
A description of the orangutan's vocal and sound repertoire, with a focus on geographic variation

Madeleine E. Hardus, Adriano R. Lameira, Ian Singleton, Helen C. Morrogh-Bernard, Cheryl D. Knott, Marc Ancrenaz, S. Suci Utami Atmoko and Serge A. Wich



Photo © Adriano Lameira

4.1 Introduction

Many studies have focused on primate sound and vocal repertoires¹ (e.g. Marler and Hobbett 1975; Steenbeek and Assink 1996; Wich *et al.* 2003b) as well as on those of birds and other animals. These have shown that there are differences regarding repertoire size, development and acoustic structure between and within call types (e.g. Brooks and Falls 1975; Kroodsma 1976; Fischer 2002). In addition, some studies reveal geographic variation in vocalizations. There are basically two ways in which such variation can occur: acoustic variation within the same call type (e.g. orangutan long vocalizations) between populations and the presence/absence of

certain call types (e.g. orangutan nest raspberry) in some populations, but not in others. A substantial body of literature has shown that there is within call type variation in acoustic structure between populations (Marshall *et al.* 1999; Mitani *et al.* 1999; Delgado 2003, 2007; Crockford *et al.* 2004). There have been far fewer studies, however, that indicate that certain calls² do not occur in all populations in which a species has been studied (e.g. van Schaik *et al.* 2003a, 2006a). It has been suggested that the absence/presence of some calls in some populations but not in others could be socially learned and thus be another component of cultural variation between populations. To be able to make detailed comparisons of calls between populations

¹ Vocalizations are distinct from sounds, since no vocal cords are involved during emission of sounds.

² Vocalizations and sounds together are hereafter called 'calls'.

careful documentation of the vocal and sound repertoires is needed. The results of such studies can inform us about the extent of possible call cultures in non-human primates, which potentially can provide an evolutionary basis for understanding certain aspects of human language.

Although studies on primate calls are currently popular, the calls of our closest relatives the great apes, have been relatively understudied. Most research on great ape vocalizations has been conducted on chimpanzees (*Pan troglodytes*) and for this species the most attention has been given to the conspicuous pant hoots (e.g. Marshall *et al.* 1999; and Mitani *et al.* 1999; Crockford *et al.* 2004), but much less to other calls (but see Mitani and Gros-Louis 1995). For the other five great ape species (*Gorilla beringei*, *G. gorilla*, *Pan paniscus*, *Pongo abelii* and *P. pygmaeus*) there is unfortunately even less information (e.g. gorillas: Harcourt and Stewart 2001; bonobos: de Waal 1988; Fruth and Hohmann 1994; orangutans: MacKinnon 1974; Rijksen 1978; Peters 2001).

MacKinnon (1974) and Rijksen (1978) were the first to publish lists of orangutan calling behaviors. At least 16 calls have been described during these early studies (Table 4.1). Thereafter, studies mainly investigated the structure and the function of long calls (Galdikas 1983; Mitani 1985a; Delgado 2003, 2007) and recently geographic variation in the presence/absence of call types (van Schaik *et al.* 2003a, 2006a). An exception is the study by Peters (2001) who examined kiss-squeaks. In addition to long calls and kiss-squeaks, orangutans make a large number of other vocalizations that range from middle range calls that can be heard up to 300 hundred meters away to very faint raspberry sounds made during nest building. The function of all these calls, except for the long call (Chapter 14) is largely unknown and basic questions on individual and contextual variation have not yet been addressed (Hardus, Wich and Lameira unpublished data).

To further our knowledge on orangutan acoustic communication, this chapter has two aims. The first is to update the list of orangutan call descriptions by MacKinnon (1974) and Rijksen (1978) with an intensive set of new recordings from the wild and captivity. The second aim is to use the updated list of the orangutan vocal and sound

repertoire to make a preliminary comparison of orangutan calls between orangutan populations. Variation between orangutan populations may be influenced by one of three factors or a combination of all or two. These factors are (1) ecology, (2) genetics and (3) social learning. The latter has been favored as being very relevant to variation in some orangutan sounds that occur at particular sites but not at others (van Schaik *et al.* 2003a, 2006a), seemingly without any clear ecological or genetic explanation.

4.2 Results

4.2.1 Orangutan vocal and sound repertoire

In total 32 calls were classified (Table 4.1, Appendix) by ear or by using acoustic parameters of 2501 analyzed calls (Box 4.1). In Box 4.2 information can be found about the field sites and data collection. All calls are described in the Appendix to this chapter with additional details. Figure 4.2 provides the spectrograms of several call types.

We classified orangutan calls in three classes based upon the distance from the caller at which these could still be heard. Short-distance calls were heard until approximately 25 m; middle-distance calls until approximately 250 m; and long-distance calls beyond 250 m. The orangutan call repertoire of our sample consists of a large number of short-distance calls compared to middle and long distance calls (see Table 4.1). All the sounds are short-distance calls, whereas vocalizations can be short-, middle- or long-distance calls.

Immatures have 10 call types of which 4 (crying and screaming, frustration scream, play ooh, soft hoot and whimper) are emitted solely by them. Nulliparous females' repertoire contains 12 call types and females with an immature have 14 call types. Adult females (nulliparous and parous) have four unique call types: mating squeals, fear squeaks, lork call, and roar. Unflanged males have the largest repertoire (17 call types), but no unique call types. Flanged males emit 15 call types of which one is unique to this age class (fast long call).

Summaries of all characteristics of vocalizations and sounds are provided in Tables 4.2 and 4.3. In these tables we followed the terminology

Box 4.1

Most recordings of wild orangutans were made in Tuanan (Mawas Reserve, Central Kalimantan, Indonesia). This site (2° 09' 06" S; 114° 26' 26" E) consists of peat swamp on shallow peat, of varying thickness, up to about 2 m. Orangutan research in Tuanan (wild orangutans) started in January 2003 and, for this study, calls were recorded until August 2005. Additional recordings were made at the Sabangau field site (wild orangutans) and at the Great Ape Trust, Des Moines (captive orangutans). We recorded vocalizations from 27

individuals; 10 flanged males, 6 unflanged males, 5 adult females, 6 adolescents and immatures. These recordings were made on days that standard observational data were being collected (van Schaik 1999, www.aim.unizh.ch/orangutannetwork.html). All vocalizations and sounds emitted by a focal orangutan and its associates were recorded opportunistically. For each recording we noted which individual vocalized, the context, time of day, distance from the recorder to the orangutan, and information on associating orangutans.

of Struhsaker (1967) for describing the acoustic properties of the basic acoustic patterns of calls, which contain the following unit elements: tonal, non-tonal, compound and mixed. For all call types maximum frequency and duration were measured. The acoustic properties of several calls made it impossible to measure more parameters. Additional parameters were measurable in six vocalization types (Table 4.3) after consulting their spectrograms. Information concerning the calls bared-teeth scream, contact uff and grinding does not exist.

Many of the calls emitted by wild orangutans have been recognized by researchers in captive orangutans. However, this study did not intensively collect calls emitted in captivity. Consequently, no comparison was made between wild and captive individuals' repertoire.

4.2.2 Geographic variation of orangutan calls

Earlier studies indicated geographical variation in the presence/absence of several orangutan call types (kiss squeak with hands, kiss squeak with leaves, raspberries, throatscrape and nestsmack: van Schaik *et al.* 2003a, 2006b). To determine if additional call types also show geographic variation, we compared different populations. For this preliminary comparison a representative audio file was made for each call type and sent to orangutan researchers to check for presence/absence at their research site. In total eight research sites (seven wild, one captive) were compared: Sumatra;

Ketambe (H. Rijksen, S. Wich, S. Utami) and Suaq Balimbing (I. Singleton, C. van Schaik), Borneo; Ulu Segama (J. MacKinnon), Tuanan (M. Hardus, A. Lameira and S. Wich), Gunung Palung (C. Knott), Kinabatangan (M. Ancrenaz) and Sabangau (H. Morrogh-Bernard), and one research facility (Great Ape Trust, USA; R. Shumaker, S. Wich).

Table 4.4 makes a preliminary comparison of orangutan calls between orangutan populations. In this table presence/absence geographic variation can be seen for several call types. For the following seven vocalizations a presence/absence variation could be seen: ahh vocalization, ahoor call, bark, complex calls, fast long call, fear squeaks and lork, where only the complex calls show a high prevalence.

One other call type, the play ooh, has also a clear presence/absence geographic variation, although only a single individual in two populations emitted this call type. In addition, several calls (frustration whine, squeak and whine) given by an adolescent female at the Great Ape Trust (Des Moines) have not been reported elsewhere.

4.3 Discussion

4.3.1 Vocal and sound repertoire size and composition

The orangutan vocal and sound repertoire in our sample consists of 32 call types that are distinguished from each other either by ear and/or by their acoustical structure. However, we may be

Table 4.1 Comparison of call types for various descriptions of the orangutan repertoire

MacKinnon (1974)	Described call types		Short-, middle- or long-distance communication?
	Rijksen (1978)	This chapter	
	<i>Wild</i>		
Play grunt	Ahh	Ahh	Short
Ahoor		Ahoor call	Middle/long
	Bared-teeth scream	Bared-teeth scream	Short
Bark	Bark	Bark	Short/middle
Complex calls		Complex calls	Middle/long
	Contact uff	Contact uff	Short
Chomp	Chomping	Chomp	Short/middle
Crying/screaming	Scream	Crying/screaming	Short
		Fast long call	Middle/long
Fear squeak	Squeak	Fear squeak	Short
Frustration scream		Frustration scream	Short
Gorkum		Gorkum	Middle
	Grumble	Grumble	Short/middle
Grumph	Grumph	Grumph	Short/middle
	Grunt (only rehabilitants)	Grunt	Short
Kiss squeak/wrist kiss	Kiss sound	Kiss squeak regular	Short/middle
		Kiss squeak hand	Short/middle
		Kiss squeak leaves	Short/middle
Long call	Long call	Long call	Long
Lork	Lork call	Lork	Long
Matings cries	Mating squeals	Mating squeals	Short
		Nest smacks	Short
		Play ooh	Short
Raspberry	Spluttering	Raspberry	Short
		Roar	Middle
		Rolling call	Short
Soft hoot/whimper	Pout moan	Soft hoot/whimper	Short
		Throatscrape	Short
	<i>Rehabilitants and captives</i>		
		Frustration whine	Short
	Grinding	Grinding	Short
		Squeak	Short
		Whine	Short

making a distinction between calls of different types when in fact the orangutans treat them as the same, or alternatively we may be classifying calls as belonging to a single type when in fact there are multiple, systematic subtypes within a class (e.g. Seyfarth *et al.* 1980; Rendall *et al.* 1999). Thus, we could have under or overestimated the number of call types and so our results should be seen as preliminary. Carefully designed playback

experiments are needed to fully resolve this issue. Additionally systematic studies of orangutan calls have not been conducted at other sites and thus most comparisons are based on researcher recall.

An important result of this study is that the number of orangutan call types is twice that of the previously reported number (MacKinnon 1974; Rijksen 1978). This indicates that continued research

Box 4.2 Acoustic measurements

Calls were recorded using a Marantz PMD222 analog cassette recorder with a Sennheiser ME 64 directional microphone and a Sony TCD-D100 digital recorder with a Sony ECM-M907 microphone. Recordings from the recorders were transferred onto a PC and digitized with a 44.1 kHz sampling rate. Quantitative analysis of calls was carried out with Raven Interactive Sound Analysis Software (1.2.1) with the following spectrogram settings: Hann window function; filter bandwidth: 124 Hz; grid frequency resolution: 86.1 Hz; and grid time resolution: 5.80 ms. All vocalizations consisted of a fundamental frequency band (F0), and some types also showed series of harmonic overtones. All vocalization measurements were taken from F0, since previous qualitative studies of vocal repertoires (Green 1975; Hohmann 1989) indicated that from the perspective of production and perception, fundamental frequency is the most robust acoustic feature (Moody *et al.* 1990). Equivalent frequency measurements in sounds were taken measuring the frequencies where the most energy was concentrated. A call was only analyzed after visual inspection of the spectrogram showed that all parts of the call were recorded well and could be analyzed. To describe the overall acoustic

structure, we determine that the following six parameters (Fig. 4.1) – one temporal parameter and five frequency parameters – were most suitable for describing the shape and acoustic structure of the calls, although not all these parameters were applicable for all call types:

Temporal Parameter

1. Call duration (s) → D

Frequency Parameters

2. Maximum frequency (Hz): the frequency with the highest amplitude of the lowest harmonic (F0) of a single unit → MF

3. Start frequency (Hz): frequency of F0 at the beginning of an unit → SF

4. End frequency (Hz): frequency of F0 at the end of an unit → EF

5. Low frequency (Hz): frequency of F0 at the lowest point of an unit → LF

6. High frequency (Hz): frequency of F0 at the highest point of an unit → HF

The spectrograms presented in this chapter are narrow-banded with filter bandwidth of 43 Hz.

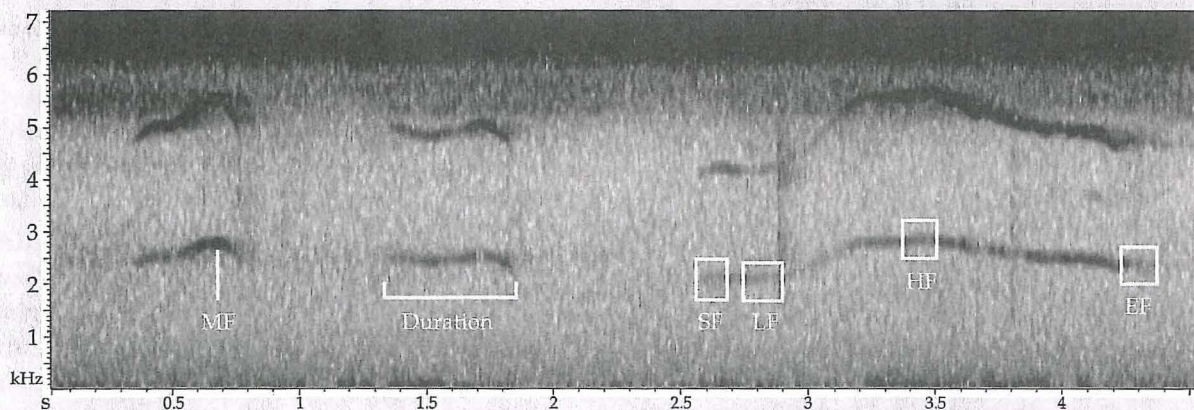


Figure 4.1 Spectrogram of two acoustical units showing the six different parameters: maximum frequency (MF), start frequency (SF), end frequency (EF), low frequency (LF), high frequency (HF) and duration (D), which are measured from each unit of a call when applicable. For these Ravens' maximum frequency function was used.

effort at one site and the inclusion of more sites will expand the repertoire and such biases should therefore be accounted for when conducting comparative studies. While acknowledging these limitations, the published number of call types for the

other great apes is: 38 call types for *P. paniscus*, 29 for *P. troglodytes* and 16 for *G. gorilla* (Fossey 1972; Goodall 1986; Bermejo and Omedes 1999), compared with 32 found here for orangutans. A recent study (McComb and Semple 2005) argued that call

Table 4.2 Features and descriptive statistics of 23 orangutan sounds and vocalizations

Call type	Basic feature *	Call	Maximum frequency (Hz) Range/mean±SD	Duration(s) Range/mean±SD	n
Ahh	Compound	V	430–3000/1014±1000	0.030–0.130/0.091±0.048	60
Ahoor call	Tonal	V	172–862/388±209	0.1–0.5/0.23±0.13	61
Bark	Non-tonal	S	258–690/425±136	0.14–0.35/0.225±0.057	15
Chomp	Tonal	S	258–517/365±67	0.14–0.34/0.194±0.037	30
Complex call	Tonal+compound+non-tonal	V	258–690/374±146	0.162–0.952/0.50±0.22	82
Gorkum	Compound	V	86–1637/270±128	0.25–2.43/0.87±0.47	145
Grumph	Compound	V	86–1723/265±82	0.15–0.51/0.224±0.08	76
Fear squeak	Mixed	V	517–690/617±65	0.07–0.15/0.1±0.03	6
Frustration Whine	Tonal	V			1
Grumble	Tonal	V	86–345/185±72	0.26–1.63/1.02±0.49	111
Grunt	Compound	S	258–431/330±68	0.11–0.22/0.15±0.035	18
Kiss squeak regular	Non-tonal	S	517–8269/3822±919	0.070–1.412/0.375±0.176	1216
Kiss squeak hand	Non-tonal	S	603–7235/2341±1357	0.046–1.506/0.396±0.217	98
Kiss squeak leaves	Non-tonal	S	172–5513/1960±1397	0.216–0.765/0.414±0.145	49
Mating squeals	Tonal	V	1292–1723/1574±122	0.063–0.196/0.146±0.04	11
Nest smack	Non-tonal	S	345–6546/1203±807	0.017–0.156/0.053±0.022	231
Play ooh	Tonal	V	258–517/338±59	0.063–0.129/0.096±0.016	41
Raspberry	Non-tonal	S	258–603	0.3–0.5	4
Roar	Compound	V	345–1809/894±573	0.31–0.63/0.46±0.11	8
Rolling call	Compound	S	172–258/241±36	0.13–0.73/0.325±0.183	10
Squeak	Tonal	V	861–3704	0.05–0.08	7
Throatscrape	Compound	S	431–1723/612±268	0.030–0.14/0.075±0.029	21
Whine	Tonal	V			1

For each call type parameters are presented with information about the range and mean with standard deviation. For these call types only maximum frequency and duration are measured, on the basis of their acoustic profile.

S, sound; V, vocalization; n, number of calls analyzed.

* Following Struhsaker 1967.

repertoire size in non-human primates is positively correlated with group size and time spent grooming, as measures of social bonding, but had orangutan call repertoire in the analyses as consisting of 10 call types. If the argued correlation between call repertoire and the measures of sociality will be upheld in future analyses, then there are two possible alternative explanations for the relatively large call repertoire of orangutans. The first is that factors other than the sociality measures used by McComb and Semple (2005) are important for the evolution of a large call repertoire in orangutans, and the second is that orangutans used to be more social than they are at present (Rijksen 1978; van Schaik 2002). Thus, the great apes species with the largest repertoire sizes live in fission–fusion societies.

Moreover, van Schaik (2002) speculates that disturbance in the past has led to a significant breakdown of orangutan societies due to human influences, which could also explain the relatively large and apparently inexplicable acoustic repertoire size in orangutans today.

The orangutan call repertoire consists of a large number of short-distance calls relative to middle- and long-distance calls. In orangutan fission–fusion society it is likely that, although infrequent, social encounters play an important role in an individual's life. The variety of interactions during such encounters substantiates the diversity of calls that individuals may employ. Moreover, orangutan sounds are only employed in short distance communication. This was expected as non-tonal signals

Table 4.3 Features and descriptive statistics of six orangutan vocalizations.

Vocalization type	Basic feature *	Maximum frequency (Hz) Range/mean±sd	Start frequency (Hz) Range/mean±sd	End frequency (Hz) Range/mean±sd
Crying and screaming	Tonal+compound response	1206–2842/2211±804	1206–1895/1637±353	1120–2067/1751±464
Fast long call	Tonal (pulse)	258–345/322±39	86–345/234±85	
Frustration Scream	Tonal	689–1723/1209±292	689–1637/1167±289	345–1637/1073±336
Long call	Tonal (pulse)	172–862/358±140	86–689/241±81	
Lork call	Tonal (pulse)	258–345/314±42	172–345/269±67	
Soft hoot/whimper	Tonal	775–3273/1813±602	775–2756/1720±505	603–3359/1683±575

Vocalization type	High frequency (Hz) Range/mean±sd	Low frequency (Hz) Range/mean±sd	Duration(s) Range/mean±sd	n
Crying and Screaming	1206–3101/2383±824	1120–1981/1550±377		7
Fast long call	258–517/379±78		0.19–0.40/0.3±0.07	2/15 pulsa
Frustration Scream	948–1723/1289±248	345–1551/1031±326		33
Long call	172–1550/569±223		0.097–2.047/0.751±0.196	62/829 pulsa
Lork call	345–603/456±73		0.95–1.19/1.08±0.08	1/17 pulsa
Soft hoot/whimper	861–3359/1915±610	603–2670/1606±506		94

For each vocalization six parameters (one temporal parameter and five frequency parameters) are presented with information about the range and mean with standard deviation. On the basis of the acoustic profile from each vocalization the six parameters could be measured.

n, number of calls analyzed.

* Following Struhsaker 1967.

(i.e. sounds) whose energy is distributed over a wide band of frequencies at any instant only permit an animal to broadcast information over small areas (Wiley and Richards 1978). Unfortunately, a division of calls according to their range, as far as we know, has not been commonly used in primate repertoire studies. Therefore, we are unable to make any comparisons on this topic.

The repertoire of adolescent and adult individuals is larger than that of immatures, who emit less than half of the calls an adolescent or adult individual does. This can be the result of the morphological changes that primates in general go through during development. Furthermore, ontogenetic factors based on social learning could also account for this contrasting difference between orangutan immatures and adults in particular. This is in accordance with the occurrence of potentially socially learned calls that do not occur in all orangutan populations as suggested in this chapter and in van Schaik *et al.* (2003a, 2006a), although the likely cases of social learning so far refer to sounds rather than vocalizations.

4.3.2 Geographic variation in orangutan calls

There are two different patterns of geographic variation in acoustic repertoires: acoustic structural variation and presence/absence variation. Both patterns of geographic variation are present in orangutan vocal and sound repertoire. Less is known about presence/absence of certain call types in some populations. Van Schaik *et al.* (2003a, 2006a) have reported that several orangutan sounds (kiss squeak with hand, kiss squeak with leaves, raspberries, throatscrape and nestsmacks) occurred in some populations, but not in others. Our results indicate that in addition to sounds, vocalizations also show presence/absence geographic variation.

Presence/absence geographic variation may be due to three main factors: ecology, genetics and/or social learning. We presume that each of these factors (or a combination) may have influenced geographic variation of orangutan calls differently. It seems unlikely that ecological factors may have contributed to the absence/presence of a call at a particular site, since the calls with geographic

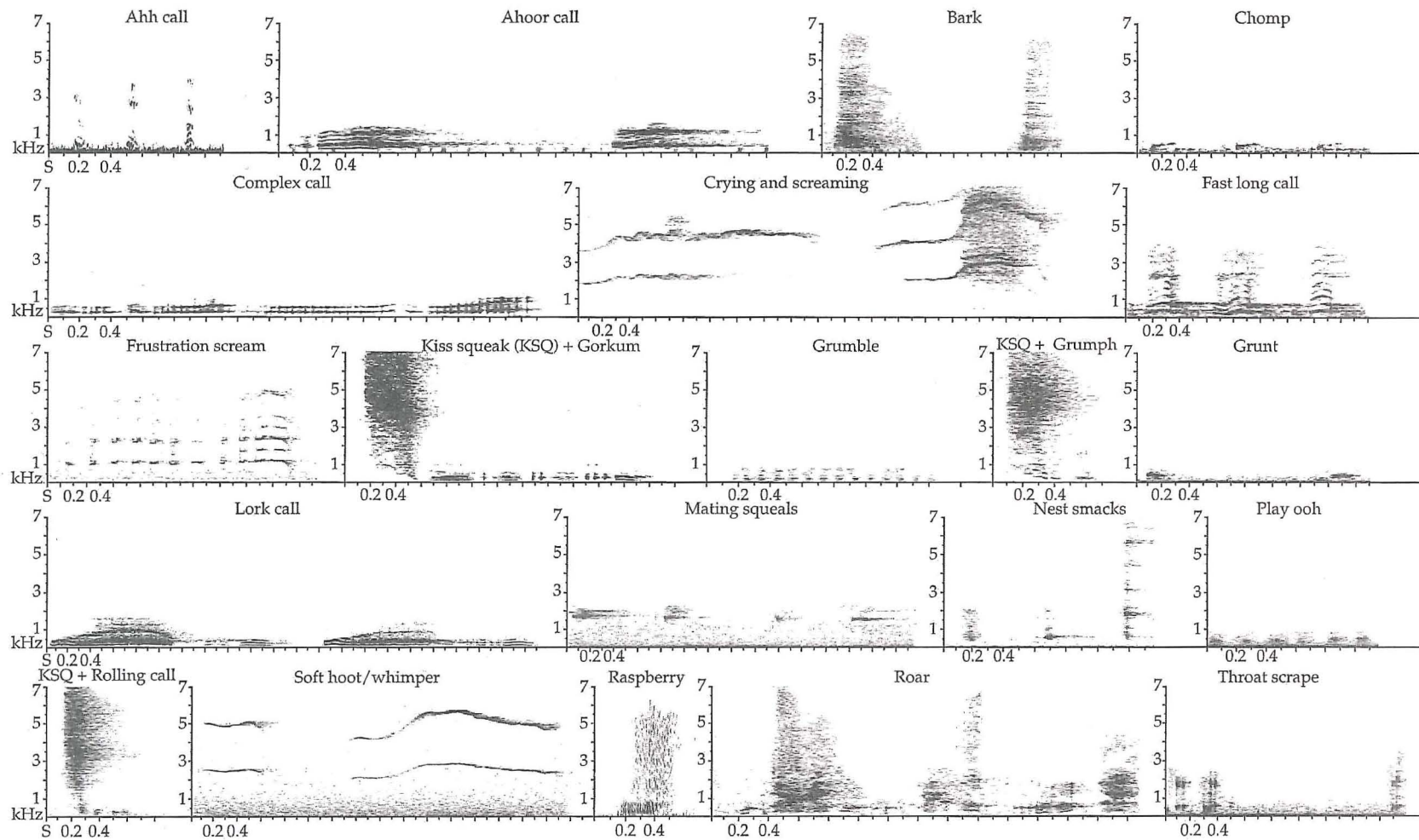


Figure 4.2 Spectrograms of certain call types, with duration on the x-axes and kHz on the y-axes.

variation reported thus far do not require any special ecological conditions to be emitted. Also, there is no evidence that genetic factors could explain entirely call presence/absence variation patterns over the studied populations (van Schaik *et al.* 2003a). Consequently, we suggest here that the presence/absence variation of vocalizations in orangutans is influenced by social learning and so these vocalizations may be cultural. However, new field work focused on calls is needed to determine whether insufficient observation time, limited sampling of the class that emits the call, and the rare context of certain calls may explain some of the presence/absence variation. The seven vocalizations with geographic variation are described below.

Call types (Table 4.4) were considered *cultural* if they were customary (shown by most or all relevant individuals) or habitual (shown by at least several relevant individuals) in at least one field site and absent in at least one other ecologically similar site (i.e. where no apparent ecological factor justifies the absence/presence of a call type, Whiten *et al.* 2001; van Schaik *et al.* 2003a, 2006a).

- *Ahh vocalization* has a low local prevalence (in Tuanan three individuals), which can be an effect of its particular context. This vocalization occurs in playing situations. Because of the confirmed or presumed absence from multiple sites, we suggest this vocalization to be cultural. In Bornean sites with similar densities as Tuanan (Sabangau, Gunung Palung), this vocalization is not recorded. If its absence were caused by the lack of opportunities for socialization, the *ahh* vocalization should be absent from all Bornean sites, given the higher sociability of Sumatran orangutans.

- The *Ahoor call* has a low local prevalence in Tuanan, emitted only by two individuals, but also described by MacKinnon (1974). This prevalence is low because this call is only emitted by males and in a particular situation (fighting context) that it is not common to observe. Nevertheless, it has never been recorded in Gunung Palung and is presumed absent in Sabangau, thus potentially a cultural call.

- Just like the *ahoor call*, the *Bark* has a low prevalence, although it occurs in several field sites.

Because of its absence in Sabangau, the bark is possibly a cultural call.

- *Complex calls* have a high prevalence. Their presence/absence variation pattern could be assumed to be a difference between Borneo and Sumatra, thus genetic. However in Malaysian Borneo it is presumed not to occur in one field site (Kinabatangan), whilst it is present in nearby Ulu Segama. Additionally complex calls have not been recorded in Suaq Balimbing (Sumatra), and may therefore be cultural.

- *Fast Long Call* is mainly made by flanged males while highly disturbed, its prevalence is low, probably as a result of its class exclusiveness and uncommon context (see description). Because of its presence in Tuanan and Tanjung Puting (Galdikas and Insley 1988) and its (presumed) absence in three other field sites in Borneo it is suggested to be cultural.

- *Fear Squeak* has a low prevalence at Tuanan but its incidence across several field sites is rather patchy. Since this call is (presumed) absent in Sabangau it is potentially cultural. Forced copulations elicited this call in the harassed female at Tuanan and these incidents also exist in Sabangau, but fear squeaks have never been heard at these sites.

- The *Lork* has a low prevalence but has been reported over nearly all field sites. Although the contexts of this call are assumed to be common and widespread, its absence at two field sites suggests that the lork may also be a cultural call.

So far, the complex call, *ahh* vocalization, *ahoor call*, *bark*, *fast long call*, *fear squeak* and *lork* are suggested to be cultural vocalizations because they are present at certain field sites, being emitted by more than one individual and they may be absent in other field sites. Nonetheless, we do not reject the possibility that some of the suggested cultural calls may become universal with the accumulation of more observation hours, in particular because focused attention on calls has been biased toward just a few sites. The play ooh is suggested to be an innovation (following van Schaik *et al.* 2006a; Ramsey *et al.* 2007), because it is only emitted by one individual at two field sites. The three vocalizations from captives seem to be innovations, but some caution is necessary as they may be the

Table 4.4 Presence/absence geographic variation of orangutan call types

Call types	Field sites								Suggestion for culture or innovation?
	Sumatra Ketambe	Borneo Ulu Segama	Borneo Tuanan	Borneo Gunung Palung	Sumatra Suaq Balimbing	Borneo Kinabatangan	Borneo Sabangau	USA Des Moines	
<i>Wild</i>									
Ahh	P	PA	P	A	P	PP	PA		Culture
Ahoor call		P	P	PP			PA		Culture
Bared teeth scream	P	PA	PA				PP		
Bark	P	P	P	PP			A		Culture
Complex calls		P	P	P	A	PA	P		Culture
Contact uff	P	PA	PA						
Chomping	P	P	P	P	PA	P	P		
Crying/ screaming	P	P	P	P	P	P	P	NA	
Fast long call		PA	P	A	P		P		Culture
Fear squeaks	P	P	P	P			PA		Culture
Frustration scream		P	P	PP	P	P	PP	NA	
Gorkum		P	P	P	P	P	P		
Grumble	P	PA	P	P	P	P	P	P	
Grumph	P	P	P	PA	P	PA	P	P	
Grunt	P	PA	P	PP			PP		
Kiss squeak regular	P	P	P	P	P	P	P	P	
Kiss squeak hand		PA	P	P			P	P	Culture
Kiss squeak leaves		PA	P	P	A		P	NA	Culture
Long call	P	P	P	P	P	P	P	P	
Lork	P	P	P	PA	P	A	P		Culture
Mating squeals	P	P	P	PP			PA		
Nest smacks	A	A	P	PA	A	A	A		Culture
Play ooh		PA	P	A	P	PP	PA	NA	Innovation
Raspberry*	P ¹	P	A	A	P	P	P	P	Culture
Roar	P	PA	P	PA			PA		
Rolling call		PA	P	P			P		
Soft hoot/whimper	P	P	P	P	P	P	P	NA	
Throatscrape	A	PA	P	A	A	PA	A		Culture
<i>Rehabilitants and captives</i>									
Frustration whine								P	Innovation ?
Grinding	P								
Squeak								P	Innovation
Whine								P	Innovation

The call types are listed in the first column. The subsequent columns refer to the status of occurrence for each call in different research sites.

P, Present; A, Absent; PP, Presumed Present; PA, Presumed Absent.

The statuses A, PP, PA can change over time after more observation hours. N.A. (Not Applicable) refers to Great Ape Trust (Des Moines) where some calls cannot occur due to lack of natural conditions or no particular age class. If the cell is empty, no information is available about the call at the specific research site.

¹Raspberries have been reported from Ketambe but never in the nesting context for which it is reported at other sites.

* These sounds at the different sites are now considered as one, but preliminary impressions from the field indicate that they might need to be separated in the future.

outcome of other learning origins, such as improvisation. The innovation status may also change with the increase of observation hours since innovations that are performed more often are more likely to reach cultural status in a population (van Schaik *et al.* 2006a). However, the possibility that these individuals are emitting this vocalization for the first time should not be discarded. Innovations are known to be at the onset of socially learned behavioral chains (culture) and the increasing number of cultural sounds and vocalizations in orangutan acoustic repertoire suggest that this phenomenon may be more frequent than assumed.

Our results suggest that not only sounds, but also vocalizations may support the cultural interpretation of geographic variation in orangutan calling behavior. This might indicate that part of the acoustic system in orangutans depends on vocal learning where vocal innovations are learned by other individuals in a population, for example songbirds, Slater (1986); killer whales, Yurk *et al.* (2002), and chimpanzees, Marshall *et al.* (1999). Because this study is heavily based on recordings from one field site (Tuanan) more research is needed to explore the extent of cultural variation in orangutan calls. Future studies at other sites may reveal more population-specific call types as well as the occurrence of shared calls not previously recognized.

Acknowledgments

We thank LIPI (Indonesian Institute of Sciences) for authorization to carry out research in Indonesia; BOS (Borneo Orangutan Survival Foundation)

for permission to work at Mawas; Universitas Nasional (UNAS) for acting as a sponsor; Jo Kolk Foundation, Foundation Fonds voor Geld-en Effectenhandel and Utrechts Universiteitsfonds for financially supporting M.E. Hardus; and Calouste Gulbenkian Foundation for financially supporting A.R. Lameira and the Netherlands Organisation for Scientific Research (NWO), Leakey Foundation, and the National Geographic Society for funding S.A.Wich. We gratefully acknowledge Peter Assink and all the students who made audio recordings (Marietta Paul, Agata Naso, Fleur Scheele and Natascha Nieuwenhuis) and film footage (Ellen Meulman) in the wild and Tine Geurts for audio recordings in Des Moines. We thank Carel van Schaik and Maria van Noordwijk for fruitful discussions and three anonymous reviewers for valuable comments on earlier drafts of the manuscript. Finding orangutans would have been impossible without the skilled tracking of the field assistants Nadi, Ganda, Linandi, Rahmat, Putra and Yandi. HM-B thanks CIMTROP and Universitas Palangkaraya for sponsoring the project. Financial support from: the L.S.B. Leakey Foundation, US Fish and Wildlife Service Great Ape Conservation Fund, The Wildlife Conservation Society (WCS), Primate Conservation, Incorporated (PCI), and the Orang-utan Tropical Peatland Project (OuTrop). Many thanks to all my assistants for help in collecting the data; and S. Husson and D. Chivers for their advice and help. All housing of apes and research at the Great Ape Trust was in accordance with current guidelines of the National Institutes of Health.

Appendix 4.1

Ahh vocalization

Exclusively heard in connection with the relaxed open-mouth facial expression that is characterized by the rather widely opened mouth whose corners may be withdrawn. Teeth may remain visible but are often covered by the lips, which may even be pulled inwards over the teeth. In any case the gums remain covered by the lips. This expression is nearly always accompanied by short rhythmic 'Ahh' vocalizations. It seems possible that this 'Ahh' vocalization is also a restrained form of the squeak vocalization (Rijksen 1978). Ahh vocalizations give also a more rasping impression than the squeak vocalization. The breathing nature is difficult to determine. During intense play with other individuals, unflanged males, females with immatures and young individuals emit this uncommon call.

Ahoor call

Uncommon threatening sound given during intimidation display, exclusively by flanged and unflanged males. Heard from subordinate individuals involved in a fight. A short inhalatory gasp is followed by an explosive exhalatory grunt (MacKinnon 1974). Ahoor calls in particular circumstances, e.g. frequently in fights, can be emitted between a grumph and a rolling call, singly or in bouts. A long ahoor call may resemble a lork pulse or the final pulse of a long call, but differs because it mainly goes up in pitch.

Bared-teeth scream

This vocalization consists of one or several very loud, high-pitched, drawn-out hoarse screams, each of which may end with a choking sound. Distinctive of this vocalization is a wide open mouth with the teeth and gums exposed. This facial-vocal display was given by animals that were attacked and bitten; in such cases the vocalizations last at least as long as the contact. In a less intensive form, the bared-teeth scream was observed during 'rapes'. On such occasions the female might show this element in connection with 'ducking', 'struggle' and 'flight' (Rijksen 1978).

Bark

MacKinnon (1974) describes the bark as a single sharp exhalation, only rarely heard from flanged and unflanged males that had heard but not seen the observer, possibly as a threat or as a warning to other animals nearby. Also reported by Rijksen (1978) as a short, loud cough-like vocalization, made with an open mouth, again rarely heard and only by suddenly startled rehabilitant orangutans. This call is regularly emitted in fight situations. It can be given before a kiss squeak, a grumph, a gorkum or singly.

Chomps

Bubbling sound emitted by flanged males and sporadically by unflanged males during pre-combat situations. Probably produced by the tongue in a partly closed mouth while making chewing movements with the jaws (Rijksen 1978). This vocalization is built up in three or two parts that will often rise in pitch. Chomps are often accompanied by nestsmacks (after or associated with breaking branches), kind of chewing sounds, grumblings and frequently followed by kiss squeaks.

Complex calls

Long string of rolling guttural noises. May be given in vigorous intimidation display (MacKinnon 1974). Males and females give this call, which is also combined with kiss squeaks. Similarly with the gorkums, the complex calls are bouts of grumphs alternated by rolling calls. The difference resides in the fact that the grumphs and rolling calls gradually change shape during the complex call. Most often rolling calls turn into long continuous strings and proportionally rise in pitch and the grumphs tend to become less audible or even absent. Bouts of gorkums and rolling calls, where they may not alternate with each other or their order may be unclear, are also named complex calls. Conversions between gorkums and complex calls may happen, where it is sometimes difficult to determine a clear separation.

Contact uff

A very soft sound, the production of which is not marked by any particular facial expression, takes

form of short repetitive expulsions of air through the nose. It seems to be a restrained 'squeak' vocalization. The contact uff can only be heard at very close range (Rijksen 1978). Rijksen (1978) only noticed it in rehabilitant orangutans, when an individual performed 'touch and smell' behavior in which it brought its nose close to the face of its partner.

Crying and screaming

Uttered by young dependent individuals who are frightened and/or separated from their mother, and by females being attacked by males (MacKinnon 1974). The screams were interspersed with cramped choking as air was inhaled (MacKinnon 1974). Crying and screaming are more intense (loudness) and higher in pitch than soft hoot/whimper. The accompanying facial expression of the screams consists of a widely opened mouth; the teeth are hardly or not exposed and the lips may be slightly protruded as if in a widened-pout expression (Rijksen 1978).

Fast long call

This uncommon call is a sequence of melodic ahoor-like pulses. These pulses may alternate and/or finish, similarly as the long call's pulses, with short bubbling sounds and they constantly go up in pitch. Emitted by highly disturbed flanged males toward observers and other orangutans (Galdikas and Insley 1988) and mainly given when the flanged males are on the ground or below five meters high.

Fear squeak

Single sharp squeaks made by frightened immatures (MacKinnon 1974) and adult females. From adult females only heard in a copulation context. This vocalization consists of a low frequency tonal unit with a distinct simultaneous noisy component at higher frequencies.

Frustration scream

It is similar to soft hoot/whimper, but the long wails were punctuated by short staccato gasps

(MacKinnon 1974). The frustration scream resembles a *gasping chicken* sound. The punctuated wails finalize with a rise in pitch. This vocalization is made in feeding situations by infants begging for food or prior to suckling. In the scale of anxiety and agitation of the infants this scream seems to lay in between soft hoot/whimper and crying/screaming.

Frustration whine

This vocalization is similar to the frustration scream in the wails punctuated by staccato gaps. However, the wails of the frustration whine are shorter and constant in tone. While staccato gaps in frustration screams are sharp, this vocalization seems to be more undulating and the tone more dull. This vocalization is given by an adolescent female before and during non-cooperative copulations with a flanged male.

Gorkum

Gorkums are bouts of grumphs alternated by rolling calls, where the throat pouch plays an important role and is swollen during emission. Grumphs, gorkums and lorks are components of a rising sequence in duration and intensity (loudness), although transition phases also occur. This middle range vocalization is mostly given after a kiss squeak, but occasionally also alone. Regularly emitted by adolescent and adult individuals of both sexes as a sign of disturbance and annoyance, and during intimidation display (MacKinnon 1974). Also made toward predators, dangerous animals to intimidate or scare them away, or toward observers. Also heard in fighting situations made by the non-dominant (unflanged) male and from a female during and after copulation. Conversions between gorkums and complex calls may happen, where it is sometimes difficult to determine a clear separation.

Grinding

This sound consists of loud grinding of the teeth, only heard from rehabilitant orangutans. While grinding its teeth, the individual's eyes were usually wide open and it repeatedly urinated and

defaecated, which could indicate fear. Grinding occurred when after its quarantine, a newly arrived rehabilitant was introduced to other rehabilitants within the confinement of the quarantine cage. Grinding also occurred in escape attempts and retreating. Schultz (1969) has described 'teeth grinding' in orangutans (Rijksen 1978).

Grumble

Mostly given by flanged males, but also heard from unflanged males and adult females. This vocalization resembles a starting engine and can only be heard from nearby. Grumbles are atonal quickly repeating vocalizations, often not interrupted by inhalation pauses (Davila Ross 2002). Most commonly given at the beginning of a long call, but can also occur as a separate vocalization in the following situations: toward observers; during copulation made by a non-dominant flanged male as a sign of excitement (Van Schaik 2004) and during a confrontation between flanged and unflanged males. Often it is clear where each individual grumble starts, but there are grumbles that are ambiguous to measure since they appear to be continuous for several tens of seconds (e.g. copulation grumbles). Copulation grumbles are softer than grumbles in other contexts.

Grumph

This inhalatory *belch-like* vocalization, that lasts up to one or two seconds, can rise up to gorkums or grumbles. Mostly given after a kiss squeak, but can also occur by itself. Made by adolescent and adult individuals of both sexes as a sign of disturbance and annoyance. It can be made toward predators or dangerous animals to intimidate or scare them away, or toward observers. Also heard in fighting situations made by the non-dominant (unflanged) male and from a female during and after copulation.

Grunt

An uncommon soft, low rasping sound. Heard from unflanged males playing and from a nulliparous female while playing with an immature, although it can also be emitted in the same contexts

as throatscapes. Also noted when an unflanged male wanted to copulate with an independent female. Possibly it is given while having contact. The grunt is not emitted in a sequence.

Kiss squeak

Common sound and given by both sexes and all ages. A sharp intake of air through pursed trumpet-like lips causes a sharp kiss sound (MacKinnon 1974). Several types of kiss squeaks occur in addition to the kiss squeak with mouth only (regular): with hand, finger(s), wrist, leaves, branch, on tree trunk, etc. Often made by unhabituated animals toward observers or when observers are too close to a habituated animal. Kiss squeaks toward other orangutans are also common as well as toward predators and other large or potentially dangerous animals. Infants have been observed practicing kiss squeaks. Acoustically these softer kiss squeaks are similar, but when involving hand, wrist or leaves the sequence of the motor acts can be uncoordinated. While vocalizations are emitted, often kiss squeaks are also given with a hoo call, bark, chomp, complex call, (fast) long call, gorkum, grumph and rolling call. Kiss squeaks with leaves, hand and regular are acoustically different from each other (Hardus and Wich unpublished data). However, in this description all the kiss squeaks are lumped.

Long call

Most frequently emitted by flanged males, but can also be made by unflanged males in transition to flanged-hood. Long calls can be spontaneous (to attract females or to discourage other males from coming into the area), as response to other long calls, as response to a falling tree or other forms of disturbance, and toward observers. Long calls consist of three parts: grumbles (the introduction), pulses (the climax) and bubbles (the tail-off). A long call always contains pulses, but not necessarily grumbles and/or bubbles. Long calls are audible up to approximately 1500 m.

Lork call

A rare vocalization that has so far only been described to be given by females and unflanged

males (MacKinnon 1974; Rijksen 1978). It is made toward other orangutans and predators. Rijksen (1978) describes lork series as unflanged male intimidation calls (only heard from rehabilitant males) and adult female advertisement call. MacKinnon (1974) states that lorks are series of call units similar to those of the long call tail-off, but that rise in pitch. MacKinnon (1974) also mentions that they are inhalatory, but they seem to be exhalatory. This vocalization resembles the long call, although the pulses are noisier. Lork series can be heard for hundreds of meters. Single lork units can also be heard after gorkums as the climax of the rising sequence grumphs- gorkums- lork and as a sign of disturbance and annoyance.

Mating squeals

The mating squeal is an exhalatory, high-pitched queaky vocalization, which is emitted by females during a copulation context. MacKinnon (1974) found them essentially the same as fear-screams but the cries were rhythmical and continued for several minutes. They are distinct from the bared-teeth scream because of their lower intensity and the different accompanying facial expression. Some females habitually cried whereas others were never heard to do so (Rijksen 1978).

Nestsmacks

A soft smack sound, resembling a feeding smack, that can be given by immatures and adults most often when making an evening nest. The sound is probably made during the breaking of twigs to improve the nest with and that are sometimes wiped through the mouth before being put down on the nest. It appears to be produced by the tongue against the palate while having a partly closed mouth. Nestsmacks have also been heard during solitary play with a twig or branch. Nestsmacks are given in two variants, one low, and the other higher in frequency. In a spectrogram the latter is usually seen as a vertical line and the former as a low big spot.

Play ooh

This vocalization was only heard from an immature when playing solitary. It resembles an *owls*

oe-oe call, but in a faster tempo. The mouth is slightly protruded while this exhalatory vocalization is emitted. Multiple and often more than three are emitted. The ooh vocalizations follow each other with an interval of approximately one-tenth of a second.

Raspberry

A faint exhalatory spluttering sound that can be given by immatures and adults when making an evening nest. The sound is probably made during the breaking of twigs to improve the nest with and that are sometimes wiped through the mouth before being put down on the nest. Raspberries occur in the same context as nestsmacks. MacKinnon (1974) refers raspberries only as calls occasionally made by juvenile animals in the same context as kiss squeaks.

Roar

A loud exhalatory roaring vocalization. The roar has only been documented twice and it was emitted by subordinate females with immature during fight with another female with immature. The subordinate female made this vocalization before and more intensively while there was contact. During the roar the immatures were not clinging.

Rolling call

This common vocalization consists of several low frequency units which resemble the ones of grumbles. Emitted by males and females during intimidation display. Can be given by itself, after or during a kiss squeak and most often as a component of gorkums alternating with grumphs. In fight situations, flanged males can produce short rolling guttural noises that resemble rolling calls, although these are just deep breaths in and out.

Soft hoot/whimper

This exhalatory *nagging* vocalization is made with the lips extended into a pout and emitted by infants. The soft hoots can be drawn out into whimpers. These vocalizations are most frequently uttered by frightened immatures. Can be made

while immatures are separated from the mother and/or in another stressful situation and in a feeding context (when begging).

Squeak

This very short rasping vocalization is emitted in bouts separated by gaps of less than 1 second. It is high toned and it may vary between a rather melodic and noisy vocalization. This is given by an adolescent female in three contexts. First, if she was begging for food from her mother; second, during new contact with familiar orangutans and human caregivers before play; third, during approach of a flanged male before they are completely accustomed to each other.

Throatscrape

A soft scraping sound from mainly adult females and rarely from unflanged males. Throatscapes are often made twice or three times in a row interspaced by 10 ms. It can be made in several situations: from a

mother before she collects her infant and most often after the infant was making soft hoot/whimpers or frustration screams; after breaking a branch during feeding or nest building; in the presence of a higher ranking individual, in particular of an adult male. Rijksen (1978) also noted this vocalization when one of two 'gnaw-wrestling' partners attempted to break off the play. The context of this vocalization suggests that it functions to express a non-agonistic attitude, sometimes submission in cases where non-dominant individuals were concerned; or possibly reassurance – in particular when a mother directed such behavior toward her offspring.

Whine

This vocalization is similar to soft hoot and whimper in its nagging tone and its drawn out whimpers. However, it is emitted in a lower frequency and it can be interspersed with chokings of air. This vocalization is given by an adolescent female before and during non-cooperative copulations with a flanged male.