Graded variation in D notes of Carolina chickadees: Do signalers vary acoustic fine structure with context and do receivers respond to that variation?

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Hwayoung Jung
May 2021

DEDICATION

This dissertation is dedicated to my parents, my little sister
and all of my mentors

ACKNOWLEDGEMENTS

When I was a high school student, one of my teachers told me that in the course of our lives, people may confront life-changing opportunities when three conditions are met: when they find what makes them enthusiastic, the intellectual and physical capability to carry on, and the right time. When I got my official PhD admission in the University of Tennessee, Knoxville, I felt one of such rare opportunities came to me and I grasped it tightly.

My advisor Dr. Todd Freeberg was the key person who allowed for me to bring this golden opportunity to my life. He was a great academic mentor that I highly look up to. Dr. Todd Freeberg has helped all of my papers develop with practical comments and edits. He also guided me to a correct path in all steps in research, from proposal to drawing final conclusions. Besides of his academic instructions, his gentle, kind and respectful personality helped me a lot to maintain my mental health during the PhD course. Whenever I attended conferences, I came across with some researchers saying they knew him and how he was nice when they saw his name on my poster, which I strongly agree.

I also greatly appreciate to all of my committee members that I admire: Dr. Gordon Burghardt, Dr. Elizabeth Derryberry, Dr. Nina Fefferman and Dr. Haileab Hilafu. You are all my heroes and I am feeling honored to have all of you as my committee members.

The dissertation research project has been funded by a NIMBioS Graduate Award and a Dissertation Support Grant from the Department of Psychology.

Harry Pepper, an Undergraduate student in the EEB department has partially assisted in data collection during many freezing winter days and helped convert part of the audio data into an Excel form. He was an extraordinary student who was beyond the level of a typical undergraduate student and I am sure someday I will see his name on a prestigious academic paper.

My lab members, Steven Kyle, Brittany Coppinger, Scott Benson, and Heather Brooks, thank you all for giving me some insightful comments on my previous drafts.

My family- my mom, dad and my little sister, I love you all till the end of my life and always thank you for giving me emotional support. My pet Guily, Homil and Robi who now all became angels in heaven, I am sorry to leave you all behind to study abroad and took so long to come back. Now it's too late, but hope to see you all again later.

BSSW members including Doyool Sunim, the mentor of my life since 2003, thank you all for helping me practice mindfulness and keep the habit of meditation that helped me being mindful during hard times, especially when I was in the final stages of my PhD work.

ABSTRACT

Animal vocal signals vary in terms of repetition rate of signal units and graded acoustic properties such as frequency and duration. Variation in the signals is known to be associated with contexts directly related to fitness, such as predation threat and food detection, and with other external factors such as group membership, geographic location of the signaler, call structure, and so on. Among animal species that produce combinatorial signals such as some primates, meerkats, and birds in the Paridae, graded acoustic properties have gotten little attention. In this dissertation, I tested whether Carolina chickadees changed their vocal signal properties in predation and food contexts and whether receivers in turn changed their behavior according to that vocal variation (i.e. tested if communication occurs in food and predation contexts). Then I also tested whether the signal varied across flocks, geographic locations, and call structures. I found the very first evidence that chickadees varied their vocalizations across food and predation contexts by changing both repetition rate and graded acoustic properties including spectral measures. However, there was no evidence that receivers behaved differently according to the signal variation. In the last study, it turned out that flock membership and call structure were accurately classified with the graded acoustic properties of notes. This dissertation demonstrates that acoustic properties of the same specific signal unit can be adjusted in two opposing fitness-related contexts, and also evidence of consistent vocal convergence regardless of food and predation contexts within flocks. Taken together, this study advances our knowledge of acoustic signal variation with various ecological factors and of receiver perception of signal variation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Significance of communication in animals

Communication is ubiquitous in all animal species. Communication takes place when individuals produce signals to influence other individuals for their own survival benefit (Burghardt, 1970), or produce cues as inadvertent by-products of their normal behavior that can also influence the behavior of other individuals (Freeberg et al., 2017). Signals tend to vary with internal characteristics such as a signaler's sex, local population, motivational status and behavioral likelihood, and with external contexts such as predation threat and food availability (Bradbury & Vehrencamp, 2011). Receivers process the signals and cues of other individuals for their own decision-making to benefit themselves (Bradbury & Vehrencamp, 2011). Therefore, communication is a crucial part to sustain life in all animals.

There have been debates regarding whether information is encoded in animal signals. There are mainly two views on whether signals contain information. Researchers such as Hailman (1977), Bradbury and Vehrencamp (2011) and Hauser (1996) argued that information is conveyed via animal signals from senders to receivers. Opposing this point of view, Burghardt (1970), Dawkins & Krebs (1978), Rendall et al (2009), Owren et al. (2010) argued that the concept of "information encoded into a signal" was problematic to understanding the function of communicating, and instead suggested that animals produce signals in order to influence other organisms to the benefit of the signalers. In the dissertation that follows, I viewed 'information' as a shorthand for variation in signals associated with either variation in the internal status of a signaler or variation in its external environment. I assessed this question of variation both in terms of signaler communicative responses to different contexts as well as in terms of receiver responses to variation in signal structure.

Signal structures

Structural variation in a signal can provide information about the state or behavior of a signaler and potentially about external events (Marler et al., 1992; Seyfarth et al., 1980;). Structural variation in signals can be discrete or can be continuously graded (Marler, 1961). Signals that are discrete are clearly distinct signal types: there is no (or minimal) intermediate form between different signal types (Marler, 1961). Under this coding scheme, signalers either repeat a single signal unit (Marler et al., 1986) or combine two or more signal units to vary the message (Bradbury & Vehrencamp, 2011). Conversely, graded signals gradually change into other types of signal, with considerable intermediate form between different signals (Bradbury & Vehrencamp, 2011; Marler, 1961). In this case, structural parameters such as duration or frequency (Hz) of an acoustic signal, or continuous variation between two different visual display postures, can vary with signaler characteristics (Sloan et al., 2005). Identifying signal units in appropriate ways is crucial to analyzing the potential diversity and complexity in animal signaling systems (Kershenbaum et al., 2016). Signal units are easily identified in cases of discrete signal types because they are separated from one another. However, in cases of graded signals, it is harder to identify different signal units.

Discrete

- **Combinatorial syntax (phonological, lexical)**

Combinatorial syntax refers to a signal structure consisting of multiple distinct signal units and can be divided into phonological and lexical syntax. Phonological syntax refers to meaningful sequences that consist of different signal units that do not have meaning when these are produced independently, and is commonly found in animal songs (Crockford & Boesch, 2005). For example, Marler (1998) pointed out that phonological syntax can be found in the song of male winter wrens (*Troglodytes troglodytes*). They produce 6-7 song types that are composed of different syllables, but the different song types do not vary in terms of their message (Marler, 1998). Besides songs, gargle calls of

black-capped chickadees (*Poecile atricapillus*) are composed of several distinct note types but the function of gargle calls of different combinations did not seem to vary (Ficken & Popp, 1992). Some combinatorial calls of some primate species likewise are classified as phonological syntax (Crockford & Boesch, 2005; Marler, 1977).

Lexical syntax refers to meaningful sequences that consist of different signal units that also have meaning when they are produced independently (Crockford & Boesch, 2005). Lexical syntax is an inherent component of structural complexity of sentences in human language. Lexical syntax usage in animal communication seems to be still debatable. Marler and Hurford argued that non-human animals do not use lexical syntax (Collier et al., 2014; Hurford, 2012; Marler, 1998). However, Crockford & Boesch (2005) argued there are some species that use lexical syntax, although lexical syntax usage is much rarer than phonological syntax. For example, banded mongooses (*Mungos mungo*) produce 'close calls' in the form of a single call when they are involved in certain activities such as digging, moving, and searching. The call can be divided into 2 parts that relate to the caller's identity and activity (Fitch, 2012). Putty-nosed monkeys (*Cercopithecus nictitans*) also use lexical syntax in their "Pyow" and "Hack" calling system (Arnold & Zuberbuhler, 2008). They mainly produce 'Pyow' calls for leopards which makes receivers travel small distance with a short latency and 'Hack' calls for eagles which makes receivers travel small distance with increased latency, but when the two distinct calls are combined to a 'Pyow-Hack' call, receivers travel far away. In non-human animal signaling systems, lexical syntax usage is mainly found in predation contexts (Arnold & Zuberbuhler, 2008; Ouattara et al., 2009a; Ouattara et al., 2009b).

Repetition

A common coding scheme with discrete signals is when a signaler repeats the same signal units, observed in a wide range of contexts. For instance, "chick-a-dee" calls from chickadees (Genus Poecile) and titmice (Genus Baeolophus)

are a representative signaling system that possesses this repeating property. The “chick-a-dee” calls consist of several note types, such as A, B, C and D notes (Sturdy et al., 2000). Any notes in the call can be present or absent and, if present, can occur more than once, making the number of possible combinations open-ended (Krams et al., 2012). For instance, the number of D notes in chick-a-dee calls increases with predation threat level (Krams et al., 2012; Templeton et al., 2005). Similar patterns appear in mammals when signalers produce calls by repeating a signal unit (reviewed in Kershenbaum et al., 2016). Besides the predation context, the repetition of a signal unit also appears in feeding contexts. Carolina chickadees (*Poecile carolinensis*) produced more D notes in their “chick-a-dee” calls when they detected a food source, and latency to take seeds was shorter to playback of calls that contained a larger number of D notes (Mahurin & Freeberg, 2008). Importantly, Wilson and Mennill (2011) found the black-capped chickadees (*P. atricapillus*) paid attention to call duty cycle, which is the proportion of time of a signal being produced within a certain time duration. Wilson and Mennill (2011) manipulated both duty cycle and the number of notes in “chick-a-dee” calls, by performing playback experiments with 4 treatments: 1) low duty cycle: 2D calls with low call rate, 2) high duty cycle: 2D calls with high call rate, 3) high duty cycle: 10 D calls with low call rate and 4) a silence as a control. Chickadee receivers showed a stronger response toward higher duty cycle calls, regardless of the note composition, indicating that the duty cycle may be as salient a signal to receivers as note composition (Wilson & Mennill, 2011).

Graded (continuous)

There is acoustically graded signal variation associated with the identity and emotional status of signalers, predation and food contexts, and the structure of call itself. Acoustically graded properties, mainly frequency (Hz) and duration varied with individual identity (Freeberg et al., 2003; Jansen et al., 2012; Mammen & Nowicki, 1981; Townsend et al., 2010;) and emotional status of signalers (Briefer et al., 2015a; Briefer et al., 2015b; Clay et al., 2015; Linhart et

al., 2015), contexts experienced by signalers (e.g. predation and food intake; Clay & Zuberbühler, 2009; Ficken, 1990; Ficken & Witkin, 1977; Slocombe & Zuberbühler, 2006; Soard & Ritchison, 2009; Templeton et al., 2005) and the discrete call structure produced by signalers (e.g. the number of a signal unit; Fischer et al., 2004; Freeberg et al., 2003). The identity of the signaler (e.g. sex, local population or group identity, individuality) varied with graded parameters within certain types of signal for a wide range of species. For example, the local population of the signaler was associated with spectral energy distribution of D notes, duration of introductory notes, peak frequency of D notes, maximum frequency of D notes, and bandwidth of D notes in black-capped chickadees (Mammen & Nowicki, 1981), and in Carolina chickadees, the duration of A notes, C note, amplitude modulation and frequency modulation in D notes varied with local population (Freeberg et al., 2003). In the banded mongoose (*Mungos mungo*), spectral and temporal parameters in a certain segment within a vocal signal predicted group identity (Jansen et al., 2012). Likewise, frequency and duration within a signal were predictive of group membership in meerkats (*Suricata suricatta*; Townsend et al., 2010).

Sex was associated with the loudest frequency of D notes (Charrier et al., 2004) and with the start frequency of A notes (Campbell et al., 2016) in black-capped chickadees. In Carolina chickadees, there was a significant interaction between the sex and the local population of the signaler in spectral and temporal properties of A notes (Freeberg et al., 2003). In crows (*Corvus brachyrhynchos*), both spectral and temporal features of the calls significantly predicted sex and individuality (Yorzinski et al., 2006). In kittiwakes (*Rissa tridactyla*), the spectral and temporal features within signals significantly predicted the sex of the signaler (Aubin et al., 2007). Also, individual signalers were identifiable by spectral and amplitude properties in giant pandas (*Ailuropoda melanoleuca*; Charlton et al., 2009), spectral and energy distribution across frequencies in spotted hyenas (*Crocuta crocuta*; Mathevon et al., 2010), spectral and pitch difference in adelie penguins (*Pygoscelis adeliae*) and gentoo penguins (*P. papua*; Aubin &

Jouventin, 2002), and temporal features (Phillips & Stirling, 2000) and spectral and energy distribution within frequencies (Charrier et al., 2002) in fur seals (*Arctocephalus tropicalis*).

Affective states are thought to be strongly associated with graded vocal structures in mammalian species. In studying emotions of animals, two main dimensions are often assessed - arousal (being emotionally excited vs calm) and valence (positive, neutral or negative emotion; Mendl et al., 2010). For example, frequency generally correlates positively with arousal level in mammals (Briefer, 2012; Zimmermann et al., 2013), including goats (*Capra hircus*; Briefer et al., 2015a), pigs (genus *Sus*; Linhart et al., 2015) and horses (*Equus caballus*; Briefer et al., 2015b), which also produced different energy distributions across frequencies according to arousal levels. For emotional valence, the fundamental frequency and/or duration of signals varied in bonobos (*Pan paniscus*; Clay et al., 2015) and horses (Briefer et al., 2015b). Hence, graded acoustic properties are associated with signaler's affective state, rather than referencing external objects or events.

Within-note properties can also vary with the note composition structure of calls. In Carolina chickadees, within-note properties in the first A or D note were associated with the number of remaining notes of the calls (Freeberg et al., 2003). This association of within-note properties with call syntax implies there might be redundancy in these calls. Redundancy is known to increase detectability of signals for receivers (Wiley & Richards, 1982) in cases when individuals may hear only part of a call, which is particularly useful in environments where signals are easily degraded (Campbell et al., 2016; Lengagne et al., 1999). Similarly, the fundamental frequency of "wahoo" calls and the duration of "hoo" calls changed across the position of signals in baboons (*Papio cynocephalus ursinus*; Fischer et al., 2004). The fundamental frequency of the "wahoo" call decreased and the duration of "hoo" call got shorter over the course of the calls in a series.

Context-specificity and functionally referential communication

The signal variation mentioned so far can be associated with different "contexts", which are individual changes or environmental stimuli accompanied with signal production (Smith, 1965), such as the presence of food and predators nearby. Smith (1963) argued that signals produced in different contexts can vary in information content. Signalers produce signals as a "message" that may convey the activity of the signaler or its affective state, etc. Receivers gain the "meaning" from both the signal and the context – the same signal or message might convey different meanings in different contexts (Smith, 1965).

For predation contexts, chickadees and titmice tend to vary the frequency and duration of note types in calls they produce, according to different predator species. For instance, chickadees varied the frequency (Hz) and duration of Z notes (black-capped chickadees: Ficken & Witkin, 1977; Mexican chickadees *Poecile sclateri*: Ficken, 1990) and the duration of the first D notes in a call sequence (black-capped chickadees: Templeton et al., 2005; Carolina chickadees: Soard & Ritchison, 2009), depending on the predator detected. However, in tufted titmice, there was no acoustic parameter variation observed according to predator species (Courter & Ritchison, 2010). Probably this is because only the duration of each note was measured and additional acoustic features such as spectral parameters and energy distribution across frequency might vary with predation contexts. Chimpanzees (*Pan troglodytes*) also produce graded alarm calls (Crockford & Boesch, 2003). Like chickadees and titmice, they have discrete calls such as pant-hoots and barks. Within the barks, individuals can vary acoustic structure in a graded way to indicate either a detected snake predator or their own hunting context (Crockford & Boesch, 2003). Similarly, chickadees are expected to have subtypes within the note types of their calls. Carolina chickadee calls have typically been classified into 6 to 7 distinct note types (6 note types: Freeberg & Lucas, 2012, 7 note types: Bloomfield et al., 2005; Freeberg, 2008), but there seems to be a need to classify more subtypes under the existing coding scheme (Lucas & Freeberg, 2007)

because of the considerable acoustic variation within some of the note types of "chick-a-dee" calls.

For food contexts, some studies (all in non-human primates) found graded fine-acoustic variation. For food contexts, both bonobos (*Pan paniscus*; Clay & Zuberbühler, 2009) and chimpanzees varied the frequency (Hz) of their calls according to the level of food preference, and chimpanzees also varied the harmonic structure and duration of calls (Slocombe & Zuberbühler, 2006). These examples showed that, like discrete signals, graded signals can vary with different contexts. If a signal variant has a strong association with a specific context, it is regarded as a context-specific signal (Crockford & Boesch, 2003), whereas there are more criteria that must be met for a signal to be considered referential (Smith, 2017). The first criterion is from the signaler's perspective, which is producing a specific signal structure in the context of a specific object or event. The second criterion is from the receiver's perspective, which requires that receivers respond uniquely and appropriately without contextual cues to the context-specific signals from signalers (Evans, 1997; Marler et al., 1992). Hence, functionally referential communication occurs when individuals produce context-specific signals that are associated with external objects or events, and elicit unique responses from receivers (Seyfarth et al., 1980; Smith, 2017; Townsend & Manser, 2013; von Frisch & Lindauer, 1956). Although some studies indicate that graded signals can vary in context-specific ways (Crockford & Boesch, 2003; Slobodchickoff et al., 1991) only a study on white-tailed ptarmigans (*Lagopus leucura*) found that they show functionally referential signals according to different predation contexts with graded signal traits, using both the production- and perception-specificity criteria (Ausmus & Clarke, 2014). Although structural distinctiveness in a signal is required to be considered a referential signal (Evans, 1997), it seems likely that more species beyond ptarmigans may use graded variation in signal units in context-specific and potentially functionally referential ways.

Gaps in the literature

There have been many attempts to examine how graded signals are associated with signaler and environmental characteristics, but there are still some gaps. First, most of the research revealed that the within-note variation in vocal signals was associated with a signaler's identity, such as sex, flock, or individuality, but only a few studies were performed regarding external contexts, and even fewer on the possibility of referentiality. To the best of my knowledge, the only study that investigated whether referential communication occurs by varying graded signal properties was the study conducted by Ausmus and Clarke (2014). Although context-specific structural distinctiveness in a signal has been documented, it is still largely unknown whether receivers react differently to that structural variation. From the previous studies, we know there are some variations in graded signal features related to a signaler's identity, emotional status, or current context. However, it is important to test whether receivers actually respond differently to the within-note variation. Appropriate playback studies that manipulate graded signal variability and control other structural variation are necessary to test this.

Second, it is notable that within-note variation is widely studied in animal species that produce a single signal type and relatively little is known for the species that produce combinatorial signals. For such species that possess combinatorial call systems, including birds in the genus *Poecile* (chickadee; Reviewed in Lucas & Freeberg, 2007) and *Baeolophus* (titmice; Hailman, 1989; Jung & Freeberg, 2017), meerkats (*Suricata suricatta;* Collier et al., 2017), and some primate and ape species (Arnold & Zuberbuhler, 2008; Casar et al., 2013; Crockford & Boesch, 2005), studies typically measured descriptive features of discrete signal units, rather than graded signal features within a single signal type. Investigating within-unit / within-note variation might expand our knowledge on how these species of animals use different coding schemes that serve different functions according to structural variation. Combinatoriality in signal structure enables signalers to produce a wide range of distinct call structures. If

graded variation within signal units that compose calls is also communicative, this work will uncover a much larger dimension of vocal complexity in these combinatorial calls (Lucas & Freeberg, 2007).

The study system

Carolina chickadees are an ideal species to address these gaps. They belong to the family Paridae and are known to possess the structurally complex signaling system of “chick-a-dee” calls. As noted above, calls are composed of distinct note types (Sturdy et al., 2000) that can be either repeated or omitted in a call structure (Krams et al., 2012). Because of their unique ability to combine the note types differently according to different contexts, these birds have become a model species in studies of acoustic communication (e.g., Courter & Ritchison, 2010). One main limitation of studies on those species is that even though their calls exhibit both discrete aspects (e.g. the number of notes and number of calls) and graded traits (e.g. duration and frequency (Hz)), the discrete note composition properties of calls have by far been the primary focus of research (Lucas & Freeberg, 2007). In the “chick-a-dee” calling system, D notes are regarded as the most salient note in calls in many contexts. Indeed, as shown in the examples discussed above, chickadees and titmice actively adjust the number of D notes in accordance with predation risk and feeding contexts (Krams et al., 2012; Mahurin & Freeberg, 2008; Templeton et al., 2005; Wilson & Mennil, 2011). Therefore, D notes seem to be an ideal signal type to test context specificity and referential usage related to graded acoustic variation.

Aims of the dissertation

To address whether acoustic properties within D notes of Carolina chickadees vary with contexts related to their survival, and how receivers respond to the variation of D notes, I conducted three studies. This research tested 1) whether senders produce different D notes in terms of within-note

acoustic structure and/or different rates of D note production according to contexts (food vs predators), 2) whether receivers respond differently to the D notes that were produced in the two different contexts, and 3) whether the acoustic structure of D notes vary by micro-geographic locations and call composition.

I assessed variation in D notes of calls with both traditional inferential statistic models and with machine learning approaches. The biggest difference between the two approaches is that traditional statistical models perform statistical hypothesis testing by estimating parameters in a particular data distribution (Valletta et al., 2017). The machine learning approach does not rely on the data distribution and its main purpose is prediction by finding patterns in the data (Valletta et al., 2017).

For example, machine learning can predict numerical values based on regression as well as categories by classification (James et al., 2013). There are multiple different models that performs classification such as random forest, SVM (Support Vector Machine), KNN (K-Nearest Neighbor), LDA (Linear Discriminant Analysis) and QDA (Quadratic Discriminant Analysis; James et al., 2013). Among these, random forest and SVM have a good reputation in general in terms of their efficacy, but random forest is generally known to be the one that results in the best classification accuracy (Hastie et al., 2009). KNN is used in some studies that classified marmoset (*Callithrix jacchus*) vocal types (Turesson et al., 2016), taxonomic classification of fish vocalization (Noda et al., 2016), and species classification of frog vocalization (Xie et al., 2016). LDA is used in vocal classification in bird and amphibian species calls but generally showed the poorest classification accuracy among other models. In some cases, however, it did outperform, implying that LDA is not always the worst method (Acevedo et al., 2009). It is also used in fish species classification by their vocalizations (Xie et al., 2016). QDA is not commonly used in animal vocalization classification, but it was used in song classification to predict bird species (McIlraith & Card, 1997). In this study, I used KNN, Random Forest, SVM with linear kernel, SVM with

radial kernel, LDA, and QDA to examine which algorithm performs the best in terms of accuracy and to obtain important variables for classification.

Significance

This work will strengthen understanding of how animals use graded signal variation in acoustic communication. It will provide much-needed information about how signalers use context-specific signals and the perceptual specificity of receivers. Furthermore, this work may broaden fundamental understanding of evolution in animal communication and cognition. The strength of this study lies in its innovative approach by focusing on graded acoustic variation, both as it is produced by signalers and how it is interpreted by receivers (Figure 1.1).

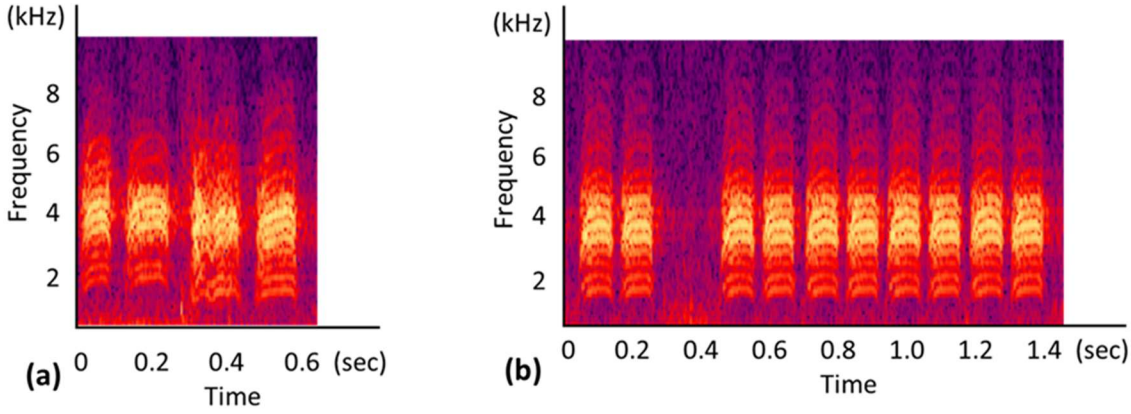

Figure 1.1 D notes that vary in terms of graded acoustic properties and repetition rate

(a) D notes of different graded properties (e.g. duration, peak frequencies, entropy, etc) (b) D notes of a different repetition rate (2 D notes vs 8 D notes)

CHAPTER I
SIGNALER'S PERSPECTIVE: SIGNAL STRUCTURE VARIATION ACROSS PREDATION AND FOOD CONTEXTS

This article was revised primarily under Dr. Todd Freeberg's suggestions. Harry Pepper, an undergraduate student in the Department of Ecology & Evolutionary Biology assisted with data collection and part of the data coding using Avisoft SAS Lab Pro software. Some statistical advice was given by OIT's Research Computing Support Group. R scripts are available at https://github.com/why-young/Dissertation/

Abstract

It is crucial for prey animals to avoid predation risk while maintaining enough food intake. Although the two factors are associated with the opposite environmental selection pressures (predation threatens survival whereas food-intake is helping survival), Carolina chickadees mainly produce the same note types ("D notes") when they call in these contexts. I tested whether there is variation in the repetition rate and graded acoustic properties in D notes across food and predator contexts. D notes were recorded in both food and predation contexts by presenting food only at the feeder and then presenting visual predator stimuli including humans wearing masks and screech owl study skins. Chickadees varied both repetition rate as well as graded acoustic properties of D notes produced in these contexts. This work suggests that chickadees might have faced ecological pressure to evolve such a complex form of vocal signaling. Perhaps controlling both graded acoustic properties within D notes, which is relatively subtle variation, and repetition rate of D notes at the same time might allow receivers to perceive and process signals more accurately for two different situations.

Introduction

Predation and food intake are two key components of survival of prey animals. They need to intake food in order to get sufficient nutrition for physical maintenance but they must also pay attention for the presence of predators nearby (Lima, 2009). Thus, there is a substantial trade-off between finding and

exploiting food and avoiding predation while taking food (Hugie, 2003). Therefore, evolution may favor individuals that successfully detect when to take food and when to avoid predator threats, and communication often facilitates these processes.

Vocal signals produced in predation and foraging contexts can appear differently due to the signaler varying components at the level of signal units (e.g. repetition rate, duty cycle) and within signal units (frequency, duration, energy distribution within the same signal unit). For example, chickadees and titmice produce a greater number of D notes in predation contexts (Courter & Ritchison, 2010; Jung et al., 2020; Soard & Ritchison, 2009; Templeton et al., 2005), and white-breasted nuthatches (*Sitta carolinensis*) increase the number of quank notes in predation contexts (Jung et al., 2020; Nolen & Lucas, 2009; Ritchison, 1983). Carolina chickadees (*Poecile carolinensis*) also produce calls that contain a greater number of D notes when they first detect food (Mahurin & Freeberg, 2009). On the other hand, variation within graded signals also commonly happens. White-tailed ptarmigans (*Lagopus leucura*) adjust graded acoustic properties such as the fundamental frequency and the dominant frequency to encode information about different predators (Ausmus & Clark, 2014). Black-capped chickadees (*Poecile atricapillus*) also vary graded acoustic properties such as the duration of D notes, inter-note intervals, the number of energy peaks, and note bandwidth according to the threat levels of different predator types (Templeton et al., 2005). Chimpanzees (*Pan troglodytes*) varied peak frequency and duration (Slocombe & Zuberbuhler, 2006), and Bonobos (*Pan paniscus*) varied absolute frequency, according to their food preferences (Clay & Zuberbuhler, 2009). These examples show how animals change their signal repetition rate and graded properties within signal types in both predation and food contexts.

Both predation and foraging contexts are important factors in survival, but it is surprising that chickadees mainly produce the same signal types in the two very different contexts – the D notes (Mahurin & Freeberg, 2009; Soard &

Ritchison, 2009). The D notes in both contexts appear to have a highly similar auditory and visual structure and are hard to be distinguish by humans' visual and auditory perception. However, no studies to the best of my knowledge tried to examine structural variation (including both graded variation such as spectral, temporal, and energy aspects as well as discrete variation) in the D notes in the two different contexts. In the recent study of great tits call, it was shown that they produce D notes with differing temporal aspects and repetition rates across foraging and predation contexts (Kalb et al., 2019). However, frequency and energy are important graded acoustic properties that cannot be neglected since in the examples above, animals commonly varied frequency and energy components (Fallow & Magrath, 2010; Ficken, 1990; Ficken & Witkin, 1977; Soard & Ritchison, 2009; Templeton, 2005).

Here, I compared within-note variation in D notes produced in contexts involving low-threat predators, high-threat predators, and food exploitation using machine learning methods. It is known that more D notes are produced in calls when Carolina chickadees detect high-risk predators (Soard & Ritchison, 2009) and when Carolina chickadees first detect food (Mahurin & Freeberg, 2008), but it is still unknown whether the D notes differ by their acoustic structures and/or by their calling rate in those contexts. I tested the following four hypotheses.

H0: No variation across contexts

This null hypothesis posits that there is no graded and no discrete signal (call rate) variation in D notes across feeding and two predator (low-threat and high-threat) contexts. Such a finding would indicate that there must be other coding schemes such as call combination and inter-note intervals associated with the contexts (Bartmess-LeVasseur et al., 2010; Templeton et al., 2005). This would represent a very different finding from what has been observed in other experimental studies with this and related species.

H1: Only graded variation across contexts

This hypothesis posits that there is graded variation in D notes, but not call rate variation, between the food and the two predation contexts. Such a finding would indicate that within-D-note variation is a salient coding scheme that can be flexibly changed according to different contexts, as shown in white-tailed ptarmigan (Ausmus & Clarke, 2014).

H2: Only discrete variation across contexts

This hypothesis posits that there is no graded variation, but there is call rate variation, associated with the food and the two predation contexts. Such a finding would indicate that, rather than adjusting graded signal properties, signalers rely on a discrete signal coding scheme, and increase call rate with increased perceived risk or with food.

H3: Both graded and discrete variation across contexts

This hypothesis posits that both graded and call rate variation are associated with the food and the two predation contexts, similar to what has been observed in black-capped chickadees (Templeton et al., 2005). Such a finding would indicate that chickadees use the most complicated forms of coding scheme among the ones suggested in this study. This sort of result would suggest that chickadees have been under social or ecological pressure to evolve to produce highly flexible acoustic variation associated with continuously changing circumstances.

Methods

General methods

There were two separate locations where I collected the data: UTFRREC (University of Tennessee Forest Resources, Research and Education Center; 36.11° N, 84.20° W; roughly 50 km^2), and NDSP (Norris Dam State Park; 36.41° N, 84.15° W; roughly 16 km^2). Each study site within a location was separated at

least 400 m from the next closest feeder site to make sure each site represented a unique flock (distance validated in Bartmess-LeVasseur et al., 2010). Each feeder site had a feeding station the birds generally exploited when stocked with black-oil sunflower seeds before experiments. Feeders were stocked with approximately 100g of black-oil sunflower seeds. All experiments were conducted between 8 am and 3 pm EST.

Data collection

Once a chickadee flock visited the feeder and started taking seeds, experiments were conducted. The recording session began with the food context for 30 min, followed by two predator stimuli at the feeder each for 5 min, (with a 15 min inter-trial baseline interval), so that the entire recording period was at least 55 min. A human observer mounted a taxidermy screech owl (high-threat predator) on a tripod, 2m away from the feeder (see Soard & Ritchison, 2009). This predator stimulus was presented for 5 min. For the other predator stimulus, a human observer wore a tiger mask (see Freeberg et al., 2014) as a low-threat predator for 5 min, also 2 m away from the feeder. To avoid pseudoreplication, I used 3 different screech owls and 4 different tiger masks. To remove the order effect between the two predator stimuli, the owl and human presentations were counter-balanced across sites. Since I was not able to start the experiment before birds took seeds on the feeder, bird calls in the feeding context were recorded earlier than predator contexts in most cases. However, for the days I did not collect complete data for all 3 contexts, I revisited the same sites again on other days and collected data of missing contexts. There are a total of 5 sites where I collected food calls later than predator contexts. When the birds did not stay for the entire experimental session, or did not produce at least 5 calls that contained D notes in a particular context, I carried out another testing session on another day at the same site. For all 3 contexts (food, owl and human), I recorded their calls with a Marantz PMD-660 recorder with 16-bit and 44.1 kHz sampling rates connected with a Sennheiser microphone (ME-66).

I initially used Avisoft SAS Lab Pro 5.2.13 (Avisoft Bioacoustics, Berlin, Germany) to extract 31 acoustic features within every D note: peak frequency, minimum frequency, maximum frequency, bandwidth, the number of peaks, entropy and harmonic-to-noise ratio in 4 different measurement locations (start, end, center and maximum amplitude), duration, internote-interval between D notes and distance to the maximum amplitude (Table 2.1).

Sampling rate was converted from 44100Hz to 22050Hz, and Fast Fourier Transformation (length of 512) was performed with the FlatTop window and overlap of 93.75%. For automated parameter measurement, batch processing for every D note was conducted. I set the range of frequency of measurement between 2 to 10kHz for better automatic detection of each D note, to remove low frequency background noise that sometimes overlapped with D notes.

To measure Inter-Observer Reliability, HJ and HP independently used Avisoft to extract 31 acoustic features in 5243 D notes. Later, HJ randomly picked 50 D notes based on random numbers generated in R version 3.6.1 (R Core Team). I performed Pearson correlation tests and found that variables measured in the maximum amplitude were reliable and only considered these 9 variables in all following analyses (peak frequency, minimum frequency, maximum frequency, bandwidth, entropy and harmonic-to-noise ratio, measured at maximum amplitude, duration, internote-interval between D notes and distance to the maximum amplitude; Table 2.2).

Table 2.1 Description of the graded acoustic properties

Definition by Avisoft SAS Lab Pro 5.2.13 (Avisoft Bioacoustics, Berlin, Germany)

Category	Measurement	Description
Frequency	Peak frequency	Frequency measured at the signal's highest energy
Frequency	Minimum frequency	Frequency measured at lowest part where the signal meets the designated threshold of amplitude
Frequency	Maximum frequency	Frequency measured at highest part where the signal meets the designated threshold of amplitude
Frequency	Bandwidth	The difference between maximum and minimum frequency
Frequency	The number of peaks	The number of peaks above a designated threshold of amplitude
Energy	Entropy	Tonality of the signal measured from 0 to 1 which is pure tone and random noise, respectively
Energy	Harmonic to noise ratio	Ratio between periodic elements of the signal and noise that quantifies noise level.
Temporal	Duration	Time between the start and end point of the signal
Temporal	Inter note-interval	Time between the end of the preceding signal and the start of next signal
Temporal	Distance to the maximum amplitude	Time between the start and where the signal reaches its maximum energy

IOR between Avisoft vs Cool Edit Pro: HJ randomly picked 30 D notes based on the random numbers generated in R version 3.6.1 (R Core Team). TF used Cool Edit Pro to measure the peak frequency, minimum frequency, maximum frequency, bandwidth, the number of peaks above -10dB, duration, break and distance to the maximum amplitude and also performed the Pearson correlation test in R version 3.6.1 (R Core Team; Table 2.2). Overall, the IOR of the number of peaks tended to be low in all measurement locations, and among the four measurement locations, variables in the “Maximum amplitude” were the highest than those measured in other locations. After excluding variables with low IOR, only 9 variables were left to be used in the actual data analyses: peak frequency, minimum frequency, maximum frequency, band width, entropy, harmonic-to-noise ratio, duration, inter-note interval, and distance to maximum amplitude.

Data analyses

All statistical analyses were performed in R version 3.6.1 (R Core Team). For hypothesis testing, I used linear mixed effects models to determine whether any acoustic properties varied across contexts. All response variables were normally distributed according to the Shapiro-Wilk normality test, which satisfies the assumption of using the linear mixed model. Random effects were sites and the interaction between site and contexts (1|site/contexts in R code; considering food, mask, and owl contexts) were nested under each site. I used the lmer() function under library “lme4”. Then I used the emmeans() function to perform post-hoc tests with Fisher’s LSD.

I used generalized linear mixed models with a negative binomial distribution and a log link function. The random effect was site (1|site) with the glmmTMB() function. To account for the time duration difference between contexts, (food context was taken for 30 min whereas mask and owl contexts were taken for 5 min), I set the time duration as an offset variable. Independent

variables were contexts with three levels (food, mask, and owl predator) and the dependent variable was the number of D notes.

As a follow-up to the hypothesis testing, I used supervised machine learning algorithms to classify contexts and determine the important acoustic properties for classification. I used KNN, random forest, SVM with linear kernel, SVM with radial kernel, LDA, and QDA to classify 3 contexts (food, mask and owl) and also 2 contexts (food and predators). Dependent variables were 9 acoustic properties and the response variable was context (food, mask and owl), and I also re-ran the same analyses with a different response variable which is context with two levels (food and predator context; combined mask and owl into predator context). Training and testing datasets were divided into a 80% and 20% ratio, respectively. The training dataset is the part of the data to be used for model fitting and the testing dataset is the part of the data to be used for unbiased evaluation (James et al., 2013).

The variable importance plot was generated from the random forest approach to show the important variables for classification, sorted in a descending order so that most important variable is placed on the top. The importance is calculated by mean decrease in “Gini index”, which is the probability of misclassifying a randomly selected data point (James et al., 2013). To evaluate classification accuracy in the test dataset, I used a hypothesis testing that tests whether classification accuracy is significantly higher than no-information rate (NIR, proportion of largest class in an observed data; Kuhn, 2008).

Table 2.2 Results of Inter-observer reliability analyses

(N/A means unmeasurable variables in Cool Edit Pro)

Measurement location	Variables measured	Correlation coefficieint	
		Between Avisoft	Between Cool Edit Pro & Avisoft
Start	Peak frequency	0.823	N/A
	Minimum frequency	0.731	N/A
	Maximum frequency	0.838	N/A
	Bandwidth	0.767	N/A
	Entropy	0.865	N/A
	Harmonic-to-noise ratio	0.562	N/A
	Number of peaks	0.319	N/A
End	Peak frequency	0.734	N/A
	Minimum frequency	0.721	N/A
	Maximum frequency	0.619	N/A
	Bandwidth	0.687	N/A
	Entropy	0.802	N/A
	Harmonic-to-noise ratio	0.610	N/A
	Number of peaks	0.370	N/A
Center	Peak frequency	0.930	N/A
	Minimum frequency	0.918	N/A
	Maximum frequency	0.931	N/A
	Bandwidth	0.932	N/A
	Entropy	0.969	N/A
	Harmonic-to-noise ratio	0.876	N/A
	Number of peaks	0.540	N/A
Maximum amplitude	Peak frequency	0.982	0.892
	Minimum frequency	0.987	0.878
	Maximum frequency	0.999	0.437
	Bandwidth	0.999	0.776
	Entropy	0.994	N/A
	Harmonic-to-noise ratio	0.893	N/A
	Number of peaks	0.642	0.468
N/A	Duration	0.913	0.913
N/A	Inter-note interval	0.755	0.925
N/A	Disttomax	0.981	0.721

Results

I collected a total of 4862 D notes throughout 25 study sites, after excluding missing variables of inter-note intervals of the first D note of a call. (The very first D note gets eliminated as these cannot take values of inter-note interval). D notes from food context were the most frequently observed (2293 D notes), followed by owl predator context (1805 D notes) and Mask (764 D notes). Out of 25 study sites, only 17 study sites had all 3 contexts each with at least 5 calls containing D notes, whereas 4 sites had 2 contexts and other 4 sites had only 1 context (Table 2.3).

Pearson correlation analyses between 9 variables yielded 36 correlation coefficients as shown below (Figure 2.1). The highest absolute value of correlation coefficient was 0.88 between maximum frequency and bandwidth, followed by 0.61 between entropy and bandwidth and 0.46 between peak frequency and maximum frequency. Except for those three major high correlation coefficients, most coefficients were under 0.45.

In the linear mixed model analyses, among the 9 dependent variables, only 2 variables (maximum frequency and bandwidth) showed significant differences between contexts (Table 2.4). Post-hoc tests showed that maximum frequency was significantly higher in the food context (5248.004 ± 12.803) than in the owl predator context (5032.810 ± 11.785; $p=0.0007$) and also higher in the mask context (5079.753 ± 20.387) compared to the owl predator context ($p=0.0292$). Bandwidth was significantly higher in the food context (1580.526 ± 13.915) than in the owl predator context (1369.152 ± 13.847; $p=0.0019$), and higher in the mask context (1506.873 ± 22.211) than in the owl predator context ($p=0.0147$; Table 2.5).

Table 2.3 The number of of D notes in each context by flock

Flock	Food	Mask	Owl	Total
flock 1	295	58	303	656
flock 2	217	35	101	353
flock 3	33	98	211	342
flock 4	168	81	82	331
flock 5	156	42	108	306
flock 6	146	76	77	299
flock 7	118	136	35	289
flock 8	68	27	125	220
flock 9	167	34	13	214
flock 10	121	5	84	210
flock 11	26	36	88	150
flock 12	74	26	41	141
flock 13	38	34	56	128
flock 14	69	28	11	108
flock 15	28	16	25	69
flock 16	37	15	17	69
flock 17	16	7	21	44
flock 18	129	N/A	47	176
flock 19	68	N/A	95	163
flock 20	123	N/A	31	154
flock 21	106	N/A	36	142
flock 22	N/A	N/A	98	98
flock 23	90	N/A	N/A	90
flock 24	N/A	10	78	88
flock 25	N/A	N/A	22	22
Grand Total	2293	764	1805	4862

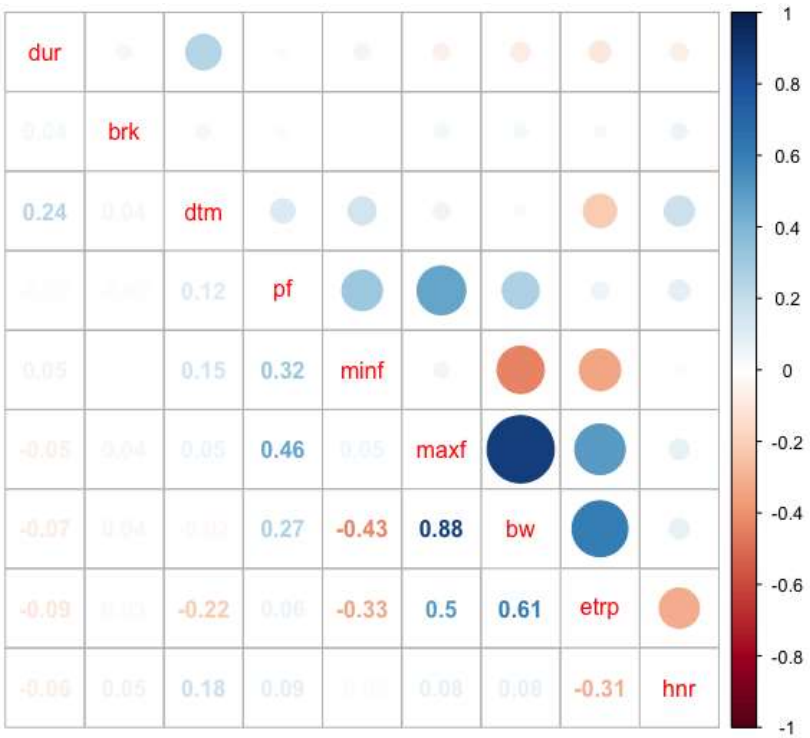

Figure 2.1 Correlation plots between 9 graded acoustic properties

dur: duration, brk: inter-note interval, dtm: distance to maximum amplitude, pf: peak frequency, minf: minimum frequency, maxf: maximum frequency, bw: bandwidth, etrp: entropy, hnr: harmonic-to-noise ratio. The blue color indicates positive and the red color indicates negative relationships with variables. The size of the circle also indicates the intensity of correlation between variables.

Table 2.4 Linear mixed model results for 9 graded acoustic properties

Response variables	Independent variables	NumDF	DenDF	F value	p-value
Duration	Context	2	40.199	1.926	0.1589
Inter-note Interval	Context	2	38.141	0.2549	0.7763
DTM	Context	2	37.833	0.0313	0.9692
Peak frequency	Context	2	40.252	2.6596	0.0823
Min frequency	Context	2	41.324	0.7241	0.4908
Max frequency	Context	2	34.233	6.0317	0.0057
Bandwidth	Context	2	34.545	5.5248	0.0083
Entropy	Context	2	42.447	2.4107	0.1019
HNR	Context	2	46.121	0.0946	0.9099

Table 2.5 Post-hoc test results of the linear mixed models

Variable	Contrast	Estimate	SE	Z ratio	p-value
Max freq	Owl – Food	-180	53.1	-3.39	0.0007
	Owl – Mask	-129.1	59.2	-2.181	0.0292
	Food – Mask	50.9	59.4	0.856	0.3919
Bandwidth	Owl – Food	-201	64.7	-3.105	0.0019
	Owl – Mask	-175.4	71.9	-2.441	0.0147
	Food – Mask	25.7	72.3	0.355	0.7226

Table 2.6 Post-hoc test results of generalized linear mixed model

Contrast	Estimate	SE	T ratio	p-value
Owl - Mask	-0.618	0.262	-2.363	0.0572
Mask - Food	1.135	0.263	4.316	0.0002
Owl - food	1.753	0.263	6.677	<.0001

The number of D notes also differed by all contexts (Table 2.6). The repetition rate of D notes per minute was significantly higher in the owl predator context (55.353 ± 11.433) than in the mask context (28.459 ± 5.285; p=0.0572) and also than the food context (9.145 ± 1.801; p=0.0002), also higher in the mask context than the food context (p<0.0001).

When classifying 3 contexts (food, mask and owl) by supervised machine learning algorithms, the classification accuracy ranged from 0.481 to 0.617. The random forest approach showed the highest accuracy (0.617) followed by SVM with radial kernel (0.589), whereas SVM with the linear kernel exhibited the lowest accuracy (0.481). All these values are much higher than the chance probability, 0.33. When classifying only 2 contexts by combining mask and owl under a “predation” context together, accuracy slightly increased with ranges from 0.587 to 0.680. However, based on the chance level of 0.5, this is not relatively higher than in classifying 3 contexts (Table 2.7).

According to the confusion matrix from the random forest approach (the best model) that classified 3 contexts, 78.2% of notes in the food context and 58.4% of notes in the owl predator context were correctly classified. However, only 29.7% of notes in the mask context, was correctly classified (Table 2.8).

Besides classification accuracy, the random forest approach assesses which variables are important for the classification accuracy based on the Gini index. In the analyses of my recorded calls, the peak frequency, followed by harmonic to noise ratio and duration, were the top three important variables that distinguished contexts, whereas bandwidth and maximum frequency were the least important variables (Figure 2.2).

Based on the variable importance plot, I performed feature selection by excluding less important variables such as minimum frequency, inter-note interval, maximum frequency and bandwidth and the classification accuracy decreased to 0.589 from 0.636 (original random forest with all features included).

Table 2.7 Machine learning algorithms results for context classification

# of classes	Machine learning algorithms	Accuracy	p-value [Acc > NIR]
3 contexts	KNN	0.578	2.729e-10
3 contexts	Random forest	0.636	< 2.2e-16
3 contexts	SVM_L	0.526	0.001734
3 contexts	SVM_R	0.610	< 2.2e-16
3 contexts	LDA	0.508	0.00335
3 contexts	QDA	0.554	1.591e-06
2 contexts	KNN	0.651	3.308e-16
2 contexts	Random forest	0.710	<2e-16
2 contexts	SVM_L	0.583	8.143e-05
2 contexts	SVM_R	0.663	<2e-16
2 contexts	LDA	0.580	0.0001357
2 contexts	QDA	0.625	6.099e-11

Table 2.8 Confusion matrix of random forest (best model)

Predicted class		Actual class		
		Food	Owl	Mask
	Food	363	139	66
	Owl	81	211	36
	Mask	20	11	43

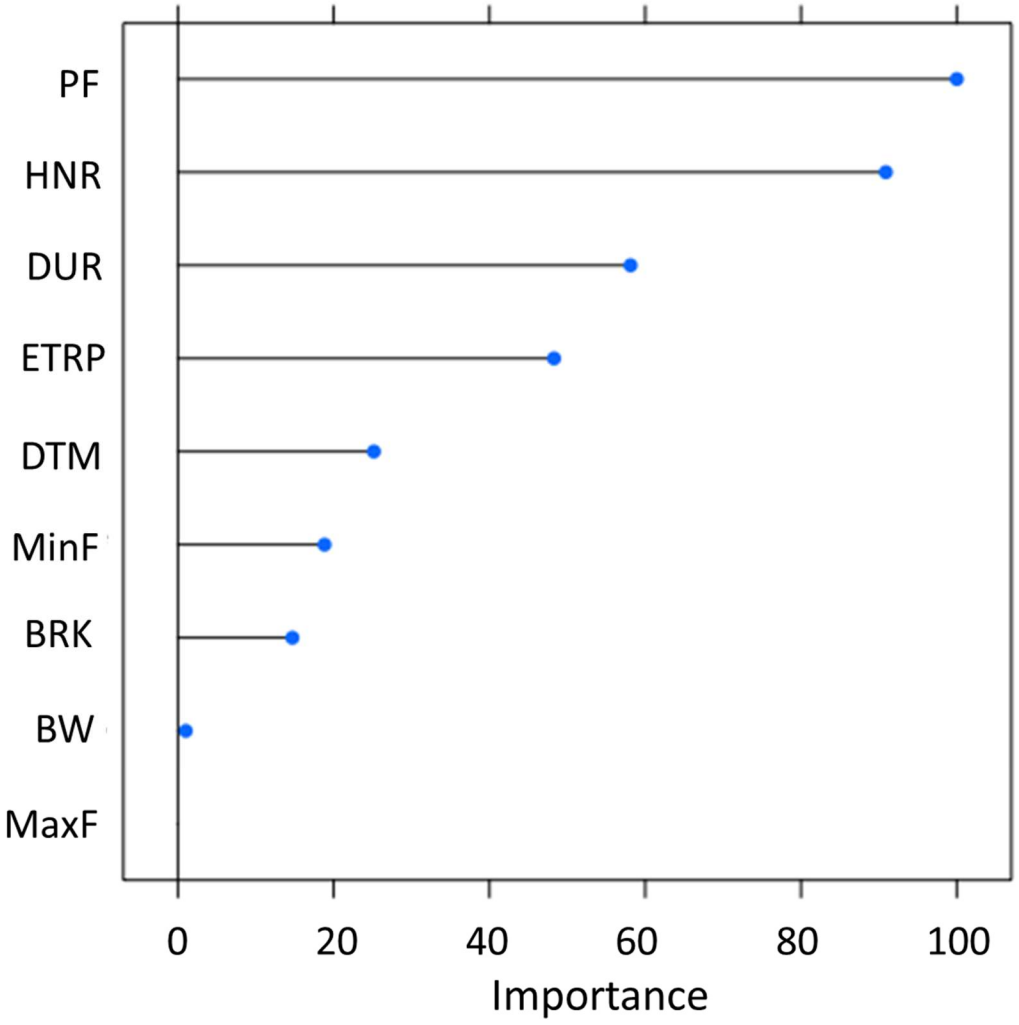

Figure 2.2 Variable importance plot in random forest classification

PF: Peak frequency, MinF: Minimum frequency, MaxF: Maximum frequency, BW: Bandwidth, ETRP: Entropy, HNR: Harmonic to noise ratio, DUR: Duration, BRK: Inter note-interval, DTM: Distance to the maximum amplitude

Discussion

This study supports the “Both” Hypothesis: Carolina chickadees varied the fine-acoustic properties, as well as the repetition rate of D notes across food and two predator contexts. This study provides the first evidence that Carolina chickadees produce D notes with different maximum frequency and bandwidth across food and predator contexts. D notes have been compared between predator contexts (Soard & Ritchison, 2009; Templeton et al., 2005) but not between food and predator contexts aside of Kalb et al. (2019). D notes produced in food and predation contexts are hard to distinguish by human auditory perception and by visual perception when the acoustic signals are represented in spectrogram form. D notes seem to function in similar ways in that both attract conspecifics and heterospecifics to the location of the signaler – mobbing calls attract more individuals to join mobbing and food calls also attract more birds to join in foraging (Mahurin & Freeberg, 2009). Receivers likely recruit to signalers producing many D notes in their calls and the context (low signaler arousal and food stimuli vs. high signaler arousal and predator stimuli) together convey meaning to those receivers (Smith, 1965). Still, there are subtle acoustic structure differences lying within D notes.

In the current study, spectral features such as maximum frequency and bandwidth differed across contexts. Similarly, Superb fairy wrens (*Malurus splendens*) also changed maximum frequency according to predation threat level (Fallow & Magrath, 2010) and chickadees including Black-capped chickadees (*Poecile atricapillus*) and Mexican chickadees (*Poecile sclateri*) produced introductory notes that are higher in frequency (Ficken, 1990; Ficken & Witkin, 1977) as predator threat increased.

According to Phillips et al., 2020, both maximum frequency and bandwidth were associated with sound propagation. In urban areas where sound tends to attenuate and reverberate more than in the rural areas, white-crowned sparrows (*Zonotrichia leucophrys*) produced songs lower in maximum frequency and bandwidth to counter sound attenuation and reverberation to make the sound

travel more stably. In this study, the maximum frequency and bandwidth both decreased in the predation context compared to the food context, which suggests chickadees produced calls to propagate with less attenuation in a predation context that might require the urgent attention and recruitment of receivers.

On the other hand, temporal aspects such as durations of notes and inter-call (or inter-note) intervals did not significantly change, unlike in previous studies (Black -capped chickadees: Templeton et al., 2005; Carolina chickadees *Poecile carolinensis*: Soard & Ritchison, 2009; Great tits *Parus major*: Kalb et al., 2019; Meerkat *Suricata suricatta*: Manser, 2001; Yellow-bellied Marmots *Marmota flaviventris*: Blumstein 2007). This difference might come from different contextual and methodological set-ups in the studies; in this study, I tried to distinguish 2 predator contexts and a food context, while other existing studies tried to distinguish only predator contexts.

The difference might also come from simple species differences; Carolina chickadees in this study did not vary any temporal measurements whereas great tits varied duration and inter-note interval in the same context comparison (food and predation contexts; Kalb et al., 2019). Perhaps the context associated with varied temporal features might be different between Carolina chickadee and great tits. For example, Carolina chickadees produced D notes with different durations in the context of social interaction (Kelley, 1988) but great tits produced D notes with different durations and inter-note intervals between food and predation contexts (Kalb et al., 2019).

Besides the graded properties of D notes, the number of D notes also appeared differently in the 3 contexts. The repetition rate of D notes was significantly higher in the predation contexts, as also shown in previous studies where species increased the number of a signal unit in their calls (Campbell's monkey *Cercopithecus campbelli*: Lemasson et al., 2010; Chaffinch *Fringilla coelebs*: Randler & Förschler, 2011; Chickadees: Templeton et al., 2005; Great tits: Kalb et al., 2019; Meerkat: Manser, 2001; Superb fairy-wrens and white-browed scrubwrens *Sericornis frontalis*: Fallow & Magrath, 2010; Krams et al.,

2012; Titmice: Courter & Ritchison, 2010). But the repetition rates in the mask and owl predator contexts were marginally different. This suggests the birds might not differentiate much between same predator contexts with varying level of threat (mask and owl) because both can be perceived as a survival threat whereas food is not. Although chickadees varied their fine-graded acoustic structure of D notes for the different contexts, they also adjusted the number of D notes, which might increase effectiveness in information transmission to the receiver. Perhaps they might need to rely more on the auditory channel for communication, considering their visual channel is limited in dense forest areas. This also goes in line with the fact that Egyptian fruit bats (*Rousettus aegyptiacus*) that live in dark caves where the visual channel is almost completely blocked, varied graded acoustic properties in their calls in multiple social interaction contexts (Prat et al., 2016).

According to the random forest model, peak frequency and the harmonic-to-noise-ratio variables appeared as the top two important variables, whereas bandwidth and maximum frequency appeared as the least important variables for classifying 3 contexts. Perhaps it is because maximum frequency appeared to have low Inter-observer reliability between measurements from two different software programs (Cool Edit Pro and Avisoft). However, using random forest with feature selection showed poorer classification performance than the original random forest that included all features, which implies that even the acoustic features that were shown as less important in the variance importance plot may nonetheless be important to the message and perhaps the meaning of these notes and calls.

On the contrary to the random forest results, linear mixed models that assessed each acoustic property showed maximum frequency and bandwidth to be the ones that differed across contexts. The difference perhaps comes from the operation logic between the two approaches. The random forest model uses the Gini index to calculate each variable's importance by measuring variance for each class. The linear mixed model does not measure such variance in a class

and it does not perform classification. Perhaps this difference in the algorithm is the major cause of the different results. Another possible explanation would be that linear mixed model considers a random effect, and accounts for the clustered nature of our data, whereas machine learning algorithms do not and considers all observations as independent.

In conclusion, Carolina chickadees produced D notes that varied in terms of graded acoustic properties and repetition rate in two predator contexts and one foraging context. Although Kalb et al. (2019) suggested that variation in the graded acoustic properties in the D notes in great tits may imply that receivers would be able to recognize the subtle difference, with graded D notes alone, it might be costly for receivers to correctly perceive such subtle variation in D notes elicited in the two opposite contexts. However, with the aid of difference in the repetition rate, it might be easier for receivers to decode the important information. In the next chapter, I performed a playback study to examine whether birds actually respond differently according to the variation of both graded and discrete properties of D notes.

CHAPTER II
RECEIVERS' PERSPECTIVE: RESPONSE TO THE PLAYBACK ELICITED FROM FOOD AND PREDATOR CONTEXTS

Abstract

To understand animal communication fully, analyses of behavior of both signalers and receivers are required. In the previous chapter, Carolina chickadees produced D notes that varied in both graded acoustic properties and repetition rate across food and predator contexts. Since the first chapter only covered the signaler's perspective, it is essential to assess whether receivers respond accordingly to that signal variation. In this study, I tested whether responses of chickadees varied across different playback treatments near feeders that manipulated both graded acoustic properties and repetition rates of D notes. Surprisingly, no evidence was found that receivers responded differently to the playback treatments. There are several possible, and non-mutually exclusive, reasons discussed: 1) sample size, 2) receivers not attending to the signal variation, 3) receivers assessing predation risk with direct rather than indirect information, and 4) receivers were already using the feeder before playbacks started. The most likely case is that the birds rather relied on actual information that they could directly assess (no actual predators near feeders) rather than paying attention to the indirect information from other individuals (call playbacks). For better understanding how chickadees perceive predation threat and signal variation, further studies are necessary.

Introduction

There are many ways to define what "animal communication" is, but one common component to all definitions is the presence of two parties: senders and receivers (Batteau, 1968; Bradbury & Vehrencamp, 2011; Burghardt, 1970; Freeberg et al., 2017; Hailman, 1977; Owren et al., 2010). The sender is the one that produces signals (or cues) that vary depending in part on its own characteristics and/or on the social and physical environmental factors it experiences, while the receiver is the one that perceives the variation in signals and changes its behaviors accordingly (Bradbury & Vehrencamp, 2011). For this

reason, studies that assess only the signaler's perspective seem to be incomplete because they do not connect receiver's responses. To assess receiver responses experimentally it is importance to use playback studies.

In previous studies that included both a sender's signal variation and conspecific receiver's response to it, receivers tended to vary their responses according to the variation in signals. For example, in predation contexts, receivers tended to behave in ways that were suitable for escaping from immediate danger (black-capped chickadees *Poecile atricapillus*: Templeton et al., 2005; great tits *Parus major minor*: Suzuki, 2014; vervet monkeys *Chlorocebus pygerythrus*: Seyfarth et al., 1980; white-tailed ptarmigan *Lagopus leucura*: Ausmus & Clarke, 2014). In food-associated situations, Carolina chickadees (*Poecile carolinensis*) visited feeders faster to calls similar in note composition to those produced by the first chickadee arriving to a new food source (Mahurin & Freeberg, 2009). In that study, chickadees that first detected the food source tended to produce more D notes in their calls, which in turn made other individuals arrive at the feeder with decreased latency. However, some studies in meerkats indicate that receivers do not always respond differently to playbacks of signals that vary reliably in different contexts (Schibler & Manser, 2007; Townsend et al., 2010).

In the previous chapter, we found that Carolina chickadees produced D notes with varying graded properties and call rates in both food and predation contexts. They produced a greater number of D notes in owl and mask (predator) contexts than in a food context. Slightly more D notes were produced in the owl than in the mask context, but the difference was more drastic between the two predator contexts and the food context. Chickadees also produced D notes with variation in their graded properties (maximum frequency and bandwidth) across owl, mask and predator contexts. According to the linear mixed model results, the difference between the owl and food contexts appeared to be the most prominent.

It would be beneficial for their survival if receivers could detect variation in signals and respond accordingly (reviewed in Smith, 2017). For example, related to the chickadee study described in the previous chapter, if a receiver misinterprets a signal produced in a predation context with one produced in a food context, it seems more likely that the receiver would be killed by predators nearby. On the contrary, if a receiver misinterprets a signal produced in food context with one produced in a predation context, using a “better safe than sorry strategy” (Haftorn, 2000), the receiver loses access to food resources, and may spend physical energy unnecessarily to stay overly alert. Or, if a receiver does not detect any variation in the two signal variants, and so does not produce a different response, that receiver also loses an opportunity to make use of information that could potentially increase its survival value. Although it seems likely that chickadees would produce different responses according to signals produced in different contexts, actual playback experiments are required to know whether receivers do vary their responses.

In previous studies, playback experiments that varied either call rate, call type, or graded differences were conducted (Ausmus & Clarke, 2014; Mahurin & Freeberg, 2009; Seyfarth et al., 1980; Suzuki, 2014; Templeton et al., 2005; Wilson & Mennill, 2011). Carolina chickadees were tested to determine whether calls elicited in food detection would attract other individuals by playing back different number of D notes in a call (Mahurin & Freeberg, 2009). Black-capped chickadees were tested by assessing whether the number of D notes in a call or the duty cycle of calls matters to receivers (Wilson & Mennil, 2011). Templeton et al. (2005) played back calls elicited from less-threatening predators and more threatening predators to test whether receivers responded differently. Similarly, Seyfarth (1980) and Suzuki (2014) played back calls elicited from different species of predators to test how receivers reacted. However, to the best of my knowledge, there is no single study that actively tested call rate and graded acoustic difference while controlling one another in animals other than in frogs (e.g., Gerhardt, 2008) and insects (e.g. Wagner & Basolo, 2006). Gerhardt

(2008) found that females preferred advertisement calls with a lower carrier frequency (one of graded acoustic properties) when pulse repetition rate was held at a specific rate in *Hyla versicolor*, but not in the *H. chrysoscelis*. Hence, it is potentially possible that receivers in birds would discriminate the subtle differences of graded acoustic properties when the calling rate is fixed. Also, some female field crickets (*Gryllus lineaticeps*) tended to show strong preferences for higher chirp rates regardless of chirp duration (Wagner & Basolo, 2006).

In this study, I tested whether Carolina chickadees responded uniquely and accordingly to the playback of D notes of conspecifics that were produced in food and predator stimulus contexts in Study 1, by varying both graded and call rate of D notes. I tested three main hypotheses against a null.

H0: No response

This null hypothesis posits that receivers will not respond to graded variation in D notes of signalers, nor to discrete variation (call rate of signalers). Such a finding (particularly the latter finding) would represent a very different finding than has been documented to date in these species (such as Wilson & Mennill, 2011; which varied call rate to test a recruitment function).

H1: Only graded variation

This hypothesis posits that chickadees will respond appropriately (i.e., more mobbing calls and less foraging in predation and no mobbing call response and more foraging in food contexts) to graded variation in D notes but will not respond to call rate variation. Such a finding would suggest receivers are able to perceive graded variation in notes and respond accordingly to that variation, and indicate they use referential calls according to feeding and predation contexts, with a graded signal system.

H2: Only discrete variation

This hypothesis posits that the birds will respond differently with regards to variation in call rate, but will not be sensitive to graded variation within D notes. Such a finding would suggest the birds rely heavily on repetition and sheer output of calling rates.

H3: Both graded and discrete variation

This hypothesis posits that chickadee receivers will respond differently both to graded variation and to call rate variation of D notes. Such a finding would suggest they use referential communication by using both graded and discrete (call rate) coding scheme.

Methods

According to the results of study 1, there was a larger difference in fine acoustic structure of D notes when comparing the owl predator context and food context as opposed to the two predator contexts. Among all models, the random forest method performed the best in terms of classification accuracy. Therefore, based on the classification results of the random forest and to simplify the playback design, we traced back which specific D notes were correctly classified as having been produced in the owl predator context or the food context, ignoring the mask context.

To decide which call type to playback, I examined which call type that includes D note was the most common. In the owl context, the [D] call type, (which represents a call comprising D notes only) was the most frequent call type. Among 351 calls, 47 calls were [D]. The most frequent length of the [D] call type was 5 (8 calls) followed by 6 (7 calls). In the food context, [D] was also the most frequent call type. Among 402 calls, 89 calls were [D]. The most frequent length of the [D] call type was 6 (16 calls). To remain consistent across all playback variants, I chose to use 5D note calls for playbacks.

To determine the typical call rate of D notes in the food and owl predator contexts, I calculated the average number of D notes/min in each context, which was 55 ± 11.43 in the owl predator context and 9.14 ± 1.80 in the food context. To simplify, I set the call rate for the owl predator playback context to be 50/min and for the food context to be 10/min.

To determine inter-note intervals between calls for my playback stimuli, I selected 82 calls that only contained 5 D notes and found the mean and standard deviation was 0.046 and 0.009, respectively. Therefore I used a 45 msec inter-note interval to make playback files. To synthesize 5D calls with only D notes that were correctly classified, I selected 50 correctly classified and good quality D notes (25 D notes for owl, 25 D notes for food, Figure 3.1).

In 5D calls, not all D notes are from the same original calls because some D notes in the same call were correctly classified while the others were incorrectly classified, or it was unknown whether it was correctly or incorrectly classified. Since the ratio of training data and testing data was 80% and 20%, respectively, only 20% of the entire D notes were known as to whether or not these were correctly classified. I ran the random forest for two times with different permutations in splitting the training and test data to get two different prediction results in order to get more correctly classified results. In Cool Edit Pro, the following 4 types of playback unit were synthesized by copying and pasting the 5D notes (either from owl or food) to repeat (either at high or low calling rate):

1) [PP] D notes that were classified as owl predator with a repeating frequency of owl predator (high repetition rate),
2) [PF] D notes that were classified as owl predator with a repeating frequency of food context (low repetition rate),
3) [FP] D notes that were classified as food context with a repeating frequency of owl predator (high repetition rate), and
4) [FF] D notes that were classified as food context with a repeating frequency of food context (low repetition rate).

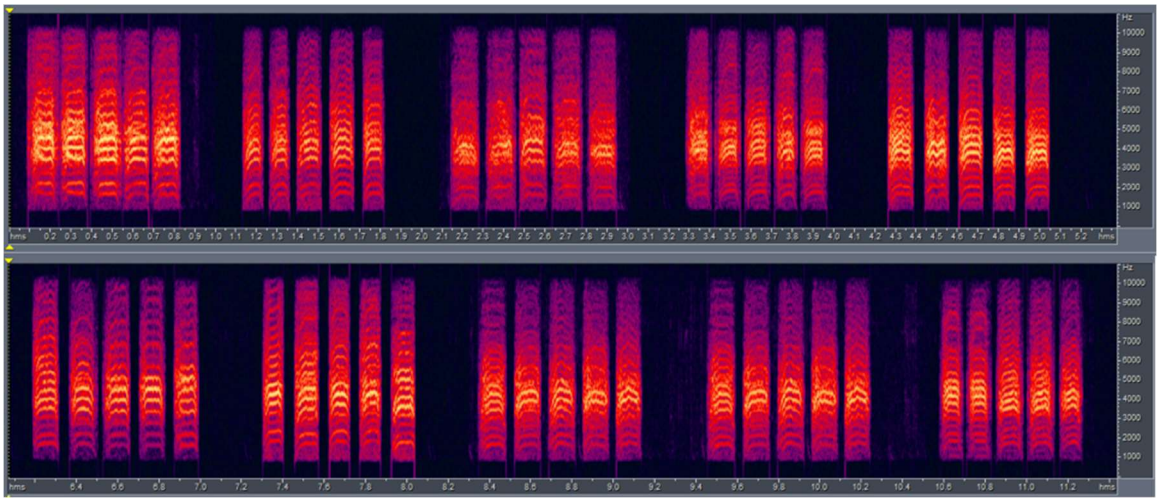

Figure 3.1 Different variety of synthesized 5D note units

All D notes in upper 5 calls were composed of D notes that were classified as owl and all D notes in the lower 5 calls were classified as food context.

Each stimulus was spaced for 2 min, followed by 3 min of silence. Hence, with the pre-stimulus period of 5 min and the 20 min of playback stimulus periods, the entire playback duration for one experiment at a site was 25 min.

Data collection

I visited 24 sites in UTFRREC and NDSP between 9:00 am to 3:30 pm (EST) from late December 2019 to early March 2020. To attract wild chickadee flocks, I stocked around 100g of black-oil sunflower seeds on a feeder (wood board (25cm * 40cm* 2cm) topped on a metal pole that was partially buried underground. Once birds were found taking seeds on the feeder, I set up all experimental devices. I placed a microphone (Sennheiser ME-64 electret microphone) connected with a digital recorder (Marantz PMD-660 recorders with 16-bit and 44.1 kHz sampling rate) under the feeder, and hung a speaker (Saul Mineroff Electronics SME-AFS portable field speaker) that was connected to a laptop PC (Asus) with an electronic cable on to a tree with rubber ties that was approximately 5m away from the feeder. A video camera (Canon VIXIA HF R80) was placed approximately 7m away from the feeder. We set the amplitude of the playback as 75dB (measured with Quest Technology 2,100 SPL meter in dBA from 1-meter distance from the speaker) right before the playback. I started each playback file with 5 min of silence as a pre-stimulus, followed by the 4 different stimulus types ([PP], [PF], [FP] and [FF]) that were ordered randomly. During the entire playback period including pre-stimulus period, I counted the number of birds for each species with binoculars (Nikon Sporter EX 8 × 42).

Data analyses

For call analyses, Cool Edit Pro was used to count the number of I (Introductory), C, H and D notes (Figure 3.2).

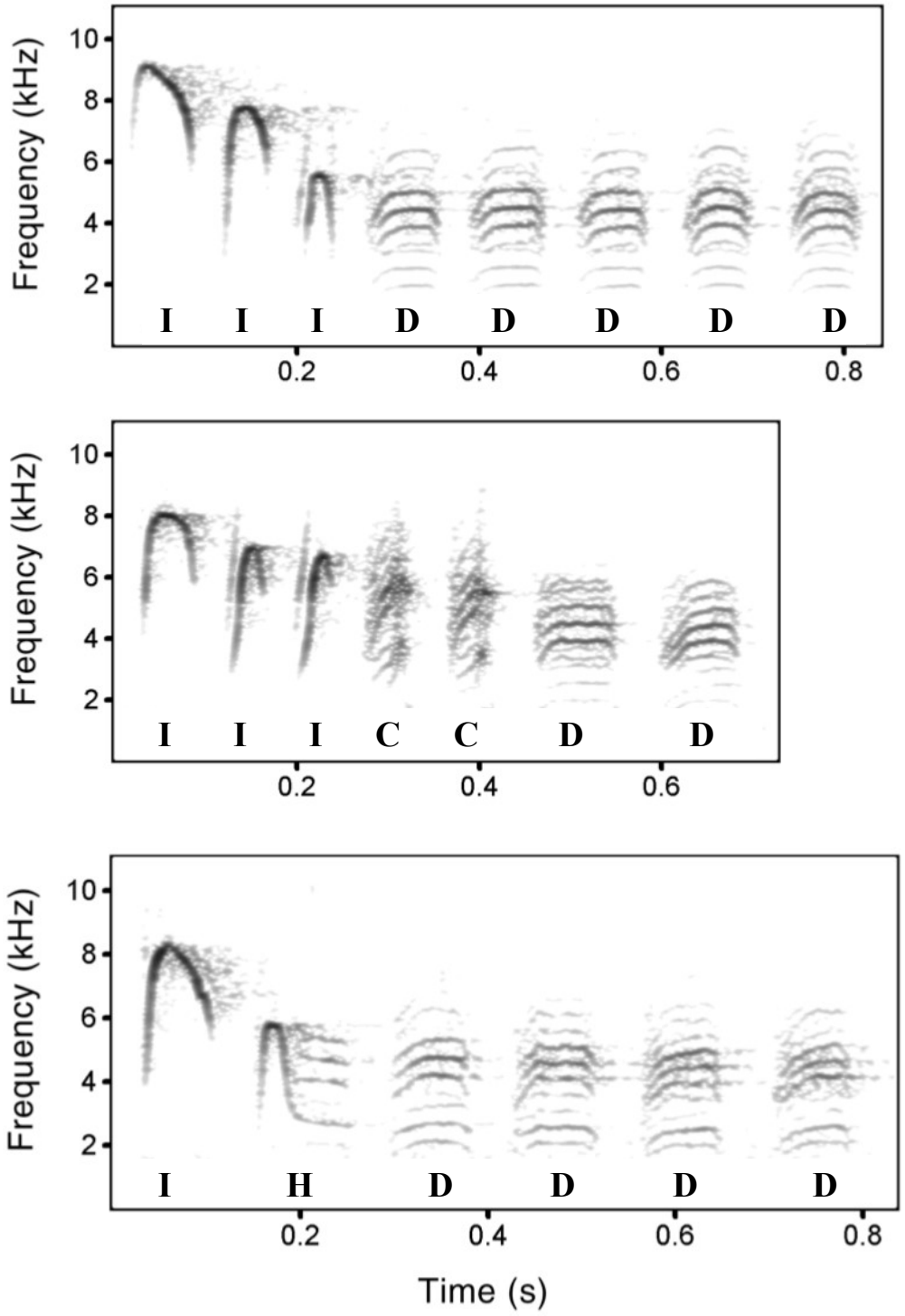

Figure 3.2 Spectrogram of I, C, H and D notes

(Adapted from Freeberg, 2008)

For seed taking behavior, I watched the video files and counted the number of successful feeder visits (i.e. the number of feeder visit where chickadees pecked and took at least 1 seed from the feeder, regardless of the number of pecking trials per visit).

For both call and feeder visit analyses, I treated the pre-stimulus period rates of behavior as a covariate. For call analyses, I first tried to perform a linear mixed model, but it did not satisfy normality assumption of the residuals, so I instead chose a generalized linear mixed model (negative binomial distribution). The independent variables were treatment (4 levels: FF, FP, PF and PP), the number of calls in the pre-stimulus period (log-transformed), the number of chickadees within 5m from the feeder, and the interaction between the treatment and the number of chickadees. Response variables were the number of I, C, H and D notes. Site was regarded as the random effect. To consider a case where there is no treatment effect and chickadees just respond to the order of the treatment, I ran another set of generalized linear model that included order (4 levels: first, second, third and fourth treatment, regardless of the type of playback), instead of treatment.

For seed taking behavior analyses, I used the linear mixed model, which satisfied the residual normality assumption. I also ran two sets of models as I did for call analyses. In the first model, independent variables were treatment, the number of chickadees and the interaction between two variables, and the number of feeder visits in the pre-stimulus period. The response variable was the number of successful feeder visits and the random effect was the site. In the second model, treatment was replaced with the order of the treatment, holding all other variables the same in the first model.

There were a total of 120 treatments (24 sites * 5 treatments per site). I generated 30 random numbers to select which audio and video files to be assessed for Inter-Observer Reliability analyses for both call and seed-taking behavior, respectively. I ran a Pearson correlation test between data coded by HJ and the subset of data independently coded by TF for the number of I, C, H

and D notes and the number of seeds taken. Reliability was high for behavior coding in this study: the correlation coefficients for all note types and the number of feeder visits ranged from 0.754 to 0.948 (median 0.935). To adjust p-values for multiple comparisons, Benjamini & Hochberg's (1995) approach was used.

Results

D notes were produced most frequently (16.033 ± 1.957 per 25 min of recording), followed by I notes (11.992 ± 2.797), but the maximum number of I notes was 283, about 3 times exceeding than that of D notes (115). None of the treatment or order effects appeared to be significant (Table 3.1, Table 3.2, Figure 3.3, Figure 3.4, Figure 3.5, Figure 3.6). In both sets of models that analyzed a treatment effect and the order of treatment effect, the number of feeder visits in the pre-stimulus period was highly predictive of the number of feeder visits during the stimulus period.

Discussion

There was no evidence that chickadees responded differently to different playback treatments or to the order of treatments. They maintained similar levels of note type production and feeder visits across all treatments.

Overall, the results are different from previous playback studies where receivers discriminated the signals of senders that were produced in functionally different social or physical environmental contexts. However, there are some studies that show meerkats (*Suricata suricatta*) did not respond differently to signals that varied with signaler identity (Schibler & Manser 2007; Townsend et al., 2010). The reasons for such mixed results may vary across studies and species, but here I suggest four possible explanations why the birds did not behave significantly differently with the playback treatment that included different calling rate and graded properties.

Table 3.1 Mixed model results for calling and seed taking behavior

CACH indicates the number of Carolina chickadees around the feeder, Pre_I, Pre_C, Pre_H, Pre_D and Pre_Seed indicate the number of I notes, C notes, H notes, D notes and Seed taken in the pre-stimulus period.

Response variables	Independent variables	Chisq	Df	p-value	Adjusted p-value
# of I notes	(Intercept)	0.0279	1	0.867	0.997
	Treatment	1.1732	3	0.759	0.997
	CACH	0.8642	1	0.353	0.889
	log(Pre_I)	0.5229	1	0.470	0.889
	Treatment:CACH	2.0877	3	0.554	0.923
# of C notes	(Intercept)	0.0714	1	0.789	0.997
	Treatment	0.1815	3	0.981	0.997
	CACH	0.0053	1	0.942	0.997
	log(Pre_C)	2.3074	1	0.129	0.645
	Treatment:CACH	0.1037	3	0.991	0.997
# of H notes	(Intercept)	0.0233	1	0.879	0.997
	Treatment	0.2398	3	0.971	0.997
	CACH	0.7111	1	0.399	0.889
	log(Pre_H)	6.181	1	0.013	0.108
	Treatment:CACH	0.045	3	0.997	0.997
# of D notes	(Intercept)	3.3368	1	0.068	0.425
	Treatment	2.3766	3	0.498	0.889
	CACH	0.4791	1	0.489	0.889
	log(Pre_D)	6.261	1	0.012	0.108
	Treatment:CACH	2.5268	3	0.470	0.889
# of seeds	(Intercept)	0.0579	1	0.810	0.997
	Treatment	3.868	3	0.276	0.889
	CACH	0.8826	1	0.348	0.889
	log(Pre_Seed)	23.1827	1	0.000	< 0.001
	Treatment:CACH	4.3207	3	0.229	0.889

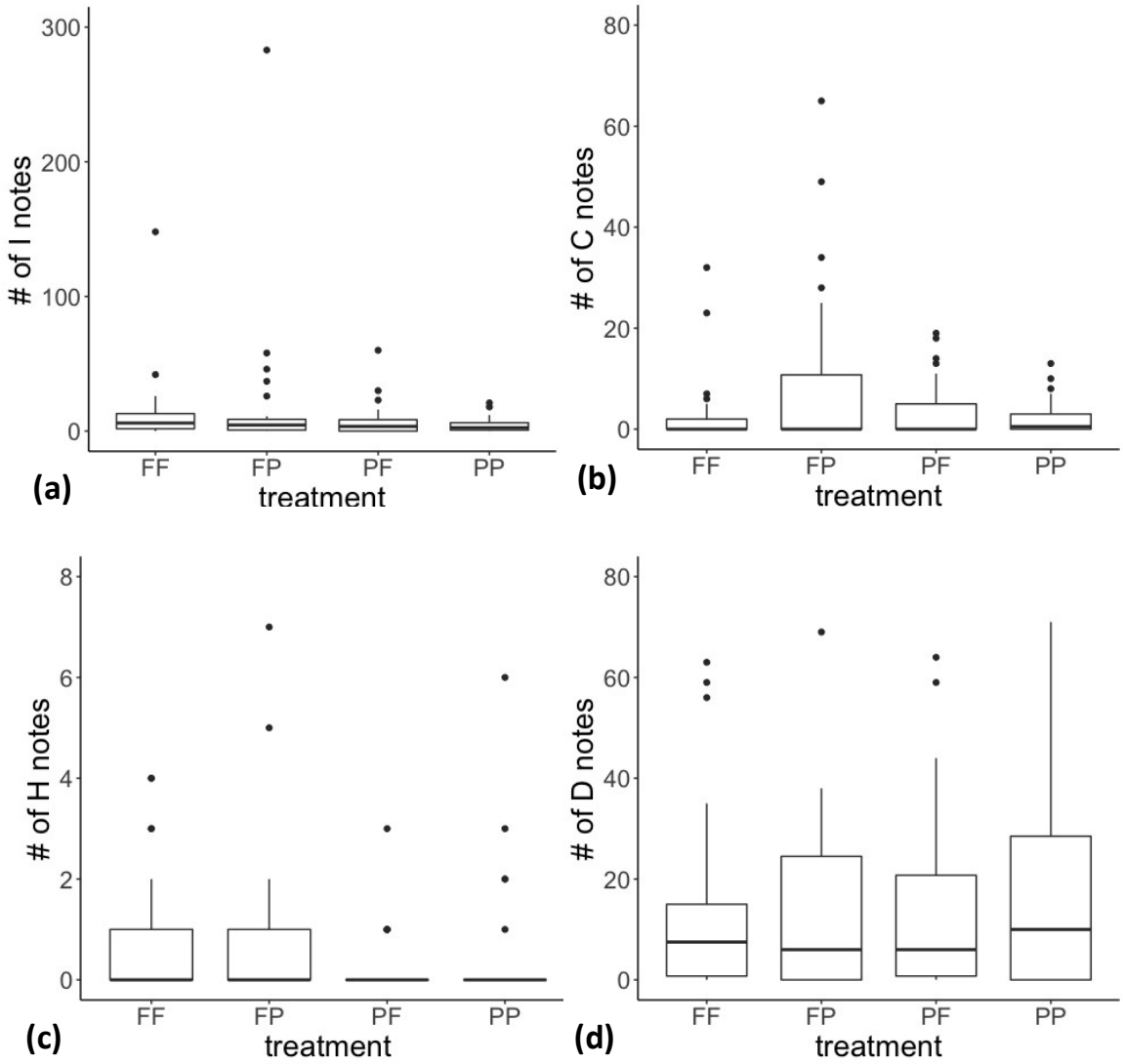

Figure 3.3 The number of each note type produced across treatments

PP: D notes classified as owl predator with a repeating frequency of owl predator (high repetition rate), PF: D notes classified as owl predator with a repeating frequency of food context (low repetition rate), FP: D notes classified as food context with a repeating frequency of owl predator (high repetition rate), FF: D notes classified as food context with a repeating frequency of food context (low repetition rate).

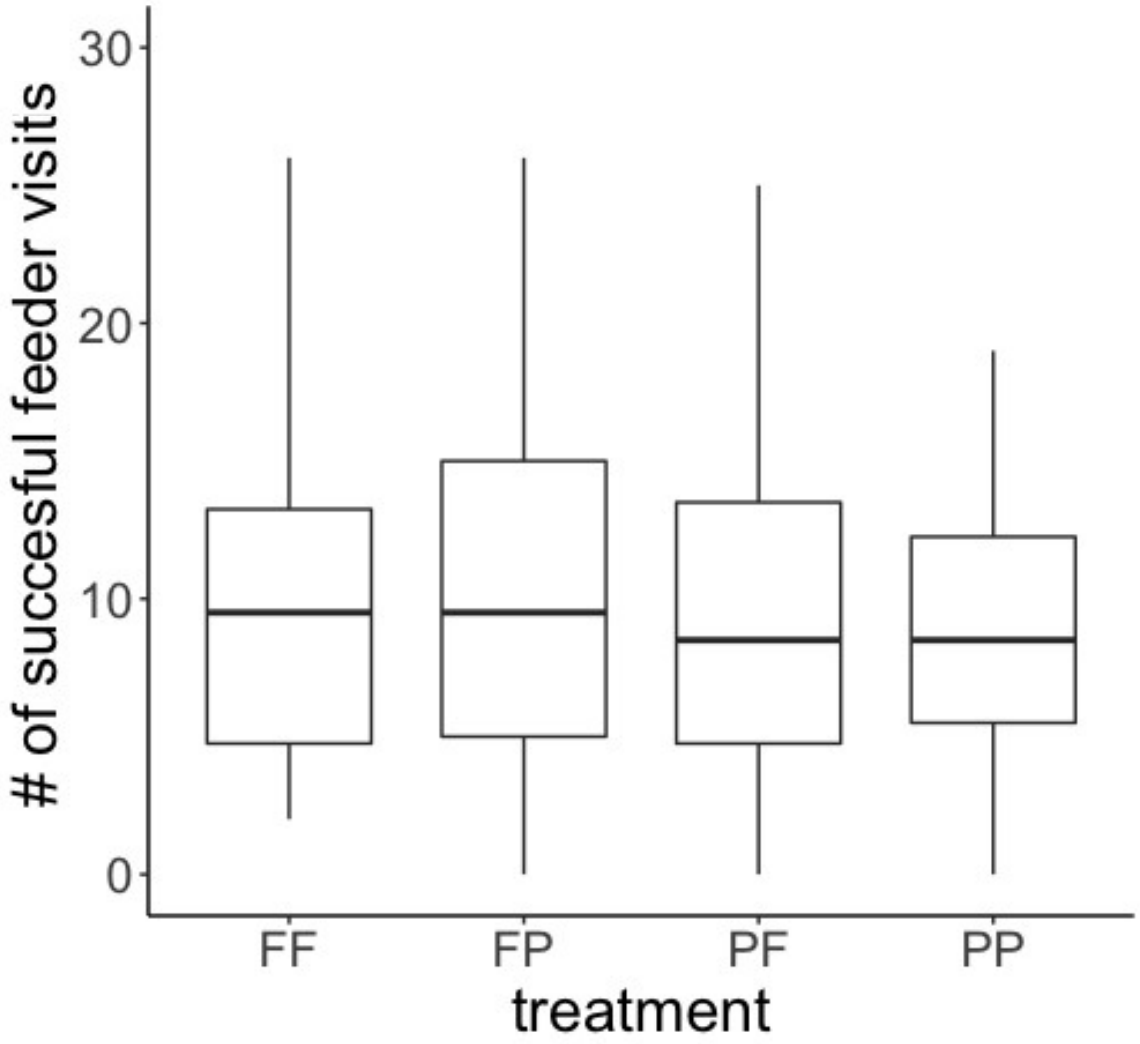

Figure 3.4 The number of successful feeder visits across treatments

Table 3.2 Mixed model results based on treatment order

Response variables	Independent variables	Chisq	DF	P-value	Adjusted p-value
# of I notes	(Intercept)	0.4465	1	0.504	0.7969
	Order	3.5498	3	0.314	0.7182
	CACH	0.183	1	0.669	0.7969
	log(Pre_I)	0.5322	1	0.466	0.7969
	Order:CACH	3.9029	3	0.272	0.7182
# of C notes	(Intercept)	0.2483	1	0.618	0.7969
	Order	11.6922	3	0.009	0.0625
	CACH	0.1183	1	0.731	0.7969
	log(Pre_C)	7.231	1	0.007	0.0625
	Order:CACH	9.6489	3	0.022	0.0917
# of H notes	(Intercept)	0.111	1	0.739	0.7969
	Order	1.443	3	0.695	0.7969
	CACH	1.0044	1	0.316	0.7182
	log(Pre_H)	6.5795	1	0.010	0.0625
	Order:CACH	1.1492	3	0.765	0.7969
# of D notes	(Intercept)	2.9015	1	0.089	0.3179
	Order	1.5692	3	0.666	0.7969
	CACH	0.1672	1	0.683	0.7969
	log(Pre_D)	5.3871	1	0.020	0.0917
	Order:CACH	1.1795	3	0.758	0.7969
# of seeds	(Intercept)	0.0023	1	0.961	0.9610
	Order	3.1489	3	0.369	0.7688
	CACH	0.3496	1	0.554	0.7969
	log(Pre_Seed)	23.6618	1	<0.001	<0.001
	Order:CACH	4.3971	3	0.222	0.6938

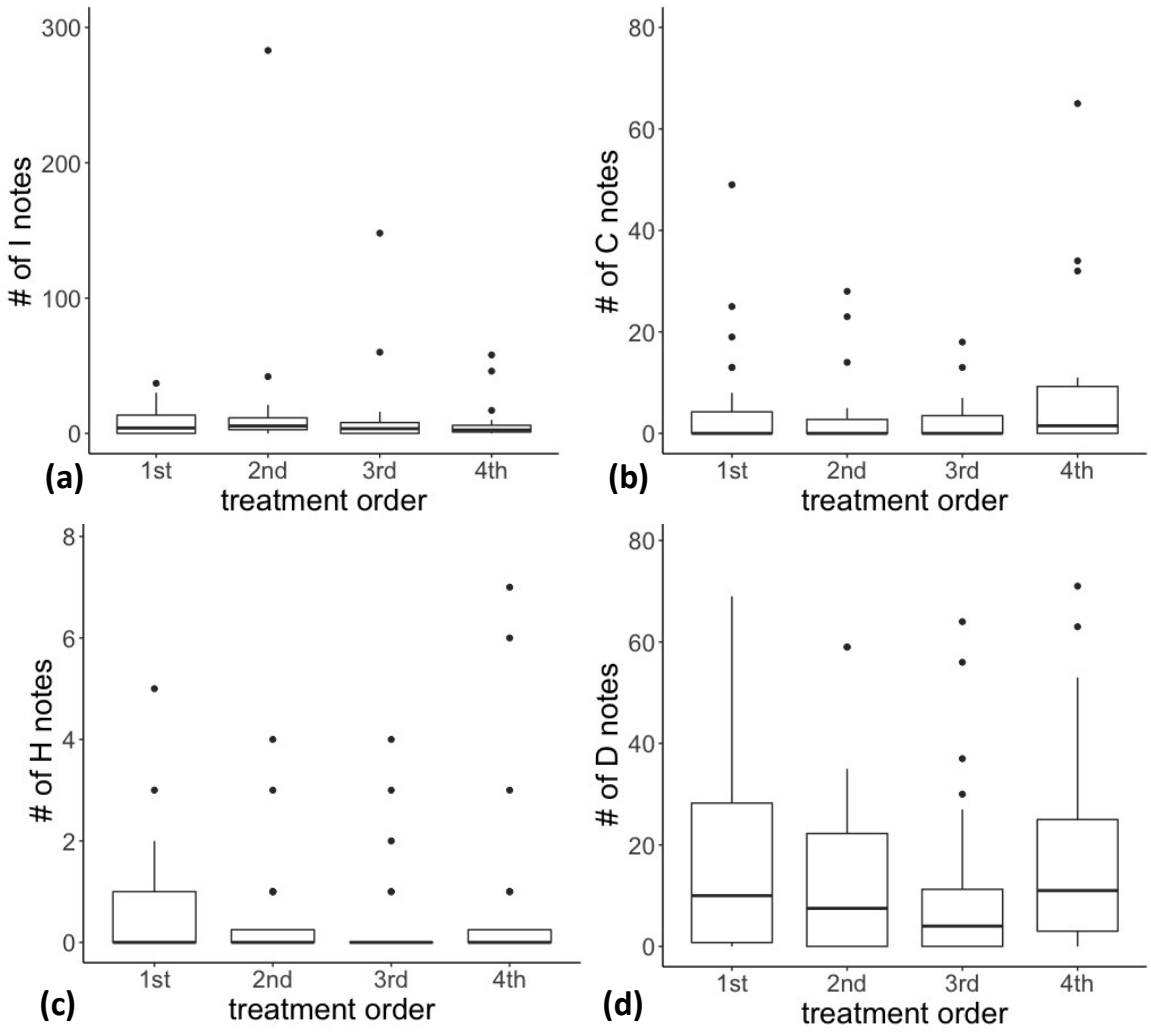

Figure 3.5 The number of each note type produced across treatment order

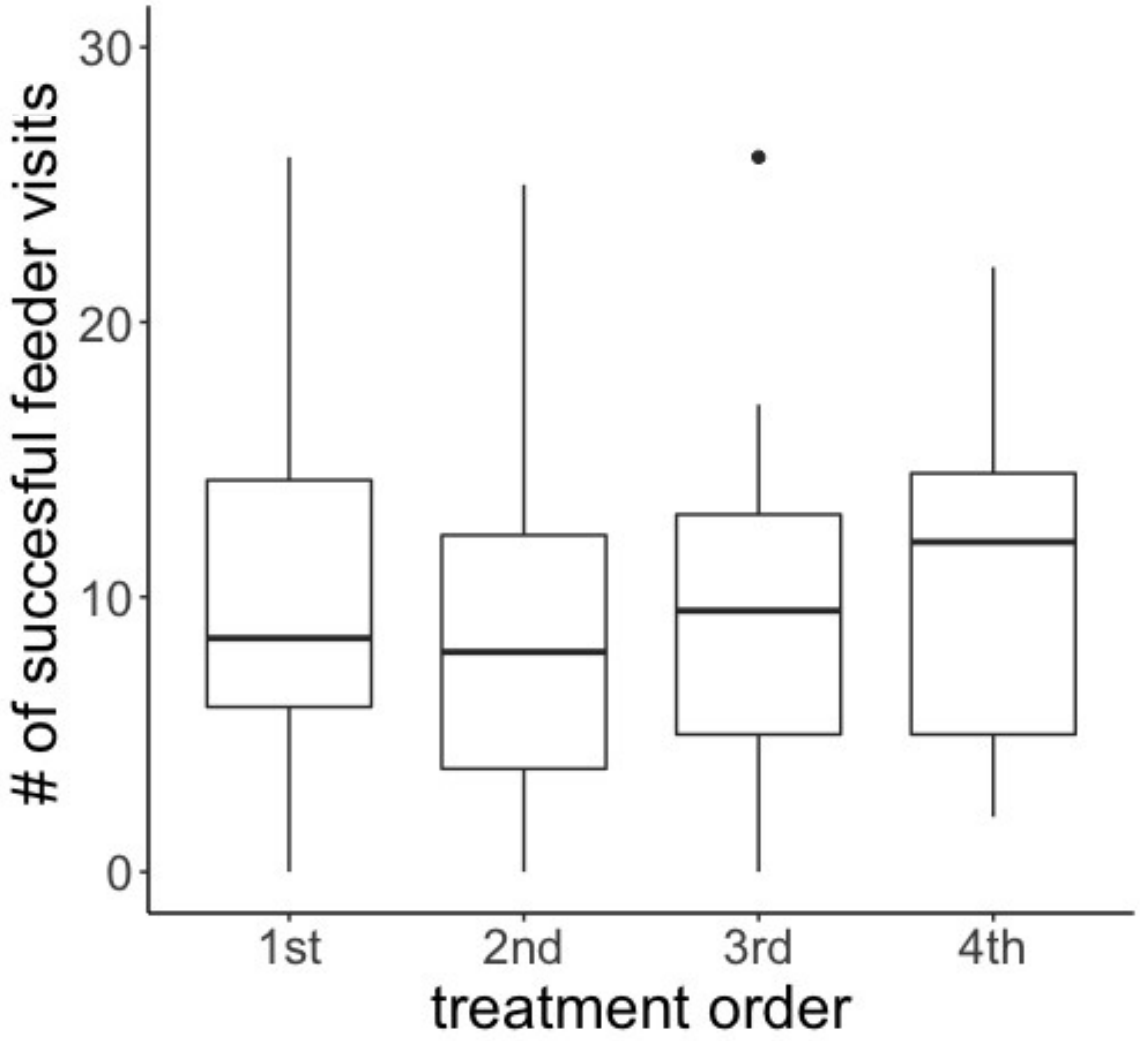

Figure 3.6 The number of successful feeder visits across treatment orders

The simplest factor to consider is the sample size. In this study, a sample size of 24 might have not been sufficient to detect behavioral differences. Meerkats did not respond differently to the different playback treatments in which the sample size was only 6 (Townsend et al., 2010). Compared to this small sample size, 24 seems to be large, but still it does not guarantee that it would be enough to detect statistical differences. However, considering the fact that Carolina chickadees in another study, in which sample size was only 12 (a half of sample size in this study), changed seed-taking behaviors according to different calls (Mahurin & Freeberg, 2009), it seems less likely that the sample size in this study was too insufficient.

A second possible explanation is that the chickadees did not vary their response to the playback types because all of these included calls with the same number of D notes (5 D notes) per call. In a previous study, Carolina chickadees discriminated calls with a greater number of D notes or with fewer D notes by varying latency to visit the feeder (Mahurin & Freeberg, 2009). Since the study did not control the total number of D notes in the playback, it is hard to know if the chickadees responded to the number of D notes per call or the total number of D notes in an entire playback period. However, there is still a chance that Carolina chickadees pay more attention to the number of D notes per call. If this is the case, it contradicts a finding in Black-capped chickadees in which birds responded to the entire number of D notes in a playback (duty cycle) rather than the fine call structure (the number of D notes per call; Wilson & Mennill, 2011).

The third possible reason was that chickadees naturally did not pay attention to the difference in the calling rate and graded properties across predators and food contexts, although signalers vary the calling rate and graded properties according to the differences in context. This does not match with the previous research in Black-capped chickadees, in which birds responded more intensively when they were exposed to playbacks that included a higher number of D notes in total (Landsborough et al. 2020; Wilson & Mennill, 2011). It is understandable that chickadees did not respond differently to the graded

variation in the D notes since graded variation is more subtle than the repetition rate of the D notes. However, it is surprising that chickadees in this study did not even differentiate their response to the variation in the total number of D notes in the playback. Future studies are required to test if this lack of response appears consistently – perhaps variation in D note number and fine acoustic structure primarily serves a recruitment function (Mahurin & Freeberg, 2009), which was not tested in this playback study.

The fourth possible explanation is that Carolina chickadees might have paid more attention to the uncertainty related to predator presence. Risk assessment is a crucial part of anti-predator behavior for prey species (Lima & Dill, 1990) and under-responding or over-responding to the predator threat is costly (Krams, 2001). Some studies found that anti-predator behavior intensity is reduced when there is more uncertainty in the predator presence. Great tits (*Parus major*) discriminated direct (visual presence of a predator; certainty in the predator presence) and indirect (only hearing conspecific mobbing without actual predators; more uncertainty) predation threat (Lind et al., 2005). A majority of individuals did not show anti-predator behavior when they only heard conspecific mobbing calls without actually seeing predators, compared to when they actually saw a predator model. Also, Red-breasted nuthatches (*Sitta canadensis*) responded more strongly to direct predator calls compared to heterospecific mobbing calls (Carlson et al., 2020).

My results, in which chickadees did not reduce the intensity of their calling or foraging responses due to indirect information sources, are atypical because in the majority of the preceding studies mentioned, prey species showed anti-predator behavior even in absence of direct (or, certain) information about predator presence. For example, there are several bird (reviewed in Smith, 2017) and mammal species including primates, Gunnison's prairie dogs (*Cynomys gunnisoni*), and meerkats (reviewed in Townsend & Manser, 2013) that use referential call systems in which animals respond to the conspecific alarm calls without direct cues of predator presence. Black-capped chickadees also changed

their responses to conspecific call playbacks that were elicited from predators with varying level of threat (Templeton et al., 2005). According to Soard and Ritchison (2009), Carolina chickadees showed stronger responses in terms of the number of individuals to approach and their approach distance to a speaker that played back *chick-a-dee* calls elicited by large predators (of less-threat, containing smaller number of D notes) and small predators (of greater-threat, containing larger number of D notes).

This indicates Carolina chickadees relied on indirect acoustic information in their study. Perhaps the chickadees from the eastern TN population in my study relied more on direct information. I suggest that acoustic stimuli might not always be sufficient to elicit anti-predatory responses from prey animals in comparison to visual stimuli of predators, as shown in Insular tammar wallabies (*Macropus eugenii*), which did not respond to auditory cues from predators while they did respond to visual predator cues (Blumstein et al., 2000). Likewise, black-capped chickadees showed more intensive anti-predator behavior to threatening visual stimuli (predator mounts) than to acoustic stimuli associated with predator contexts (conspecific mobbing calls; Arteaga-Torres et al., 2020). Further research would be beneficial to investigate to which extent Carolina chickadees discriminate visual and auditory predator stimuli, and what factors drive variation in predator stimulus sensitivity in different populations.

A last possible explanation is that different experimental settings from previous studies might have resulted in unexpected results. In this study, chickadees were already around the feeders even before playback, whereas earlier playback experiments were performed when birds were not around the feeder (Mahurin & Freeberg 2009; Wilson & Mennill, 2011). If calls in playback in this study function to recruit conspecific individuals, then birds that were already in close distance would not be motivated to further change their behavior.

In conclusion, Carolina chickadees did not vary their responses to treatment or to playback order in this study. The most likely possible reason is that they detected no visual or auditory cues about predator presence and relied

more on direct, personal information rather than the indirect and more uncertain information of the synthesized conspecific call playbacks. Further research is required to better understand how Carolina chickadees perceive predation risk and acoustic properties variation, and how it differs from other species from various taxa. To deal with the lack of response in Carolina chickadees, I suggest studying playback responses towards D note variation associated with food and predation contexts in the tufted titmice. Titmice also produce similar D notes although they tend not to produce D notes as much as chickadees do.

CHAPTER III
ACOUSTIC SIGNAL VARIATION ACROSS NON-CONTEXTUAL FACTORS: GEOGRAPHIC LOCATION, FLOCKS, AND CALL STRUCTURE

Abstract

Acoustic properties of vocal signals may vary with sender or signal properties like social groups and call structure. Individuals in the same group tend to produce vocal signals that are similar with other group members, which may facilitate social bonds. This vocal convergence is commonly found in mammals and birds. On the contrary, redundancy in animal vocalization has not been widely studied. Although it was previously found that some graded acoustic properties in Carolina chickadee note types were associated with call structure, it is still unknown that whether call length can be predicted by the graded acoustic properties of D note in the call regardless of positions. Carolina chickadee flocks were fairly accurately classified in both food and predation contexts, despite low chance level agreement, which implies vocal convergence in the chickadees. Also, the length of [D] calls was significantly associated with entropy and harmonic-to-noise ratio of D notes, which suggests the evidence of redundancy in the call that may help birds to recognize calls easily if parts of the calls were masked by noise.

Introduction

Vocal convergence is a phenomenon where the acoustic properties of individual vocalizations become more similar when the individuals are in the same social group, and is widely observed in diverse mammalian and avian species (reviewed in Tyack, 2008). For example, there are some observational studies that found vocal parameters in a group to appear differently by groups (Banded mongoose *Mungos mungo*: Jansen et al., 2012; Black-capped chickadees *Poecile atricapillus*: Mammen & Nowicki, 1981; Carolina chickadee: Freeberg et al., 2003; Chimpanzee *Pan troglodytes*: Crockford et al., 2004; Greater spear-nosed bats *Phyllostomus hastatus*: Boughman, 1997; Meerkat *Suricata suricatta*: Townsend et al., 2010). Besides observational studies, some experimental studies also show the evidence of vocal convergence. Female

Budgerigars (*Melopsittacus undulatus*) that were initially unfamiliar and had distinct contact calls developed similar contact calls after they were caged together (Hile & Striedter, 2000). Black-capped chickadees are known to possess flock-distinctive graded acoustic characteristics, but when they are caged together for a certain period of time, vocal parameters of different individuals became more similar with other birds in the cage (Mammen & Nowicki, 1981; Nowicki, 1989).

In Meerkats, contact calls of 10 groups were relatively accurately classified by group using various measurements in frequency, duration, and the spectral energy distribution and even after controlling for dominance status and sex within the group (Townsend, 2010). Analysis of frequency, duration and spectral energy distributions in Banded mongoose calls revealed that there is a specific position (the noisy part of a call) that varied across groups (Jansen, 2012). Amplitude modulation and frequency modulation of D notes were found to differ among two local populations of Carolina chickadees (Freeberg et al., 2003). Frequency and temporal measurements in calls of spear-nosed bats also differed by groups (Boughman, 1997). Similarly, frequency and temporal measurements in Chimpanzee's pant hoot calls were found to differ by community (Crockford et al., 2004)

Despite the evidence for microgeographic variation within certain call types, it is still unknown whether there would be an interaction with situational contexts (for example feeding and predator threat) or whether geographic variation would be maintained even in the different contexts (but see Jansen et al., 2012). Also, in chickadees, it appeared that only few flocks or local populations (2 in Freeberg et al., (2003), 5 in Mammen and Nowicki (1981)) were examined to assess possible differences among groups that were separated by about 5km and <1.5km (Freeberg et al., 2003; Mammen & Nowicki, 1981). Studies with a larger number of independent flocks (after Bartmess-Levasseur et al., 2010) are needed to make stronger claims about microgeographic variation.

Graded structural variation associated with variation in note or syllable composition of calls has gotten relatively less attention so far. There are only two studies that actually addressed this topic and revealed the graded acoustic properties of a signal unit vary with the structure of a call that contains the signal unit in Carolina chickadee (Freeberg et al., 2003) and in baboons (*Papio cynocephalus ursinus*; Fischer et al., 2004).

However, it is still unknown whether call length, which might imply the intensity of the motivation of the signaler (Hailman, 2008) can be predicted with graded acoustic properties regardless of D note positions in a call. Therefore, there is a need to assess whether redundancy occurs in the calls with different length controlling the call type.

In this chapter, I tested signal and signaler-related variation of graded acoustic properties of D notes in calls produced in food and predator contexts. I tested whether there was within-note variation in D notes according to natural conditions such as flock membership and note composition of calls, using the data from study 1. I tested three alternative hypotheses against a null.

H0: No graded variation

This null hypothesis posits that there is no within-note variation with flock membership and call structures. If this hypothesis holds true, it would indicate that graded properties in D notes are not associated with flock or study location, nor with the note composition of calls.

H1: Call structure

This hypothesis posits that within-note variation only exists across different call structures / note compositions. For example, I tested whether the call length (the number of D notes) of the call type containing only D notes can be predicted by graded acoustic properties. This hypothesis predicts that within-note variation in D notes would change with call structure, showing the similar

kind of redundancy found in a different population of Carolina chickadees (Freeberg et al., 2003).

H2: Flock membership

This alternative hypothesis posits that within-note variation exists across different flock identities such as variation in a) flocks and/or b) separated locations that include multiple flocks.

H3: Both call structure and flock membership

This alternative hypothesis combines H1 and H2. Within-note variation can be observed across both call structure and flock membership.

Methods

Data collection

I used all the same audio data recorded from study 1.

Data analyses

- **Within-note variation in local population**

For all classification analyses, 6 Supervised machine learning approaches including KNN, random forest, SVM with radial kernel, SVM with linear kernel, LDA and QDA were used.

The predictor variables were the 9 acoustic properties measured within D notes (Table 2.1), and the dependent variable was location with 2 levels (UTFRREC and NDSP), contexts (food, mask and owl contexts) and flocks (11 levels in food-only, 5 levels in owl-only and 18 levels when including all contexts). When I split the data into flock level, the sample size got too reduced at some flocks to include them. I only included flocks in which I had obtained more than 100 D notes (Table 4.1, Table 4.2 and Table 4.3).

Table 4.1 11 Flocks and the number of D notes within the food context

(N: NDSP, U: UTFRREC)

Locations	Flocks	# of D notes
N	**WAR**	**295**
N	**WCP**	**217**
N	**WCK**	**168**
U	**J**	**167**
N	**ECB**	**155**
U	**AR2**	**146**
U	**N**	**129**
N	**WBR**	**123**
U	**C**	**121**
U	**K**	**118**
N	**WDP**	**106**
N	ECP	89
U	UV1	74
N	ESM	69
U	UV2	68
U	O	68
U	AR1	38
U	UV3	36
N	ET	32
U	I	28
U	F	26
N	W1	16

Table 4.2 Flocks and the number of D notes within in owl context

Locations	Flocks	# of D notes
N	**WAR**	**303**
N	**ET**	**206**
U	**O**	**125**
N	**ECB**	**108**
N	**WCP**	**101**
N	EDM	98
U	UV2	95
U	F	88
U	C	84
N	WCK	82
N	WXC	78
U	AR2	77
U	AR1	56
U	N	47
U	UV1	41
N	WDP	36
U	K	35
N	WBR	31
U	I	25
N	WCB	22
N	W1	21
U	UV3	17
U	J	13
N	ESM	11

Table 4.3 Flocks and the number of D notes within in mask context

Locations	Flocks	# of D notes
N	**WAR**	**656**
N	**WCP**	**353**
N	**ET**	**335**
N	**WCK**	**331**
N	**ECB**	**305**
U	**AR2**	**299**
U	**K**	**289**
U	**O**	**220**
U	**J**	**214**
U	**C**	**210**
U	**N**	**176**
U	**UV2**	**163**
N	**WBR**	**154**
U	**F**	**150**
N	**WDP**	**142**
U	**UV1**	**141**
U	**AR1**	**128**
N	**ESM**	**108**
N	EDM	98
N	ECP	89
N	WXC	88
U	I	68
U	UV3	68
N	W1	44
N	WCB	22

Since the overall sample size was not large compared to the other sets of machine learning methods, I divided the training and test data in to 60% and 40% respectively (which was 80% and 20% in all other cases).

Such supervised algorithms were used to classify 2 locations (UTFRREC and NDSP) in a) food context only, b) mask context only, c) owl context only and d) all contexts altogether without controlling contexts, and to classify flocks within locations.

Significance of classification accuracy was measured in comparison with the baseline (NIR: no information rate) which is expressed as “ACC>NIR” in the results (Zohair, 2019). Variable importance plot from random forest was created to visualize which variables were important in classification in a descending order. It is based on the “Gini index” which measures probability that an observation is misclassified.

I performed a series of linear mixed models by using the lme4 package to determine whether each acoustic variable (that was used as an independent variable in the previous machine learning approach) differs by location and context and also whether two-way interactions between location and contexts exist.

There are two main assumptions to run linear mixed models which are normal distribution of the residuals and the homogeneity of variance. To test such assumptions, I used the Shapiro test and if the statistic exceed 0.9, then I assumed that the residuals in the model follow the normal distribution (see Erskine et al., 2012). All models passed this assumption except the one in which the response variable was peak frequency. After I log-transformed peak frequency, the residuals appeared to follow the normal distribution as the statistic of the Shapiro test exceeded 0.9.

Dependent variables for these mixed model analyses were the 9 acoustic variables (duration, inter-note interval, distant to max, peak frequency, minimum frequency, maximum frequency, bandwidth, entropy and harmonic-to-noise ratio. Independent variables were location with two levels (UTFRREC and NDSP),

contexts with three levels within (food, mask and the screech owl predator) and two-way interaction between location and the contexts. Since the repeated-measure design was used in each flock to measure the effect of contexts within each flock, I used 1|flock/context (which is (1|flock) + (1|flock:context)) as a random effect instead of (1|flock). When the model did not converge, I used either (1|flock:context) or optimizer “Nelder_Mead”.

Variable selection procedures were performed in an automated way by using the bidirectional step() function in the same lme4() package. When the step() did not work because of convergence issues, I manually performed the backward selection by gradually removing the most non-significant effect from the full models. Since the level of the contexts are three, further post-hoc tests were conducted by using emmeans() with Fisher’s LSD adjustment.

- **Within-note variation in call structure**

To predict the length of calls in the same call type, I first examined which call type was most frequent. Among 4851 D notes, the number of D notes belong to the [I][D] call type was the most common (1585) followed by the [D] call type (1060, Table 4.4; call types in a pair of brackets indicate note composition and its order consisting the call type regardless the number of repetitions. For example, call type [I][D] indicates any calls that consist of “I note” followed by “D note”, such as IIDDDD, ID, IIIID and so on). Although the [I][D] call type occupied the highest frequency, the number of I notes vary within the [I][D] calls. Hence, I chose the [D] call type to predict the length of [D] calls (Table 4.4).

I used linear mixed models to predict the length of the D calls with 9 acoustic properties (peak frequency, minimum frequency, maximum frequency, bandwidth, entropy and harmonic-to-noise ratio, measured at maximum amplitude, duration, internote-interval between D notes and distance to the maximum amplitude), location, context and interaction between location and context.

Tabel 4.4 The number of each call type

Call types	Count
ID	**1585**
D	**1060**
ICD	595
IHD	582
HD	363
CD	110
ICHD	9
CHD	4
HID	2
Overlap	541

Results

Within-note variation comparing two local populations

Table 4.5 shows all results from 6 different machine learning approaches in predicting local population (UTFRREC vs NDSP). Overall, random forest outperformed the other models in terms of classification accuracy and the accuracy was much higher (around 80%) than the chance level (50%).

Variance important plots (Figure 4.1) show duration and peak frequency are the main variables that were most important, whereas bandwidth and inter-note interval were less important variables.

Flock classification in different contexts

Table 4.6 shows all results of 6 machine learning methods that predicted flocks in different contexts and Table 4.7, 4.8, and 4.9 show the confusion matrices obtained from random forest algorithm. Overall, the prediction accuracy was highest using random forest and classification accuracy was much higher than chance level.

The variable importance plot from random forest appeared very similar in all three cases and indicated duration and peak frequency were the two most important variables for classification (Figure 4.2).

Differences in acoustic properties across locations and contexts

Table 4.10 and Table 4.11 represent full models and reduced models that predict all 9 acoustic properties, respectively. In all cases, there was no significant main effect of location and no significant two-way interaction between location and context. Among 9 linear mixed models, 2 models appeared to include the main effect of contexts. Table 4.12 presents the results of the post-hoc analyses.

Table 4.5 Machine learning algorithms results for classifying locations

Comparing locations between UTFRREC and NDSP

Contexts	Machine learning algorithms	# of classes	Chance level	Accuracy	p-Value [Acc>NIR]
Food	KNN	2	0.5	0.753	<2E-16
Food	RF	2	0.5	0.794	<2E-16
Food	SVM L	2	0.5	0.659	4.81E-06
Food	SVM R	2	0.5	0.755	<2E-16
Food	LDA	2	0.5	0.652	1.77E-05
Food	QDA	2	0.5	0.707	2.43E-11
Mask	KNN	2	0.5	0.757	4.32E-09
Mask	RF	2	0.5	0.796	4.44E-12
Mask	SVM L	2	0.5	0.553	0.28523
Mask	SVM R	2	0.5	0.803	1.23E-12
Mask	LDA	2	0.5	0.566	0.1859
Mask	QDA	2	0.5	0.697	1.30E-05
Owl	KNN	2	0.5	0.794	1.43E-14
Owl	RF	2	0.5	0.803	8.17E-16
Owl	SVM L	2	0.5	0.733	2.46E-07
Owl	SVM R	2	0.5	0.786	2.16E-13
Owl	LDA	2	0.5	0.731	4.44E-07
Owl	QDA	2	0.5	0.756	1.33E-09
All contexts	KNN	2	0.5	0.709	<2E-16
All contexts	RF	2	0.5	0.768	< 2.2E-16
All contexts	SVM L	2	0.5	0.684	4.40E-13
All contexts	SVM R	2	0.5	0.734	< 2.2E-16
All contexts	LDA	2	0.5	0.675	1.91E-11
All contexts	QDA	2	0.5	0.683	7.17E-13

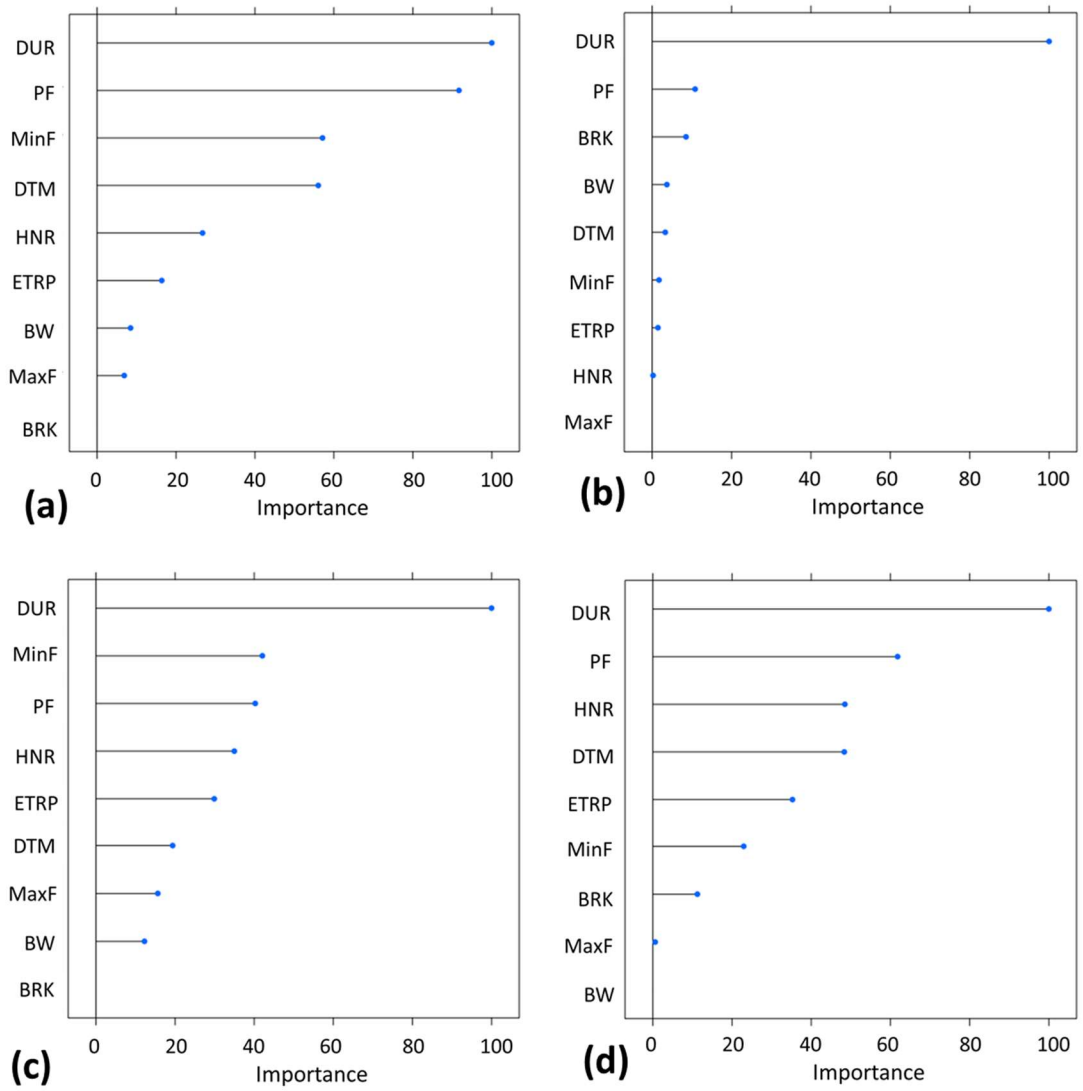

Figure 4.1 Variable importance plot from random forest that classified location

(a) in food context only (b) in mask context only (c) in owl context only and (d) without controlling contexts

PF: Peak frequency, MinF: Minimum frequency, MaxF: Maximum frequency, BW: Bandwidth, ETRP: Entropy, HNR: Harmonic to noise ratio, DUR: Duration, BRK: Inter note-interval, DTM: Distance to the maximum amplitude

Table 4.6 Machine learning algorithms results for classifying flocks

Contexts	Machine learning algorithms	# of classes	Chance level	Accuracy	P-Value [Acc>NIR]
Food	KNN	11	0.091	0.513	< 2.2e-16
Food	RF	11	0.091	0.639	< 2.2e-16
Food	SVM L	11	0.091	0.310	< 2.2e-16
Food	SVM R	11	0.091	0.295	< 2.2e-16
Food	LDA	11	0.091	0.476	< 2.2e-16
Food	QDA	11	0.091	0.473	< 2.2e-16
Owl	KNN	5	0.2	0.733	< 2.2e-16
Owl	RF	5	0.2	0.843	< 2.2e-16
Owl	SVM L	5	0.2	0.567	< 2.2e-16
Owl	SVM R	5	0.2	0.623	< 2.2e-16
Owl	LDA	5	0.2	0.736	< 2.2e-16
Owl	QDA	5	0.2	0.786	< 2.2e-16
All contexts	KNN	18	0.056	0.411	< 2.2e-16
All contexts	RF	18	0.056	0.531	< 2.2e-16
All contexts	SVM L	18	0.056	0.172	0.0717
All contexts	SVM R	18	0.056	0.221	8.89E-12
All contexts	LDA	18	0.056	0.356	< 2.2e-16
All contexts	QDA	18	0.056	0.375	< 2.2e-16

Table 4.7 Confusion matrix of 11 flocks classification in the food context

Predicted flocks	Actual flocks										
	AR2	C	ECB	J	K	N	WAR	WBR	WCP	WCK	WDP
AR2	26	0	3	3	0	4	3	0	1	4	2
C	0	37	2	2	11	2	6	0	2	1	3
ECB	6	0	21	3	0	0	2	2	1	4	1
J	5	1	9	49	0	1	3	1	1	5	1
K	3	4	1	1	32	5	2	0	0	0	1
N	5	1	3	0	1	34	1	2	3	5	0
WAR	3	1	3	7	4	1	87	6	21	2	4
WBR	2	0	4	2	1	0	0	42	0	1	0
WCP	1	6	0	0	2	1	8	0	48	2	0
WCK	3	2	5	1	4	0	2	0	6	44	6
WDP	0	0	0	0	0	1	3	1	2	1	26

Table 4.8 Confusion matrix of 5 flocks classification in the owl context

Predicted flocks	Actual flocks				
	ECB	ET	O	WAR	WCP
ECB	30	1	5	0	1
ET	1	73	6	3	0
O	3	3	37	0	1
WAR	8	7	5	110	7
WCP	0	1	0	1	34

Table 4.9 Confusion matrix of 18 flocks classification in all contexts

Predicted flocks		Actual flocks																	
		AR1	AR2	C	ECB	ESM	ET	F	J	K	N	O	UV1	UV2	WAR	WBR	WCP	WCK	WDP
	AR1	25	0	0	0	0	0	1	0	0	1	1	3	1	3	0	3	0	1
	AR2	0	61	0	8	1	3	4	1	3	7	0	0	3	13	0	0	2	3
	C	1	0	42	4	0	4	5	1	10	0	4	0	2	2	0	2	1	1
	ECB	1	3	2	58	4	9	2	9	2	3	2	4	2	2	4	2	12	1
	ESM	0	2	0	1	13	0	4	0	0	1	0	1	0	0	1	0	3	1
	ET	0	6	1	10	1	84	4	5	3	8	8	5	2	5	2	1	3	3
	F	0	0	1	0	1	2	9	1	2	0	1	1	0	3	0	0	0	1
	J	1	6	0	4	0	6	0	62	1	2	3	0	2	5	2	2	4	0
	K	0	4	12	3	0	1	5	2	63	8	2	4	8	4	0	6	3	0
	N	1	4	1	0	4	3	1	0	3	29	0	2	2	2	0	1	1	0
	O	2	0	2	4	0	6	4	0	2	0	46	1	2	3	0	4	0	3
	UV1	0	1	1	0	0	1	2	0	1	1	0	9	2	3	0	1	0	0
	UV2	0	1	0	1	0	0	0	1	1	0	3	5	4	1	0	1	1	2
	WAR	7	20	0	14	4	11	4	6	6	8	9	7	14	189	13	11	15	8
	WBR	1	3	1	14	0	1	1	0	0	0	0	0	3	4	25	0	1	0
	WCP	4	6	8	0	0	0	4	0	8	0	2	7	5	26	0	100	2	3
	WCK	0	5	1	16	9	9	5	2	6	6	1	5	9	10	2	6	83	9
	WDP	0	3	0	0	1	0	0	4	0	0	1	0	1	3	0	0	1	26

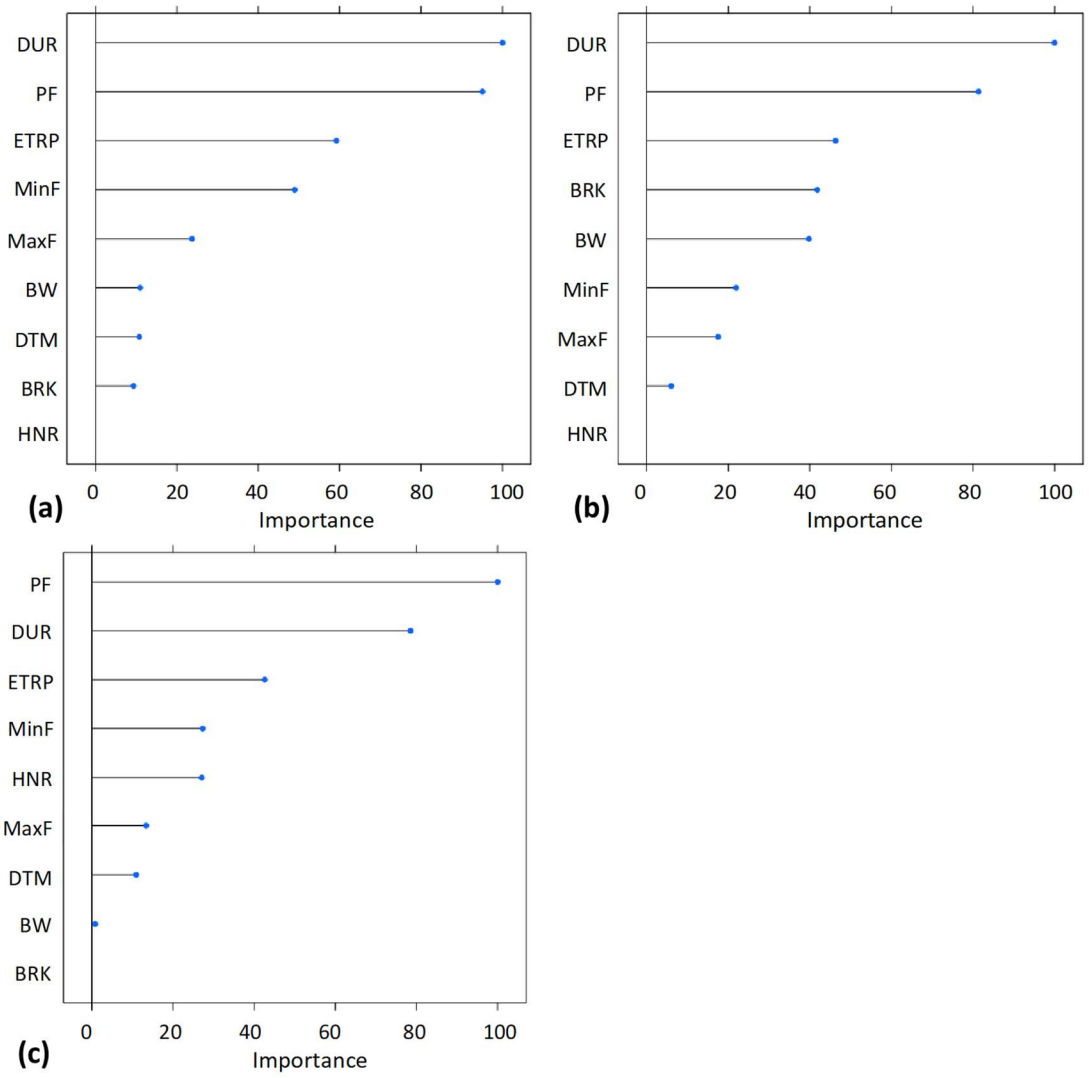

Fig 4.2 Variable importance plot from random forest that classified flocks

(a) 11 flock in food context only (b) 5 flocks in owl context only and (c) 18 flocks in all contexts

Maximum frequency differed across contexts and was higher in the food context (5247.118 ± 12.711) than in the owl context (5030.123 ± 11.704; p=0.0007), and higher in the owl context than in the mask context (5078.848 ± 20.332; p=0.0292). Bandwidth appeared smaller in the owl context (1371.322 ± 13.843) than in the food context (1582.729 ± 13.896; p=0.0019) and in the mask context (1507.163 ± 22.224; p=0.0147).

Structural variation

Table 4.13, and Table 4.14 represent the full model and the reduced model, respectively. There were no significant main effects of location or context, and no interaction between location and context. The reduced model included entropy and harmonic-to-noise ratio. Call length significantly decreased as entropy and harmonic-to-noise ratio increased.

Discussion

Graded acoustic properties of D notes varied across location and flock, which supported hypotheses 3. Among 6 machine learning methods, random forest tended to outperform the other approaches in terms of classification accuracy in all cases and it was beyond the chance level, suggesting there is a non-random pattern occurring according to location and flock. In classifying location, the classification accuracy was above chance level in each separate context and including all contexts. However, there was no significant location effects in the mixed model when both location and context were included as independent variables to predict variation in the graded acoustic properties. This might imply that chickadee calling is associated more by the context rather than location because context including food and predation are associated with the birds' survival value.

Table 4.10 Linear mixed models results for graded acoustic properties

Response variables	Independent variables	Num DF	Den DF	F value	p-value
Duration	Location	1	23.837	2.1260	0.1579
	Context	2	38.578	1.6542	0.2045
	Location:Context	2	38.578	0.8651	0.4290
Inter-note interval	Location	1	23.658	0.3118	0.5818
	context	2	36.243	0.1883	0.8292
	Location:context	2	36.243	1.8760	0.1678
Distomax	Location	1	22.367	0.0002	0.9897
	context	2	35.900	0.0405	0.9604
	Location:context	2	35.900	1.2591	0.2961
Log(peak frequency)	Location	1	23.517	0.3761	0.5456
	context	**2**	38.864	2.8356	0.0709
	Location:context	2	38.864	0.6407	0.5324
Minimum frequency	Location	1	23.678	0.1197	0.7324
	context	2	39.766	0.7317	0.4874
	Location:context	2	39.766	0.2842	0.7541
Maximum frequency	Location	1	22.582	0.3871	0.5401
	context	**2**	**33.276**	**6.2154**	**0.0051**
	Location:context	2	33.276	0.7845	0.4646
Bandwidth	Location	1	22.502	0.1790	0.6762
	context	**2**	**33.983**	**5.6943**	**0.0074**
	Location:context	2	33.983	0.9572	0.3941
Entropy	Location	1	23.551	0.0044	0.9475
	context	2	41.023	2.5650	0.0892
	Location:context	2	41.023	1.1608	0.3233
Harmonic-to-noise ratio	Location	1	23.893	1.5422	0.2263
	context	2	43.213	0.0981	0.9068
	Location:context	2	43.213	0.4027	0.6710

Table 4.11 Reduced models from linear mixed models

Response variables	Fixed effects	Num DF	Den DF	F value	p-value
Max Frequency	Context	2	34.233	6.032	0.00570
Bandwidth	Context	2	34.545	5.5248	0.00828

Table 4.12 Post-hoc tests results of linear mixed models

Response variables	Contrast	Estimate	SE	Z ratio	p-value
Max Frequency	Owl-Food	-180	53.1	-3.390	0.0007
	Owl-Mask	-129.1	59.2	-2.181	0.0292
	Food-Mask	50.9	59.4	0.856	0.3919
Bandwidth	Owl-Food	-201	64.7	-3.105	0.0019
	Owl-Mask	-175.4	71.9	-2.441	0.0147
	Food-Mask	25.7	72.3	0.355	0.7226

Table 4.13 Linear mixed model results for call structure

Independent variables	Estimate	SE	DF	t value	p-value
(Intercept)	9.47E+00	1.07E+00	3.49E+01	8.875	1.79E-10
Duration	-9.95E-02	1.56E-01	9.77E+02	-0.638	0.5239
Inter-note interval	1.04E-01	1.31E-01	1.02E+03	0.799	0.4247
Distance to maximum amplitude	1.10E-01	1.40E-01	1.03E+03	0.786	0.4320
Peak frequency	-3.01E-01	1.48E-01	1.03E+03	-2.029	0.0427
Minimum frequency	-9.88E+00	8.10E+00	1.03E+03	-1.219	0.2231
Maximum frequency	2.02E+01	1.63E+01	1.03E+03	1.242	0.2146
Bandwidth	-2.23E+01	1.82E+01	1.03E+03	-1.229	0.2194
Entropy	-4.25E-01	1.80E-01	1.03E+03	-2.355	0.0187
Harmonic to noise ratio	-3.27E-01	1.60E-01	1.02E+03	-2.047	0.0409
LocationUTFRREC	-1.12E+00	1.46E+00	3.28E+01	-0.767	0.4488
Contextfood	-4.87E-01	9.30E-01	1.87E+01	-0.524	0.6065
Contextmask	-1.49E+00	1.42E+00	2.34E+01	-1.049	0.3049
LocationUTFRREC*contextfood	9.26E-03	1.25E+00	1.63E+01	0.007	0.9942
LocationUTFRREC*contextmask	2.36E+00	1.78E+00	2.21E+01	1.325	0.1986

Table 4.14 Reduced model results for call structure

Independent variables	Estimate	SE	DF	t value	p-value
(Intercept)	8.7358	0.5998	21.891	14.566	9.61E-13
Entropy	-0.3613	0.1376	1039.436	-2.625	0.00878
HNR	-0.2934	0.148	1035.5904	-1.982	0.04777

Surprisingly, flocks within locations showed strong classification accuracy, especially considering low chance level (1/11 in food context, 1/5 in owl context, 1/18 in all contexts). In the food context and the predator context, and in all contexts combined, accuracy was comparable with the one from other analyses where chance level was 0.5 or 0.33. There are some possible explanations for this high accuracy in flock classification.

First of all, perhaps the flock difference was largely caused by individual differences. Mammen and Nowicki (1981) found that 3 frequency measurements, 6 temporal measurements, and the number of introductory notes all differed across individual black-capped chickadees, and the difference between individuals was clearer than between flocks because only 1 frequency measurement and 2 temporal measurements differed by flocks. In this study, I did not band individual birds so it is not clear if this is the case. However, I personally observed that 2 or more chickadees were calling in most study flocks, rather than just one bird that kept calling. This would make it less likely that individual differences alone are producing the call variation across flocks.

Second, the other possibility is vocal convergence, in which individual birds adjust their graded vocal parameters to match with the other individuals in the same flock. There are some benefits of vocal convergence in calls of various taxa of animals including mammals and birds, suggested by Tyack (2008): 1) group cohesion, 2) exclusive resource access and 3) affiliation to integrate new members. These functions are not mutually exclusive and these all seem to be useful in maintaining social bonds within group members. For highly social species, individuals are likely to stay in close proximity, recognize group members, and get access to food sources jointly (Wilkinson & Boughman, 1998). Chickadees are a highly social bird species. Especially during the fall and winter, which was my study period, they form mixed-species flocks and lead flock movement as a nuclear species (reviewed in Mostrom et al., 2020) and also signal at newly discovered food sources to flock members nearby (Mahurin & Freeberg, 2009). Mammen and Nowicki (1981) who found evidence of vocal

convergence in black-capped chickadees, suggested that vocal convergence would be beneficial for chickadees to quickly recognize flock members, especially when they are under predation threat, when they could gain more time to forage, and when birds of different flocks encounter.

Besides locations and flocks, the graded acoustic properties tended to vary with the different signal structures (length of the call). Entropy and HNR significantly predicted call length in [D] calls with different numbers of D notes. These results imply that redundancy occurs in calls that consist of D notes – fine acoustic structure of D notes can predict note position and call length. Since chickadees inhabit forests with vegetation cover that can block or distort both visual and acoustic signals and with sounds produced by a wide range of other animal species, it could be beneficial for receivers to be able to predict the entire call structure by hearing only part of the call. The number of notes in a call is associated with threat level (Soard & Ritchison, 2009) and motivation to recruit other individuals in a flock when food is discovered (Mahurin & Freeberg, 2009). Although it needs to be experimentally tested, chickadees might be able to predict predator-related threat level or a newly discovered food source by hearing only part of a call, if they can perceive graded acoustic differences produced in different contexts by signalers.

Freeberg et al. (2003) also found evidence of signal redundancy in Carolina chickadees. Graded acoustic properties of the first A and D notes predicted the call structure, including the call length. In this study, graded acoustic properties in any D notes regardless of the position in [D] call type predicts the call length, suggesting the calls of these birds might possess stronger redundancy than originally found. If graded acoustic properties in any D notes, not just in the very first D notes in calls can predict call length, it would strengthen redundancy by increasing resilience from noise masking part of calls. For example, receivers might be able to predict the call length by hearing only part of a “DDDDDD” call even if noise masks the first few notes of the call.

In this study, the common important graded acoustic properties that classified location and flock were peak frequency and duration, according to the random forest. Specifically, duration was also the common variable that varied with geographic location in black-capped chickadees (Mammen & Nowicki, 1981), and with call structure in Carolina chickadees (Freeberg et al., 2003). In social interactions, D notes with shorter duration were associated with escaping from interactions, mostly produced by females while longer duration of D notes tended to be produced by males in agonistic interactions (Kelley, 1988). All these findings together suggest that D note duration might serve multiple functions, but future research is needed to examine if receivers actually respond to this subtle variation, and whether other graded properties of non-D notes are associated with duration of D notes because it is unlikely that only duration of D notes vary with the multiple situations.

But other than this temporal measurement, frequency variables did not match with the previous studies because peak frequency was the only frequency variable for classifying geographic difference in my study, whereas fundamental frequencies, bandwidth (black-capped chickadees: Nowicki, 1989), maximum frequency (black-capped-chickadees: Mammen & Nowicki, 1981), and frequency and amplitude modulation (Carolina chickadees: Freeberg et al., 2003) were important variables that appeared differently across geographic regions.

Therefore, it seems that most variables investigated in different studies differently predicted geographic location, flocks, and call structures. Perhaps it was because of the species difference (black-capped chickadees, Mexican chickadees and Carolina chickadees), but even when compared with the previous studies on Carolina chickadees, little evidence was found that common variables predicted the geographic location and call structure (Freeberg, 2003).

In conclusion, although there are some confounding factors such as individuality in this study, there were clear differences in graded acoustic properties by flocks, location, and call structure, which imply possible group convergence and signal redundancy. Future work needs to examine whether

there are individual differences and whether receivers pay attention to graded acoustic variation from flocks and call structure.

CONCLUSION

Animal communication is a ubiquitous activity in which senders produce signals related to variation in their own internal states or behavioral tendencies, or to external situations, and these signals can make receivers change their behaviors (Bradbury & Vehrencamp, 2011). One of the main modalities for communication is sound. Animal acoustic signals vary in the form of repetition rates of the signal units and graded acoustic properties within the signal units. This variation is commonly associated with individual identity, geographic location, motivation and affective state, call structure itself, and contexts that affect survival such as feeding and predation. In particular, prey animals need to perceive predation risk while they are foraging for food (Lima, 2009). For this purpose, animals often communicate about detected predation threat and food source availability (Bradbury & Vehrencamp, 2011; Clay et al., 2009).

There are a number of studies that focused on comparing signal variation only in predation contexts and a far fewer number of studies that examined signal variation only in food contexts, and research comparing signal variation between predation and food contexts is rare (Kalb et al., 2019). In particular, graded properties in acoustic signals have rarely been investigated in comparison to signal repetition rates associated with predation and food contexts in species that produce combinatorial signals, such as some primates (Arnold & Zuberbuhler, 2008; Casar et al., 2013; Crockford & Boesch, 2005), Meerkats (*Suricata suricatta;* Collier et al., 2017), titmice (Baeolophus; Hailman, 1989; Jung & Freeberg, 2017), and chickadees (Poecile; Reviewed in Lucas & Freeberg, 2007).

In this dissertation, I questioned whether Carolina chickadees actually "communicated" (i.e. signalers produced varying signals that, in turn, made receivers change their behavior) with regards to these two quite different contexts: food and predation. I also assessed whether the signals varied across "stable" individual factors such as flock membership and geographic location that are not directly related to the contexts associated with their survivorship (food and predation). In chapter 1, I focused on the signaler's perspective. I asked

whether Carolina chickadees *produced chick-a-dee* calls that varied in terms of graded acoustic properties within D notes and/or the repetition rate of D notes in food and two predator contexts. In the following chapter 2, I focused on the receiver's perspective. I asked whether chickadees responded differently to the signals that were produced in the food and predator contexts by varying their foraging or anti-predatory calling behavior. In chapter 3, I tested whether graded acoustic properties of chickadee vocalization changed across non-contextual factors such as flock membership and geographic location.

Signal variation

Both chapters 1 and 3 showed that Carolina chickadee D notes varied across contexts that were associated with their immediate survival and with non-contextual factors related to the signaler. In Chapter 1, I found the first evidence that Carolina chickadees varied both repetition rate and graded acoustic properties of D notes, by varying spectral measurements across food and two predator contexts. Both temporal and spectral features contributed to accurate classification of the contexts via machine learning approaches. The results indicate that the birds adjusted both graded acoustic properties and repetition rate simultaneously in the production of calls in these contexts.

It is notable that chickadees varied acoustic structures including graded features and call rate in a single note type, D notes. However, this is not the only signal that chickadees produce in these calls. Other than D notes, I, C, and H notes could have been used for different contexts (e.g. producing I notes in food and D notes in predation without changing acoustic properties in D notes). Since I did not measure non-D notes and only focused on the variation of D notes, it is unknown whether chickadees vary non-D notes across different contexts. However, it is true that the Carolina chickadees varied a single signal unit (D note) for different contexts. In black-capped chickadees that are known to produce similar note types (A,B,C and D notes), D notes are the lowest in frequency and are found to travel the farthest distance (Proppe et al., 2010).

Considering that both food and predation are directly associated with birds' survival value, it might be beneficial for chickadees to produce D notes with different acoustic properties for stable information transfer to be detected better for other individuals at a distance from signalers.

In Chapter 3, the chickadees also produced D notes that varied across geographic location, flock membership, and call structure. In particular, graded acoustic properties of D notes were highly accurately classified according to flock membership by machine learning approaches. These findings suggest that vocal convergence occurred within flocks or across flocks within locations, or both. Vocal convergence is thought to facilitate social bonds among individuals within the same group (Tyack, 2008). Also, graded acoustic properties of D notes predicted call length, which can be explained by signal redundancy (Hailman, 2008). This finding implies that receivers may be able to predict the entire call length when only part of a call is perceived. Future playback studies could test whether receivers discriminate this acoustic variation related to flock, location, or call length.

Receiver responses

Although I found evidence that chickadees produced D notes that varied in terms of both graded properties and repetition rate, there was no evidence in Chapter 2 that receivers responded differently according to that variation. One of the likely reasons for the lack of receiver response variation was that Carolina chickadees in the study may have relied more on direct information (no predator cues) than on indirect information (playback treatments). There are other recent cases where individuals responded less to indirect than to direct information (great tits *Parus major*: Lind et al., 2005; red-breasted nuthatches *Sitta canadensis*: Carlson et al., 2020). Also some prey animal species tended to show weaker responses to acoustic stimuli than to visual stimuli related to risk (Arteaga-Torres et al., 2020; Blumstein et al., 2000). However, other than these

studies, most animals responded to playbacks even without direct information (reviewed in Smith, 2017; Townsend & Manser, 2012).

Future directions

If Carolina chickadees in our population actually assess predation risk more by direct information (but see Soard & Ritchison, 2009), unlike the majority of other species tested (reviewed in Smith (2017) and Townsend and Manser (2012)), and not by conspecific D notes that were elicited by predator stimuli as discussed above, this population represents a very different one in terms of sensitivity to predator stimuli. Further testing of Carolina chickadee populations that differ in predation pressure would inform our understanding of how predator selection pressure drives sensitivity to direct and indirect signals and cues related to risk. Similar species that also produce more D notes in both food and predation contexts might also be informative for this sort of playback study. For example, tufted titmice also produce a series of rapid D notes when they first detect food on feeders (my personal observation) as well as in predation contexts (Courter & Ritchison, 2010). In the wild, when tufted titmice produced D notes, I often observed other conspecifics get closer to the calling individual. This species might make it easier for researchers to detect variation in behavioral responses to potentially subtle variation in call playbacks. One limitation in studying tufted titmice, though, is that they do not tend to call as much as Carolina chickadees (my personal observation).

Also, comparing the diversity of acoustic signal patterns between species would broaden our understanding of the potential variation in production and reception of signal variation in multiple contexts. For example, the total note types in tufted titmice are fewer than Carolina chickadees (4 note types in Tufted titmice; Owens & Freeberg, 2007, 7 note types in Carolina chickadees; Freeberg, 2008), which implies the titmice might need to adjust graded acoustic properties variation in their D notes more widely than chickadees.

In this study, I was not able to consider the individual identity of each chickadee. Instead, individual chickadees in the same study site were all lumped into the level of flocks. However, there might have been individual variation in signal production as shown in black-capped chickadees (Mammen & Nowicki, 1981). Also, in previous studies of Carolina chickadees, the duration of D notes tended to be longer in males than females (Kelley, 1988). Hence, it is still unclear the extent to which call variation in food and predator contexts in this study would still be maintained after controlling for individual differences and flock composition with sex, dominance rank, and age. Also, it would be meaningful to investigate individual differences (mainly due to age and sex) in vocal characteristics elicited in the same context. Vervet monkeys (*Cercopithecus aethiops*) tended to produce predator-specific calls that were similar among individuals with different age and sex classes (Seyfarth et al., 1980). It is more likely that chickadees also produce similar graded acoustic properties and repetition rates of D notes regardless of their age and sex, because predators are an immediate threat to any individuals. Still, the role of sex, age, and dominance status on variation in calling in different contexts remains unknown until it is actively tested.

Future studies taking machine learning approaches should include techniques that enhance classification accuracy in two ways: using deep learning and using various ways of acoustic feature selection. In this study, machine learning methods were applied. Although there are plenty of previous studies already used multiple machine learning models to analyze vocalization variation (Acevedo et al., 2009; Turesson et al., 2016; Xie et al., 2018), most studies in animal communication still rely heavily on traditional statistical analyses. However, machine learning enables researchers to find important patterns and to classify various factors including individual identity and vocal signals that cannot be easily found by using classical statistics (Valleta et al., 2017). Deep learning is one area of study within the field of machine learning, but provides more powerful results with maximized prediction accuracy compared to traditional machine

learning methods, because deep learning considers non-linearity in the data (Dargan et al., 2019). The major difference between deep learning and machine learning is that, unlike in machine learning where human researchers need to explicitly select features to classify, deep learning itself learns which features tend to lead to accurate classification (Bermant et al., 2019; LeCun et al., 2015). For this reason, there are increasing numbers of bioacoustics studies using deep learning in various taxa: killer whales (*Orcinus orca*; Bergler et al., 2019), beluga whales (*Delphinapterus leucas*; Zhong et al., 2020), ultrasonic sound classification in rodents (Coffey et al., 2019), and flight calls from multiple bird species (Salamon et al., 2017).

If deep learning were combined with the acoustic feature extraction scheme that has been used in human speech recognition such as MFCC (Mel Frequency Cepstral Coefficients), it may show other important patterns in signal variation that have not been found in this study. MFCC is the most popular feature extraction technique (Asda et al., 2016; Dave, 2013) because of its good performance in acoustic classification (Asda et al., 2016; Bouchakour & Debyeche, 2018; Lee et al., 2008). The MFCC method is also commonly used in extracting features in animal vocalization in various taxa (Gradišek et al., 2017; Noda et al., 2016; Prat et al., 2016; Salamon et al., 2017; Xie et al., 2015). Since MFCC does not take into account temporal variation in signals, including MFCC measurements with temporal measurements such as duration might allow us to detect more novel patterns in animal vocalization that were previously unseen in my study that only used classical acoustic variables.

Concluding remark

This study provided the first evidence that Carolina chickadees produce D notes differently across food and predation contexts in terms of the repetition rate and the graded acoustic properties (maximum frequency and bandwidth) within D notes. Although this study failed to find evidence that receivers responded to that signal variation, it raised some possible reasons for the lack of response and

pointed to next potential steps such as tests with closely related species. Identifying individuals, experimentally controlling flock composition, and using techniques to increase deep learning classification accuracy with measurements from different feature extraction methods will help increase our understanding of this call system, one of the most structurally complicated communication systems known (Krams et al., 2012).

References

Acevedo, M. A., Corrada-Bravo, C. J., Corrada-Bravo, H., Villanueva-Rivera, L. J., & Aide, T. M. (2009). Automated classification of bird and amphibian calls using machine learning: A comparison of methods. Ecological Informatics, 4(4), 206-214.

Arnold, K., & Zuberbühler, K. (2008). Meaningful call combinations in a non-human primate. Current Biology, 18(5), R202-R203.

Arteaga-Torres, J. D., Wijmenga, J. J., & Mathot, K. J. (2020). Visual cues of predation risk outweigh acoustic cues: a field experiment in black-capped chickadees. Proceedings of the Royal Society B, 287(1936), 20202002.

Asda, T. M. H., Gunawan, T. S., Kartiwi, M., & Mansor, H. (2016). Development of Quran reciter identification system using MFCC and neural network. Indonesian Journal of Electrical Engineering and Computer Science, 1(1), 168-175.

Aubin, T., & Jouventin, P. (2002). How to vocally identify kin in a crowd: the penguin model. Advances in the Study of Behavior, 31, 243-278.

Aubin, T., Mathevon, N., Staszewski, V., & Boulinier, T. (2007). Acoustic communication in the Kittiwake Rissa tridactyla: potential cues for sexual and individual signatures in long calls. Polar Biology, 30(8), 1027-1033.

Ausmus, D. M., & Clarke, J. A. (2014). Mother knows best: functionally referential alarm calling in white-tailed ptarmigan. Animal cognition, 17(3), 671-679.

Bartmess-LeVasseur, J., Branch, C. L., Browning, S. A., Owens, J. L., & Freeberg, T. M. (2010). Predator stimuli and calling behavior of Carolina chickadees (Poecile carolinensis), tufted titmice (Baeolophus bicolor), and white-breasted nuthatches (Sitta carolinensis). Behavioral Ecology and Sociobiology, 64(7), 1187-1198.

Basolo, A. L., & Wagner Jr, W. E. (2006). Genetic Variation in Maternal Investment Patterns in Platyfish Xiphophorus maculatus. Zebrafish, 3(3), 339-345.

Batteau, D. W. 1968. The world as source; the world as sink. In Freedman, S.)., ed. The Neuropsychology of Spatially Oriented Behavior. Homewood, Il!., Dorsey Press, pp. 197·203.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal statistical society: series B (Methodological), 57(1), 289-300.

Bergler, C., Schröter, H., Cheng, R. X., Barth, V., Weber, M., Nöth, E., ... & Maier, A. (2019). ORCA-SPOT: An automatic killer whale sound detection toolkit using deep learning. Scientific reports, 9(1), 1-17.

Bermant, P. C., Bronstein, M. M., Wood, R. J., Gero, S., & Gruber, D. F. (2019). Deep machine learning techniques for the detection and classification of sperm whale bioacoustics. Scientific reports, 9(1), 1-10.

Bloomfield, L. L., Phillmore, L. S., Weisman, R. G., & Sturdy, C. B. (2005). Note types and coding in parid vocalizations. III: The chick-a-dee call of the Carolina chickadee (Poecile carolinensis). Canadian Journal of Zoology, 83(6), 820-833.

Blumstein, D. T., Daniel, J. C., Griffin, A. S., & Evans, C. S. (2000). Insular tammar wallabies (Macropus eugenii) respond to visual but not acoustic cues from predators. Behavioral Ecology, 11(5), 528-535.

Bouchakour, L., & Debyeche, M. (2018). Improving continuous Arabic speech recognition over mobile networks DSR and NSR using MFCCs features transformed.

Boughman, J. W. (1997). Greater spear-nosed bats give group-distinctive calls. Behavioral Ecology and Sociobiology, 40(1), 61-70.

Bradbury, J. W., & Vehrencamp, S. L. (2011). Principles of animal communication, 2nd edn Sunderland. MA: Sinauer Associates.

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. (2015). Segregation of information about emotional arousal and valence in horse whinnies. Scientific reports, 4, 9989.

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

Burghardt, G. M. (1970). Defining “communication.”. Communication by chemical signals, 1, 5-18.

Campbell, K. A., Hahn, A. H., Congdon, J. V., & Sturdy, C. B. (2016). An investigation of sex differences in acoustic features in black-capped chickadee (Poecile atricapillus) chick-a-dee calls. The Journal of the Acoustical Society of America, 140(3), 1598-1608.

Carlson, N. V., Greene, E., & Templeton, C. N. (2020). Nuthatches vary their alarm calls based upon the source of the eavesdropped signals. Nature communications, 11(1), 1-7.

Cäsar, C., Zuberbühler, K., Young, R. J., & Byrne, R. W. (2013). Titi monkey call sequences vary with predator location and type. Biology letters, 9(5), 20130535.

Charlton, B. D., Keating, J. L., Rengui, L., Huang, Y., & Swaisgood, R. R. (2009). Female giant panda (Ailuropoda melanoleuca) chirps advertise the caller's fertile phase. Proceedings of the Royal Society B: Biological Sciences, 277(1684), 1101-1106.

Charrier, I., Bloomfield, L. L., & Sturdy, C. B. (2004). Note types and coding in parid vocalizations. I: The chick-a-dee call of the black-capped chickadee (Poecile atricapillus). Canadian Journal of Zoology, 82(5), 769-779.

Charrier, I., Mathevon, N., & Jouventin, P. (2002). How does a fur seal mother recognize the voice of her pup? An experimental study of Arctocephalus tropicalis. Journal of Experimental Biology, 205(5), 603-612.

Clay, Z., & Zuberbühler, K. (2009). Food-associated calling sequences in bonobos. Animal Behaviour, 77(6), 1387-1396.

Clay, Z., Archbold, J., & Zuberbühler, K. (2015). Functional flexibility in wild bonobo vocal behaviour. PeerJ, 3, e1124.

Coffey, K. R., Marx, R. G., & Neumaier, J. F. (2019). DeepSqueak: a deep learning-based system for detection and analysis of ultrasonic vocalizations. Neuropsychopharmacology, 44(5), 859-868.

Collier, K., Bickel, B., van Schaik, C. P., Manser, M. B., & Townsend, S. W. (2014). Language evolution: syntax before phonology?. Proceedings of the Royal Society B: Biological Sciences, 281(1788), 20140263.

Collier, K., Townsend, S. W., & Manser, M. B. (2017). Call concatenation in wild meerkats. Animal Behaviour, 134, 257-269.

Courter, J. R., & Ritchison, G. (2010). Alarm calls of tufted titmice convey information about predator size and threat. Behavioral Ecology, 21(5), 936-942.

Crockford, C., & Boesch, C. (2003). Context-specific calls in wild chimpanzees, Pan troglodytes verus: analysis of barks. Animal Behaviour, 66(1), 115-125.

Crockford, C., & Boesch, C. (2005). Call combinations in wild chimpanzees. Behaviour, 142(4), 397-421.

Crockford, C., Herbinger, I., Vigilant, L., & Boesch, C. (2004). Wild chimpanzees produce group-specific calls: a case for vocal learning?. Ethology, 110(3), 221-243.

Dargan, S., Kumar, M., Ayyagari, M. R., & Kumar, G. (2019). A survey of deep learning and its applications: A new paradigm to machine learning. Archives of Computational Methods in Engineering, 1-22.

Dave, N. (2013). Feature extraction methods LPC, PLP and MFCC in speech recognition. International journal for advance research in engineering and technology, 1(6), 1-4.

Dawkins, R., & Krebs, J. R. (1978). Animal signals: information or manipulation. Behavioural ecology: An evolutionary approach, 2, 282-309.

Erskine, R. J., Bartlett, P. C., Byrem, T. M., Render, C. L., Febvay, C., & Houseman, J. T. (2012). Association between bovine leukemia virus, production, and population age in Michigan dairy herds. Journal of dairy science, 95(2), 727-734.

Evans, C. S. (1997). Referential signals. In Communication (pp. 99-143): Springer.

Fallow, P. M., & Magrath, R. D. (2010). Eavesdropping on other species: mutual interspecific understanding of urgency information in avian alarm calls. Animal Behaviour, 79(2), 411-417.

Ficken, M. S. (1990). Acoustic characteristics of alarm calls associated with predation risk in chickadees. Animal Behaviour.

Ficken, M. S., & Popp, J. W. (1992). Syntactical organization of the gargle vocalization of the black-capped chickadee, Parus atricapillus. Ethology, 91(2), 156-168.

Ficken, M. S., & Witkin, S. R. (1977). Responses of black-capped chickadee flocks to predators. The Auk, 94(1), 156-157.

Fischer, J., Kitchen, D. M., Seyfarth, R. M., & Cheney, D. L. (2004). Baboon loud calls advertise male quality: acoustic features and their relation to rank, age, and exhaustion. Behavioral Ecology and Sociobiology, 56(2), 140-148.

Fitch, W. T. (2012). Segmental structure in banded mongoose calls. BMC biology, 10(1), 98.

Freeberg T. M., Book D.L., Jung H & Kyle S.C. 2017. Communication, cues, and signals. In: Shackelford TK, Weekes-Shackelford VA, editors. Encyclopaedia of evolutionary psychological science. doi:10.1007/978-3-319-16999-6_2728-1

Freeberg, T. M. (2008). Complexity in the chick-a-dee call of Carolina chickadees (Poecile carolinensis): associations of context and signaler behavior to call structure. The Auk, 125(4), 896-907.

Freeberg, T. M., & Lucas, J. R. (2012). Information theoretical approaches to chick-a-dee calls of Carolina chickadees (Poecile carolinensis). Journal of Comparative Psychology, 126(1), 68.

Freeberg, T. M., Krama, T., Vrublevska, J., Krams, I., & Kullberg, C. (2014). Tufted titmouse (Baeolophus bicolor) calling and risk-sensitive foraging in the face of threat. Animal cognition, 17(6), 1341-1352.

Freeberg, T. M., Lucas, J. R., & Clucas, B. (2003). Variation in chick-a-dee calls of a Carolina chickadee population, Poecile carolinensis: identity and redundancy within note types. The Journal of the Acoustical Society of America, 113(4), 2127-2136.

Gerhardt, H. C. (2008). Phonotactic selectivity in two cryptic species of gray treefrogs: effects of differences in pulse rate, carrier frequency and playback level. Journal of Experimental Biology, 211(16), 2609-2616.

Gradišek, A., Slapničar, G., Šorn, J., Luštrek, M., Gams, M., & Grad, J. (2017). Predicting species identity of bumblebees through analysis of flight buzzing sounds. Bioacoustics, 26(1), 63-76.

Haftorn, S. (2000). Contexts and possible functions of alarm calling in the willow tit, Parus montanus; the principle of ‘better safe than sorry’. Behaviour, 137(4), 437-449.

Hailman, J. P. (1977). Optical signals: animal communication and light. Indiana U Press.

Hailman, J. P. (1989). The organization of major vocalizations in the Paridae. The Wilson Bulletin, 305-343.

Hailman, J. P. (2008). Coding and redundancy: Man-made and animal-evolved signals. Harvard University Press.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: data mining, inference, and prediction. Springer Science & Business Media.

Hauser, M. D. (1996). The evolution of communication. MIT press.

Hile, A. G., & Striedter, G. F. (2000). Call convergence within groups of female budgerigars (Melopsittacus undulatus). Ethology, 106(12), 1105-1114.

Hugie, D. M. (2003). The waiting game: a “battle of waits” between predator and prey. Behavioral Ecology, 14(6), 807-817.

Hurford, J. R. (2012). The origins of grammar: Language in the light of evolution II (Vol. 2). Oxford University Press.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning (Vol. 112, p. 18). New York: springer.

Jansen, D. A., Cant, M. A., & Manser, M. B. (2012). Segmental concatenation of individual signatures and context cues in banded mongoose (Mungos mungo) close calls. BMC biology, 10(1), 97.

Jung, H., & Freeberg, T. M. (2017). Variation in chick-a-dee calls of bridled titmice (Baeolophus wollweberi): Frequent use of non-combinatorial calls in a combinatorial calling system. Ethology, 123(11), 835-842.

Jung, H., Sherrod, A., LeBreux, S., Price, J. M., & Freeberg, T. M. (2020). Traffic noise and responses to a simulated approaching avian predator in mixed-species flocks of chickadees, titmice, and nuthatches. Ethology, 126(6), 620-629.

Kalb, N., Anger, F., & Randler, C. (2019). Great tits encode contextual information in their food and mobbing calls. Royal Society Open Science, 6(11), 191210.

Kelley, K. C. (1988). Duration of the Dee Component of the" chickadee" Call and Agonistic Behavior in the Carolina Chickadee Parus Carolinensis (Doctoral dissertation, Graduate School of Arts and Sciences, University of Pennsylvania).

Kershenbaum, A., Blumstein, D. T., Roch, M. A., Akçay, Ç., Backus, G., Bee, M. A., . . . Caesar, C. (2016). Acoustic sequences in non-human animals: a tutorial review and prospectus. Biological Reviews, 91(1), 13-52.

Kershenbaum, A., Blumstein, D. T., Roch, M. A., Akçay, Ç., Backus, G., Bee, M. A., . . . Caesar, C. (2016). Acoustic sequences in non-human animals: a tutorial review and prospectus. Biological Reviews, 91(1), 13-52.

Krams, I. (2001). Communication in crested tits and the risk of predation. Animal Behaviour, 61(6), 1065-1068.

Krams, I., Krama, T., Freeberg, T. M., Kullberg, C., & Lucas, J. R. (2012). Linking social complexity and vocal complexity: a parid perspective. Phil. Trans. R. Soc. B, 367(1597), 1879-1891.

Kuhn, M. (2008). Building predictive models in R using the caret package. Journal of statistical software, 28(5), 1-26.

Landsborough, B., Wilson, D. R., & Mennill, D. J. (2020). Variation in chick-a-dee call sequences, not in the fine structure of chick-a-dee calls, influences mobbing behaviour in mixed-species flocks. Behavioral Ecology, 31(1), 54-62.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. nature, 521(7553), 436-444.

Lee, C. H., Han, C. C., & Chuang, C. C. (2008). Automatic classification of bird species from their sounds using two-dimensional cepstral coefficients. IEEE Transactions on Audio, Speech, and Language Processing, 16(8), 1541-1550.

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.

Lengagne, T., Aubin, T., Lauga, J., & Jouventin, P. (1999). How do king penguins (Aptenodytes patagonicus) apply the mathematical theory of information to communicate in windy conditions?. Proceedings of the Royal Society of London B: Biological Sciences, 266(1429), 1623-1628.

Lima, S. L. (2009). Predators and the breeding bird: behavioral and reproductive flexibility under the risk of predation. Biological reviews, 84(3), 485-513.

Lima, S. L., & Dill, L. M. (1990). Behavioral decisions made under the risk of predation: a review and prospectus. Canadian journal of zoology, 68(4), 619-640.

Lind, J., Jöngren, F., Nilsson, J., Alm, D. S., & Strandmark, A. (2005). Information, predation risk and foraging decisions during mobbing in Great Tits Parus major. Ornis Fennica, 82(3), 89-96.

Linhart, P., Ratcliffe, V. F., Reby, D., & Špinka, M. (2015). Expression of emotional arousal in two different piglet call types. PloS one, 10(8), e0135414.

Lucas, J. R., & Freeberg, T. M. (2007). Information and the chick-a-dee call: communicating with a complex vocal system. Ecology and behavior of chickadees and titmice. Oxford University Press, Oxford, 199-213.

Mahurin, E. J., & Freeberg, T. M. (2008). Chick-a-dee call variation in Carolina chickadees and recruiting flockmates to food. Behavioral Ecology, 20(1), 111-116.

Mammen, D. L., & Nowicki, S. (1981). Individual differences and within-flock convergence in chickadee calls. Behavioral Ecology and Sociobiology, 9(3), 179-186.

Manser, M. B. (2001). The acoustic structure of suricates' alarm calls varies with predator type and the level of response urgency. Proceedings of the Royal Society of London. Series B: Biological Sciences, 268(1483), 2315-2324.

Marler, P. (1961). The logical analysis of animal communication. Journal of Theoretical Biology, 1(3), 295-317.

Marler, P. (1977). The structure of animal communication sounds. — In: Recognition of

Marler, P. (1998). Animal communication and human language. The origin and diversification of language, 1, 19.

Marler, P., Dufty, A., & Pickert, R. (1986). Vocal communication in the domestic chicken: II. Is a sender sensitive to the presence and nature of a receiver? Animal Behaviour, 34, 194-198.

Marler, P., Evans, C. S., & Hauser, M. D. (1992). Animal signals: motivational, referential, or both. Nonverbal vocal communication: Comparative and developmental approaches, 66-86.
Marler, P., Evans, C. S., & Hauser, M. D. (1992). Animal signals: motivational, referential, or both. Nonverbal vocal communication: Comparative and developmental approaches, 66-86.
Mathevon, N., Koralek, A., Weldele, M., Glickman, S. E., & Theunissen, F. E. (2010). What the hyena's laugh tells: Sex, age, dominance and individual signature in the giggling call of Crocuta crocuta. BMC ecology, 10(1), 9.

McIlraith, A. L., & Card, H. C. (1997, May). Bird song identification using artificial neural networks and statistical analysis. In CCECE'97. Canadian Conference on Electrical and Computer Engineering. Engineering Innovation: Voyage of Discovery. Conference Proceedings (Vol. 1, pp. 63-66). IEEE.

Mendl, M., Burman, O. H., & Paul, E. S. (2010). An integrative and functional framework for the study of animal emotion and mood. Proceedings of the Royal Society B: Biological Sciences, 277(1696), 2895-2904.

Mostrom, A. M., Curry, R. L., & Lohr, B. S. (2020). Carolina chickadee: Poecile carolinensis. Birds of North America.

Noda, J. J., Travieso, C. M., & Sánchez-Rodríguez, D. (2016). Automatic taxonomic classification of fish based on their acoustic signals. Applied Sciences, 6(12), 443.

Nolen, M. T., & Lucas, J. R. (2009). Asymmetries in mobbing behaviour and correlated intensity during predator mobbing by nuthatches, chickadees and titmice. Animal Behaviour, 77(5), 1137-1146.

Nowicki, S. (1989). Vocal plasticity in captive black-capped chickadees: the acoustic basis and rate of call convergence. Animal Behaviour, 37, 64-73.

Ouattara, K., Lemasson, A., & Zuberbühler, K. (2009). Campbell's monkeys concatenate vocalizations into context-specific call sequences. Proceedings of the National Academy of Sciences, pnas-0908118106.

Ouattara, K., Zuberbühler, K., N'goran, E. K., Gombert, J. E., & Lemasson, A. (2009). The alarm call system of female Campbell's monkeys. Animal Behaviour, 78(1), 35-44.

Owens, J. L., & Freeberg, T. M. (2007). Variation in chick-a-dee calls of tufted titmice, Baeolophus bicolor: note type and individual distinctiveness. The Journal of the Acoustical Society of America, 122(2), 1216-1226.

Owren, M. J., Rendall, D., & Ryan, M. J. (2010). Redefining animal signaling: influence versus information in communication. Biology & Philosophy, 25(5), 755-780.

Phillips, A. V., & Stirling, I. (2000). Vocal individuality in mother and pup South American fur seals, Arctocephalus australis. Marine Mammal Science, 16(3), 592-616.

Phillips, J. N., Rochefort, C., Lipshutz, S., Derryberry, G. E., Luther, D., & Derryberry, E. P. (2020). Increased attenuation and reverberation are associated with lower maximum frequencies and narrow bandwidth of bird songs in cities. Journal of Ornithology, 1-16.

Prat, Y., Taub, M., & Yovel, Y. (2016). Everyday bat vocalizations contain information about emitter, addressee, context, and behavior. Scientific Reports, 6, 39419.

Proppe, D. S., Bloomfield, L. L., & Sturdy, C. B. (2010). Acoustic transmission of the chick-a-dee call of the Black-capped Chickadee (Poecile atricapillus): forest structure and note function. Canadian journal of zoology, 88(8), 788-794.

Randler, C., & Förschler, M. I. (2011). Heterospecifics do not respond to subtle differences in chaffinch mobbing calls: message is encoded in number of elements. Animal behaviour, 82(4), 725-730.

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

Ritchison, G. (1983). Vocalizations of the white-breasted nuthatch. The Wilson Bulletin, 440-451.

Salamon, J., Bello, J. P., Farnsworth, A., & Kelling, S. (2017, March). Fusing shallow and deep learning for bioacoustic bird species classification. In 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 141-145). IEEE.

Schibler, F., & Manser, M. B. (2007). The irrelevance of individual discrimination in meerkat alarm calls. Animal Behaviour, 74(5), 1259-1268.

Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980). Vervet monkey alarm calls: semantic communication in a free-ranging primate. Animal Behaviour, 28(4), 1070-1094.

Sloan, J. L., Wilson, D. R., & Hare, J. F. (2005). Functional morphology of Richardson's ground squirrel, Spermophilus richardsonii, alarm calls: the meaning of chirps, whistles and chucks. Animal Behaviour, 70(4), 937-944.

Slobodchikoff, C., Kiriazis, J., Fischer, C., & Creef, E. (1991). Semantic information distinguishing individual predators in the alarm calls of Gunnison's prairie dogs. Animal Behaviour, 42(5), 713-719.

Slocombe, K. E., & Zuberbühler, K. (2006). Food-associated calls in chimpanzees: responses to food types or food preferences? Animal Behaviour, 72(5), 989-999.

Smith, W. J. (1963). Vocal communication of information in birds. The American Naturalist, 97(893), 117-125.

Smith, W. J. (1965). Message, meaning, and context in ethology. The American Naturalist, 99(908), 405-409.

Smith, C. L. (2017). Referential signalling in birds: the past, present and future. Animal Behaviour, 124, 315-323.

Soard, C. M., & Ritchison, G. (2009). ‘Chick-a-dee’calls of Carolina chickadees convey information about degree of threat posed by avian predators. Animal Behaviour, 78(6), 1447-1453.

Sturdy, C. B., Phillmore, L. S., & Weisman, R. G. (2000). Call-note discriminations in black-capped chickadees (Poecile atricapillus). Journal of Comparative Psychology, 114(4), 357.

Suzuki, T. N. (2014). Communication about predator type by a bird using discrete, graded and combinatorial variation in alarm calls. Animal Behaviour, 87, 59-65.

Templeton, C. N., Greene, E., & Davis, K. (2005). Allometry of alarm calls: black-capped chickadees encode information about predator size. Science, 308(5730), 1934-1937.

Townsend, S. W., & Manser, M. B. (2013). Functionally referential communication in mammals: the past, present and the future. Ethology, 119(1), 1-11.

Townsend, S. W., & Manser, M. B. (2013). Functionally referential communication in mammals: the past, present and the future. Ethology, 119(1), 1-11.

Townsend, S. W., Hollén, L. I., & Manser, M. B. (2010). Meerkat close calls encode group-specific signatures, but receivers fail to discriminate. Animal Behaviour, 80(1), 133-138.

Turesson, H. K., Ribeiro, S., Pereira, D. R., Papa, J. P., & de Albuquerque, V. H. C. (2016). Machine learning algorithms for automatic classification of marmoset vocalizations. PloS one, 11(9), e0163041.

Tyack, P. L. (2008). Convergence of calls as animals form social bonds, active compensation for noisy communication channels, and the evolution of vocal learning in mammals. Journal of Comparative Psychology, 122(3), 319.

Valletta, J. J., Torney, C., Kings, M., Thornton, A., & Madden, J. (2017). Applications of machine learning in animal behaviour studies. Animal Behaviour, 124, 203-220.

von Frisch, K., & Lindauer, M. (1956). The" language" and orientation of the honey bee. Annual review of entomology, 1(1), 45-58.

Wagner Jr, W. E., & Basolo, A. L. (2007). The relative importance of different direct benefits in the mate choices of a field cricket. Evolution, 61(3), 617-622.

Wiley, R. H., & Richards, D. G. (1982). Adaptations for acoustic communication in birds: Sound transmission and signal detection. Pages 132-181 in Acoustic communication in birds, vol. 1. (D. E. Kroodsma and E. H. Miller, Eds.). Academic Press, New York

Wilkinson, G. S., & Boughman, J. W. (1998). Social calls coordinate foraging in greater spear-nosed bats. Animal Behaviour, 55(2), 337-350.

Wilson, D. R., & Mennill, D. J. (2011). Duty cycle, not signal structure, explains conspecific and heterospecific responses to the calls of black-capped chickadees (Poecile atricapillus). Behavioral Ecology, 22(4), 784-790.

Xie, J., Towsey, M., Truskinger, A., Eichinski, P., Zhang, J., & Roe, P. (2015, April). Acoustic classification of australian anurans using syllable features. In 2015 IEEE tenth international conference on intelligent sensors, sensor networks and information processing (ISSNIP) (pp. 1-6). IEEE.

Xie, J., Towsey, M., Zhang, J., & Roe, P. (2018). Frog call classification: a survey. Artificial Intelligence Review, 49(3), 375-391.

Yorzinski, J. L., Vehrencamp, S. L., McGowan, K. J., & Clark, A. B. (2006). The inflected alarm caw of the American crow: differences in acoustic structure among individuals and sexes. The Condor, 108(3), 518-529.

Zhong, M., Castellote, M., Dodhia, R., Lavista Ferres, J., Keogh, M., & Brewer, A. (2020). Beluga whale acoustic signal classification using deep learning neural network models. The Journal of the Acoustical Society of America, 147(3), 1834-1841.

Zimmermann, E., Leliveld, L. M. C., and Schehka, S. (2013). Toward the evolutionary roots of affective prosody in human acoustic communication: a comparative approach to mammalian voices, in The Evolution of Emotional Communication: From Sounds in Nonhuman Mammals to Speech and Music in Man, eds E. Altenmuller, S. Schmidt, and E.

Zimmermann (Oxford: Oxford University Press), 116–132. doi: 10.1093/acprof:oso/9780199583560.003.0008

Zohair, L. M. A. (2019). Prediction of Student's performance by modelling small dataset size. International Journal of Educational Technology in Higher Education, 16(1), 27.

VITA

Hwayoung Jung was born in Tokyo, Japan in 1985 and raised in an animal-loving family. She started her first academic coursework in animal behavior in 2011 at Kyung Hee University and received MS of Biology. After graduation she worked at the National Institute of Ecology in South Korea. She left the job and became a PhD student in the Experimental Psychology program in the University of Tennessee, Knoxville in 2015.