**Annotation manual**

In this project, we look at voice descriptions of birds in field guides.

**Examples**
An example is the following description of the vocalizations of the Dusky Flycatcher (*Empidonax oberholseri*) from the National Geographic Birds of North America guide:

(1) Calls include a *wit* note, softer than Gray Flycatcher; a mournful *deehic*, heard on breeding grounds. Song has several phrases: a clear *sillit*; an upslurred *ggrrreep*; another high *sillit*, often omitted; and a clear, high *pweet*.

Here are a few more examples, the Common Redstart (*Phoenicurus phoenicurus*) from the RSPB Handbook of British Birds:

(2) Female quack like Blue-winged. Male in display gives a dry, chattering or rattling 'gredek gredek…'; vaguely reminiscent of Northern Shoveler, unlike the whisle of Blue-winged.

And the Cinnamon Teal (*Anas cyanoptera*) from the Sibley North American Bird Guide:

(3) Call a rather sweet, warbler-like *hooweet*. Song a loud warble ending with distinctive mechanical jangle.

And the Blackbird (*Turdus merula*) from the Collins BTO Guide to British Birds:

(4) Utters a harsh and repeated *tchak* alarm call, often at dusk or if it is disturbed by a prowling cat. Male is an excellent songster with a rich and varied repertoire.

And the Little Owl (*Athene noctua*) from the Collins Gem Birds fieldguide:

(5) Penetrating yelps and ***poop*** whistles.

Finally, an example of a description of the Crested Francolin (*Dendroperdix sephaena*) from the Guide to the Birds of the Kruger National Park:

(6) "Its call, typically a well-synchronised duet by the male and female ('*tina-turner, tina-turner ...*'), is one of the characteristic sounds of the park."

**Digitization**
Our first step was to digitize all voice sections by hand-typing them in from the written field guides. This proved to be more efficient than using any OCR technology because we are only dealing with small sections from entire books, many of which would have to be scanned first. The only data that was not entered this way is the All About Birds data, which comes from a website curated by the Cornell Ornithology lab.

As the examples (1)-(6) above show, onomatopoeias and warblish (Sarvasy, 2016) are highlighted using in different typological means in different fields. For example, the Collins Gem Birds fieldguide (5) represent onomatopoeias using italics and boldface. Most other field guides used only italics, as in examples (1), (3), (4), and (6). Some field guides used single quotes, as in (2). To ensure machine readability and facilitate later extraction of onomatopoeias, the typographic highlighting was represented using double quotes. Cases that were obviously warblish (multi-syllable expressions such as "teacher-teacher", "whit-poor-will," or "who-cooks-for-you"), were often bracketed using **. The file `fieldguides_metadata.csv` in the `data` folder preserves information about how each field guide managed typographic highlighting of onomatopoeias.

**Goals**

We have several goals with this dataset, and knowing about these goals is important to understand our annotation procedures. In the following discussion, I will refer to 'sound object' as a cover term for a bird's calls, songs, or other sounds they may make (e.g., beak clacking, wing flapping).

*Goal 1: Verbal descriptors*
First, we are interested in the *verbal descriptors* of sound objects. There are many questions that pose themselves here, such as: what adjectives are used to describe these sound objects, and how do these compare to known tendencies in "synesthetic metaphor" (Winter, 2019)? And which adjectives are over- or under-represented compared to descriptions of musical language (Wallmark, 2019)? These questions are impossible to investigate with the raw data because the texts contain a lot of information that does not directly relate to sound itself, but to the context in which a bird produces a sound, such as "from high perch," or "when flushed," or "during breeding season." So, one goal will be to erase all language that does not directly relate to sound.

*Goal 2: Semantic space of verbal descriptors*
Second, we are interested in reconstructing the underlying semantic space of the verbal descriptors using some form of distributional semantic analysis such as Latent Semantic Analysis (LSA). These approaches look at which words tend to co-occur with each other, and beyond that, whether even if words never occur next to each other, whether they occur in similar contexts (e.g., synonyms). These questions are impossible to investigate with the raw data because it is currently coded at the level of entire entries, which means that several acoustically distinct types of sound objects a bird produces (e.g., a harsh alarm call versus a soft and clear song) are collapsed into the same unit of analysis. This will grossly distort the semantic space as alarm calls and songs are very different types of sound objects that should not be lumped together. So, another goal will be to separate different types of sound objects the same bird produces.

*Goal 3: Syntactic annotation of onomatopoeias*
Third, we are interested in what syntactic context onomatopoeias occur in, and specifically, whether like ideophones, onomatopoeias are less syntactically integrated and can assume different syntactic positions (cf. Dingemanse & Akita,

2017). Later, we also want to investigate to what extent certain syntactic positions are more or less preferred as a function of how conventionalized an onomatopoeia is. Example (1) shows how sometimes onomatopoeias can be pre-nominal ("*wit* note") or act as noun-like elements themselves ("mournful *deehic*"). There are many other types of constructions, such as quotative constructions (e.g., "sounds like X"), or onomatopoeias that are completely separated from sentences, e.g., by following a semicolon after a complete sentence. These syntactic expressions need to be annotated.

*Goal 4: Comparisons*
Fourth, we are interested whether there are comparisons, such as in example (1) ("softer than the Gray Flycatcher"), example (2) ("Female quack like Blue-winged"; "vaguely reminiscent of Northern Shoveler, unlike the whisle of Blue-winged"), and example (3) ("warbler-like"). In other cases, sound objects are often compared to mechanical objects, such as the sound of two stones clacking together, or the sound of a squealing door; or they are compared to other animals, such as insects, cats, and pigs; or they are compared to musical instrument, such as flutes or trombones. Here, several questions pose themselves: what types of sounds are bird vocalizations most often compared to? And later, we want to investigate which birds are compared to which other birds, to investigate the network of comparisons within the bird structure using some network statistics, potentially phylogenetically informed. We are also potentially interested in the types of comparison constructions that occur most frequently, such as "sounds like…", "resembles…", "X-like," or "X-er than …"

## Major steps

With these goals in mind, there are several major steps of data preparation:

### ***First: Data cleaning***
First, typos need to be fixed, such as "whisle" in example (2), which should be "whistle." For this, the whole text needs to be read in context. While doing that, the annotator should look out for bracketing errors in particular: it cannot be the case that a double quote that highlights an onomatopoeia is not *doubly* present, once before the onomatopoeia, once after. For example …

    "krek
    krek"

… would both have to be fixed to:

    "krek"

Importantly, please ensure that any edits of double quotes surrounding onomatopoeias involve straight quotes (as above) and *not* formatted curly quotes. That is, "krek" not "krek". As an additional step, trailing spaces should be deleted, e.g., " krek " to "krek". And if a comma or sentence final dot or semicolon is part of the onomatopoeia, this should be moved out. For example, this string:

> Males sing a tinkling, warbling song,  " see, see, see me, see me, hear me, hear me, see. "

Should become:

> Males sing a tinkling, warbling song, "see, see, see me, see me, hear me, hear me, see".

However, please do leave in instances where a dot or hyphen may mark anything meaningful. For example, "kek-kek-kek-" or "kek-kek-kek…" could indicate that the onomatopoeia is repeated. Likewise, if an exclamation mark is part of the onomatopoeia, this could also be relevant as it may indicate emphasis or intonation.

A final step of the data cleaning procedure is that we want to flag *clear* cases of warblish. Following Sarvasy (2016), we take "warblish" to be the use of existing non-onomatopoeic words to imitate bird vocalizations, such as the California quail, which is said to say "*Chicago!*". We *sometimes* flagged clear cases of warblish with ** instead of the double quotes we use for onomatopoeias. However, this decision was made late in the process, and the All About Birds data from the Cornell Ornithology lab has not been processed this way at all. For the purposes of this annotation, a *clear* case of warblish is any multi-syllabic or multi-word expression that is also an English word, such as *teacher*, *Chicago*, or *who-cooks-for-you?*. Single-syllable cases such as "sing" often provide too little evidence to know whether this was intended to be warblish (read as an imitation with 'sing' being the English word) or onomatopoeia (read as an imitation with 'sing' understood as a nonce-word imitation). We will later look at how many of these cases that we did not explicitly flag as warblish involve existing English words. However, we need to exclude obvious cases of warblish from our future work on onomatopoeia since this could mess up our phonological analysis as in onomatopoeias, there often is a direct correspondence between grapheme and phoneme that is not present in warblish.

### *Second: Setting boundaries*
After the first data preparation step ("***Data cleaning***") is complete, we need to prepare the data for Goal 2, for which we need distinct types of sound objects that are described in the same entry separated from each other. We will use semicolons for this, so that I can write a script that decomposes the text in pieces separated by semicolons. So, boundary setting will involve placing semicolons between each sound object. This may involve replacing existing dots, commas, or other punctuation marks with semicolons that are used to separate sentences anyway.

Example (1), repeated here, looks already complete because the particular author of this field guide used a lot of semicolons to separate particular phrases. However, notice that there is no semicolon between "a mournful *deehic*" and "a clear *sillit*", which would mean that my script would lump these two sound objects together.

(1)     Calls include a *wit* note, softer than Gray Flycatcher; a mournful *deehic*, heard on breeding grounds. Song has several phrases: a clear *sillit*; an upslurred *ggrrreep*; another high *sillit*, often omitted; and a clear, high *pweet*.

So, the corrected version needs to be:

(1c)    Calls include a *wit* note, softer than Gray Flycatcher; a mournful *deehic*, heard on breeding grounds; Song has several phrases: a clear *sillit*; an upslurred *ggrrreep*; another high *sillit*, often omitted; and a clear, high *pweet*.

As another demonstration, consider example (2) (now with the typo in "whisle" corrected after stage 1 of data preparation):

(2)     Female quack like Blue-winged. Male in display gives a dry, chattering or rattling 'gredek gredek…'; vaguely reminiscent of Northern Shoveler, unlike the whistle of Blue-winged.

Here we have the opposite problem. The comparison "vaguely reminiscent of Northern Shoveler, unlike the whistle of Blue-winged" is separated by a semicolon by this author, but a contextual reading actually clearly reveals that the whole phrase is modifying what came previously, the "dry, chattering or rattling 'gredek gredek…'". Therefore, this semicolon needs to be deleted, as in the corrected version given below:

(2c)    Female quack like Blue-winged; Male in display gives a dry, chattering or rattling 'gredek gredek…', vaguely reminiscent of Northern Shoveler, unlike the whistle of Blue-winged.

Please note also the replacement of the dot after the first sentence with a semicolon. Replacing sentence-final dots with semicolons will be very common, but please be aware that sometimes the same sound object can be described across multiple sentences, in which case no semicolon should be placed. There unfortunately is no way around reading the entire description in context.

Example (5) repeated here poses a different case:

(5)     Penetrating yelps and ***poop*** whistles.

Notice that it is relatively obvious from context that the "penetrating yelps" and "poop whistles" describe two separate sound objects (otherwise, why would the author use "and"?). The corrected version for this should be:

(5c)    Penetrating yelps; and ***poop*** whistles.

The placement of the semicolon with respect to the "and" doesn't really matter as much because we will later exclude function words anyway.

### *Third: Sound-only language*
To be continued.

### *Fourth: Syntactic annotation of onomatopoeias*
To be continued.

### *Fifth: Annotation of comparisons*
To be continued.

**References**

Dingemanse, M., & Akita, K. (2017). An inverse relation between expressiveness and grammatical integration: On the morphosyntactic typology of ideophones, with special reference to Japanese. *Journal of Linguistics*, *53*(3), 501–532. https://doi.org/10.1017/S002222671600030X

Sarvasy, H. (2016). Warblish: Verbal mimicry of birdsong. *Journal of Ethnobiology*, *36*(4), 765–782. https://doi.org/10.2993/0278-0771-36.4.765

Wallmark, Z. (2019). A corpus analysis of timbre semantics in orchestration treatises. *Psychology of Music*, *47*(4), 585–605. https://doi.org/10.1177/0305735618768102

Winter, B. (2019). *Sensory linguistics: Language, perception, and metaphor*. John Benjamins.