In Parncutt, R. & Sattmann, S. (Eds). *Proceedings of ICMPC15/ESCOM10*. Graz, Austria: Centre for Systematic Musicology, University of Graz. 389-394.

Behavioral and Neurophysiological Effects of Singing and Accompaniment on the Perception and Cognition of Song

Yke P. Schotanus[1], Lynn S. Eekhof[2], Roel M. Willems[2]

[1] *Institute for Cultural Inquiry, ICON, Utrecht University, The Netherlands*

[2,3] *Centre for Language Studies, Radboud University Nijmegen, Nijmegen, The Netherlands*

[1]schotschrift.teksten@planet.nl

Abstract

In order to further investigate the effects of singing and accompaniment on the processing of language, a classroom experiment reported earlier was followed by an EEG experiment, using the same materials. 24 participants listened to four songs, each in one of four versions: spoken, sung a cappella, complete (sung with accompaniment), or vocalized (sung a cappella on 'lala'). During listening, EEG was measured, and after each song, a questionnaire was filled out. Behavioral results suggest that singing supports cued word recall, even after just one exposure, and focus on the lyrics. Furthermore, an accompaniment supports positive affect and appreciation of voice quality, and decreases seriousness. A preliminary EEG data analysis reveals that out-of-key notes elicit a slightly larger ERAN and N400 than in-key notes, a smaller P2 and N5, and a larger P600, followed by a larger late negativity ('N1400'). However, the larger ERAN and P2 are not visible in all conditions; the larger N400 is only significant in the condition complete, the larger P600 predominantly in the condition vocalized, and the late negativity only in the condition a cappella. These differences lead to the conclusion that the processing of in-key and out-of-key notes interacts with the presence of interpretable lyrics, indicating that music might affect the meaning of words or vice versa. The interaction between the processing of these notes and the presence of an accompaniment is more difficult to interpret.

Introduction

The processing of song lyrics is thought to be affected by the song's music. On the one hand, music is thought to increase arousal and attention, which could support language processing. On the other hand, violations of both rhythmic and musical expectancies within the melody can hamper language processing (Kunert, 2017; Gordon et al., 2010), although they might also accentuate linguistic events (Schotanus, 2015).

To date, interactions between language and music have been tested with deviant notes or chords in short standardized stimuli only, not in complete songs in which notes or chords can have different functions, can be repeated, and followed by other unexpected notes. Furthermore, the role of accompaniment is seldom taken into account.

In a classroom study, in which 274 pupils reacted to different versions of four different songs, Schotanus (2016) has found that pupils rate sung lyrics as more intelligible and comprehensible than spoken lyrics, and that they rate accompanied song versions (either sung or spoken) both emotionally and aesthetically more positive, and easier to focus on than unaccompanied ones. Furthermore, their ratings of the emotional meaning of the lyrics were more in line with the author's intentions if they were sung and accompanied. Finally, the voice quality was rated less positive, and the singing less pure in song versions without accompaniment.

These results were interpreted to be the result of Musical Foregrounding, and thus to support the Musical Foregrounding Hypothesis (MFH; Schotanus, 2015), according to which music functions as a foregrounding device (Miall & Kuiken, 1984), obstructing the processing of sung words, and thereby accentuating them as well, if the listener is able to overcome the obstruction. In this study, the processing of difficulties in the music, especially on-beat silences or loud rests (London, 1993; Honing et al., 2009) and of out-of-key notes (ooks), was hypothesized to be more demanding in versions without accompaniment. In accompanied versions, loud rests would be absent, and harmonies would not just be implied but actually sound. As a result, foregrounding would be maximum in a-cappella versions, possibly too much so for the listener. Hence the enhanced concentration, tonal perception, and valence in accompanied versions, but also the enhanced processing of the lyrics in sung versions in general.

The aim of the current study is to replicate the results of the classroom experiment and to test whether there is neurophysiological evidence that a cappella versions are more demanding than accompanied ones, and that music affects the meaning of words.

Earlier research on the processing of ooks (and other unexpected tones or chords) has shown somewhat confusing results. In several studies (see Kim, 2013, for a review) both in-key notes (iks) and ooks elicit an Early right anterior negativity (ERAN; associated with a preconscious detection of an unexpected event) and an N5 (associated with the meaningful integration of this event). The amplitude of both ERAN and the N5 are usually larger for unexpected notes and chords, and can also be modulated by position. Furthermore, the ERAN for single notes tends to be earlier than the ERAN for chords. However, other researchers found a larger P2 instead of an ERAN (associated with the violation of memory-based expectations) for ooks (Choi et al, 2014), a bipolar ERAN (Sammler et al., 2012), or a right frontal N350 combined with a large P600 (Patel et al., 1998) (associated with the integration of syntactical anomalies). The P600 might sometimes mask an N5 (Koelsch et al., 2005).

Importantly, these ERPs have shown to be largely independent of musical experience, attention, and repetition, although these factors do influence the amplitudes. For example, both P2 and ERAN are usually larger in experienced musicians, and both ERAN and N5 are somewhat smaller if listening to the music is not the participant's main task, or if a particular note or note sequence is repeated during an experimental trial (Choi et al., 2014; Kim, 2013).

In line with these findings, we expected to find ERANS or P2s and N5s for iks, and the same ERPs with larger amplitudes, plus a P600 or an N350 for ooks. We expected the smallest amplitudes for complete (accompaniment effect) and the largest for vocalized conditions (attention effect).

In Parncutt, R. & Sattmann, S. (Eds). *Proceedings of ICMPC15/ESCOM10*. Graz, Austria: Centre for Systematic Musicology, University of Graz. 389-394.

Finally, the N400, associated with the processing of linguistic meaning, but sometimes affected by music as well (Koelsch et al., 2004) might be larger for ooks, but absent in vocalized versions, due to the absence of lyrics.

Method

Participants

24 participants (18 women) were recruited from the subject pool of the Radboud University Nijmegen to take part in this experiment in exchange for money (*n* = 22) or course credit (n = 2). Participants were 19 to 37 years old (*M* = 24;4, *SD* = 4.8) native speakers of Dutch without a history of neurological problems or hearing deficits. All participants except one were right-handed. They were either still in university (*n* = 21), or had finished university a few years prior to the experiment (*n* = 3). Self-reported musical sophistication, measured by the Gold MSI (Bouwer et. al, in preparation (translation); Müllensiefen et al., 2014) varied from 36 to 111 (*M* = 67.7; *SD* = 19.1 (scale maximum = 136)), musical experience (an MSI subscale) from 7 to 41 (*M* = 22,0; *SD* = 11.9 (scale maximum = 49)).

Materials

Ecologically valid stimuli, namely four complete cabaret songs in Dutch, were used for the experiment (all stimuli and questionnaires available online (Schotanus, 2017)). Each song contained several occurrences of ooks and (different kinds of) loud rests in the melody part. The frequencies of these events per song can be found in Table 1 below. For the EEG analysis of the ooks, an equal amount of iks of similar length was selected in similar metric positions (see Figure 1). Thus, possible confounding effects of metrical position, note length, musical context, and repetition are eliminated. All EEG markers were time-locked using Praat (Boersma & Weenink, 2015).

All songs were pre-existing but seldom-performed cabaret songs composed and sung by the first author, a male baritone. Using these songs ensured that there were enough ooks and loud rests in each song for EEG-measures and practically excluded the possibility that participants had prior knowledge of the songs. A disadvantage might be that the songs are written in the popular tradition which is not strictly tonal and in which several out-of-key notes are might not be perceived as out-of-key (Temperley, 2001). However, the musical expectations of Western people (musician or not) have shown to be largely in line with the rules of tonal music (see Huron, 2006, for a review). Furthermore, a classically-schooled musician has judged all the target notes as 'feeling out-of-key within the context'**.**

The recordings were made using a Neumann TLM 103 microphone and an Avalon VT 737 SM amplifier. In order to ascertain that all the singing was in tune, the voice-treatment software Waves Tune, Renaissance Vox compression and Oxford Eq were used to perform vocal pitch correction.

Table 1: distribution of musical foregrounding events over the four songs.

Event: / Song:	1	2	3	4
Out-of-key notes	38	29	36	36
Loud rest	94	79	66	55

Figure 1. Out-of-key notes and in-key-control notes in the verse part of one of the songs (WM).

Accompaniments were improvised according to the chord scheme of each song on a keyboard and recorded using ProTools 10 by Christan Grotenbreg, a professional musician.

All songs are of approximately the same complexity. In a separate online survey, 42 participants recruited via Amazon Mechanical Turk between 24 and 66 years of age (Mean 36.6; SD 10.5; 48.7 % male) rated the complexity of all accompaniments and vocalized voice parts. Mean complexity ratings ranged between 4.1 and 4.9 on a 7-point scale. A regression on these ratings with song, song version, and their interaction showed no significant effects. The effect of song was just trending to significance (*p* = .066). Lyrics were found to be of the B2-level according to the CEFR tool (Council of Europe, 2011; Velleman & Van der Geest, 2014; Stichting Accessibility, 2014). Nevertheless, there are clear differences in tempo and content, which might affect appreciation. Two songs are relatively upbeat (about 110 BPM) and have love-related lyrics, while the other two are somewhat slower (approximately 80 BPM), and have lyrics about dealing with life's threats. Hence, in our statistical analyses, we used random intercepts for songs.

All songs were recorded in four different conditions: spoken, a cappella, vocalized (sung a cappella on 'lalala'), and complete. In the complete versions, the sung lyrics are presented together with a harmonic accompaniment, with a musical event at every beat. The isolated voice part of this recording was used for the a-cappella version. Furthermore, both the vocalized and the spoken version are sung or recited in accordance with the timing of the complete version. That is, they were recorded while the 'singer' listened to the accompaniment, and started each phrase at approximately the same time as the sung lyrics would have started in the song. As a result, all tracks in all conditions had approximately the same duration (± 2 seconds), excluding a possible confound of rate of presentation when comparing the spoken and the musical conditions. However, the speech in the spoken version is quite slow and riddled with silences, which might create foregrounding effects in the spoken version as well.

In order to assess appreciation, perception, comprehension, and recall, questionnaires were designed. For each condition, there was a questionnaire with 7-point Likert scales assessing participants' judgments of the lyrics, music, performance, listening experience and voice of the singer. Furthermore, there were song specific questionnaires containing three very simple comprehension questions to assess whether participants had listened to the lyrics. In addition, there were several questions to assess comprehension of the lyrics (to be reported elsewhere), and a fill-in-the-blank test to measure verbatim recall. Six sentences were selected, evenly distributed over the full length of the song. In half of these sentences, a word that was supported by rhyme (alliteration, assonance or end rhyme) was left out, whereas in the other half neutral words were left out. The questionnaires specific to lyrics were not filled out after the vocalized condition.

Procedure

The experiment took place in a small, sound-proof booth with a desk and a computer screen in front of the participant from which the stimuli were played. Another screen was present in the booth but invisible to the participant, to conduct the recording of the EEG data. Stimuli were pseudo-randomly assigned to participants in such a way that every participant heard every song and every condition exactly once.

Before the experiment started, participants were asked to sign informed consent and fill out the Gold MSI and another questionnaire (not discussed here). During the experiment, songs were presented over Sennheiser HD 215 headphones using Presentation software (v. 17.1, Neurobehavioral Systems) at a sound level that was judged to be optimal by each individual participant after the presentation of several beeps. The start of each song was indicated with a beep and a fixation cross accompanied each song until the end in order to make sure participants would not erroneously think the song had finished during silences in the versions without an accompaniment.

After each song, participants manually filled out the questionnaires about the track they had just heard. The complete experiment, including the set-up of the EEG, took about 100 minutes.

Data recording and analysis

Behavioral data were analyzed in RStudio (version 1.1.383; R Core Team, 2017). For dimension reductions of the questionnaires, we used the *psych* and *GPArotation* packages (Revelle, 2017; Bernaards & Jennrich, 2005), for linear mixed models the *lme4* package (Bates et al., 2012).

EEG data were recorded in Brain Vision Recorder with a 64 channel system with active electrodes. Four electrodes were used for the vEOG and hEOG, and two electrodes were placed on the mastoids. An additional electrode was placed on the forehead as a ground. Data was recorded at a sampling rate of 500 Hz with a low cutoff filter of 0.016 Hz and a high cutoff, anti-aliasing filter of 1000 Hz.

All EEG data were analyzed using the Fieldtrip MATLAB toolbox (Oostenveld, Fries, Maris & Schoffelen, 2011). For analysis, the continuous EEG signal was cut into epochs starting 500 ms before and ending 2000 ms after the onset of the target note. The data were re-referenced to the mean of the two mastoids. In addition, detrending and a low-pass filter of 70 Hz were applied. After the manual rejection of any large and obvious disturbances of the signal, an Independent Component Analysis was run on the data to remove artifacts such as eye blinks and noise. We focused on time windows related to the components of interest. We compared iks and ooks in each condition between 100-175 (early ERAN), 175-225 (later ERAN and P2); 300-400 (N350 and onset N400); 400-500 (N5); and 600-900 (P600). Furthermore, we compared iks between conditions a cappella and the other conditions between 100-225; 300-400 and 400-500. Post hoc, after a visual inspection of the waveforms we tested the significance of a few unexpected but salient ERPs.

Results

We will present a preliminary analysis of a part of our data.

Questionnaires

After a Principal Axis Factoring analysis with oblique rotation (direct oblimin) on the Likert-scale items occurring in all questionnaires concerning song versions with text, four factors with eigenvalues over 1 were retained: Positive affect (representing joyfulness, sensitivity, energeticness, intelligibility, comprehensibility, and the opposite of boringness and tiringness); Seriousness (representing sadness, sensitiveness, heaviness, strikingness of formulations, and the ability to make you think); Voice quality (representing naturalness, relaxedness and pleasantness of the ‘singer’s’ voice); and Strikingness of textual features (representing funniness and, again, strikingness of formulations). Eigenvalues were 3.32, 3.01, 2.57 and 1.80 respectively.

For each component separately, a linear mixed model was constructed with condition as a fixed effect and random intercepts for participants and songs. Likelihood Ratio Tests were used to compare the models with and without the fixed effect of condition. This revealed that condition had a significant effect on all components: Positive affect ($\chi^2(2) = 14.58$, $p = 0.0007$), Seriousness ($\chi^2(2) = 14.90$, $p = 0.0006$), Voice quality ($\chi^2(2) = 7.08$, $p = 0.03$) and Textual features ($\chi^2(2) = 15.80$, $p = 0.0004$).

Post-hoc Tukey tests were performed to compare the means of the three conditions. For Positive affect, the complete versions were scored significantly higher than the spoken ($p = 0.001$) and a-cappella versions ($p < 0.001$). Spoken and a cappella did not differ significantly ($p = 0.99$). For Seriousness, the score for the complete versions was significantly lower than the score for the spoken one ($p < 0.001$). The score for the a-cappella versions was higher, but not significantly higher, than the score for the complete ones ($p = 0.096$). Similarly, the score for the a-cappella versions was lower, but not significantly lower, than the score for the spoken ones ($p = 0.082$). For Voice quality, the score for the complete versions was significantly higher than the scores for the spoken ($p = 0.041$) and a-cappella versions ($p = 0.047$). Again, the difference between the spoken and a cappella versions was not significant ($p = 0.999$). Finally, for Strikingness of textual features, the complete versions were scored significantly higher than the spoken ones ($p < 0.001$).

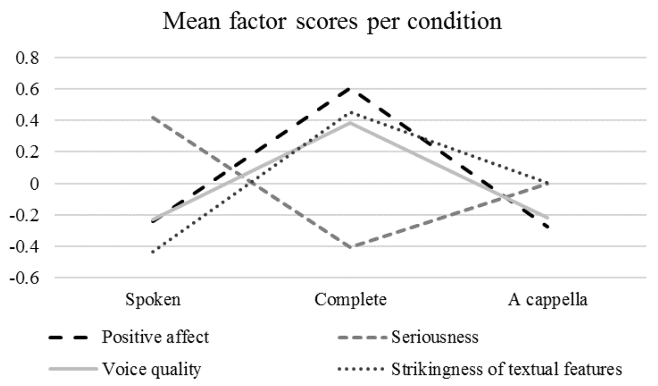

Figure 2. Mean factor scores per condition.

Compared to the complete versions the a-cappella versions again scored lower but this difference only approached significance (p = 0.07). The same was observed for the difference between the spoken and the a-cappella versions (p = 0.08).

Recall

Figure 3 shows the mean scores per condition for the fill-in-the-blank recall test. As there were six recall trials per song, the maximum total score was 6. Of these six trials, half were supported by rhyme, and half were not. One song trial was excluded because the participant scored below chance level on the comprehension check questions (score < 2).

A linear mixed model was fitted to the recall data to assess the effect of condition on total recall. Condition was added as a fixed effect, with random intercepts for participants and songs. A Likelihood Ratio Test showed that condition had a significant effect on total recall ($\chi^2(2) = 7.20$, $p = 0.027$). A post-hoc Tukey test revealed that only the difference between the spoken and the a-cappella versions was significant (p = 0.015), with superior recall in the a cappella condition. The differences between the spoken and the complete versions (p = 0.41) and between the complete and the a-cappella versions (p = 0.28) were not significant, neither was the effect of rhyme ($\chi^2(1) = 0.14$, $p = 0.71$).

ERPs

Iks. The ERPs elicited by iks do not differ significantly. They all show an ERAN somewhat earlier than expected (around 100 ms after target onset (to), see figure 4), followed by a P2 and an N5 peaking relatively late in the conditions complete and a cappella. We did not expect this difference and hence we did not measure the significance of it yet. In our target time windows, 300-400 and 400-500 ms after to, there were no significant differences between a cappella and complete or a cappella and vocalized. However, in the time window 100-225 ms after to four electrodes show a significant negativity for vocalized compared to a cappella, representing a larger and later ERAN for vocalized ($p < 0.05$, see figure 4).

Ooks. The ERPs elicited by ooks show significant differences compared to those elicited by iks in all conditions. However, these differences vary per condition and are often visible in another time window than expected. Furthermore, most differences are significant only on a 0.05 level which might cause problems with multiple comparisons. Nevertheless, it is interesting to report them.

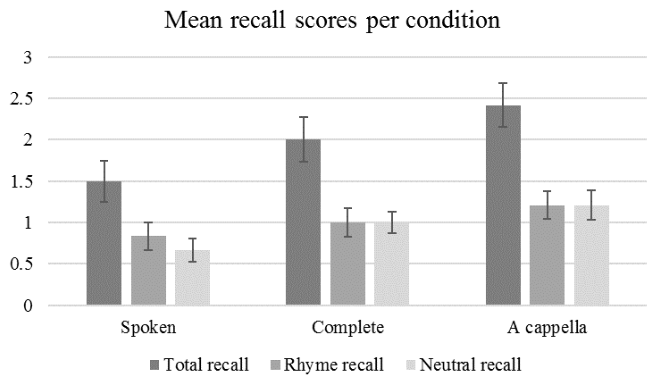

Figure 3. Mean recall scores per condition.

Time window 100-175; 175-225; 100-300. In the two time windows associated with the ERAN, the only difference which reaches significance is a small one between a cappella iks and a cappella ooks in the time-window 100-175. Three electrodes show a significant negativity for ooks ($p < 0.05$, see figure 4). However, visual inspection of the waveforms revealed that the ERANS often peak earlier and the P2s later. Hence, a post hoc analysis of the time windows 100-300 showed a significant negativity for all ooks compared to all iks ($p < 0.05$), which seems mainly due to a cappella ooks, where the P2 is almost absent.

Time window 300-400; 400-500. In the time-window 300-400, the waveforms of both a cappella and complete show a negative tendency for ooks. This difference is significant only for 5 centro-parietal electrodes in the condition complete ($p < 0.05$, see Figure 4, bottom). In the condition vocalized on the other hand, ooks show a positive tendency in this time window, which becomes significant between 400 and 500 on eleven centro-parietal electrodes ($p < 0.05$, see figure 4, bottom). In the conditions a cappella and complete on the other hand ooks show a positive shift in this time window.

*Time windows 600-900; 500-1000***.** Within the time window 600-900 all conditions together ($p < 0.05$), and the condition vocalized ($p < 0.05$) show a significant positivity for ooks across the whole scalp. In the complete condition 3 electrodes show such a significant positivity ($p < 0.05$). As the difference curves indicated that the P600 peaks either earlier or later (partly due to a late N5 peak) we also tested the time window 500-1000. In this time window, the positivities for both vocalized and all conditions together are significant at a higher level ($p < 0.005$). A cappella shows both positivities and negativities at various electrodes, hence on average there were no significant effects.

Time windows 1350-1450, 1350-1800. In all conditions, the ook-curve shows a late negative shift. In the a-cappella condition, this leads to a significant negativity peaking around 1400. This negativity is significant over the whole scalp between 1350 and 1450 ($p < 0.05$); and at 8 electrodes between 1350 and 1850 ($p < 0.05$), see figure 4 upper half.

Discussion

The effect of musical context and musical complexity on the processing of song lyrics was tested in an experiment with both offline behavioral measures and online EEG measures. The results of important parts of the data confirm and extend the results of an earlier classroom study, and might be interpreted as support for the MFH.

In Parncutt, R. & Sattmann, S. (Eds). *Proceedings of ICMPC15/ESCOM10*. Graz, Austria: Centre for Systematic Musicology, University of Graz. 389-394.

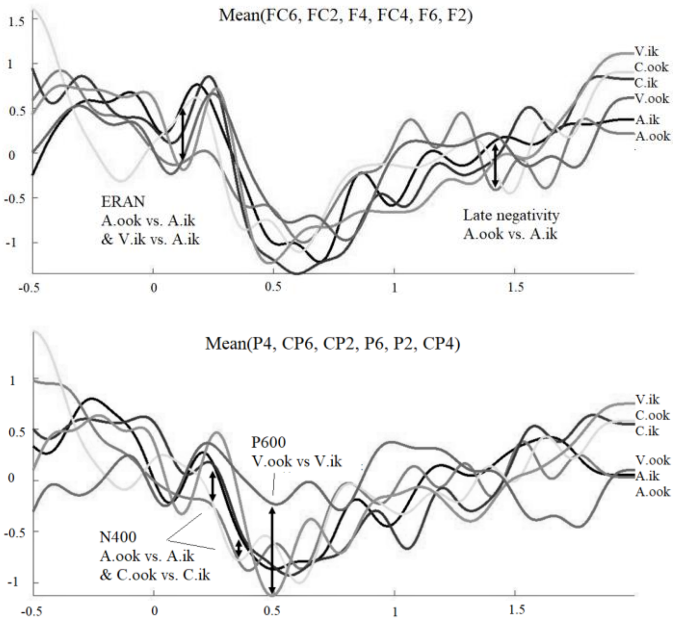

Figure 4. Grand average waveforms of ERPs elicited by iks and ooks at right-anterior (above) and right-centro-parietal electrodes (below) in the conditions: Vocalized (V), A cappella (A) and complete (C). Arrows show salient early and late right-anterior negativities (above), N400s and P600s (below).

Behavioral results

Complete versions were scored significantly higher on Positive affect, Appreciation of voice quality and Strikingness of textual aspects, and significantly lower on Seriousness than their spoken counterparts. Furthermore, recall was significantly better after hearing a cappella version than a spoken version, which proves that music can support recall without repetition. These results directly support the MFH's claim that music, via musical foregrounding, can have beneficial effects on the listener by increasing appreciation for and perception and cognition of song lyrics.

The results also seem to confirm the complex predictions concerning a cappella singing. We hypothesized that the foregrounding is maximum in a cappella music which might make hard to overcome. And indeed in the condition a cappella the scores for Appreciation of voice quality and Positive affect are lower than in the condition complete. At the same time, however, cued word-recall is enhanced, which indeed indicates a maximum effect of foregrounding. Furthermore, the scores for the text centered factors Seriousness and Strikingness of textual features indicate a slightly enhanced focus on the lyrics in this condition.

One important difference with the classroom study is that in the current study intelligibility and comprehensibility (worked up in Positive affect) do not seem to be enhanced in a cappella singing compared to speech. The fact that it was enhanced in the classroom study because a classroom is much more demanding for intelligibility and concentration than a soundproof booth.

ERPs

In spite of differences in musical-syntactic saliency, length and metric position of ooks, and of the repetitive structure of the music, the ERPs elicited by ooks differ on average from those elicited by iks. This indicates that even in a musical context which is not strictly tonal, the schematic expectations of Western listeners concerning tonality are stronger than veridical expectations. However, while the ERPs elicited by iks are roughly similar across conditions, the differences between the ERPs elicited by iks and ooks are not. These results are just partly in line with our predictions.

Iks. We hypothesized the occurrence of an ERAN and an N5 in all conditions. Furthermore, we expected the ERAN in the condition complete to be later and smaller than in the unaccompanied conditions, and largest in the condition vocalized. At the same time, the N400 was hypothesized to be smaller in this condition. Indeed, we did find an ERAN and an N5 in all conditions, however, the only significant difference we found, was a larger ERAN in vocalized compared to a cappella, which is in line with the expectation that the presence of lyrics reduces attention for the music.

Ooks versus iks. We expected larger ERANs, P2s, N400s, N5s, and P600s for ooks than for iks in all conditions, and possibly an N350. Furthermore, we expected the amplitudes of these ERPs to be smallest in the condition complete, and largest in the condition vocalized, although the N400 was expected to be the lowest in that condition.

ERAN. Just a cappella ook shows a significantly larger ERAN. Furthermore, unexpectedly, the P2 is reduced in the conditions a cappella and complete, resulting in a significant negativity for ooks in general between 100-300ms after to, which might be interpreted as part of a build-up towards the N400.

N350, N400, N5. In the time window 300-400 both a cappella and complete show negativities, while vocalized does not. However, only the negativity in the condition complete is significant. Nevertheless, the waves for both a cappella ook and complete ook show a clear peak around 375 ms after to., As this peak is more salient in centro-parietal electrodes, instead of right anterior electrodes (where Patel (1998) found the N350), we think it might rather be interpreted as a larger N400 than as an N350. In the time window 400-500, which is the onset of the N5 we found an unexpected positivity for ooks in all conditions. Although this positivity is significant only in the condition vocalize, it is unmistakable that the ooks curve in a cappella and complete has taken a positive turn in this time window as well. This is a puzzling result. It seems to indicate that ooks are interpreted as linguistically but not musically meaningful and that the onset of the P600 reduces the N5 immediately.

N5, P600. All conditions together, and the conditions complete and vocalized separately, show a significant P600 for ooks though it is just marginally significant in the condition complete. What is puzzling, however, is that at several electrodes this P600 does affect the onset but not the amplitude of the N5 at its ultimate peak (around 600ms after to in the condition complete, and around 750ms in the condition vocalized ook). As a result the P600 peaks at unexpected moments. Possibly, if both P600 and N5 are relatively strong. This is the result.

Late negativity. It is unclear what kind of neural activity the late negativity visible predominantly in the condition a cappella, and possibly somewhat late in the condition vocalized represents.

Overall. Obviously, the absence of interpretable lyrics plays a part in the deviant processing of vocalized ooks. However, it is striking that it did not affect the processing of vocalized ik in such a salient way as well. Possibly, as expected, the focus on the language has alleviated the effect of the ooks in the conditions a cappella and complete. However, the effect of ooks is not just smaller in these conditions, above all it is different. Notably, the ERAN for ooks was largest in the condition a cappella, while the N400 for ooks was larger in the condition complete. Because of the N400 for ooks in a cappella and complete, it is more reasonable to think that the ooks are interpreted as more meaningful in the presence of language.

Conclusion

Both behavioral and EEG data show significant differences between musical conditions. The behavioral data support our hypothesis that singing would enhance lyric processing. It even enhances cued word-recall after a single exposure of a song. Furthermore, as expected, an accompaniment enhances the appreciation of voice quality and positive valence and suppresses seriousness. In line with our expectation that music might affect the meaning of words, the processing of iks and ooks seem interacts with the presence of lyrics. Ooks seem to be interpreted as being more meaningful in conditions with lyrics. Evidence for our expectation that an accompaniment would ease the processing of lyrics is more ambiguous, although the ERAN for ooks in the condition a cappella is larger than in the condition complete. Finally, it is remarkable that several components (ERAN, N400, P600) seemed to be earlier than expected, while others (P2, N5) were later.

Acknowledgments. The first author would like to thank his supervisors for their comments; his colleagues from UiL OTS for their help with Lime Survey; Christan Grotenbreg for recording the stimuli and digitally manipulating the recordings; and NWO, the Dutch Government, and SG het Rhedens for granting him the opportunity to avail himself of a PhD scholarship for teachers. The current paper is partly based on an internship project of the second author.

References

Bates, D., Maechler, M., Bolker, B. & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48.

Bernaards, C. A. & Jennrich, R. J. (2005). Gradient Projection Algorithms and Software for Arbitrary Rotation Criteria in Factor Analysis. *Educational and Psychological Measurement, 65*, 676-696.

Boersma, P. & Weenink, D. (2015). *Praat: doing phonetics by computer* [Computer program]. Retrieved, 28 October 2015 from http://www.praat.org/

Bouwer, F., Schotanus, Y. P., Sadakata, M., Mullensiefen, D & Schaefer, R. (in preparation) *Measuring musical sophistication in the low countries; validation of a Gold MSI translation in Dutch*.

Choi, I., Bharadwaj, H. I., Bressler, S., Loui, P., Lee, K. & Shinn-Cunningham, B. G. (2014) Automatic processing of abstract musical tonality. *Frontiers in Human Neuroscience,* 8, 988.

Gordon, R. L., Magne, C. L. & Large, E. W. (2011). EEG correlates of song prosody: a new look at the relationship between linguistic and musical rhythm. *Frontiers in Psychology,* 2, 352.

Honing, H. J., Ladinig, O. Winkler I. & Háden G. (2009). Is beat induction innate or learned? Probing emergent meter perception in adults and newborns using event-related brain potentials (ERP). *Annals of the New York Academy of Sciences*, nr. 1169: *The Neurosciences and Music: Disorders and Plasticity*, p. 93-96.

Huron, D. (2006) *Sweet anticipation*. Cambridge, MA: MIT Press.

Kim, J. N. (2013) *Online processing of tonal melodies: effects of harmonic expectations* (doctoral dissertation). Evanston, IL.

Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T. C., Friederici, A. D. (2004). Music, language, and meaning: Brain signatures of semantic processing. *Nature Neuroscience 7*, 302-307.

Koelsch, S., Gunter, T. C., Wittfoth, M. & Sammler, D. (2005). Interaction between syntax processing in language and in Music: an ERP study. *Journal of Cognitive Neuroscience*, 17(10), 1565-1577.

London, J. M. (1993) Loud rests and other strange metric phenomena (or, meter as heard). *Music Theory online*. Http://mto.societymusictheory.org/issues/mto.93.0.2/mto.93.0.2london.art.

Miall, D. & Kuiken, D. (1994). Foregrounding, Defamiliarization, and Affect Response to Literary Stories. *Poetics*, 22, 389-407.

Müllensiefen, D., Gingras, B., Musil, J., & Stewart L. (2014). The Musicality of Non-Musicians: An Index for Assessing Musical Sophistication in the General Population. *PLoS ONE,* 9(2). Questionnaire retrieved from https://www.gold.ac.uk/music-mind-brain/gold-msi/download/.

Oostenveld, Oostenveld, R., Fries, P., Maris E. & Schoffelen J-M. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience, 11*.

Patel, A., Gibson, E., Ratner, J., Besson, M. & Holcomb, P. J. (1998). Processing syntactic relations in language and music: an event-related potential study. *Journal of Cognitive Neuroscience,* 10(6), 717-733.

R Core Team (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Revelle, W. (2017). Psych: Procedures for Psychological, Psychometric, and Personality Research (version 1.7.8) [Computer software]. Evanston, IL: Northwestern University.

Sammler D., Koelsch, S., Ball, T., Brandt, A., Grigutsch, M., Huppertz, H.-J., Knösche, T. R., Wellner, J., Widman, G., Elger, C., Friederici, A. D., Schulze-Bonhage, A. (2012). Co-localizing linguistic and musical syntax with intracranial EEG. *NeuroImage*, 64, 134-146.

Schotanus, Y. P. (2015) The musical foregrounding hypothesis: How music influences the perception of sung language, Ginsborg, J., Lamont, A., Philips, M. & Bramley, S. (Editors) *Proceedings of the Ninth Triennial Conference of the European Society for the Cognitive Sciences of Music*, 17-22 August 2015, Manchester, UK.

Schotanus, Y. P. (2016/submitted) *Both singing and accompaniment support the processing of song lyrics and change their meaning.* (Based on) Y.P. Schotanus, Music supports the processing of song lyrics and changes their contents: Effects of melody, silences, and accompaniment (Abstract), In: *Proceedings of the 14th International Conference on Music Perception & Cognition (ICMPC14*), July 2016, San Francisco. Zanto, T.P. (Ed.). Adelaide: Causal Productions, p. 361. (Full text available on Researchegate and AcademiaEdu, and in Schotanus, 2017).

Schotanus, Y.P. (2017). Supplemental materials for publications concerning three experiments with four songs, hdl:10411/BZOEEA, DataverseNL Dataverse, V1

Temperley, D. (2001) *The cognition of basic musical structures*. Cambridge, MA/London, UK: MIT.