Creating Meaningful Melodies from Text Messages

Florian Alt, Alireza Sahami Shirazi, Stefan Legien, Albrecht Schmidt
University of Duisburg-Essen
Pervasive Computing and User Interface Engineering
Schützenbahn 70, 45117 Essen, Germany
{florian.alt, alireza.sahami, albrecht.schmidt}@uni-due.de
stefan.legien@stud.uni-due.de

Julian Mennenöh
University of Duisburg-Essen
Marketing and Retail
Universitätsstraße 12, 45141 Essen, Germany
julian.mennenoeh@uni-due.de

ABSTRACT
Writing text messages (e.g. email, SMS, instant messaging) is a popular form of synchronous and asynchronous communication. However, when it comes to notifying users about new messages, current audio-based approaches, such as notification tones, are very limited in conveying information. In this paper we show how entire text messages can be encoded into a meaningful and euphonic melody in such a way that users can guess a message's intention without actually seeing the content. First, as a proof of concept, we report on the findings of an initial online survey among 37 musicians and 32 non-musicians evaluating the feasibility and validity of our approach. We show that our representation is understandable and that there are no significant differences between musicians and non-musicians. Second, we evaluated the approach in a real world scenario based on a Skype plug-in. In a field study with 14 participants we showed that sonified text messages strongly impact on the users' message checking behavior by significantly reducing the time between receiving and reading an incoming message.

Keywords
Sonority, text sonification, instant messaging, SMS

1. INTRODUCTION
SMS, instant messaging, and email are ubiquitous communication channels, which are widely used. To make users aware of incoming messages, communication tools use different types of notification. While synchronous communication tools (e.g., chat clients) mainly use visual clues (highlighting the application's window), asynchronous communication tools (e.g., email clients or mobile phones) often make use of audio notifications. However, such notifications do neither convey the content nor the intention of a message. Recent work focused on augmenting audio notifications with information about the sender, e.g., by associating a specific ring tone with a caller. In this work, we present an approach, as to how notification can be used to convey more detailed information about a message's content such as its intention, included keywords, or the precise wording.

Many communication applications and platforms, on mobile devices as well as on PCs allow for intercepting incoming messages, analyzing the content, and determining a message's intention based on included statements, emoticons, keywords, and punctuations. Instant messaging clients (e.g., Skype), mail clients (e.g., Gmail), and mobile phones (S60 series) offer APIs (Application Programming Interface) to process incoming messages. Hence, a musical representation of a message can be created in such a way, that the message's intention is indicated to the user while at the same time preserving his privacy in contrast to reading out the message loudly.

We investigate if users can understand a message's intention and how this impacts on their message checking behavior. Our results show that users check messages more quickly if they understand the content type. Further, people might eventually even be able to not only guess a message's intention but to understand the whole content by learning the musical representation of words frequently used. The past time success of Morse code, which can be used both for visual and audio encoding of messages, shows the feasibility of such approaches.

This paper makes the following contributions: (1) We present a concept and algorithm for the transformation of text messages into euphonic melodies in such a way that the intention of a message can be communicated without reading the message. (2) We report on a survey among 69 participants evaluating the feasibility and validity of the approach. (3) In a two-week field study among 14 participants we explored the impact of a message's sonification on the users' behavior in the real world based on both qualitative and quantitative findings.

2. RELATED WORK
Communication of information in non-verbal ways has been subject to research for more than 150 years, one of the most prominent examples being the invention of the Morse code in the early 1840s. The popularity of this rhythm-based character encoding system can be explained by its ability to be read by humans without any decoding device and its high learnability.

In the following we focus on more recent research projects on sonification that is the use of non-speech audio to convey information. An important research field is the sonification of different types of data. MUSART is a sonification toolkit, which produces musical sound maps to be played in real-time [9]. Walker et al [17] presented the Audio Abacus, an application for transforming numbers into tones. The Sonification Sandbox [16] allows users for creating auditory graphs from several sets of data. An important application area is sonification for blind and visually impaired users. In this context Petrucci et al [11] showed how to use sonification in auditory web browsers to allow visually impaired users to explore spatial information by means of an audio-haptic interface.

Quite a number of research projects focused on the sonification of synchronous and asynchronous messaging. In [1] a commercial IM client is augmented with a tool that estimates the type of

NIME2010, June 15-18, 2010, Sydney, Australia.

instant messages and modifies the salience of those, which deserve immediate attention. The results reveal that modifying notifications can create a benefit for the users. Issac et al [8] introduced Hubbub, a sound-enhanced mobile instant messaging client, aiming at increasing background awareness by providing audio clues. In [14] the impact of abstracted audio preview of SMS is investigated. The authors report that this way of previewing affects the reading and writing behavior of users – however the approach focuses on a simple notification tone only. Shake2Talk [4] is a mobile messaging system that allows users for creating messages through gestures and send them to each other via their mobile phones.

Another research area is *aesthetics of sonification*. Inspired by the AIM (Arts in Multimedia) project, Babble online [6] sonifies browsing activity, trying to communicate information both clearly and in a well-composed and appealing way. Song et al [15] present mapping strategies derived from an analysis of various sound attributes, allowing for better representing and accessing information from complex data sets.

Finally, research has looked into the creation and functionality of *auditory icons*, so-called earcons, in computer interfaces [5]. Brewster [2] studied the use of earcons and evaluated whether they provide effective means for communicating information.

In contrast to our approach the presented projects focus either on auditory clues or the transfer of very small chunks of information rather than on the content of a message itself.

3. FROM A MESSAGE TO A MELODY

Text strings can be easily converted into a melody by mapping characters to tones. However, such a trivial approach completely ignores the music's power to express feelings and emotions and to confer intentions. In the following we show how to encode more than just characters into a melody.

3.1 Sonority

A study carried out in [13] revealed that users of mobile phones intending to send a melody to a friend or partner, do care a lot about how the message is going to sound like on the receiver's phone. Hence, the foremost task when transforming a text message to a melody is to define the mapping of characters to tones in such a way that a harmonic melody is created from whatever message. To do so, we map our tones to a pentatonic scale. A pentatonic scale can be created by combining five quint-related tones, meaning that one selects a tonic keynote and takes its four neighbors (in clockwise order) on the quint circle. Figure 1 gives as an example for a pentatonic C major scale consisting of the notes C, D, E, G, and A. Thus, a euphonic melody can be created from arbitrary text strings. Pentatonic scales can also be created in minor (a very prominent example is Gershwin's *Summertime*, based on a F# minor pentatonic scale).

In order to enhance the quality of our melody further we decided not to just randomly map characters to the pentatonic scale but additionally considered the frequency and the position of a character in the text. In the German language, each syllable contains at least one vowel. Thus, we decided to map vowels to the three tones of the tonic keynote's triad, also taking into account the average occurrence probability – hence *e* as the most frequent vowel (17.40%) is mapped to the tonic keynote, *i* (7.55%) and *a* (6.51%) to the third, and *u* (4.35%) and *o* (2.51%) to the quint. Further, for the consonants, we analyzed the frequency as an ending character. Based on the results we mapped the most frequent ending characters *n* (21.0%) and *r* (13.0%) to the tonic keynote, *t* (10.3%) and *s* (9.6%) to the third. The other consonants are mapped randomly.

Figure 1: Quint circle and C major pentatonic scale.

Character	Note (C major)	Note (a minor)
H	d	b
E	c	a
L	e	c
O	g	e
!	c / e / g	a / c / e

Table 1: Example for the mapping of a sample word in C major and a minor tonality

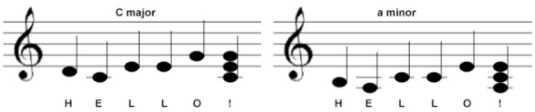

Figure 2: Melodies in C major and a minor

3.2 Intention of a Message

Messages are sent for specific purposes, such as for coordination, exchanging information (positive, negative), or expressing feelings and emotions (happy, sad). To take this into account, we analyze the content of each message for the occurrence of punctuation marks, keywords and emoticons. We use a simple mechanism to detect a message's intention and then accordingly transform it into a melody. To reduce complexity, we focus on the distinction between major and minor scales only**.** Table 1 gives an overview on the mapping of different intentions, given the phrase "HELLO". One option would be to not only transform the music into major and minor but also consider the association of certain keys with specific moods (e.g., flat / sharp keys) – however key-mood associations are invalid for modern (digital) equal temperament keyboards [12].

3.3 Punctuations

Punctuations are not only used to indicate the end of a sentence but also to specify the type of the statement. Sahami et al [14] showed how question marks can be used for creating abstracted audio previews in SMS. We consider both aspects in the following way: whenever a sentence ends with a punctuation mark, we insert a predefined triad into the melody. We use the triad's type to indicate the type of sentence (Table 2). Whereas we use a seventh chord to represent a question, regular triads are used for a point or exclamation point – however, triads are adapted to the intention of the message (major or minor).

3.4 Greeting and Leave-Taking Phrases

Text messages often begin with a greeting (hi, hey) and end with a leave-taking phrase (cu, regards). Missing phrases indicate that other messages were previously sent back and forth. Since the beginning often already conveys the message's intention, we also transform this phrase into a chord.

We deliberately used a simple mechanism to detect the message's intention. There is a body of work that looks into text mining and text understanding - however this was not the center of our research. By having a simple mechanism, we hope to increase the learnability of the sonification and the relation to messages. More sophisticated approaches for understanding intentions could be included in the algorithm easily.

	Characters	Intentions	Mapping
Emoticons	:-) :) ;-) ;)	positive	C
	:-(:(	negative	a
Punctuation (selection)	?	interrogative	C^7
	! .	declarative	C
Keywords (selection)	when/where	question	C^7
	yes, ok	positive	C
	no, sorry	negative	a

Table 2: Intention and according mapping of emoticons, punctuations, and keywords to chords (C: C major, a: a minor, C7: C major seventh chord)

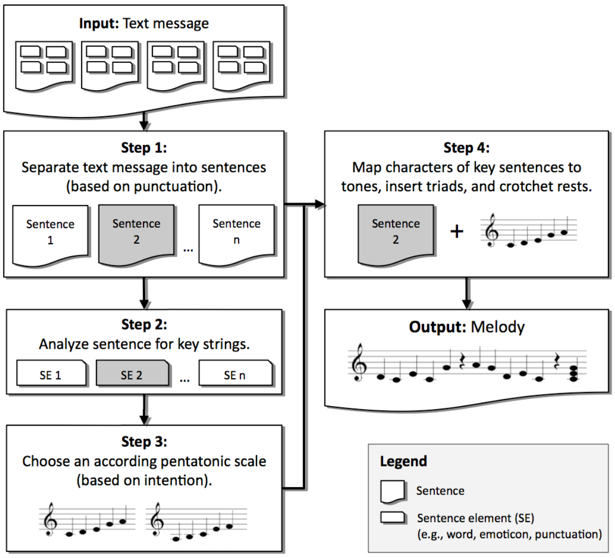

Figure 3: Visualization of the Sonification Algorithm

4. SONIFICATION ALGORITHM

Based on the approach described in the previous chapter, we use the algorithm depicted in Figure 3 to create a melodic representation from arbitrary message strings.

The algorithm takes a message string as input and separates it into sentences by analyzing it for punctuation marks (Step 1). Each sentence is analyzed for hints (key strings) that reveal its intention. Such hints include emoticons, keywords, and punctuation marks (Step 2). Based on this analysis we choose a corresponding pentatonic (major or minor) scale (Step 3) to which later the single characters of each word are mapped (Step 4). Besides mapping single characters to keys, we also create triads and tetrads for keyword, emoticons, and punctuation marks, varying between root position and 1^{st} and 2^{nd} inversion, depending on the intention. Spaces are transformed into crotchet rests.

The current version of our script supports text sonification in German language only. However, this approach could be extended to other languages. For a comparable sonification, an analysis of the language as explained above is required.

5. PROOF OF CONCEPT

In an initial online study we tried to prove the feasibility and validity of our approach. Therefore we implemented an AJAX-based web application, which reads a text message, sends it to a PHP-based sonification script and creates a local MIDI file from the returned XML code. The local MIDI file can then be played back using a media player (e.g., Flash or Quicktime).

The scope of the survey was to reveal whether our approach allowed the participants to understand the intentions of *the encoded messages*. We formulated the following hypotheses.

H1: Users can guess form hearing the melody only, if a message contains positive content, negative content, or a question.

Users may be able to determine a message's intention. This might affect their behavior in that they want to check certain messages immediately whereas they want to finish their current task first before checking other messages' content

H2: Users with a musical knowledge will learn the message intention faster.

We expected people with medium/strong musical skills to experience less difficulty understanding our musical representation.

5.1 Online Survey

We ran an online survey in summer of 2009 over a period of four weeks. People were recruited from music forums, mailing lists, facebook and from university mailing lists.

First, we collected *personal data* on gender, age, profession, and musical knowledge. We asked if the participants played any instrument and had them rate their musical skills on a 5-Point Likert scale (1=Beginner, 5=Professional). Second, we were interested in the *understandability of our mapping* and in the users' *association between intention of the message and sound of the melody*. Therefore we auditioned the sonification of 3 real text messages encoded with our tool to the participants. We used a piano melody for the representation. The melodies were 10-12 seconds in length. The 3 melodies included the sonification of one message with positive content, one with negative content, and one question (random order). We then asked the users (1) to associate the melody with one of three provided text messages ("no answer" was also an actual choice), hence revealing which intention, the melody messages in their opinion had, and (2) if the melody for them sounded happy, neutral, or sad. Third, we let people try out the algorithm with their own messages using a web application. We asked them for their personal opinion, privacy concerns, and if they would use the tool.

5.2 Results

In total, 69 persons completed our online survey (54 male) with an average age of 27.7 years. The participants were mainly students (40) and employees (23). 37 of the participants played a musical instrument, the most popular instruments were piano / keyboard (17), guitar (15), drums (7), and base (6).

5.2.1 Interpreting a Message Intention

As depicted in Figure 4 we found out that questions were correctly interpreted by 65.2% of the participants. Messages containing positive (40.6% correct answers) and negative content (43.5% correct answers) were more difficult to distinguish. However, these results are well above the 25% random choice.

For those participants who could correctly link the played sounds to a message we further investigated to which general intention they linked the melody (e.g., did they associate a happy sounding mapping based on a major scale with positive content, a sad-sounding mapping based on a minor scale with negative content). Here the results show that 83.3% considered questions to sound neutral, 84.6% considered the negative messages to sound sad, and 68.2% considered positive messages to sound happy. This is a strong indicator that people who are able to distinguish between negative messages, positive messages, and questions associate the sonification in the way we intended, hence making it very understandable.

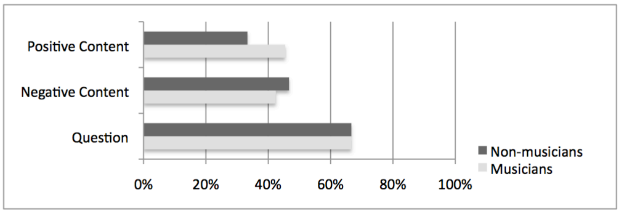

Figure 4: Comparison of correct answers between musicians and non-musicians.

			% correct answers per group	
	χ2-Value	p	non-musicians	musicians
question	0.004	0.940	0.656	0.649
neg. message	0.28	0.597	0.469	0.405
pos. message	0.235	0.628	0.375	0.432

Figure 5: Comparison of understandability among musicians and non-musicians (χ2-Test)

5.2.2 Musicians vs. Non-Musicians
Answers between musicians and non-musicians did not reveal a significant difference. Figure 4 shows that for all melodies both groups produced comparable results. To test the hypothesis, that the understandability is not influenced by the experience in playing an instrument, we used a Pearson's χ^2-test of independence for each message type. To compare the overall mean of correct answers, we used an Analysis of Variances (ANOVA).

The ANOVA shows, that the musicians' mean of correct answers (2.54) is lower than the other group's mean (2.81). This difference is not significant (p = 0.34, F-Value 0.924, df 1; 67). Hence, it is likely that a random effect caused the differences.

The χ^2-test of independence for each question is based on 2x2-matrixes, to compare correct and wrong answers for each question. The lowest amount of cases in one matrix field was 9. For valid results a value of at least 5 is required ("rule of five"). The significance level (Figure 5) shows, that giving the correct answer never depends on the fact, that a person is musician or not.

5.2.3 Qualitative user feedback
We got numerous interesting hints and suggestions for improvement from the participants. Suggestions included the adaptation of different music genres, different instruments (which could match the receiver's music taste, be mapped to the gender, or be used to distinguish different intentions), adding variations in the tempo of the music, and inclusion of sequences from popular songs indicating the intention. Several users stated that they would prefer shorter musical representations.

5.3 Discussion
The results show that almost half of the survey's participants could, without any learning, understand the intention of a message. Most understandable were questions, followed by negative and positive messages. The majority of the participants associated a message sonification with the envisioned intention. This is a strong indicator supporting *H1*. We believe that the understandability can be significantly enhanced, if people use message sonification over a longer period of time. We discovered that musicians did not perform better than non-musicians. Since both groups perform similarly we reject *H2*.

6. FIELD STUDY (skypeMelody)
The online survey gave us a good understanding of the issues and challenges related to message sonification. Yet, no evidence was found that results are either representative or true in the real world. Hence we implemented skypeMelody, a Skype plug-in, which we tested in a real world setting and conducted a study over the course of two weeks.

6.1 Prototype
The Java-based Skype Plug-in is implemented in a similar way as the web application used for the online study. It reads incoming text messages, transforms them into XML-conform MIDI and plays back the melody. For intercepting incoming text messages, users have to once authorize the connection to Skype.

With regard to the users' comments from the online study, we decided to decrease the length of the sonified message. Many participants stated that the message representations were too disruptive when played in full length. However, since we believe users might be able to mid-term learn to understand an entire message, we did not simply crop the message in length but filtered only those sentences including no keywords.

We defined in total 24 keywords. The keywords were derived from the online study were we asked the participants to enter each 2 short text messages including a question, a positive content, and a negative content. In total we analyzed 414 messages.

To gather quantitative data we also included a logging functionality into the Skype plug-in, which allowed us for storing certain types of information in an external database.

6.2 Study Design
In the field study we mainly focused on changes in the user behavior and the understandability as well as learnability of our representation. Users were asked to install skypeMelody and continue their regular Skype behavior (no specific tasks were given during the study). We used an initial questionnaire to evaluate demographics, text-messaging behavior (amount of conversations per week, average length of conversations, communication partners and situations), and musical experience.

To measure how the use of skypeMelody influenced their behavior we asked the users to fill in questionnaires after each week. In these questionnaires we were mainly interested in how easily users could distinguish different types of messages and if they checked incoming messages quicker or later than normally. Additionally users had to fill in a system usability scale (SUS).

6.3 Data Collection and Cleaning
For each message we stored two time-stamped records each including the hashed user id and a message id to later associate both records with each other. We used the *Received Event Record* for storing the current status of the user (online, away), the message type (type 1: positive, type 2: negative, type 3: question), the amount of keywords, the message length, and a list of all enrolled key words. The *Read Event Record* consisted of the reading time (allowing for calculating a receive-to-read time) and the message id (required for associating both records).

In total we collected 2533 data records. For consistency reasons we excluded 66 records were only the read event was registered, and 119 where more than two events occurred per each message (Skype's message IDs are not unique). We also performed a semantic check of the data. We excluded one message containing all 24 keywords, and 23 messages were the recipients' user status was set to "AWAY" so that we could not be sure that they received the sonified message. We finally removed 64 outlier messages where the receive-to-read-time lay beyond a threshold of 120 seconds (assuming read-events after more than 120 seconds were not caused by the sonification).

In the end 2170 records (representing 1085 messages) were used for the analysis described in the next section.

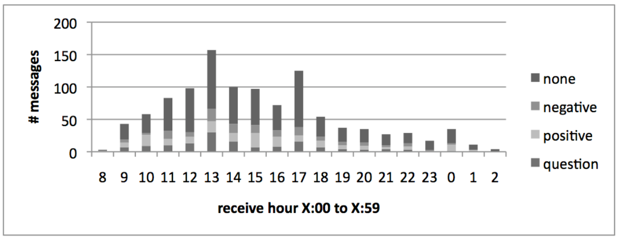

Figure 6: Distribution of received messages over the day. Peaks are between 1pm - 2pm and between 5pm - 6pm.

	time to check in sec.			
msg type	question	positive	negative	none
mean	8.03	5.87	6.95	11.50
std_dev	20.15	15.19	17.29	21.22
n	138	165	124	658
F-Value=4.979, df= 3, 1081 (p < 0.01)				

Figure 7: ANOVA of receive-to-read time based on message intention.

		week		F-Value	df	p
		1*	2*			
time to check for each message type	question	5.868	9.376	0.99	1, 136	0.322
	positive	2.391	8.365	6.415	1, 163	0.012
	negative	7.442	6.597	0.072	1, 122	0.789
	none	10.15	12.44	1.857	1, 656	0.173

Figure 8: ANOVA of receive-to-read time based on weeks.

message-type	Cluster								sum
	Cluster-1		Cluster-2		Cluster-3		Cluster-4		
question	83	93.3%	55	22.0%	0	0.0%	0	0.0%	**138**
positive	4	4.5%	161	64.4%	0	0.0%	0	0.0%	**165**
negative	2	2.2%	34	13.6%	88	100.0%	0	0.0%	**124**
none	0	0.0%	0	0.0%	0	0.0%	658	100.0%	**658**
sum	89	100.0%	250	100.0%	88	100.0%	658	100.0%	**1085**
#, % corr.	83	93.3%	161	64.4%	88	100.0%	658	100.0%	**990**

Figure 9: Clustering of Message Types.

6.4 Sample

6.4.1 Demography

We recruited 14 participants for the study from our courses, from forums, and via Facebook. Participants were mainly students (2 employees), making the sample rather homogenous but representing a main target group for such an application.

6.4.2 Questionnaire

Twelve participants had on average more than 10 text-based Skype conversations per week. Their most important conversation partners were friends (72%), colleagues (61%), partners (40%), and family members (40%). The most important situations in which they used Skype were after work (80%), on weekends (75%), and also during work (60%). The main purpose of the conversations included side conversations (80%), discussion of complex problems (70%), and dating (50%). Our participants used Skype to a large extend for short conversations, e.g., to schedule the time to go for lunch together (61%).

6.4.3 Log File Analysis

Each user received on average 20.47 messages per day. Out of the 1085 messages, 658 contained no keywords, 124 were negative, 165 were positive, and 138 were questions. The most common keywords were "yes" (153), "?" (113), "not" (89), "where" (50), ":)" or ":-)" (35) and ";)" or ";-)" (31). The average number of keywords per sonified message was 1.54 (overall mean: 0.61). Figure 6 depicts the distribution of the received messages on a daily basis (data aggregated over 14 days). Peaks can be found around lunchtime and in the after hours.

6.5 Results

In the following we analyzed the results of the study in order to obtain qualitative and quantitative data on (1) changes in the user behavior based on the sonification and (2) the understandability and learnability of our sonification algorithm.

6.5.1 Messaging Behavior

To assess the effect on the users' message checking behavior, we compared the receive-to-read time for different aspects.

First, we compared the *receive-to-read time based on the message intentions*. The results in Figure 7 show that the users checked non-sonified messages most slowly. From the sonified messages, questions required the most time until they were checked. Positive messages were checked faster than negative messages. The results are significant ($p<0.01$). Second, we analyzed *differences between week 1 and 2* of the study. We found out that for all message types the mean time increased. We believe that this is the result of a curiosity effect (people got more used to the sonification in week 2). However, the increase in time is only significant for positive messages ($p<0.05$).

Subjective user feedback from the questionnaire revealed that surprisingly the perceived receive-to-read time decreased between week 1 and 2. We assume that this effect of "false perception" is a result of the users' adaption to the system.

6.5.2 Understandability / Learnability

We used a two-step algorithm for clustering the messages according to the combination of occurred keywords in order to verify, whether our separation of message types was distinct. We use the two-step algorithm for its good performance with discrete values. As can be seen in Figure 9, the results are 4 clusters (columns), which exactly fit the self-chosen classification algorithm (rows). 91.2% of all messages were clustered like our algorithm would have done it. According to [9] the cluster solution can be considered to be "good" (separation accuracy: 0.7). The visualization of the results shows that only cluster 2 (positive messages) lacks precision beyond 90%. This is explained by the fact that messages containing positive keywords are very likely to also contain other keywords.

Results from the online survey show that the understandability of the different message types increased between week 1 and 2 (results based on 5-Point Likert scale, 1=not understandable at all, 5=very understandable). Participants could recognize questions best (mean=3.7, increase=14.6%), followed by positive messages (mean=3.2, increase=28.9%) and negative messages (mean=2.8, increase=20.8%). The results are not significant.

6.5.3 System Usability Scale (SUS)

In both questionnaires we asked the participants to fill in an SUS [3]. The score for week 1 was 72.72 (sd=12.5), for week 2 82.65 (sd=7.1). Considering the small sample, it is rather unlikely that differences are the result of a random effect ($p=0.108$). We evaluated the reliability of the scale using the Cronbach's Alpha measure, representing the inter-item correlation of all answers for each questionnaire (high values represent a consistent opinion of the participants over all questions). The alpha values were 0.544 for week 1 and 0.539 for week 2 (a minimum value of 0.5 is required, to justify the combination of all answers to an index). The results certainly don't provide empirical evidence, yet they give some good indications.

6.6 Discussion

The study reveals a significant difference in the receive-to-read time not only between sonified and non-sonified messages but also among the different message types. Interestingly, time is maximal for questions. We assume that users tend to finish their task at hand before checking a message containing a question since such messages require a certain level of attention and an effort to reply. In contrast, messages containing positive or negative news seem to be more interesting to users so that they want to check immediately. The results from the questionnaire are not significant – however there is a strong indication that the users were more comfortable with the sonification in the second week of the study.

7. ACCEPTANCE AND PRIVACY

From the online survey we found that participants considered it to be essential that the sonification matches the intention of the user: even though 71.4% of the participants liked the overall idea only 19.7% of the participants would use it if the intention was not obvious. Semi-structured interviews after the field study revealed that prevalent, user-centered requirements are aesthetically pleasing sounds and customizability.

From a privacy point of view we discovered that only 19.0% of the survey's participants were afraid of a very strong or strong influence on the privacy considering that the content of a message might be understandable to other persons familiar with the encoding (5-Point Likert scale, 1=very strong influence, 5=no influence at all). In contrast 75.2% of the users felt that reading out the message loudly would strongly influence their privacy.

8. CONCLUSION

In this paper we showed how text messages can be mapped to a meaningful and euphonic melody and how this impacts on the users' text messaging behavior. Firstly, we reported on the results of an online study with 69 participants verifying the feasibility and validity of our approach. Secondly, we presented skypeMelody, a prototype implementation to evaluate impacts on the real user behavior and discussed results from a two-week field study among 14 participants.

The main findings of both studies are as follows: (1) There is no significant difference between musicians and non-musicians when it comes to understanding the intention of a sonified message (*H2*). However, it is crucial to use intuitive and easy-to-distinguish musical elements (chords, keys). (2) The sonification has a significant influence on users' message checking behavior. Users check positive and negative messages fastest, questions are checked later (probably after finishing their current task). (3) A cluster analysis showed that our algorithm produces good results regarding the mapping of a message's intention to a musical representation. However, even though our results indicate a good understandability, we cannot yet provide empirical evidence on if and how easy users could learn to understand the precise wording of a text message.

For future work we plan to extend our approach to other means of communication (e.g., email, SMS). Though this might seem to be a straightforward approach, we believe that the properties of different ways of communication might have a strong influence on acceptance and usage, which requires further investigation. The asynchronous character of SMS or the fact that users carry their mobile phones at any time might be important factors. Further work may include the application in psychoanalysis and as a tool for visually impaired users.

9. REFERENCES

[1] Avrahami, D. and Hudson, S.E. QnA: augmenting an instant messaging client to balance user responsiveness and performance. CSCW (2004), 515–518.

[2] Brewster, S., Wright, P. and Edwards, A. 1993. An evaluation of earcons for use in auditory human-computer interfaces. In Proc. of CHI '93, NY, 222-227.

[3] Brooke, J. 1996. SUS: A 'quick and dirty' usability scale. In Jordan, P.W. Usability Evaluation in Industry, pp. 189-194. London: Taylor & Francis.

[4] Brown, L.M., Sellen, A., Krishna, R. and Harper, R. Exploring the potential of audio-tactile messaging for remote interpersonal communication. In Proc. of CHI 2009, Boston, USA, 1527–1530.

[5] Gaver, W. 1993. Synthesizing Auditory Icons. In Proc. of CHI '93, ACM, NY, 228-235.

[6] Hankinson, J. C. and Edwards, A. D. 1999. Designing earcons with musical grammars. SIGCAPH Comput. Phys. Handicap., 65, 16-20.

[7] Hansen, M. H. and Rubin, B. 2001. Babble Online: Applying Statistics and Design to Sonify the Internet. In Proc. of ICAD 2001, Helsinki, Finland, pp. 10-15.

[8] Isaacs, E., Walendowski, A. and Ranganthan, D. 2002. Hubbub: a sound-enhanced mobile instant messenger that supports awareness and opportunistic interactions. In Proc. of CHI '02. ACM, New York, 179-186.

[9] Joseph, A.J. and Lodha, S.K. 2002. MUSART: Musical Audio Transfer Function Real-Time Toolkit. In Proc. of ICAD 2002. Kyoto, Japan.

[10] Kaufman, L. and Rousseeuw, P. 1990. Finding groups in data. an introduction to cluster analysis. New York: Wiley.

[11] Petrucci, L. S., Harth, E., Roth, P., Assimacopoulos, A. and Pun, T. WebSound: a generic Web sonification tool, and its application to an auditory Web browser for blind and visually impaired users, ICAD 2000, Atlanta, US, 6-9.

[12] Powell, J. and Dibben, N. 2005. Key-mood association: A self- perpetuating myth. In: Musicae Scientiae, VOL 9; NUMB 2, Belgium, 289-312.

[13] Sahami Shirazi, A., Alt, F., Schmidt, A., Sarjanoja, A., Hynninen, L., Häkkilä, J. and Holleis, P. 2009. Emotion sharing via self-composed melodies on mobile phones. In Proc. of MobileHCI'09, Bonn, Germany, p. 222-225.

[14] Sahami Shirazi, A., Sarjanoja, A., Alt, F., Schmidt, A. and Häkkilä, J. 2010. Understanding the Impact of Abstracted Audio Preview of SMS. In Proc. of CHI 2010, Atlanta, US

[15] Song, H. and Beilharz, K. 2008. Aesthetic and auditory enhancements for multi-stream information sonification. In Proc. of DIMEA '08, vol. 349, NY, 224-231.

[16] Walker, B. and Cothran, J. 2003. Sonification sandbox: a graphical toolkit for auditory graphs. In Proc. of ICAD'03.

[17] Walker, B., Lindsay, J. and Godfrey, J. 2004. The Audio Abacus: Representing a wide range of values with accuracy and precision. In Proc. of ICAD2004, Sydney, Australia.