I have read the PhD Thesis of Pavel Linhart, "Song based rival assessment in songbirds", which I found scientifically sound, coherent as a whole, and pertinent facing the state of the art in the area of animal communication. This thesis used two species as models to investigate the communication function of acoustic traits, and it successfully addressed what are arguably the most interesting questions on the function of signal design that these models can answer.

In the absence of major criticism, the opponent's report comments on methodological or interpretative aspects that we might discuss during the thesis defence.

Sincerely,
Gonçalo Cardoso

Opponent's report

Chapter 1

Introduction
Elegant, covering the main background in an accessible way, and making a good case for the pertinence of the thesis goals. Overall friendly to the reader, briefly explaining the purpose of each subsection at the onset when appropriate.

Factors affecting RHP

Body size and vocalization frequency
Page 9 - I'd like to hear more about why some of these studies found an inverse relation between body size and pitch.

Do males use the information on aspects of RHP in vocalizations during male-male interactions?
In page 11 and at least another subsection, you chose to give the state of the art at the time the PhD began, which is useful to understand the novelty of the project. But it also raises interest in knowing what significant studies happened in the meantime.

Other factors affecting vocalization pitch
It would be useful to back up the statement that "in some species correlation between body size and syrinx size is missing". Has this been shown?

Other possibilities for RHP assessment from the vocalizations
Good review of the functional role of song length, and of the overlooked questions.
Paragraph on song performance lacks references.
Good point about the likely perceptual salience of syllable rate.

Overview and synthesis
The different predictions for the two species are very elegant, based on what is perceptually salient in their songs. As far as I am aware, I don't think this idea has been well articulated in the literature: trying to explain or to predict differences in the

communication role of song traits among species based on what is perceptually salient in their different songs.

It seemed to me a very valuable idea, perhaps useful to try and synthesise otherwise inconsistent results in the literature. Yet, the results on syllable rate on chiffchaffs contradict this possible principle... Maybe there are some universal principles in animal communication (e.g. function of syllable rate), and some which are species-specific (function of pitch or duration)?

Chapter 2

Abstract
I think the critical point here is not being a "small species" but a species with determinate growth and small variation in size among individuals.

Introduction
The prediction that was made in Chapter 1 (that minimum song frequency should correlate with body size in this species because of the decreasing song pitch throughout the song, reaching a low frequency for its body size) is missing in this introduction. I thought it was very elegant and that it should be given prominence here (and maybe in the abstract also).

Prediction 2 is sensible. But in chapter 3 you show that for some traits focusing on differences between the playback and the focal male is meaningful, and for other traits it is better to simply analyse the absolute value of the song trait in the playback. Why not to have made these simpler analyses here too?

Methods
Methods for the individual song and behavioural traits in tables 1 and 2 were not explained.

It would be useful to have information about song repertoires (or lack thereof), or about within-individual variation in song parameters, in order to understand the representativeness of this recording protocol.

I agree that PCA is useful to reduce the number of these correlated frequency measurements. But at the same time, given that the willow warbler has this interesting song structure of decreasing pitch, it would be interesting to understand if the minimum frequency is particularly informative regarding body size or receiver responses. Thus, why not to have looked at minimum frequency directly, and to contrast its results with those of other frequency measurements?

I understand that PC2 is later interpreted as minimum frequency. Why to have analysed behavioural responses only in relation to this PC2? For the same reason as above, contrasting this result with responses to peak frequency (PC1) would have been interesting.

Very good sample size for the experiments.

Too brief description of statistical models in the methods makes understanding of the results more difficult. The confounding variables were not explained.

The statistical models in Table S1 tell us that tarsus length, not body mass, is related to song frequency. What may this mean biologically? Is mass not related to tarsus length?

Discussion
The explanation and interpretation of predicted frequency and range based on Wällschlager's paper is confusing.

Good point, raising the question about whether the size-frequency allometry is due to an acoustic constraint, or other reasons, especially since minimum frequency appears not to be a stronger predictor of size than the other frequency traits (see the arguments in Hall et al. 2013 Plos ONE). This result could have been brought to the discussion.

It is not clear if closer approach and lower activity are both being interpreted as low intensity approaches? Closer approach should be a high intensity aggressive response, I would think. These interpretations should be clarified, maybe even in the methods, so that the results are easier to follow.

It would be useful to explain here what this "mixed strategy" is and how it helps interpret the results. Otherwise, it is not clear how "approach less" is consistent with an interpretation of increased aggressiveness.
The "mixed strategy" in chapter 3 is "attack more but approach less". Here it is slightly different: "move more but approach less". Is it safe to interpret "move" as aggressive?

Chapter 3

Introduction
Why were the prediction regarding frequency phrased only in relation to body size, and the prediction regarding length phrased in relation to body size and mass?

Methods
Good level of detail in the description of recordings, and general acoustic measurements (though it could have been explained for non-specialist readers what the automatic measurements tool of Avisoft does), playback, behaviour (though it is unclear if when the birds sits on the loudspeaker and pecks repeatedly these would be counted as independent attacks) and statistics.

I am used to call "peak frequency" to what is here called "cumulative peak frequency"...

A bit strange to count syllables rather than measuring song length; seems a little old-fashioned... It is irrelevant, though.

I used this type of manipulations of sound frequency in the past. I would now have preferred to manipulate frequency multiplicatively, which is in line with how animals

perceive sound, and so as not to destroy harmonic relations between sounds. I don't think this will matter for the conclusions, though.

Results

I am not used to see Bonferroni correction applied to the multiple parameters of a single multiple regression model. I am used to see it applied only when separate models are run on identical hypotheses. Is this conservatism necessary? Instead couldn't the F statistic of the whole model be seen as the single hypothesis being tested, and the F statistics of the individual parameters simply to assess what drove the full model to be significant?

The interpretation of the results from the AICc model selection is indeed the most plausible, but not significantly better than the alternatives in the interval of 2AICc from the best model. Since the hypothesis of interest is whether birds react to absolute vs. relative differences in song traits, wouldn't it be cleaner to focus the interpretation in the AICc difference between models using absolute vs. relative measurements, but otherwise identical?

Discussion

Small variation in body size among individual chiffchaffs is invoked as an explanation for the small R^2 of the allometry with frequency. How does the degree of variation in body size among willow warblers compare to this, and could it help explain the larger R^2 in chapter 2?

Page 72 mentions that chiffchaff alter song length "in both directions" in response to challenges. This is puzzling; why "in both directions"?

Conclusions

Agree that the behavioural data complements well the result on allometry, in suggesting that despite the correlation not being strong it may influence communication.

Chapter 4

Methods

Good detail on playback stimuli, acoustic measurements, experiments and behaviour; might give a little more detail on morphological measurements.

Very concise and informative Fig. 1.

The one-sample t-test makes intuitive sense, but some people advice against using statistics (r) as if they were data.

Why was it chosen to analyse separately each response behaviour, while in the previous chapters a PCA was used to reduce them?

Results

Clear, and quite exhaustive!

Discussion
The results are interpreted in the scope of the context, predictive and response criteria for aggressive signals. When discussing the predictive criterion, you are quite careful (e.g. "We could not answer the question whether syllable rate during the playback predicts attack. However […]", etc). Do you really think this, or is this caution due to reviewer criticism? How important is it that the signal (syllable rate) was measured before or during the playback, in order to meet the predictive criterion?

Good arguments about syllable rate not being an index signal in this species.

Pavel Linhart: Song based rival assessment in songbirds
Opponent's review

This thesis comprises four chapters, with the main aims following the title. Two closely related species of warblers, each showing distinct vocalization, were used as model species in this thesis, the Chiffchaff and the Willow warbler. Both species are easy to handle, respond strongly to playback and are easily accessible in the field. Both species have drab and inconspicuous colouration with no, or only weak, sexual dimorphism; hence male vocalization is likely to be the most important trait used for sexual signalling. Thus, I believe the choice of species was indeed a good one.

While I am not an expert on avian vocalization, this study addresses areas that I am familiar with in that, in addition to warbler song properties, it also contributes to a long lasting discussion about mechanisms of sexual selection, fighting ability and honesty of signalling through sexual ornamentation (in this case to other males in the population, the RHP). I believe the study contributes significantly to this topic. Two chapters of the thesis have already been published in high-ranking quality international journals focused on animal behaviour. This alone could stand as proof of the high quality research presented in this thesis. As these two chapters have already passed through the review process of an international community of experts, below I will focus on reviewing the two remaining chapters of the thesis, the introduction and the second chapter (manuscript submitted to Animal Behaviour).

To be honest, I was a bit disappointed after reading the introductory chapter. It has become something of a rule recently that introductory chapters tend to be the weakest part of the doctoral theses. In this particular case, I got the feeling that the 21 pages of introduction (plus 8 pages of references) were written very quickly, perhaps under time pressure. I will not specifically focus on the typographic errors occurring in the text, or on the fact that the English language tends not to be so good; especially in the thesis overview and chapter synthesis (the English would definitely benefit from editing by a native speaker). Rather, I would focus firstly on the list of references. Interestingly, the reference list comprises some 100 sources, of which only 12 were published between 2010 and 2014. I have two working hypotheses, and some less likely ideas, why this should be so. Either the field of animal vocalization is dead and there are no recent relevant source materials available – in which case why is this so, and where does the author feel the future of similarly oriented research lies? Or the author did not consider the recent literature valid – in which case, why? The third possibility, i.e. that the author did not follow the literature over recent years, is unlikely. Although the author's CV was not available to me at the time of writing this review, I deduced that the PhD was initiated in 2007. There have been a lot of papers published over the period just prior to the initiation of this thesis (29 between 2004 and 2007 if I counted

well) and these were cited in Chapter one. If the introduction to this thesis were written seven years ago, therefore, it would be perfectly up-to-date. My Devil's advocate feeling (hopefully incorrect) is that the author used texts written when preparing the PhD topic in 2007, since when they have been waiting in a drawer, and just added some new sources here and there when undertaking the final write-up.

Secondly, while the thesis deals with song in warblers, to my taste the introductory chapter focuses too much on vocalization from the very beginning. I would rather see a deeper discussion on honesty of quality signalling via secondary male ornamentation, the evolution of armament and signalization of male fighting ability, etc. When it comes to vocalization, the author could have provided more examples, and perhaps also have included insect models.

Chapter 2 comprises a manuscript that has already been submitted to a prestigious behavioural journal. In this chapter, the author provides support for an interesting idea that song pitch indicates body size in a small passerine, and is used as a signal in male-male contests. While it has previously been shown through comparative analysis that song pitch is associated with body size in birds across species with substantial variation in body size, within-species studies have been less convincing. Hence, the author's finding that similar processes could work within species, and be used in intraspecific communication, is extremely interesting. In this case, it was tarsus length that was associated with song pitch.

I have several questions/comments here:
1) It seems that only tarsus length was used as a proxy for male size; however, body mass was also available. Why was body mass omitted from the analysis? My feeling is that body mass should also be associated with song pitch and, moreover, would serve as a useful proxy for allometry in size between the contesting males. In many bird species, and other animals, it is usually body mass that determines the winner in contests. What was the relationship between body mass, tarsus length and song property?
2) The relationship between song property and male traits could be dependent on a third, unmeasured variable such as overall male quality/condition, which in turn may be associated with early developmental condition, etc. It is a shame that the authors did not use any alternative measures of male condition/quality that could have been associated with male willingness to fight or respond to a strong competitor. It would also have been nice to have some data on testosterone levels, etc.
3) Playback of 10 males recorded in a previous year was used to estimate size-pitch allometry in male contests – how many of these birds returned to the same place in 2008?
4) The most important factor in all behavioural studies is the question of trait repeatability and consistency over time. In this case, repeatability of song traits (song pitch) must be high as tarsus length does not change over time (hence song should also not change over time). Do you have any data supporting the idea that song trait (important in your study) is stable over time and consistently reflects male tarsus length? As many birds return to the study area year after year, it should be possible to obtain this kind of data.

5) Could there also be a relationship between song pitch and age or experience of males?
6) In some species, birds have dialects that are distinguishable even within a single population (I worked with rosefinches for a while and these represent a good example). Is this also true for the warbler species you studied?
7) Did you also follow nests of birds? Could an individual breeding history for each male be used in the analysis? Did the individual phase of breeding play a role in male response?

Despite my previous comments, I feel the overall quality of the thesis is high. As I said previously, I did not have author's CV to hand when writing this review; however, by checking the Web of Science, I noticed that there have been several other papers published by the author that he did not include in his Doctoral thesis. Together with the overall high quality of the thesis, this also indicates that Pavel Linhart is a competent field biologist and behavioural ecologist, already skilled in scientific thinking and methods of work. My final evaluation of the thesis, therefore, is positive and, in my opinion, the author fully deserves the Doctoral title. Overall, the text fulfils all the international criteria of a successful doctoral thesis and I am happy to recommend it for defence.

Studenec, October 11th 2014

Tomáš Albrecht