). Vocalizations as measures of pain in monkeys. *Pain*, *26*(3), 393-407.

Craig, K. D., & Grunau, R. V. E. (1993). Neonatal pain perception and behavioral measurement. *Pain research and clinical management*, *5*, 67-105.

Darcy, A. E., Hancock, L. E., & Ware, E. J. (2008). A descriptive study of noise in the neonatal intensive care unit ambient levels and perceptions of contributing factors. *Advances in Neonatal Care*, *8*(3), 165-175.

Datta, S., & Sturtivant, C. (2002). Dolphin whistle classification for determining group identities. *Signal processing*, *82*(2), 251-258.

Dessureau, B. K., Kurowski, C. O., & Thompson, N. S. (1998). A reassessment of the role of pitch and duration in adults' responses to infant crying. *Infant Behavior and Development*, *21*(2), 367-371.

Durier, V., Henry, S., Martin, E., Dollion, N., Hausberger, M., & Sizun, J. (2015). Unexpected behavioural consequences of preterm newborns' clothing. *Scientific reports*, *5*.

Fuller, B. F., & Horii, Y. (1986). Differences in fundamental frequency, jitter, and shimmer among four types of infant vocalizations. *Journal of communication disorders*, *19*(6), 441-447.

Ganchrow, J. R., Steiner, J. E., & Daher, M. (1983). Neonatal facial expressions in response to different qualities and intensities of gustatory stimuli. *Infant Behavior and Development*, *6*(4), 473-484.

Gaspardo, C. M., Miyase, C. I., Chimello, J. T., Martinez, F. E., & Linhares, M. B. M. (2008). Is pain relief equally efficacious and free of side effects with repeated doses of oral sucrose in preterm neonates? *PAIN®*, *137*(1), 16-25.

Gauthaman, G., Jayachandran, L., & Prabhakar, K. (1984). Olfactory reflexes in newborn infants. *Indian journal of pediatrics*, *51*(4), 397-399.

Gibbins, S., Stevens, B., Beyene, J., Chan, P. C., Bagg, M., & Asztalos, E. (2008). Pain behaviours in extremely low gestational age infants. *Early human development*, *84*(7), 451-458.

Ginsburg, G. P., & Kilbourne, B. K. (1988). Emergence of vocal alternation in mother-infant interchanges. *Journal of Child Language*, *15*(2), 221-235.

Grunau, R. V., Johnston, C. C., & Craig, K. D. (1990). Neonatal facial and cry responses to invasive and non-invasive procedures. *Pain*, *42*(3), 295-305.

Gustafsson, E., Levréro, F., Reby, D., & Mathevon, N. (2013). Fathers are just as good as mothers at recognizing the cries of their baby. *Nature Communications, 4*, 1698.

Hadjistavropoulos, H. D., Craig, K. D., Grunau, R. V., & Johnston, C. C. (1994). Judging pain in newborns: facial and cry determinants. *Journal of Pediatric Psychology*, *19*(4), 485-491.

Kelly, T., Reby, D., Levréro, F., Keenan, S., Gustafsson, E., Koutseff, A., & Mathevon, N. (2017). Adult human perception of distress in the cries of bonobo, chimpanzee, and human infants. *Biological Journal of the Linnean Society*, *120*(4), 919-930.

Kent, R. D., & Murray, A. D. (1982). Acoustic features of infant vocalic utterances at 3, 6, and 9 months. *The Journal of the Acoustical Society of America*, *72*(2), 353-365.

Koopmans-van Beinum, F. J., & van der Stelt, J. M. (1986). Early stages in the development of speech movements. In *Precursors of early speech*, Lindblom, B., & Zetterström, R. (eds.). New York, Stockton Press, 37-50.

Koutseff, A., Reby, D., Martin, O., Levrero, F., Patural, H., & Mathevon, N. (2017). The acoustic space of pain: cries as indicators of distress recovering dynamics in pre-verbal infants. *Bioacoustics*, *27*(4), 1-13.

Lasky, R. E., & Williams, A. L. (2009). Noise and light exposures for extremely low birth weight newborns during their stay in the neonatal intensive care unit. *Pediatrics*, *123*(2), 540-546.

Laufer, M. Z., & Horii, Y. (1977). Fundamental frequency characteristics of infant non-distress vocalization during the first twenty-four weeks. *Journal of Child Language*, *4*(2), 171-184.

Lemasson, A., & Hausberger, M. (2011). Acoustic variability and social significance of calls in female Campbell's monkeys (*Cercopithecus campbelli campbelli*). *The Journal of the Acoustical Society of America*, *129*(5), 3341-3352.

Lemasson, A., Ouattara, K., Bouchet, H., & Zuberbühler, K. (2010). Speed of call delivery is related to context and caller identity in Campbell's monkey males. *Naturwissenschaften*, *97*(11), 1023-1027.

Lemasson, A., Remeuf, K., Rossard, A., & Zimmermann, E. (2012). Cross-taxa similarities in affect-induced changes of vocal behavior and voice in arboreal monkeys. *PloS One*, *7*(9), e45106.

Lemasson, A., Remeuf, K., Trabalon, M., Cuir, F., & Hausberger, M. (2015). Mares prefer the voices of highly fertile stallions. *PloS One*, *10*(2), e0118468.

Mampe, B., Friederici, A. D., Christophe, A., & Wermke, K. (2009). Newborns' cry melody is shaped by their native language. *Current biology*, *19*(23), 1994-1997.

Marx, G., Leppelt, J., & Ellendorff, F. (2001). Vocalisation in chicks (*Gallus gallus dom*.) during stepwise social isolation. *Applied Animal Behaviour Science*, *75*(1), 61-74.

McCowan, B., & Reiss, D. (1997). Vocal learning in captive bottlenose dolphins: a comparison with humans and nonhuman animals. In *Social influences on vocal development*, Snowdon, C. & Hausberger, M. (eds). Cambridge University Press, 178-207.

Michelsson, K., Christensson, K., Rothgänger, H., & Winberg, J. (1996). Crying in separated and non-separated newborns: sound spectrographic analysis. *Acta Paediatrica*, *85*(4), 471-475.

Morison, S. J., Holsti, L., Grunau, R. E., Whitfield, M. F., Oberlander, T. F., Chan, H. W., & Williams, L. (2003). Are there developmentally distinct motor indicators of pain in preterm infants? *Early Human Development*, *72*(2), 131-146.

Morton, E. S. (1977). On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *The American Naturalist*, *111*(981), 855-869.

Oller, D. K., & Yeni-Komshian, G. (1980). The emergence of the sounds of speech in infancy. *Child phonology*, *1*, 93-112.

Owens, M. E., & Todt, E. H. (1984). Pain in infancy: neonatal reaction to a heel lance. *Pain*, *20*(1), 77-86.

Porter, F. L., Miller, R. H., & Marshall, R. E. (1986). Neonatal pain cries: effect of circumcision on acoustic features and perceived urgency. *Child development*, *57*, 790-802.

Portfors, C. V. (2007). Types and functions of ultrasonic vocalizations in laboratory rats and mice. *Journal of the American Association for Laboratory Animal Science*, *46*(1), 28-34.

Prechtl, H. F. (1974). The behavioural states of the newborn infant (a review). *Brain research*, *76*(2), 185-212.

Reissland, N., & Stephenson, T. (1999). Turn-taking in early vocal interaction: a comparison of premature and term infants' vocal interaction with their mothers. *Child: care, health and development*, *25*(6), 447-456.

Rendall, D., Owren, M. J., & Ryan, M. J. (2009). What do animal signals mean? *Animal Behaviour*, *78*(2), 233-240.

Richard, J. P. (1991). Sound analysis and synthesis using an Amiga micro-computer. *Bioacoustics*, *3*(1), 45-60.

Sarnat, H. B. (1978). Olfactory reflexes in the newborn infant. *The Journal of pediatrics*, *92*(4), 624-626.

Scheiner, E., Hammerschmidt, K., Jürgens, U., & Zwirner, P. (2002). Acoustic analyses of developmental changes and emotional expression in the preverbal vocalizations of infants. *Journal of Voice*, *16*(4), 509-529.

Scherer, K. R., & Kappas, A. (1988). Primate vocal expression of affective state. In *Primate vocal communication*, Todt, D., Goedeking, P., & Symmes, D. (eds.). Berlin, Springer-Verlag, 171-194.

Schinya, Y., Kawai, M., Niwa, F., Myowa-Yamakoshi, M. (2014). Preterm birth is associated with an increased fundamental frequency of spontaneous crying in human infants at term-equivalent age. *Biology Letters, 10,* 20140350.

Sèbe, F., Poindron, P., Andanson, S., Chandeze, H., Delval, E., Despres, G., Toporenko, G., Bickell, S., & Boissy, A. (2012). Bio-acoustic analyses to assess emotion in animals: Acoustic patterns are linked to behavioural, cardiac and hormonal responses of ewes to the separation from their lambs. *Bioacoustics, 21*(1), 54.

Steiner, J. E. (1979). Human facial expressions in response to taste and smell stimulation. *Advances in child development and behavior*, *13,* 257-295.

Stevenson, M. B., Ver Hoeve, J. N., Roach, M. A., & Leavitt, L. A. (1986). The beginning of conversation: Early patterns of mother-infant vocal responsiveness. *Infant behavior and Development*, *9*(4), 423-440.

Taylor, A. M., Reby, D., & McComb, K. (2009). Context-Related Variation in the Vocal Growling Behaviour of the Domestic Dog (*Canis familiaris*). *Ethology*, *115*(10), 905-915.

Van Beek, Y., Hopkins, B., & Hoeksma, J. B. (1994). Development of communicative behaviors in preterm infants: The effects of birthweight status and gestational age. *Infant behavior and development*, *17*(2), 107-117.

Westrup, B. (2007). Newborn Individualized Developmental Care and Assessment Program (NIDCAP)—family-centered developmentally supportive care. *Early human development*, *83*(7), 443-449.

Young, G., & Décarie, T. G. (1974). An ethology-based catalogue of facial/vocal behaviors in infancy. *ETS Research Report Series*, *1974*(2).

Zeskind, P. S., & Collins, V. (1987). Pitch of infant crying and caregiver responses in a natural setting. *Infant Behavior and Development*, *10*(4), 501-504.

Figures

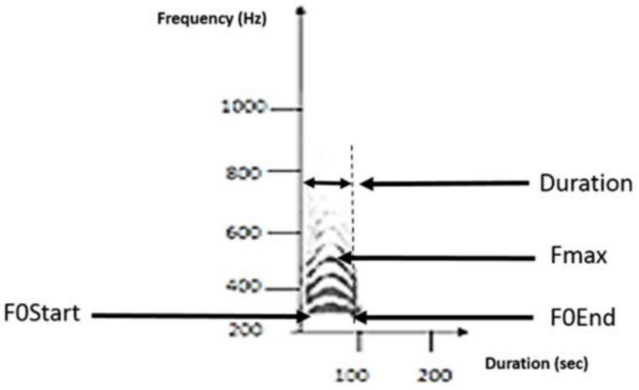


Fig. 1. Acoustic parameters measured on vocalisation spectrograms. We measured the fundamental frequency pattern (duration, frequency at the beginning (F0Start) and at the end (F0End) and the dominant frequency (Fmax, *i.e.* frequency with the highest energy)).

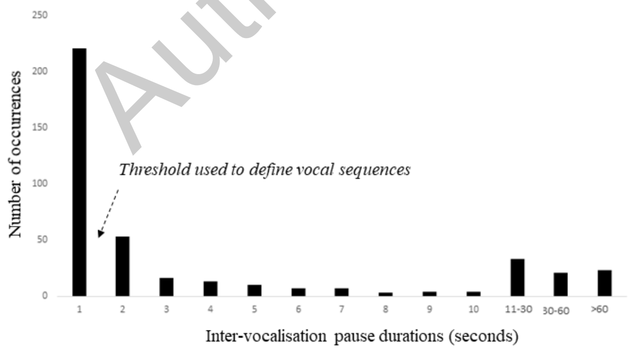


Fig. 2. Frequency distribution of inter-vocalisation pause durations.

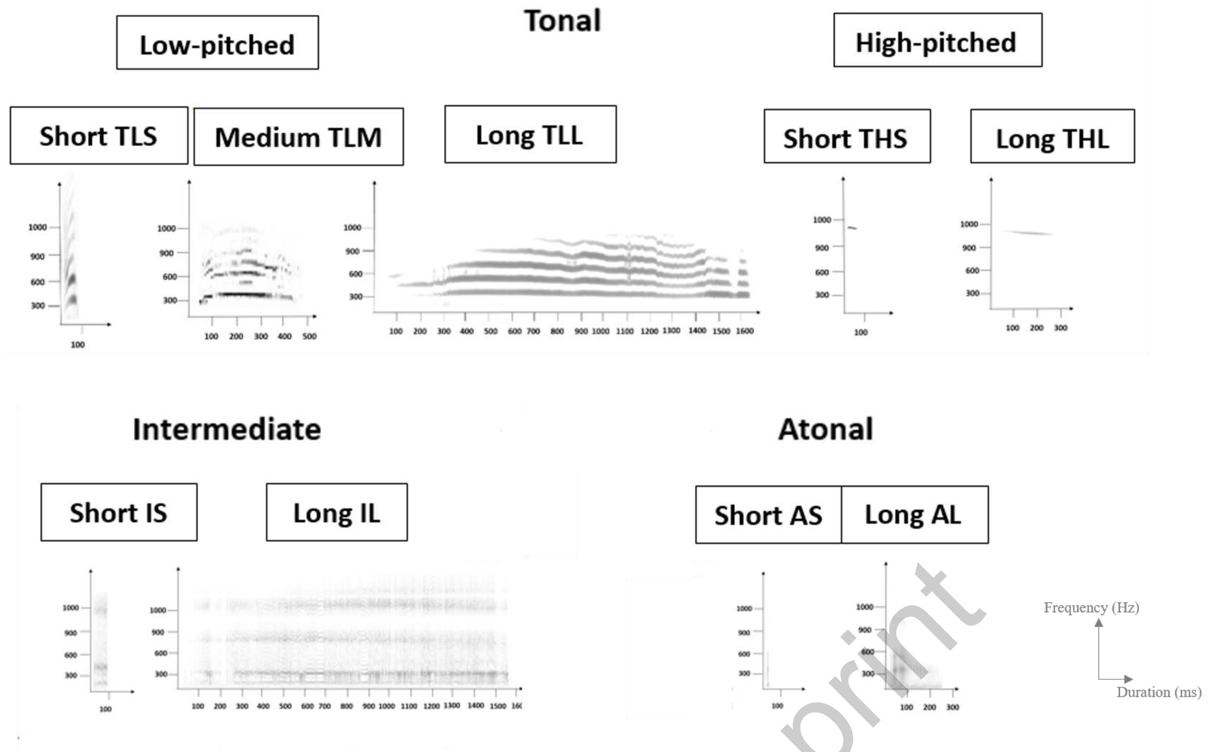
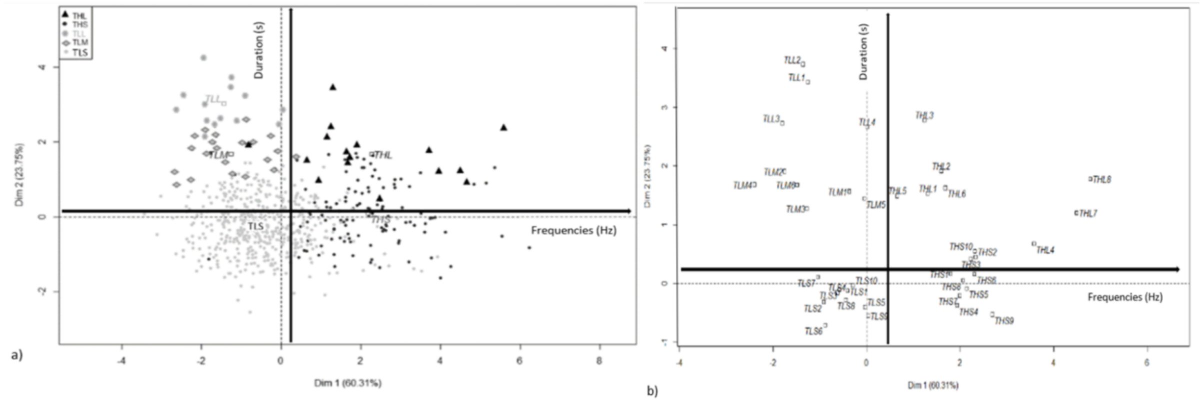


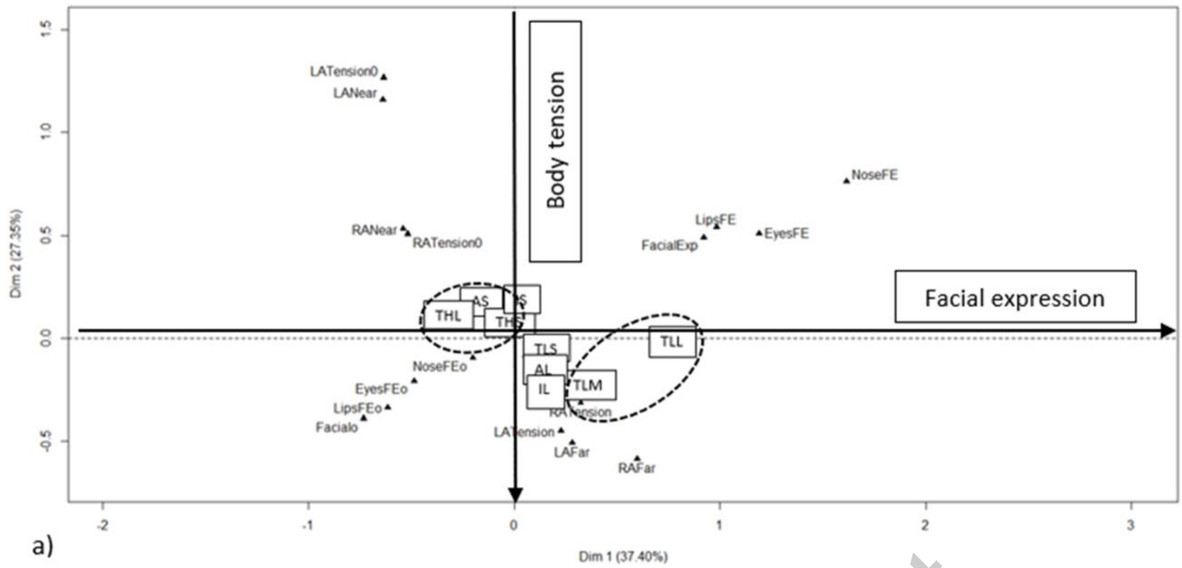
Fig. 3. The nine vocal types: Tonal Low-pitched Short (TLS), Tonal Low-pitched Medium (TLM), Tonal Low-pitched Long (TLL), Tonal High-pitched Short (THS), Tonal High-pitched Long (THL), Intermediary tonal/atonal Short (IS), Intermediary tonal/atonal Long (IL), Atonal Short (AS), Atonal long (AL). Atonal Short (AS) and Atonal Long (AL) vocalisations were always respectively single and multiple unit sounds. Intermediate Short (IS) and Intermediate Long (IL) vocalisations lasted respectively less or more than 60 ms.



c)

	Axis 1	Axis 2
F0Start	0.9204	0.0882
F0End	0.9206	0.0892
Fmax	0.7951	0.1442
Duration	-0.2908	0.9559

Fig. 4. The Principal Component Analysis (PCA) run with the different acoustic measures: a) differentiating the five tonal vocal types (with the representation of the barycentre of each type of vocalisation), THL THS TLL TLM TLS; the barycentres of the different vocalisations are labelled “THL”, “THS”, “TLL”, “TLM”, “TLS”; b) showing a high consistency between the subjects (with the representation of the individual barycentres for each type of vocalisation). The barycentres of the different vocalisations for each subject are labelled “THLX”, “THSX”, “TLLX”, “TLMX”, “TLSX” with “X” the number of the subject, accordingly; and c) contribution coefficient of each parameter on each axis (See Figure 3 and Table 2 for definitions of vocal code names and acoustic parameters). Post-hoc Anova using the PCA coordinates confirmed the acoustic differences between these vocal types. Indeed, there was a gradient in the fundamental frequencies from the highest in THS and THL, intermediate values for TLS, to the lowest in TLM and TLL ($P>0.05$: THS-THL & TLM-TLL, $P=007$: TLS-TLL, $P=0.01$: TLS-TLM, $P<0.01$: all other comparisons). Concerning sound duration, we found a gradient, TLL was longer than TLM and THL (not different from each other, $P>0.05$), that in turn were longer than TLS, again longer than THS ($P<0.001$ for all significant differences).



b)

	Items	Axis 1	Axis 2
Facial expression	FacialExp	12.52	4.87
	Facialo	9.96	3.87
	EyesFE	13.75	3.44
	EyesFEo	5.64	1.41
	LipsFE	12.42	5.14
	LipsFEo	7.77	3.21
	NoseFE	9.62	2.94
	NoseFEo	1.19	0.37
Arm Tension	RATension	2.14	2.80
	RATension0	3.44	4.50
	RAFar	5.66	7.50
	RANear	5.12	6.79
	LATension0	3.53	19.25
	LATension	1.25	6.85
	LAFar	1.82	8.26
	LANear	4.15	18.80

Fig. 5. a) Factorial Component Analysis (FCA) run with vocal type, facial expression and arm tension data; b) Contributions of behaviours to the two FCA axes (See Figure 3 and Table 1 for definitions of vocal and behavioural code names). Arrows indicate the direction of the effect of the behaviours.

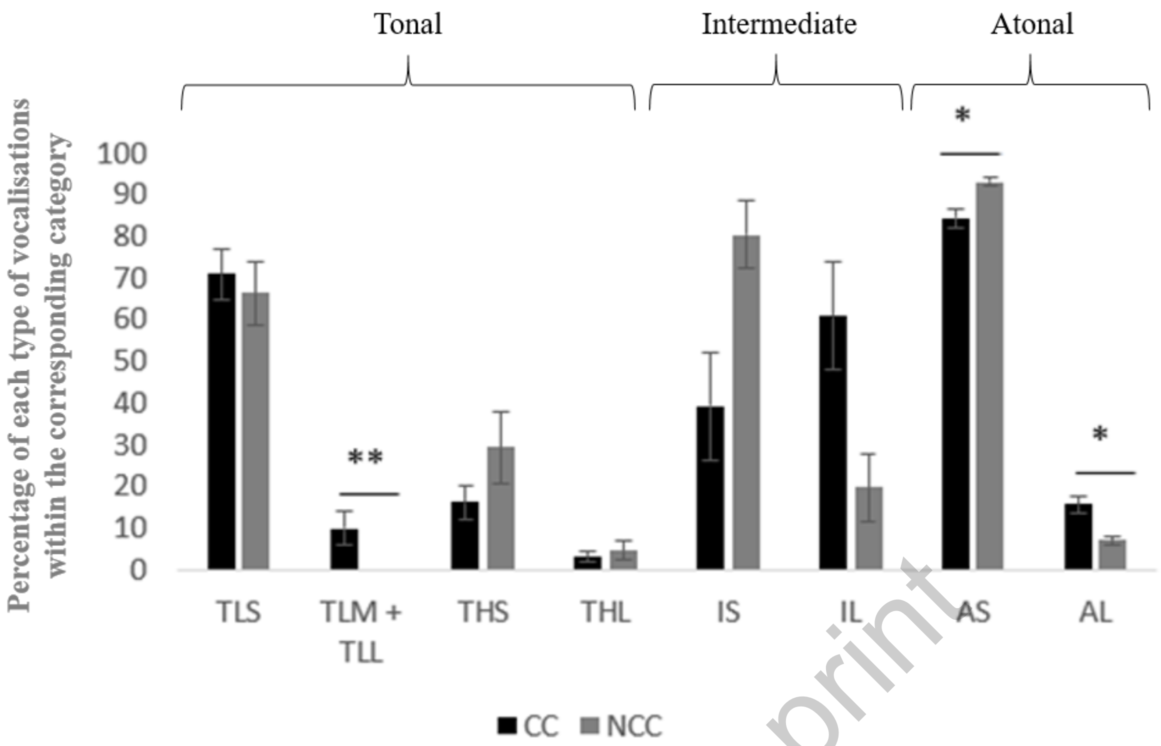


Fig. 6. Vocal production (percentage of each vocal type, in proportion) in relation to clothing constraints (CC: Clothing Constraints; NCC: No Clothing Constraints). Percentage of occurrences of the different vocal types were calculated within each vocal category (tonal/intermediate/atonal). See Figure 3 for definitions of vocal types.

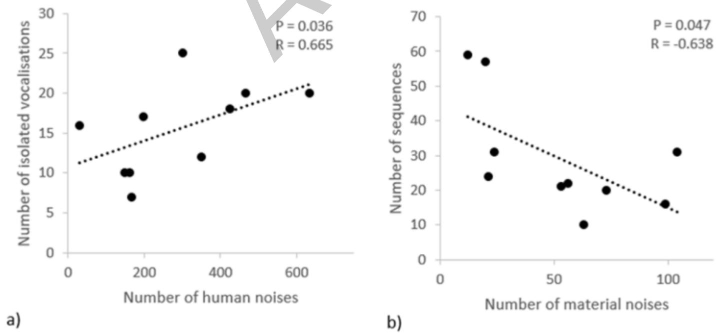


Fig. 7. a) Isolated vocalisations in relation to human noises. b) Vocal sequences in relation to material noise. One symbol corresponds to one infant. Spearman correlations.

Behavioral repertoire			Code	Definition
Facial expression	Lips	Presence	LipsFE	Smile, pout, opened, stretched or tight lips
		Absence	LipsFO	Relaxed lips
	Eyes	Presence	EyesFE	Frown or squint
		Absence	EyesFO	Relaxed eyes
	Nose	Presence	NoseFE	Frown or moving
		Absence	NoseFO	Relaxed nose
	Face	Presence	FacialExp	At least one of the item of the face showing an expression (lips, eyes or
		Absence	FacialO	Any item of the face showing an expression
	Eyes posture	Opened	OpenedEyes	Eyes opened more than the half of the eyelid
		Semi	SemiEyes	Eyes opened less than the half of the eyelid
		Closed	ClosedEyes	Eyes totally closed
Posture and Body tension	Right hand	Tension	RHTension	Tensed fingers
		Tight	RHTight	Clenched fist
		Opened	RHOpened	Opened fingers without tension
		Closed	RHClosed	Fist
		Semi	RHSemi	At least one finger, but not all, opened or closed
	Left hand	Tension	LHTension	Tensed fingers
		Tight	LHTight	Clenched fist
		Opened	LHOpened	Opened fingers without tension
		Closed	LHClosed	Fist
		Semi	LHSemi	At least one finger, but not all, opened or closed
	Right arm	Tension	RATension	Arm-forearm angle greater than 90°
		No tension	RATension O	Arm-forearm angle less than 90°
		Far	RAFar	Rigth hand localised below chest level
		Near	RANear	Rigth hand localised above chest level
	Left arm	Tension	LATension	Arm-forearm angle greater than 90°
		No tension	LATension O	Arm-forearm angle less than 90°
		Far	LAFar	Left hand localised below chest level
		Near	LANear	Left hand localised above chest level
Movement and General activity	Eyes	Presence	EyesM	Eyes moving
		Absence	EyesO	Eyes immobile
	Head	Presence	HeadM	Head moving
		Absence	HeadO	Head immobile
	Arms	Presence	ArmsM	Arms moving
		Absence	ArmsO	Arms immobile
	Hands	Presence	HandsM	Hands moving
		Absence	HandsO	Hands immobile
	Feet	Presence	FeetM	Feet moving
		Absence	FeetO	Feet immobile

Table 1. Behavioural repertoire, associated codes and definitions (these categories are defined following Young and Decarie, 1974; Sarnat, 1978; Gauthaman et al., 1984; Morison et al., 2003; Durier et al., 2015).

		F0Start (Hz)	F0End (H z)	F0Max (Hz)	Duration (s)
TLS	**Average**	352.57	336.80	578.79	61.94
	S.D.	154.75	145.73	393.42	38.35
TLM	**Average**	311.43	287.48	520.22	320.17
	S.D.	122.52	107.72	310.35	123.61
TLL	**Average**	266.69	275.38	693.69	1082.56
	S.D.	53.44	82.46	466.13	690.38
THS	**Average**	1021.17	917.54	1143.33	57.78
	S.D.	509.09	527.23	475.25	42.98
THL	**Average**	1312.56	1080.13	1331.13	221.06
	S.D.	921.69	691.75	736.03	199.54

Table 2. Acoustic characteristics of tonal vocal types (mean +/- s.d.): frequency (F0Start: fundamental frequency at the beginning; F0End: fundamental frequency at the end; Fmax: dominant frequency) and Duration. N=1447

	TLS	TLM	TLL	THS	THL	IS	IL	AS	AL	Isolated voc.	Vocal seq.	Voc. within a seq.
NB1	54	4	0	4	3	3	16	41	9	16	16	122
NB2	75	8	9	5	3	3	14	46	13	17	31	174
NB3	56	2	1	17	0	12	9	54	11	10	22	166
NB4	11	2	4	5	2	1	8	60	12	18	24	89
NB5	29	0	0	10	1	4	1	72	5	20	31	105
NB6	80	3	0	33	2	10	10	##	12	25	57	282
NB7	27	0	0	4	2	3	1	36	3	7	20	70
NB8	106	4	2	19	1	15	3	98	13	20	59	261
NB9	13	0	0	7	2	2	1	60	3	10	21	80
NB10	13	0	0	12	0	6	0	10	1	12	10	46

Table 3. Vocal production of our 10 subjects (NB1 to NB10): number of vocalisations per type, number of isolated vocalisations, number of vocal sequences and number of vocalisations emitted within a sequence. See Figure. 3 for definition of vocal type codes.