Jeffrey N. Howard and Darcee L. Datteri

Wichita State University, Wichita, KS, USA

MELODIC TEMPORAL PATTERN AND RHYTHMIC SYLLABLE PATTERN OF SONG TITLE IN UNFAMILIAR MUSIC

5.1. Introduction

Quite frequently we find ourselves singing along with a favorite song – often without awareness of the potential relationship between the melody we sing, and the language we use to express the lyrics. This study looks to further understanding on this relationship by looking specifically at melodic temporal patterns in instrumental music and the syllable patterns of song titles. Essentially, this investigation is an attempt at empirical verification of a song-writing technique commonly known as a 'hook'. A hook can be defined as a definitive, repetitive, or dynamically obvious portion of a song or melody designed to stick in the mind of the listener, thus creating an anchor point in the listeners' mind that defines the song or aids in recognition. An example of a rather famous 'hook' is the guitar part starting out on the Beatles song "Day Tripper". This guitar part hook repeats on occasion within the song and creates a strong defining characteristic of the song's identity.

Additionally, song hooks come in other forms, such as the repeated mention of the song title in the chorus of the song as the tune progresses. Interestingly, matching a song title to a melody portion of the song chorus causes this song-title lyric and the melody to have the same rhythmic or 'temporal pattern,' as it is known in the field of music cognition. By applying the definition of the aforementioned song chorus hook, one could assume that since the melody and the song title lyric within the chorus follow the same pattern, perhaps the same area of the brain processes them. However, if this were true, and if one were to remove the song-title lyric from the melody, and simply make it available as the title of the song, would such a connection remain? Is there still some remaining connection at an underlying level, due to the identically matching chorus melody line pattern, and the syllable pattern of the song title? This, in fact, was the research question for this investigation.

Although it would seem obvious that lyrics set to a melody within a song will be processed simultaneously, it is not necessarily so unless one can show how to measure them *separately and independently*, and arrive at an empirical corroborating result.

The idea of matching lyrics to melodies is not a new one and most likely has existed for hundreds, if not thousands of years. Only recently has the field of music cognition begun to delve deeper into the world of cognitive/music processes of the brain to better understand the relationship of each to the other. Much of the current research centers upon timbre and expectancy, rhythm and expectancy, and pitch variation studies. There is also a strong focus on language and music in the form of 'musical syntax' and perhaps most exciting, is the ongoing search for a system of musically syntactic relationships (rules) that vary amongst styles of music much in the way syntax varies in differing languages.

5.1.1. Psychology, Music, and Language

From an historical perspective, research in music psychology has focused on two specific areas. The first, in the tradition of Helmholtz (1954), places its primary emphasis upon perception as related to musical pitch, and a second in the tradition of Seashore (1938) that places emphasis on performance (Krumhansl, 1991). This inquiry will focus on the first of these two lines of research, via perception as related to melodic temporal patterning, and the metered syllable structure of language as married to lyrical content.

The lyric/melodic combination is known as *linguistic/musical prosody* and is defined by Ladd and Cutler (1983) as the sound patterns of language above the level of the individual phonemes, and is examined through acoustic variables such as frequency changes across an utterance (Palmer & Kelly, 1992). An additional definition of linguistic or musical prosody is cited by Gerard and Auxiette (1992), in accordance to the *Dictionnaire de linguistique Larousse* (Dubois *et al.*, 1973), defining musical prosody as the superimposition of music over words, (or words over music) submitted to rules mainly based on the correspondence between strong musical beats and weak syllables. In its direct and parallel relation to music, rhythm of discourse is created by the succession of consonants and vowels, the alteration of short and long, weak and strong syllables, the distribution of accents, and the number, duration, and location of pauses (Gerard & Auxiette, 1992). As offered by Gfeller (1990), the organization of musical materials is linked through repetition to a referential image, and over repeated encounters, connotations become habitual and automatic. Gfeller (1990) indicates this is known as, according to Radocy and Boyle (1979), the "Darling, they're playing our song" phenomenon in which a particular selection or style of music, through classical conditioning, becomes associated with a particular feeling.

According to Krumhansl (1991) and Palmer and Krumhansl (1987), the current body of modern research most generally reflects the theories of Helmholtz and attempts to answer questions such as whether musical phrase structure can be described in terms of tonal and rhythmic hierarchies arising from the pitch and temporal information. There is diversity of directions in this modern research literature addressing issues such as: Recognition of melody as a function of melodic contour (i.e., the pattern of ups and downs from note to note) and pitch intervals, and the relationship to familiarity with the musical stimuli (Dowling & Bartlett, 1981); tonal structures in perception and memory (Krumhansl, 1991); rhythm and pitch in music cognition (Krumhansl, 2000); recognition of melody,

harmonic accompaniment, and instrumentation by musicians and non-musicians (Wolpert, 1990); dynamic interaction between the musical event and the listener's knowledge of the underlying regularities in tonal music (Krumhansl & Castellano, 1983); cerebral dominance in musicians and non-musicians (Bever & Chiarello, 1974). By no means is this list a defining and limiting factor on lines of research combining psychology and music – that is, a full review of the scope of this research field would entail a complete literary work in and of itself (see, for example, Deutsch, 1999).

5.1.2. Hypotheses and Pre-Hypothesis

The foundation of this study is predicated upon the idea that song title/lyrics ('linguistic prosody' or 'metrical structure of language' as it is known) and melody will show a commonality in processing when presented independently. Thus, the first hypothesis derived for this study theorizes that a specific available song *title* will be chosen as "more representative" of a song when that title contains a syllable pattern that matches melody pattern occurrences within the chorus of the song. Secondly, it is hypothesized that song titles chosen which meet the criterion of rhythmic syllable pattern matching the melodic temporal pattern of the chorus of the song, will garner higher 'representativeness' ratings than chosen titles that do not meet the criterion.

5.2.1. Method: Participants

Eighty-three student participants (11 males, 72 females) at Wichita State University were utilized for this experiment. They were recruited from psychology courses as volunteers or were given class credit for participation. The average participant age was 26.2 years. Additionally, participants were assigned to musician and non-musician categories, post-experimentally, via gathering of personal musical data. Musicians were defined as those persons having had 24 months or more of formal music training inclusive of all formal and informal educational level school courses, or participation in music professionally for pay or remuneration, for an equivalent 24 month time period. Training criteria was inclusive of both voice and instruments.

5.2.2. Development of Materials and Apparatus

Original stereo channel musical selections were composed using a MIDI (Musical Instrument Digital Interface) system powered by Cakewalk Pro-Audio 9 music sequencing software. This software is specialized in recording keyboard input from MIDI based synthesizer/piano type keyboards, and storing it as timed data representations that identically emulate the keystroke data of all parts as originally played. The sequencing software was used to create 16 individual instrumental parts (tracks) for each of two versions of an original smooth jazz composition. These compositions were recorded and then converted into an MP3 file that allows audio wave data to be stored more efficiently using a data compression format. Each MP3 file version of the song was four minutes and twenty seconds in length.

The music was composed using multi-timbral synthesizer modules driven by the sequencing program on a Gateway Performance 450 MHz Pentium III computer. Sound mixing was done with a Mackie 1202 –VLZ 12 channel mic/line mixer with the sampling rate for all recording at a 44.1 kHz 320 kbps bit rate. Both versions of the composition were completely identical except for the independent variable, the temporal pattern of the melody within the chorus of each version (see Figure 1).

Figure 1 Version I vs. Version II melodic temporal patterns as varied within the chorus of the song.

This was achieved by writing only one version of the song and then adding the alternate version of the melody on a separate channel within the software environment. This would allow exact identical copies of the song to be recorded by simply muting one of the chorus-melody self-contained channels at a time, while allowing the other chorus-melody to play during the recording/mastering process. Additionally, within the software environment it was possible to adjust all parameters of musical dynamics including tempo, note on and note off, keystroke velocities, modulation, MIDI controller messages, timbre, and pitch information. This insured that all parameters other than temporal pattern were held constant outside of the intended independent manipulation of the chorus melody pattern. Thus the musical stimuli consisted of two identical compositions in all aspects of musical dynamics, save for the independent variable. The two compositions were composed as original musical pieces in order to control for familiarity with the music, and to add a highly complex and more realistic range of simultaneous musical stimuli to compete with the independent variable, and to facilitate support of ecological validity.

Participants listened to the compositions on closed-ear Aiwa HP-X222 brand lab headphones, as played from MP3 audio file format on a Dell Dimension 1GHZ Pentium-3 computer with 256MB RAM memory operating under Microsoft Windows XP-Pro operating system. All compositions were presented within the framework of a Microsoft Visual Basic Program specifically written to encompass demographic information gathering, stimulus presentation, and debriefing procedures. The music files themselves

were actually played by the Visual Basic 6.0 Windows Media Player object via software code command within the active program environment. This MP3 file/computer software playback arrangement removed reliance on physical transport mechanisms of variability such as mechanically driven CD players, and provided a more stable and less cumbersome non-mechanical playback method. This method also allowed for the simultaneous ease of presenting informed consent, instructions, tasks, rating and scoring, and debriefing from the computer screen. Additionally, the computer facilitated greater control of which composition the subject was to listen to, while avoiding burdensome tasks such as switching between CD disc trays, or removing and reinserting discs.

The song titles themselves were constructed from originally devised vowel-consonant-vowel nonsense syllable combinations (e.g., Ebbinghaus, 1913), one of which was developed with a syllable pattern to match the melodic temporal pattern that played within the chorus of the each song version (Table 1).

Table 1
Nonsense syllable song titles

Rik dif sar dek waf sic tum def rop dum tor
Gak tef vam dok mig bur reg suf
Lin nif dir cux fug miv bef sij pos
Pir tos gef

5.2.3. Procedure

Participants were given the required informed consent and instructed that they were going to participate in a music selection and rating task. They were seated at individual computer PC lab stations and instructed to read and follow the instructions on the computer screen. Participants were instructed to listen closely to the musical selections while reviewing the song title selections as provided within their assigned experimental condition. Participants were then asked to put on the headphones to begin the task and shown how to adjust the headphone volume control for individual listening comfort. All participants listened to their assigned song version once each for their respective experimental condition. To control for confounds due to experimenter bias toward musicians, participants confidentially completed a questionnaire within the software environment, regarding their musical proficiency and specifically polling them on their experience as a musician. This questionnaire was utilized to determine their placement regarding musician or non-musician status for data analysis. All participants were randomly assigned to listen to either version I or version II of the song.

The experiment consisted of a title-matching task. Each version of the song (version I and version II) was presented with four title choices. One title meeting the defined matching criteria was available for both versions I and II. The remaining three song title selections did not have this matching characteristic. Participants selected one song title from the four choices that they thought best represented the song they listened to. After selecting their one song title, participants rated the title they had chosen for its level of "representativeness" to the song using Likert scale ratings from one to ten (1 = 'Not very

representative' and 10 = 'Highly representative'). Additionally, the order of the titles as presented to participants on the computer screen was randomized by the program by presenting one of four possible list orderings of the titles. Debriefing statements, in accordance with the true nature and intent of the experiment, were provided to participants following the completion of all tasks.

5.3. Results

Descriptive statistics pertinent to the total overall representativeness ratings of the titles as rendered by participants, and years of musical experience accumulated, are presented in Table 2. The operational definition of 'musician' vs. 'non-musician', and group assignment as such, were derived from this 'musical experience' category data. Being that musical experience was reported by participants in whole number amounts, and the mean years of musical experience was fractional, it was decided that 14 years of accumulated musical experience or greater across all categories would operationally define a participant as a member of the 'musician' group.

Table 2
Descriptive Statistics

	N	Min	Max	Mean	SD
YRS Musical Experience	83	0	66	13.23	15.50
Overall 'RATINGS'	83	1	10	6.11	2.33

Table 3 shows the results of participants' song title choice with respect to the song version they were randomly assigned to listen to (version I or version II). Incorrect song titles were collapsed down to constitute the choice titled as "other". A 2 x 2 Chi square analysis, $\chi^2(1, N = 83) = 3.354$, $p = .067$, indicated a marginally significant difference between versions regarding correct or incorrect title choice. Due to collapsing of categories, probability favors "other" as a choice by three to one. This stringent analysis structure and result favors titles not containing the matching condition and the results do not offer support for the hypothesis that the matching condition would elicit more choices for the song title/melodic temporal pattern matching structure. Actual dispersement of title choices by all participants, according to the randomly assigned version they listened to, can also be seen in Table 4.

Table 3
Correct/Incorrect title choice decision given song version

CHOICE	VERSION I	VERSION II	Total
Total	42	41	83
Correct	7	14	21
Other	35	27	62

Table 4
Actual titles and frequency of titles chosen by song version

Song title:	Version I	Version II
**Rik dif sar dek waf sic tum def rop dum tor	12	14
*Gak tef vam dok mig bur reg suf	7	9
Lin nif dir cux fug miv bef sij pos	12	8
Pir tos gef	11	10
Totals	42	41

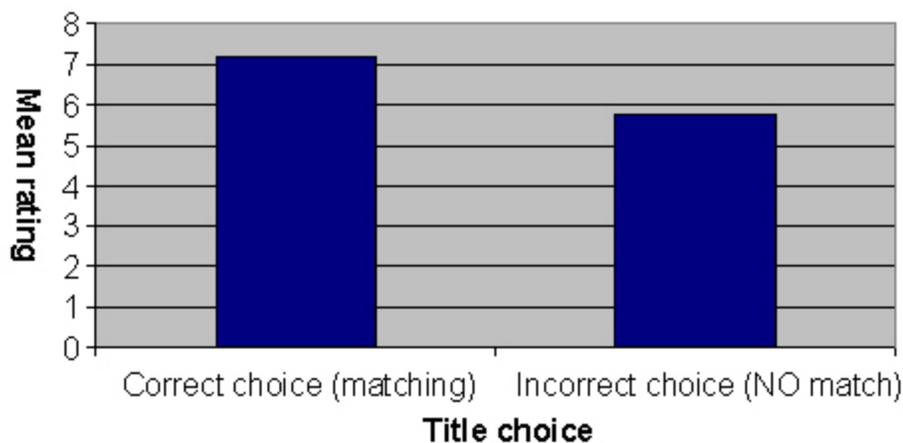*correct title choice when listening to song Version I
**correct title choice when listening to song Version II

Table 5 presents participants' choice of song title with respect to category membership defined as "musician" or "non-musician," as derived from post-experiment questionnaire data. Chi square analysis indicated no significant difference, $\chi^2(1, N = 83) = .323$, $p = .570$, with respect to participant choice by category status.

Table 5
Song title choice made by Musician status

CHOICE	Musicians	Non-musicians	Total
Correct	7	14	21
Other	25	37	62
Total	32	51	83

Additionally, the data were analyzed for potential support of the hypothesis that titles chosen meeting the criterion of melodic temporal pattern matching syllable pattern of song title (correct choice), will elicit higher "representativeness" ratings from participants (Figure 2). An independent samples t-test, $t(81) = 2.54$, $p = .013$, did support this hypothesis. Also, it must be added that although equivalency of group representation is absent, equality of variances was present within this analysis as per Levene's equality of variance test, $F = .190$, $p = .664$.

Figure 2 Descriptive statistics for ratings by 'correctness' of choice.

An independent samples t-test was also conducted comparing the mean ratings between song version I and song version II participants, t (81) = -1.769, p = .081, which indicated a marginal significance just below Alpha level .10. Descriptive statistics regarding this marginal outcome can be seen in Figure 3.
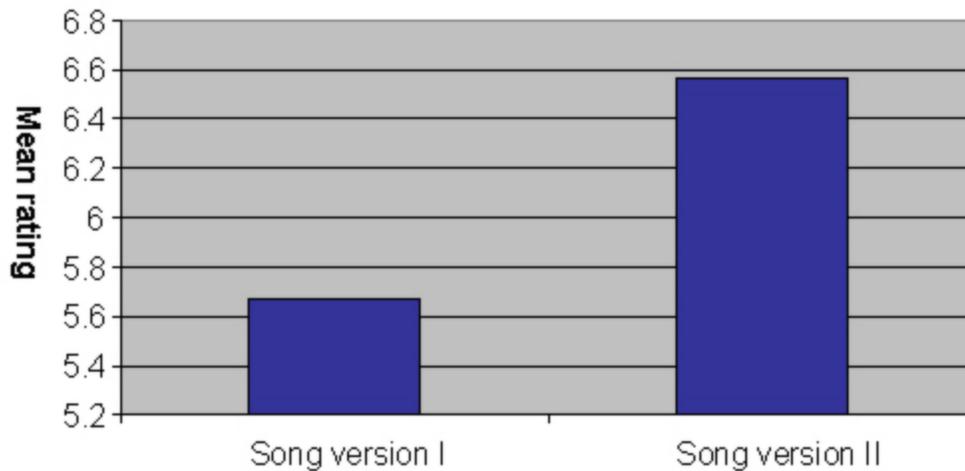

Figure 3 Descriptive statistics for ratings for each song version.

Additionally, it was of post hoc interest as to whether or not acquired dataset variables could have predictive power regarding group classification. Upon examination of the data, it was determined to use the variables of rating, age, musicianship status, and version of song listened to, to determine if these variables could effectively predict whether or not a case could be accurately classified as a predicted member of a 'correct choice' or 'incorrect choice' group. Discriminant analysis (DA) was employed to answer this question, $?^2$ (4, N = 83) = 9.323, p = .054, Wilks' Lambda = .889, which indicated near significance, and that nearly 89% of the variability could not be explained by group differences on the 4 variables of the analysis. The structure matrix result of the DA can be seen in Table 6. It is important to note that the first two variables in the structure matrix are the most important and are likely the only two which can reliably discriminate among groups (Tabachnik & Fidell, 2001).

Table 6
Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions; Variables ordered by absolute size of correlation within function.

	Function 1
RATING	.798
VERSION	.580
MUSICIAN	.177
AGE	-.121

Classification results from the DA indicate that the model correctly classified 63.9% of the original grouped cases using the 4-variable model. This indicates that the four variables of rating, song version listened to, age, and musicianship status, appear to have a reasonable ability to predict group classification/membership regarding 'correct' choice vs. 'incorrect' choice (only the first two of which can be deemed reliable). However, the discriminant analysis test in general along with the t-test comparison of ratings between assigned song versions, were derived post-hoc as per consultation of the data set, and although the discriminant analysis is nearly significant below the alpha .05 level, the credibility of this test is weakened due to post-hoc hypothesis derivation.

5.4. Discussion

The intent of this study was to indicate support for commonality in processing of song lyrics/'linguistic prosody' (i.e., metrical structure of language) with melodic temporal pattern, via language/music processing areas of the brain, and that song lyrics and melody can be shown to retain some commonality in processing when presented independently. The result from the Chi square analysis was not significant in revealing support for this, thus leaving the investigative pre-hypothesis that song title/lyrics and melody will retain equivalency in processing when presented independently, as unfounded. However, the hypothesis regarding higher ratings being rendered by those who chose the song title that included the matching syllable pattern, was in fact supported by the t-test analysis. This indicates that at the least, there is some elevated confidence level on the part of those choosing correctly, that surpasses the confidence level of those who chose the incorrect title. If one looks at the musicianship status of those making correct versus incorrect choices, musicians only chose correctly half as much as non-musicians, but still rated their choices higher. This may be due to the fact that musical training elicits elevated confidence in one's ability to accurately perform in a music related task. One would have to think this to be a relatively normal occurrence as one would likely expect, for example, that a trained juggler would rate themselves higher on a task involving juggling motor skills, than would a person with less or no juggling experience at all.

The DA provided a surprising verification that perhaps there was something between the two versions of the songs that indeed was worthy of differential responding on the part of participants. A structure matrix as generated by discriminant analysis will generally account for the majority of its variability within the listing of the first two functions of the matrix, with variables subsequent to the first two variables listed accounting for a minimal amount of variability. The ratings as rendered by participants accounted for the greatest amount of variability with respect to the structure matrix of the discriminant analysis. The serendipitous finding here is that the songs differed as per the data but not as per the hypothesis. Version II of the song was for some reason garnering higher ratings exclusive of the hypothesized matching temporal patterns of song title and melodic line. Additionally, 89% of the variability in the DA was not explained by differences between groups – alluding to version differences that were perhaps perceptively different albeit inconsistent with the hypothesis. Thus, factors such as the ability for one melodic temporal pattern to stand out over another melodic temporal pattern may be an answer to an outcome such as this. Perhaps one melodic pattern was perceived as simply more

active or attractive based on its pitch range rather than any of its linguistic or syntactically related characteristics as related to the available title choices. It must also be emphasized that responses to available song titles constructed of nonsense syllable content, as was the case in this experiment, could perhaps differ from song title choices containing semantic content. The possibility exists that the elicitation of responses to nonsensical rhythmic content simply derives from a rhythmic recognition mechanism independent of that of actual language processing mechanisms activated by semantic conveyance.

By no means should these results be construed to indicate that there is no effect of this sort when in fact the detectability of such an effect may be subtle enough to require a much greater sample size. Additionally, perhaps there is a certain specific degree of "obviousness" that must be attained by a melodic temporal pattern so as to appeal to the mechanisms of the language processing system.

5.5. Conclusion

Music and language relationships are clearly an interconnected and interrelated phenomenon of human perceptual processing. Their seeming interconnectedness as well as their obvious interconnectedness warrants a great deal more investigation as the potential discovery of common pathways in stimulus processing always leads to greater understanding of the brain and brain functions. Future lines of research should continue to address this type of music-language link – but perhaps with larger samples allowing greater focus on uncovering what appears to be an existent but delicate and elusive relationship.

References

Bever, T., & Chiarello, R. (1974, August). Cerebral dominance in musicians and non-musicians. *Science, 185*, 537-539.

Deutsch, D. (Ed.) (1982). *The psychology of music.* New York: Academic Press.

Deutsch, D. (Ed.) (1999). *The psychology of music* (2nd ed.). San Diego: Academic Press.

Dowling, W. J., & Bartlett, J. (1981, Spring). The importance of interval information in long-term memory for melodies. *Psychomusicology*, 30-49.

Dubois, J., Giacomo, M., Guespin, L., Marcellesi, C., Marcellesi, J-B., & Mevel, J-P. (1973). *Dictionnaire de linguistique.* Paris: Larousse.

Ebbinghaus, H. (1913). Memory. A Contribution to Experimental Psychology. New York: Teachers College, Columbia University (Reprinted Bristol: Thoemmes Press, 1999).

Gerard, C., & Auxiette, C. (1992). The processing of musical prosody by musical and nonmusical children. *Music Perception, 10* (1), 93-126.

Gfeller, K. (1990). Music as communication. In Unkefer, R. (Ed.), *Music therapy in the treatment of adults with mental disorders.* New York: Schirmer Books.

Helmholtz, H. L. F. (1954). *On the sensations of tone as a physiological basis for the theory of music,* ed. & transl. A.J. Ellis, New York: Dover.

Krumhansl, C. (1991). Music psychology: Tonal structures in perception and memory. *Annual Review of Psychology, 42,* 277-303.

Krumhansl, C. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin, 126,* (1), 159-179.

Krumhansl, C., & Castellano, M. (1983). Dynamic processes in music perception. *Memory & Cognition, 11,* (4), 325-334.

Ladd, D. R. & Cutler, A. (1983). Models and measurements in the study of prosody. In Cutler & D. R. Ladd (Eds.), *Prosody: Models and measurements* (pp.1-10). Berlin: Springer-Verlag.

Palmer, C., & Kelly, M. (1992). Linguistic prosody and musical meter in song. *Journal of Memory and Language, 31,* 525-542.

Palmer, C., & Krumhansl, C. (1987a). Independent temporal and pitch structures in determination of musical phrases. *Journal of Experimental Psychology: Human*

Perception and Performance, 13, (1), 116-126.

Radocy, R., & Boyle, D. (1979). *Psychological foundations of musical behavior.* Springfield, IL: Charles C. Thomas.

Seashore, C. (1938). *Psychology of music.* New York: Dover Publications, Inc. Tabachnik, B.G., & Fidell, L.S. (2001). *Using multivariate statistics, 4th ed.* Needham Heights, MA: Allyn & Bacon.

Tabachnik, B.G., & Fidell, L.S. (2001). *Using multivariate statistics, 4th ed.*

Wolpert, R. (1990). Recognition of melody, harmonic accompaniment, and instrumentation: Musicians vs. non-musicians. *Music Perception, 8,* (1), 95-106.

Jeffrey N. Howard is a 4th-year doctoral student in Human Factors Psychology at Wichita State University, where he teaches a variety of undergraduate courses. His primary research interests are music-cognition, audiovisual perception, and cross-sensory modality investigations. He holds masters degrees in clinical and experimental psychology, with a bachelors degree in radio-television production-engineering, and has 10 + years experience in the radio-television field. As a self-taught keyboardist, guitarist, and drummer he has written, produced, and engineered his own smooth jazz CD entitled "Walking On The Moon". One of his primary directives is combining his music technology and software programming skills to create unique presentation environments to explore human processing of cross-sensory stimulus combinations. His personal web page can be visited at www.jazzjeff.com.

Darcee L. Datteri is an Assistant Professor of Psychology and the Psychology Department Undergraduate Coordinator at Wichita State University. She teaches courses in research methods, cognition, states of consciousness, and music cognition. Her primary research interests include music cognition and audiovisual perception, and has published articles in the social and cognitive psychology domains. She holds a BS in psychology from St. Ambrose University, and a masters and PhD in general experimental psychology from Texas Christian University.